ENGL LING 200 INTRODUCTION TO LINGUISTICS



Sarah Harmon
Cañada College

LibreTexts™

Cañada College

ENGL LING 200 Introduction to Linguistics

Combination of *Essentials of Linguistics* by Catherine Anderson, plus additional lectures from Sarah Harmon and other instructional videos

remixed by Sarah Harmon, Ph.D.
Adjunct Professor of Spanish and Linguistics
Cañada College

Remix date: October 2021
Latest edits: August 2022

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 06/15/2026

About This Remix

This remix is a product of the work done at LibreFest 2021 in July 2021. It aligns with ENGL/LING 200, Introduction to Linguistics: A Survey at Cañada College. These courses are articulated with the University of California at Davis and at Irvine, and eligible for IGETC 4A, as well as UC-B.

This remix includes work from a few sources. The bulk of the work is from *Essentials of Linguistics* by Catherine Anderson. I have also included a large number of my lectures on a number of topics, especially adding chapters on sociolinguistics, historical linguistics, language acquisition, and cognitive linguistics. I've included some optional lectures and discussions by other linguists who post their videos on YouTube, adding richness to this text. In this way, there are 10 chapters, covering the material that is typically taught in an introductory linguistics course at 2-year and 4-year universities in the United States.

I will be adding content along the way, so the remix will continue to be updated regularly. If there are any questions or comments, or if you have suggestions for material, please send them to me at harmons@smccd.edu.

--Sarah Harmon, Ph.D.
20 May 2022

TABLE OF CONTENTS

Licensing

0: What is this course about?

- 0.1: Linguistics is a science
- 0.2: Linguistics as a social science
- 0.3: The resources used for this course

1: Introduction- What is language?

- 1.1: What is linguistics?
- 1.2: Linguistic definitions
- 1.3: Hallmarks of human language
- 1.4: Arbitrariness and ongoing changes
- 1.5: Language and thought
- 1.6: Language universals
- 1.7: Signed languages
- 1.8: Next steps and references

2: Sounds, Part 1- Phonetics

- 2.1: Introduction to Phonetics
- 2.2: Articulators and Airstream Mechanisms
- 2.3: Voicing, Resonance and Articulation
- 2.4: IPA and Charts
- 2.5: Sound Classes and Hierarchy
- 2.6: All About Vowels
- 2.7: References

3: Sounds, Part 2- Phonology

- 3.1: Comparing Sounds and Distribution
- 3.2: Assimilation and Dissimilation
- 3.3: Other Phonological Rules
- 3.4: Analyzing Phonological Data
- 3.5: References

4: Words- Morphology

- 1: Back to the Arbitrary
- 2: Morphological Definitions
- 3: Morphemes
- 4: Affixation and Other Morphological Processes
- 5: Morpho-phonology
- 6: Parts of Speech and Word Formation
- 7: Creation of New Lexicon
- 8: Morpho-syntax
- 9: Analyzing Morphological Data
- 10: References

5: Phrases- Syntax

- 5.1: Syntax Terminology
- 5.2: Word Order and Lexical Categories
- 5.3: Phrase Structure Rules, X-Bar Theory, and Constituency
- 5.4: Analyzing a PS Tree
- 5.5: Grammatical Dependencies
- 5.6: Grammaticality
- 5.7: References

6: Meaning- Semantics and Pragmatics

- 6.1: Arbitrariness and Compositionality
- 6.2: Lexical Semantics
- 6.3: Phrasal Semantics
- 6.4: Semantic Features
- 6.5: Pragmatics Definitions
- 6.6: Grice and the Cooperative Principle
- 6.7: Reading Between the Lines
- 6.8: Deixis
- 6.9: References

7: Language and Society- Sociolinguistics

- 1: Sociolinguistics definitions
- 2: Geographic dialects
- 3: Appalachian American English (AE)
- 4: African American English (AAE)
- 5: Other Social Dialects
- 6: Gender and Class Dialects
- 7: Prestige and Politics
- 8: Social Norms
- 9: Languages in Contact
- 10: Indigenous North American Languages (Optional)
- 11: References

8: Language Change- Historical Linguistics

- 1: Language Change Facts and Definitions
- 2: Reconstructions and Analysis
- 3: Indo-European and Romance Phonological Reconstruction
- 4: Stages of English
- 5: Morphological Stages
- 6: Syntactic and Lexical Changes
- 7: Typology
- 8: Language Families
- 9: Writing Systems
- 10: Arabic and Romance- A Comparison (Optional)
- 11: References

9: Learning Languages- Language Acquisition

- 9.1: Universals and Definitions
- 9.2: Child Language Acquisition Theories
- 9.3: Child Language Acquisitions Stages

- 9.4: Child Language Acquisition Linguistics
- 9.5: Child Language Acquisition Bilingualism
- 9.6: Second Language Acquisitions Theories and Components
- 9.7: Second Language Acquisition Teaching Methods
- 9.8: References

10: Language Processing and the Human Brain- Cognitive Linguistics

- 10.1: Human Language versus Animal Communication
- 10.2: Evolution of Human Language
- 10.3: Language Processing and Parsing
- 10.4: Language Processing in the Brain
- 10.5: Aphasia
- 10.6: Autonomy of Language
- 10.7: References

Index

Glossary

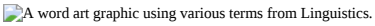
Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

0: What is this course about?



To kick things off, a brief mini-chapter on how to wrap your mind about linguistics and this course.

0.1: Linguistics is a science
0.2: Linguistics as a social science
0.3: The resources used for this course

0: What is this course about? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

0.1: Linguistics is a science

Linguistics is Science, from Catherine Anderson's *Essentials of Linguistics*.



Video Script:

Linguistics is one of those subjects that not many people have heard of, so you might well be wondering exactly what it is.

The simplest definition of Linguistics is that it's **the science of language**.

This is a simple definition but it contains some very important words. First, when we say that linguistics is a science, that doesn't mean you need a lab coat and safety goggles to do linguistics. Instead, what it means is that the way we ask questions to learn about language uses a scientific approach.

The scientific way of thinking about language involves making systematic, **empirical observations**. There's another important word: empirical means that we observe data to find the evidence for our theories.

All scientists make empirical observations: botanists observe how plants grow and reproduce. Chemists observe how substances interact with other. Linguists observe how people use their language.

A crucial thing to keep in mind is that the observations we make about language use are NOT value judgments. Lots of people in the world — like your high school English teacher, various newspaper columnists, maybe your grandparents, and maybe even some of your friends — make judgments about how people use language. But linguists don't.

A short-hand way of saying this is that linguists have a **descriptive approach** to language, not a prescriptive approach.

We describe what people do with their language, but we don't prescribe how they should or shouldn't do it.

This descriptive approach is consistent with a scientific way of thinking. Think about an entomologist who studies beetles. Imagine that scientist observes that a species of beetle eats leaves. She's not going to judge that the beetles are eating wrong, and tell them that they'd be more successful in life if only they eat the same thing as ants. No — she observes what the beetle eats and tries to figure out why: she develops a theory of why the beetle eats this plant and not that one.

In the same way, linguists observe what people say and how they say it, and come up with theories of why people say certain things or make certain sounds but not others.

In our simple definition of linguistics, there's another important word we need to focus on: linguistics is **the science of human language**.

There are plenty of species that communicate with each other in an impressive variety of ways, but in linguistics, our job is to focus on the unique system that humans use.

It turns out that humans have some important differences to all other species that make our language unique.

First, what we call the **articulatory system**: our lungs, larynx & vocal folds, and the shape of our tongue, teeth, lips, nose, all enable us to produce speech. No other species can do this in the way we can, not even our closest genetic relatives the chimpanzees, bonobos, and orangutans.

Second, our **auditory system** is special: our ears are sensitive to exactly the frequencies that are most common in human speech. There are other species that have similar patterns of auditory sensitivity, but human newborns pay special attention to human speech, even more so than synthetic speech that is matched for acoustic characteristics.

And most important of all, our **neural system** is special: no other species has a brain as complex and densely connected as ours with so many connections dedicated to producing and understanding language.

Humans' language ability is different from all other species' communication systems, and linguistics is the science that studies this unique ability.

Check Yourself

? Exercise 0.1.1

What does it mean to say that Linguistics is a science?

- The field consists of a set of true facts that can be proven objectively.
- The field uses the scientific method to determine objective rankings of language quality.
- The field uses empirical observations to develop theories of language behaviour.

Answer

"The field uses empirical observations to develop theories of language behaviour."

The reason: Linguistics is a social science, based in empirical observations.

? Exercise 0.1.2

Each of the following sentences represents something someone might say about language. Which of them illustrates a descriptive view of language?

- The use of quotative *like* in sentences such as, "She was like, I can't believe you did that!" began to enter Canadian English with the generation of speakers born in 1971.
- The song "I can't get no satisfaction" should really mean that "I can get some satisfaction" because two negatives always make a positive.
- In a phrase like, "the people who the bride invited to the wedding," it's proper to use *whom* rather than *who*.

Answer

"The use of quotative *like* in sentences such as, "She was like, I can't believe you did that!" began to enter Canadian English with the generation of speakers born in 1971."

The reason: It's an objective description of language without associating a 'prescribed' value of what is 'right' or 'proper'.

? Exercise 0.1.3

Which of the following kinds of data would a linguist be likely to observe?

- Which method is most effective to help a child stop stuttering.
- Whether Korean includes tones that change the meaning of words.
- How many undergraduates can correctly use the words *affect* and *effect* in their essays.
- If second-language speakers can pronounce English words correctly.

Answer

"Whether Korean includes tones that change the meaning of words."

The reason: This is an empirical observation of what happens in a language. The others are all subjective or prescriptive in some way.

0.1: Linguistics is a science is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.1: Linguistics is Science** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

0.2: Linguistics as a social science



Video Script

Okay, so what is linguistics what is a social science up to this point you have read and or watched Catherine Anderson's explanation. And she says that linguistics is a science and I think that's a great place to start, but it is important to note that linguistics is really a social science. Now for some of you, this might be your very first social science course so let's talk about what makes social sciences and specifically linguistics different from let's say chemistry, physics, biology, those sciences.

Now for some of you, this might be your very first social science course so let's talk about what makes social sciences and specifically linguistics different from let's say chemistry, physics, biology, those sciences. So, what is a social science well to start off let's look at what a standard definition for a social science is and I'm going to go and use the Wikipedia definition. It's a fairly straightforward definition; in most every dictionary you'll see something very similar: Social science is a branch of science devoted to the study of societies and the relationships among individuals within those societies.

There are a number of disciplines that fall under social science; linguistics is one, anthropology, sociology, some of psychology certainly falls under social science. History frequently is put into social sciences, as well as economics part of geography, certainly cultural geography. Definitely when we talk about rhetoric and communication, it frequently falls under social science. Philosophy can be considered a social science at times; it's also a humanity. Same thing with history, but they kind of border. Linguistics is solidly in the social science category because we devote our studies to aspects of the human condition, right?

Well let's take that and kind of understand it a little bit more, because there's something that makes social science a little unique compared to the physical and biological sciences. Social sciences use scientific approaches to observe and catalog human behavior. I think that's a better, more tangible definition for a social science.

So, what does it mean to have a social science mindset, and that means, what does it take to be a social scientist? Any flavor, it doesn't matter. Well, the first thing you have to do is think about descriptivism. Descriptivism is that you describe what you observe; you observe something, and you describe it objectively. You do not want to import any biases that you may have or that your discipline may have with respect to that observation or analysis; you really want to keep it straightforward. You want to think critically about what you are analyzing, and you want to use those tools that we have to analyze the aspects of humanity that interests you. It doesn't matter if we're talking language, or societies and how they form, or the history of these groups or how they work their fun, financial or monetary areas. What is really important, though, is that this critical analysis is not just linear. It is creative; you think outside the box. Of course, you have to know what is inside the box first before you can think outside the box and that's what it means to approach things with a scientific mindset in general, but especially with respect to human or social settings.

The last thing that's really important, and I would say is the main thing that separates social sciences from the behavioral as well as the physical and biological sciences, is that in the social sciences, we use inductive reasoning. This means that we use the data to analyze the behavior or phenomena we look at a to analyze be and then we look at be to analyze see. This is the opposite of the physical biological and behavioral sciences, they tend to use deductive reasoning, they go from the general to the specific we tend to go, specific and then pull out to the general. Now that doesn't make one approach better or worse than the other, it just is a difference.

So, now that we know what a social science is let's think about what is linguistics. If you look it up in a dictionary or anywhere, you'll basically get a definition that says, linguistics is the scientific study of language. Certainly, that is true, but what does that mean? To unpack it a little bit I would argue, and I would think most of my fellow linguists would also agree, that really what we're trying to say is that linguistics is the analysis of language and all of its components. And when we say that we're not just talking about the words or even the sounds or even meaning. We're looking at how they all combined together, because we do not use just a single lexicon by itself a single word by itself; we use it in combination with something else a larger phrase within a larger context. That means that we are analyzing how human beings use the tool of language, as well as the actual language itself. It's like if you were describing how a carpenter uses a hammer. You don't just describe the hammer you describe how the carpenter uses it as well.

So if that's what linguistics is what isn't linguistics and what frequently gets conflated with linguistics and it's important to note that these are things these concepts that I'm bringing up or studies. They have tangible routes to linguistics, but they are not exclusively linguistics. The first one that people often conflate with linguistics is etymology or the history or derivation of a given term or lexicon or word. We're going to use this terminology interchangeably later down the road. As linguists, we do not solely study the history of a word and how it got formed in a language. It certainly uses some of what we use in linguistics, and when we get to morphology in particular will really kind of unpack that. And even for somebody like myself, I am an historical linguist. I am going to look at the etymology of a word, but that's not all that I'm doing. I'm looking at the context; I'm looking at how it changes or doesn't change over time. Etymology is just a very small sliver of what some people might do, but not all. The other thing is linguistics is not a study of dictionaries or study of grammar, although certainly the folks who work on dictionaries and grammars are frequently linguists. But that is only a small part of what they are doing. Dictionaries and grammars are a type of catalog as to what a language does. But not entirely; it's a fixed document as it were, it's not a living language, which is what linguistics is all about.

As we go through any social science course, but especially in linguistics course, there are certain things that all students must keep in mind. They must actively do these throughout the course in order to succeed. In this case, success is not just getting an A; success is understanding the material and applying it to a future use. The most crucial thing I would argue, is that you have to keep an open mind. That's a phrase that gets thrown around quite a bit, but in the social sciences, it is imperative to do that, you must be objective. You must be descriptive, and that means you have to keep an open mind. You must observe what it is that you see and think about all of the possibilities that could be connected with that not just one or two possibilities. The last piece is to see patterns, that is what we human beings do very well, we see patterns, we find patterns and behaviors in colors in anything and then that is how we remember them. Well, language is no different; language has patterns and it's really important to find the pattern whenever you're analyzing language.

Finally, let's talk about some of the resources that we're going to use now to start off, probably the primary and most important resource here. The textbook that we're going to be using is *Essentials of Linguistics* by Catherine Anderson. It originally was published through eCampus Ontario on Pressbooks, but it has also been imported into LibreTexts. It is an open educational resource, meaning that it is free to use, and it is online, that means anyone can use it, and anyone can view it. You don't have to pay anything for it; don't worry, Catherine Anderson most probably got some compensation for her work, that could have come via sabbatical or grant, or it could have come through a number of arenas. Catherine Anderson is a linguist and a professor at McMaster University, which is in Hamilton, Ontario, Canada. Her main specialties are language acquisition: Child language acquisition, but even second language acquisition. She also has a keen interest in how bilingualism comes into play. She does a number of courses at McMaster University and then developed this OER, open educational resource, as a way to have a proper textbook based off of her lectures. Her work is absolutely outstanding and I would not use it if I didn't think it was up to par.

But there are supporting traditional textbooks that I have used over the years that others have used over the years and I can't say that any of my lectures would be complete without mentioning that these textbooks have been used at some point in my career. The main one is introduction to language by Robert Rodman, Victoria Fromkin and Nina Hyams; it is the gold standard with respect to textbooks. It is also very expensive, just like gold is, which is why I don't use it so much. *Language Files* is something I do use.

Language Files is a collection of some lecture material, but most importantly exercises and in fact the exercises that we're going to use in this class for homework do come from *Language Files.* You do not need to purchase it. *Language Files* is produced by the Linguistics Department at the Ohio State University in Columbus, Ohio. And then, of course, the other one that is frequently used the other textbook is *How Languages Work: An Introduction to Language and Linguistics* and it is an edited volume by Carol Genetti. A very good collection and, in some cases when I needed to get a better definition for something I would use one of these three textbooks.

That being said, we are not using those textbooks. The information that is in those textbooks, as far as the general information, is something that is widely used but also general knowledge. I find in a lot of cases Catherine Anderson's explanations are superior; they are broken down in a way that I think are more tangible. The other reason that I want to bring these traditional textbooks up is because, if you take a linguistics course elsewhere, you might run across those and certainly in my previous courses, I used book two, if not all three of those. Throughout each chapter or module, you will see that there are a number of resources that I'm going to pull from. Some of these are resources that I have used and some of them are common across well all of these textbooks. It's important to use those resources to show you them in case you have any lingering questions or you want to find out more information about the topic, then I will have those resources available for you.

With that this is your introduction to the introduction, you are now ready to go forward and really start studying linguistics.

0.2: Linguistics as a social science is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

0.3: The resources used for this course

What are the types of resources that have been used to build this course/book?

Primary Textbook

Catherine Anderson's *Essentials of Linguistics*, 2nd edition (2024). This second edition was lead by Catherine Anderson and includes significan contributions from her team: Bronwyn Bjorkman, Derek Denis, Julianne Doner, Margaret Grant, Nathan Sanders and Ai Taniguchi.

Other Textbooks

I've taken information from a couple of other textbooks and adapted it into my lecture notes. They are commonly used textbooks--and very expensive.

- *Introduction to Language* by Robert Rodman, Victoria Fromkin, and Nina Hyams
- *Language Files* from the Ohio State University Linguistics Department
- *How Languages Work: An Introduction to Language and Linguistics*, edited by Carol Genetti

Other resources

- Various articles, as listed at the end of each chapter
- Various YouTube videos, especially from Anthony Pym (sociolinguistics)

0.3: The resources used for this course is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

1: Introduction- What is language?

1.1: What is linguistics?
1.2: Linguistic definitions
1.3: Hallmarks of human language
1.4: Arbitrariness and ongoing changes
1.5: Language and thought
1.6: Language universals
1.7: Signed languages
1.8: Next steps and references

1: Introduction- What is language? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.1: What is linguistics?

1.1.1: What is a language? from Sarah Harmon



Video Script

So, what exactly is a language, and that is what we're going to work on with this section of the text, well, you can see here in the Oxford English dictionary. Their definition of a language is a pretty straightforward concept: A system of spoken or written communication used by a particular country people community, etc., typically, consisting of words used within a regular grammatical and syntactic structure. Then you notice that it also includes information about animal communication, facial expression and body language, as well as computer language. It has a number of definitions. Even if we go to what is considered the American dictionary, the Merriam Webster, again their standard first definition is pretty similar to the Oxford English Dictionary, that it's the words, the pronunciation and the methods of combining them, used and understood by a community. Pretty interesting, but it also includes other information about other vocal emissions, shall we say, methods of communication, involving gestures and body language. It also includes animal communication; it includes computer language, and includes a number of things. But neither of these definitions quite tracts on to what a linguist considers a language to be.

For a linguist a language is a human form of communication, which includes the phonetics, the phonology, the morphology, the syntax, the semantics, and the discourse context, as presented by a given speech community. It is a living form of communication, which changes across speech communities and over time. That is our definition of a language. It may not be ***your*** first definition of a language, and perhaps your first definition of a language is going to include a lot of other things, but as we go through this chapter, as well as the course as a whole, you will start to understand what a linguist views as a language. And remember, our job as linguists, as social scientists, is to observe to describe and to analyze.

From 1.2 Mental Grammar of Anderson's *Essentials of Linguistics*



Video Script

We know now that Linguistics is the scientific study of human language. It's also important to know that linguistics is one member of the broad field that is known as **cognitive science**.

The cognitive sciences are interested in what goes in the mind. And in linguistics, we're specifically interested in how our language knowledge is represented and organized in the human mind.

Think about this: you and I both speak English. I'm speaking English right here on this video and you're listening and understanding me. Right now I've got some idea in my mind that I want to express. I'm squeezing the air out of my lungs; I'm vibrating my vocal folds, and I'm manipulating parts of my mouth to produce sounds. Those sounds are captured by a microphone and now they're playing on your computer. In response to the sound coming from your computer speaker or your headphones, your eardrums are vibrating and sending signals to your brain, with the result that the idea in your mind is something similar to the idea that was in my head when I made this video.

There must be something that your mind and my mind have in common to allow that to happen: some **shared system** that allows us to understand each other's ideas when we speak. In linguistics, we call that system the **mental grammar** and our primary goal is to find out what that shared system is like.

All speakers of all languages have a mental grammar: the shared system that lets speakers of a language understand each other. In Essentials of Linguistics we devote most of our attention to the mental grammar of English, but we'll also use our scientific tools and techniques to examine some parts of the grammars of other languages.

We'll start by looking at **sound systems**: how speakers make particular sounds and how listeners hear these sounds. If you've ever tried to learn a second language you know that the sounds in the second language are not always the same as in your first language. Linguists call the study of speech sounds **phonetics**.

Then we'll look at how the mental grammar of each language organizes sounds in the mind; this is called **phonology**.

We will examine the strategies that languages use to form meaningful words; this is called **morphology**.

Then we take a close look at the different ways that languages combine words to form phrases and sentences. The term for that is **syntax**.

We also look at how the meanings of words and sentences are organized in the mind, which linguists call **semantics**.

These five things are the core pieces of the mental grammar of any language: they're the things all speakers know about a language. All languages have phonetics, phonology, morphology, syntax and semantics in their grammars.

These five areas are also the core subfields of theoretical linguistics. Just as there are other kinds of language knowledge we have, there are other branches of the field of linguistics, and we'll take a peek at some of those other branches along the way.

Check Yourself

Exercise 1.1.1

Newspaper headlines occasionally have unexpectedly funny interpretations. One example is: *Two cars were reported stolen by the police yesterday*. Which part of your mental grammar leads to the possibility that the police could have done the stealing or the reporting in this headline?

- Phonetics.
- Phonology.
- Morphology.
- Syntax.
- Semantics.

Answer

"Syntax

The reason: There are multiple interpretations because of the prepositional phrase 'by the police'--it could describe a couple of different elements. That leads it to be a structurally ambiguity--therefore, it's syntax-related.

Exercise 1.1.2

Newfoundland English has some characteristic differences to standard Canadian English. The following sentences are grammatical in Newfoundland English: *I eats toast for breakfast every day. You knows the answer to that question*. What part of the mental grammar of Newfoundland English is different to Canadian English in these examples?

- Phonetics.
- Phonology.
- Morphology.
- Syntax.
- Semantics.

Answer

"Morphology"

The reason: The difference is the -s that is attached to the verb. That is a suffix, and so this is a morphologically-related issue.

Exercise 1.1.3

Speakers of American English often notice that Canadians' pronunciation of the words *about* and *house* has a vowel produced higher in the mouth than the American English versions of these words.

What part of the mental grammar of American English is different to Canadian English in these examples?

- Phonetics.
- Phonology.
- Morphology.
- Syntax.
- Semantics.

Answer

"Phonetics"

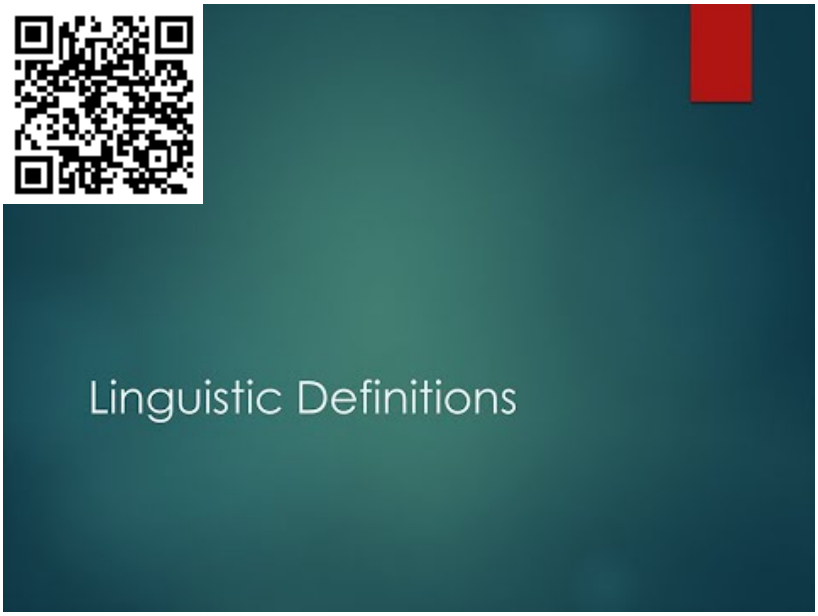
The reason: Those different sounds refer to different vowels in the phonemic inventory of the two dialects--in other words, there is a different inventory of sounds. That's a phonetic-based issue.

1.1: What is linguistics? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.2: Mental Grammar** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

1.2: Linguistic definitions

1.2.1 Linguistic Definitions, by Sarah Harmon



Video Script

Let's talk about linguistic definitions, certain definitions that we use in linguistics which may be new to you or just a little bit different than what you knew before. You already read in Catherine Anderson's text certain definitions that I'm not going to go through too much; she covered already the exact definitions of morphology, syntax, semantics, phonetics, and phonology, and certainly she went into a little bit more about the other areas of linguistics. That is something I'm not going to cover too much here. What I would like to cover more is certain basic terms that we use in linguistics and perhaps they're very different than what you knew prior to this course.

To start with, let's talk about the word 'word'. We actually don't use that term very much in linguistics; we use **lexicon** or we use **term** or **terminology**, but we certainly do not use the term 'word'. The reason is because it doesn't actually describe what happens in a given language. For example, we will see in morphology, and then again in semantics, a lot of talk about arbitrariness, how you can say a term like 're(a)d' and depending on the context that could be a verb in the past tense, like 'I read the book', or it can mean a noun or an adjective that's a color. Either way, you don't know, that so that's why the term 'word' is a little confusing.

It's also confusing because in many languages throughout both North and South America, and may Australian languages, and in a number of languages around the world, a 'word' technically has so many different components to it. It could be a full verb phrase, and it could be a full sentence even, but because it is a whole concept unit as a 'word'. That term doesn't exactly do it for us; we need something a little more specific. When we get back to morphology we'll come back to that term.

The other really important definition has to do with the concept of the language versus a dialect. These are two totally different things, and they are not at all similar to what the colloquial or every day definition is. A **language** in linguistics does not refer to anything that has to do with a political boundary; it has to do with a concept called **mutual intelligibility**, meaning Language A and Language B maybe related, but speakers of Language A cannot easily understand speakers of Language B. A great example is Spanish, Italian and French; these are all Romance languages. They are very similar in so many ways, but they certainly have differences in sound and they definitely have differences in how they construct aspects of their lexicon and in their phrases.

The other one, the one that probably frustrates most linguists, the most is when a political entity, like a government, decides that one of the languages spoken in its country is the language of the country and then calls it that. For example, in China, they only talk

about the ‘Chinese language’, which is highly inaccurate. There is no one Chinese language; there are 200 or so Chinese languages. The standard language of the government happens to be Mandarin, but if you speak Cantonese, Hunanese, Shanghainese and so many of the others, you cannot understand Mandarin alone. You need instruction in it. Therefore, there is no ‘Chinese language’ For those students who say, “Oh, I speak Chinese,” I will be correcting you every time; you speak Mandarin, you do not speak Cantonese or anything else necessarily. Also, if you go to like eastern Africa, specifically Kenya, there's Kikuyu, there's Swahili, there's a number of other languages. Those are all languages; they are related, but they're different languages.

Dialects, however, are **mutually intelligible**. A great example is, if you take somebody from California and you have them go to London, or Ontario, Canada, or go to Sydney, Australia, or Johannesburg, South Africa, or Christchurch, New Zealand. That same Californian is going to be able to understand English at all of those places. Now it may be difficult at times and there might be terms used that are not familiar, maybe pronunciations that are not familiar. But those are all different dialects of English in those regions and that's what we call mutual intelligibility.

For example, Norwegian and Swedish are technically to dialects of the same language. Even though we might now still even in linguistics call it Norwegian or Swedish, what we're really saying is: that dialect of that Northern Germanic language. There are times this gets a little hairy so, for example in the Balkans, technically speaking Serbian and Croatian are the same language; they're just two different dialects spoken by two different ethnic groups. Certainly, given the history of those two ethnic groups, along with the others in the Balkan Peninsula, it gets a little dicey sometimes for these folks to be to hear that they speak the same language. It should be noted that particularly with those two dialects, they are closer or more distant depending on the relations between those two ethnic groups; when relations are good and they're intermingling, there's a lot more similarity between those two dialects. If there's much more tension, like there has been for the last couple generations, then there are more differences between those two. To this point right now, they are still dialects of the same language.

It is really important to recognize that sometimes a person's history or geopolitics might play a factor in what they consider a different language or different dialect. But linguistically, from a social science perspective, what we observe and what we take note of and catalog is that languages are not mutually intelligible, versus dialects are.

A couple other definitions to talk about, and one has to do with grammar. What is grammar? In truth, there are three different types of grammar that we can talk about. We can talk about **descriptive grammar**, which is what we use in linguistics. Again, we describe what people do with a given language or dialect. We do not associate any values with it, meaning, we do not say that Language A or Dialect A is better than B; we don't do that. That is something that you do in **prescriptive grammar** and when I think prescriptive grammar I always think of that English teacher or language arts teacher who always told you, “You can't say this; you must say that.” Prescriptive grammar is how one should talk in a given situation, especially pushing you towards a standard and even a prestige dialect or language. Therefore, it is a prescribed value to a given dialect. Again, that is not what we do in linguistics, so my challenge to you through this course is when I asked you to talk about what you hear and what you see, is to be descriptive, to just observe and analyze and not say that one is better than another, or that you should do something, or should not do something.

Now, this is a little different than what we do when we teach a language. When we teach a language, we use what is called a **teaching grammar**. Think of any foreign language class that you may have been in. That's really a combination of a descriptive grammar and a prescriptive grammar. When we teach you a language, we want you to produce good, solid dialogue or sentences, etc, and there are certain norms that must be followed at all times. But we also describe what frequently happens, and especially in a case like an English class, a Spanish class, a French class, an Arabic class, where you have multiple dialects. Even in the case of Arabic, with multiple forms and multiple languages. When we teach you a language, we also teach you a little bit about the culture and we talk about the differences between Dialect A, Dialect B and so on and so forth.

This is the basis of a few definitions, as we go through the rest of the course, we're going to expand on what is a grammar. We're going to expand on that ‘word’ versus ‘lexicon’ element. And we're going to expand on what is a language versus a dialect and all about mutual intelligibility.

1.2: Linguistic definitions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.3: Hallmarks of human language

Creativity and Generativity in Anderson's *Essentials of Linguistics*



Video Script

Probably the most fundamental property of human language is **creativity**. When we say that human languages are creative, we don't just mean that you can use them to write beautiful poems and great works of literature.

When we say that human language is creative, we mean a couple of different things:

First, every language can express any possible concept.

That notion might surprise you at first. I often see magazine articles or blog posts that talk about supposedly untranslatable words that exist in other languages but that don't exist in English. A quick search online leads me to these gems:

Kummerspeck is the German word for excess weight gained from emotional overeating.

In Inuktitut, ***iktsuarpok*** is that feeling of anticipation when you're waiting for someone to show up at your house and you keep going outside to see if they're there yet.

And in Tagalog, ***gigil*** is the word for the urge to squeeze something that is irresistibly cute.

So if you believe that kind of article, it might seem like some concepts are restricted to certain languages. But think about it: Just because English doesn't have one single word that means "the urge to squeeze something cute" doesn't mean that English-speakers can't understand the concept of wanting to squeeze something cute. As soon as I described it using the English phrase "the urge to squeeze something cute" you understood the concept! It just takes more than one word to express it! The same is true of every language: all of the world's languages can express all concepts.

The other side of the creativity of language is even more interesting. Every language can generate an infinite number of possible new words and sentences.

Every language has a finite set of words in it. A language's vocabulary might be quite large, but it's still finite. And every language has a small and finite set of principles for combining those words.

But every language can use that finite vocabulary and that finite set of principles to generate an infinite number of sentences, new sentences every single day.

Likewise, every language has a finite set of sounds and a finite set of principles for combining those sounds. Every language can use those finite resources to generate an infinite number of possible new words in that language.

Because human languages are all capable of generating new words and generating new sentences, we say that **human grammar is generative**.

Remember that when we use the word “grammar” in linguistics, we’re talking not about the prescriptive rules that your Grade 6 teacher tried to make you follow, but about **mental grammar**, the things in our minds that all speakers of a language have in common that allow us to understand each other. Mental grammar is generative.

The final, and possibly the most important thing to know about the creativity of language is that it is governed by systematic principles. Every fluent speaker of a language uses systematic principles to combine sounds to form words and to combine words to form sentences. In *Essentials of Linguistics*, we’ll use the tools of systematic observation to discover what these systematic principles are.

Check Yourself

Exercise 1.3.1

What does it mean to say that mental grammar is generative?

- Every child inherits a mental grammar through their genes.
- Some languages can generate concepts that are impossible to express in other languages.
- The principles of mental grammar allow us to form completely novel sentences, and to understand them when we hear them.

Answer

"The principles of mental grammar allow us to form completely novel sentences, and to understand them when we hear them."

The reason: A mental grammar is just that, a concept that we have of how to put together words and phrases in a given language. Primarily, it's for our native langauge(s), but it can also include those languages in which we are fully fluent. This mental grammar helps us to both produce and comprehend langauge.

Exercise 1.3.2

The systematic principles of English phonology generate some word forms but not others. Which of the following words could be a possible word in English?

- Klaff.
- Fkal.
- Flakf.
- Sflak.

Answer

"Klaff"

The reason: When we think about possible combinations in English, we think about what could possibly be used together--what sounds could possibly work in the language. We'll get more into this in Chapters 2 and 3, but for now it's more of an issue of what could be possible. The other options are 'hard to pronounce' from an English point of view.

Exercise 1.3.3

The systematic principles of English syntax generate some sentences but not others. Which of the following sentences is not possible in English?

- You ain’t going nowhere.
- Herself have wrote these excellent book.
- Sam said she needed to speak to Chris and me.
- This expedition aims to boldly go where no expedition has gone before.

Answer

"Herself have wrote these excellent book."

The reason: There are a number of issues, both morphological and syntactical, with the sentence.

- 'Herself' is a reflexive pronoun, and has to refer back to someone--but as the subject of the sentence, that can't be done, as there's nothing before it to set up the reference.
- The verb 'have' is not a third person singular form--what we would expect for a 'She' subject.
- The past participle should be used in this verb construction, because this is a perfective form. But 'wrote' is not the past participle of 'to write'; it would be 'written'.
- 'These' implies a plural noun, but it is paired with 'book', which is singular.

Hallmarks of Human Language, from Sarah Harmon



Video Script

What is a language? For now, when we talk about a language, we're talking about a human language, and there are certain components that are found in human languages. They are consistent and they are unique, at least in their sum, that human languages all have certain components to them. With respect to animal communication, we'll cover that at a later point.

What are these attributes of human language? These hallmarks of human language? There are six in total. And it kind of goes in order, meaning that probably the most important aspect of human language is arbitrariness. **Arbitrariness** means that a given sign and a given meaning are arbitrarily connected. That means that there is no reason whatsoever that this thing is called a 'cup' or a 'mug' or a 'bottle'. In fact, the fact that we can use three different terms in mainstream American English for just the same thing, shows the arbitrariness. There is no reason whatsoever that this thing is called a 'telephone'; it just happens to be called telephone. Now I would also argue that arbitrariness may be the hardest of the areas to understand; in fact, it's the next section of this chapter is devoted to arbitrariness, and we're going to come back to arbitrariness in a number of chapters in this course. It is fundamental to what a human language is; there's no reason why anything is given a specific name. It just happens. Either we as a speaker of a language decide that's the name it should be, or we're trying to connect it to something else, but even that connection is arbitrary in and of itself. We will come back to that one.

Displacement. You might think of with respect to physical sciences about displacing water like you see in the picture. So **displacement** is the fact that we can use language to talk about things that are not in front of us. We can talk about anything at all.

It doesn't have to be right in front of us. It can be something tactile but just not in front of us; it can be something abstract; it can be something hypothetical. We can use language to describe all of it.

Cultural transmission is pretty much what you think it is, meaning that we use language to transmit culture from one person to another. Culture is not just music, art and literature; when we talk about culture, we talk about anything that is a human artifact. Remember, this is a social science.

The next one is duality. **Duality** is the fact that we can use a limited amount of lexicon or terms and a limited amount of phrase possibilities, as far as the construction of a phrase in a language. But we can create an infinite number of statements, questions, thoughts, declarations, hypotheses, anything. As we go through this course, it's really interesting to look at what a given language does. All languages have peculiarities, but they all have certain things in common as well. And one of these is duality.

Productivity or **creativity**, they're used interchangeably. Productivity is the word I tend to use; creativity is perfectly fine. It's the fact that we talk about things that we have never thought about before. We create, we produce with language. We do it in a way that is inventive. Again, humans seem to be pretty unique in this.

Last is **reflexivity.** We don't just talk about things around us; we talk about ourselves.

As we'll discuss later down the road with respect to these attributes, we can find a lot of them in a number of types of animal communication. But to this point in time, human beings are the only ones that have a type of communication that encompasses all six of these attributes. It's not to say that other animal languages don't. It's that we don't understand, or that it's possible but we just haven't figured out their code yet. We'll talk more about this when we get to human language, processing and how the brain works. In the next section we're going to focus on the main one, the biggest element, with respect to human language: arbitrariness.

1.3: Hallmarks of human language is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.3: Creativity and Generativity** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

1.4: Arbitrariness and ongoing changes

1.4.1 Arbitrariness, from Sarah Harmon



Video Script

So why focus on arbitrariness?

There's a pretty big reason. For one, as I said, it's potentially the most complicated, most abstract of all of the hallmarks or attributes of human language. But there's another reason: arbitrariness underlines everything with respect to language. Every component of language is arbitrary. What exactly does that mean?

Arbitrariness is the fact that speakers of a given language arbitrarily assigned a word or a term or a lexicon to a meaning. We'll come back to this when we get to morphology and syntax and semantics; we will actually cover quite a few ways.

I'll give you an example: every human being is born pretty much with two of these things. So what do we call them? Well, it's English you say ‘hand’ and I’ve written it in the language as well, Romanized alphabet and I’ve given you the international phonetic alphabet transcription; we will cover more about that pretty soon. So, in English, we call this a ‘hand’. In Spanish, we say ‘mano’. In French we say ‘main’. In Twi, which is a Western African language in the Niger Congo family, it's ‘nsa’. And then Russian, ‘ruka’. All for this thing that we all possess to have. So, why would that happen? Why would human beings give a different name with a different pronunciation to the things that we have? Why? Pretty interesting, right?

How about the term for this dwelling that we live in, now, granted that dwelling could be in lots of different shapes and sizes, but most every culture has a concept of a ‘house’ and again that first term is in English, but in Spanish, you say ‘casa’. In French you say ‘maison’. In Twi you say ‘awdang’. And in Russian, you use ‘dom’. And what's really curious is the fact that you have Spanish and French, very close Romance languages, but two totally different terms that they took from Latin. And in fact the Latin word for house, ‘domus’, is what you see in Russian, which is a cousin to Spanish and French— they're all Indo European languages—but Russian uses a Latin word, even though it is a Slavic language. Kind of interesting, no? For the same concept we use different terms, that is an example of arbitrariness.

There's no reason that a term for the most essential thing in life (air), the second most essential thing in life (water), the third most essential thing in life (food). Air, water, food. Those are three terms that are not shared with any other language, that the three most important things that we need as human beings and languages, do not even share those terms. That is arbitrariness at its core. Of all the things that are possible to have a name wouldn't ‘air’, ‘water’ and ‘food’ be candidates to have the same term and all human languages. But they aren't.

We can even talk about arbitrariness when it comes to **onomatopoeia**. This is a graph that I don't even remember where I took it; I've had it for so long. This is a graph of a number of onomatopoeic sounds across languages. Onomatopoeia is the encapsulating have a sound in a lexicon or a term a word, if you wish. You can see the sounds: the dog barking, a rooster crowing, a cat meowing, lowing, all of these, bomb exploding sneezing, the sound of a clock or any kind of clockwork mechanism. Notice that, while they are similar, they are not the same across all of these languages, English, German, French, Spanish, Hebrew, Hindi, Mandarin, Japanese, and Greek. In the case of the animal sounds while there is some similarity noticed that sometimes, you will have different numbers of syllables. You will have other sounds thrown in these are all written in international phonetic alphabet or IPA. We'll come to that soon enough, but just notice, even if you don't understand the symbols, just notice that they're all different across the different languages. Even in cases where languages are sister languages—English and German our sister languages, French and Spanish are sister languages—even they do not have the same symbols for these sounds. Curious, no?

As we go through the whole course we will talk about arbitrariness, because it is by far and away the most curious, if not arbitrary aspect of human language.

1.4.2 From 1.5 Language Change in Progress in Anderson's *Essentials of Linguistics*



Video Script

We saw in the last unit that every language changes over time. In this unit, we'll look more closely at some of the changes that are happening in English right now.

Every part of a language's grammar can change, but some of these changes are faster than others, and some are more noticeable. The lexicon is the vocabulary of a language — what words are in the language. New words enter English all the time as new technologies and concepts emerge, and dictionary editors like to publish lists of the new words they've added. This list shows a handful of words that were added to the Oxford English Dictionary in 2020: *beardo*, *awesomesauce*, *mentionitis*, *self-isolate*, *PPE* (for personal protective equipment), and, *thirsty*? Surely the word *thirsty* was in the Dictionary before 2020!? Yes, it was, with the meaning of "wanting something to drink", but a new meaning was added in 2020 — if you don't know that new meaning you might want to look it up.

A kind of language change that happens more gradually is in the sound system. Here are two English words, and I want you to think about whether both these words sound the same to you. For me, these two words sound the same — they're both [waɪn]. But for some English-speakers, the second word, the one that's spelled with a "wh", has a voiceless sound at the beginning, so it's pronounced [ʍaɪn]. When I say, "some English speakers", who do I mean?

Jack Chambers, a linguist at the University of Toronto, conducted a large-scale study of Canadian English and how it has changed over time. One part of his study asked people to say words that started with a "wh": words like *where*, *whine*, *whale*, *wheel*. Then

he analyzed his findings according to how old his participants were. In this graph, I've plotted people's decade of birth along the x-axis. So when Chambers was doing his research, people who had been born before 1920 were in their 80s. You can see that for this age group, more than half of them have the voiceless [ʍ] at the beginning of these wh-words. So for this group, *whine* sounds different from *wine*. For the next younger group, the pattern is about the same, but for each successively younger group, the proportion of people who pronounce a wh-word with a voiceless [ʍ] drops off. So for the people who were in their late teens and early twenties when Chambers interviewed them, only about 10% pronounced *whine* differently from *wine*. By looking at this snapshot across different age groups, we can get a picture of how Canadian English has changed over time.

Languages might also change in their morphology and syntax, though these changes tend to happen very slowly indeed. Let's look at a couple of changes that are in progress right now. The first one I want to look at has to do with the word *because*. Suppose I start a sentence like this, "Alex took an umbrella because…" and I ask you to finish it. You might finish it by appending another whole sentence, "because it was raining". Or you might choose a prepositional phrase that starts with *of*, "because of the rain". Both of these options have been available in English syntax for centuries. But a new option is emerging. If you're young, or if you spend a lot of time online, you might finish this sentence just, "because thunderstorm", with just a plain old noun phrase. This change seems to have started on Craigslist in 2011, with an ad for a car that was "completely stripped inside because race car", and now forms the title of the book *Because Internet*, in which linguist Gretchen McCulloch documents the ways that the internet has changed how we use language.

The last change I want to talk about is also happening in the morphosyntax of English, in the pronoun system. But first we need to look at a change that happened hundreds of years ago. In the 16th century, English used to have two ways of saying "you". If you were talking to a group of people, you'd say *you* just like we do now. But if you were talking to just one person, you'd address them as *thou* or *thee*, as in, "What classes art thou taking this term?" or "I really like thy new haircut". By the 17th century, *thou* and *thee* had all but disappeared and were only reserved for conversations with people you're very close to. So the word *you* was used for both singular and plural. In modern English, we don't have *thou* or *thee* at all unless we're trying to be funny or old-fashioned. But it can be pretty useful to have a way of distinguishing between singular and plural, so some varieties of spoken English have other plural forms, like *y'all* or *you guys* or *youse*. Maybe your variety of English has one of these.

So that change in the pronoun system happened hundreds of years ago without incident. These days a different change is happening, this time to the third person pronoun *they*. For centuries, *they* has been used as a plural pronoun, to refer to a group of people, as in, "The children said **they** played soccer all afternoon". And it's also very common to use *they* when we don't know how many people are involved. You might hear someone say, "Whoever was in here, he or she or they made a big mess" but it would sound very formal and stuffy. The same is true if you're talking about one person whose identity you don't know, or if it just isn't relevant — maybe I'm telling one of my colleagues, "One of my students told me **they** were locked out of their email". There's only one student, but their identity isn't relevant to the story, so I just refer to them as *they*. This so-called singular *they* has also been in English for centuries — you can see that it's documented as far back as the fifteenth century, in contexts that are really clearly singular: *each of them*, *a man*, *a person*. The change that's in progress right now is to use *they* for a single person whose identity we do know, either because they're non-binary and use *they/them* pronouns or because we're choosing not to specify their gender. When I poll students, that is, people in their 20s, I usually find that about half of them have this specific-singular-*they* in their mental grammar, and about half don't. So it's a change that's unfolding right now.

As always, when language changes, some prescriptivists get quite uptight about it. The *Chicago Manual of Style* tells people "it is still considered ungrammatical", and the *AP Stylebook* tells you it's "acceptable in limited cases" but they'd really prefer if you didn't use it. And then there are the extremely crabby folks like this author who claimed it hurt her ears and burned her eyes, poor thing! But no matter how much the prescriptivists yell, specific-singular-*they* is getting used more and more widely. In 2015 the American Dialect Society voted it the Word of the Year; the Globe & Mail added it to its style guide in 2017, and it was the Merriam-Webster Dictionary's Word of the Year in 2019. And linguists are paying attention to how this part of English grammar is changing. Bronwyn Bjorkman found that English-speakers with a conservative grammar didn't use *they* in this way, but those with an "innovative" grammar did. Lauren Ackerman has proposed that the more trans and non-binary friends you have, the likelier your grammar is to have specific-singular-*they*. Kirby Conrod found in their dissertation that older people were less likely to use it and younger people were more likely, and Lex Konnelly just published a paper tracking the three stages of grammatical change that are unfolding.

I said earlier that this change is happening no matter how much it bothers the prescriptivists. No one can stop language from changing. But can we make it happen faster? After all, grammar is people — if everyone woke up tomorrow and started calling a dog a "blimlimlim", the dictionaries couldn't stop us! There is good research that shows that misgendering people does real harm.

One way to make it less likely that non-binary people will be misgendered is for the language to change to include specific-singular-*they*. And the way that language changes is for people to change how they use it. If you already have specific-singular-*they* in your grammar, use it as much as you can! And if you'd like to change your own mental grammar, Kirby Conrod gives some good advice — slow down, listen to people who use it in their own language, and practice! The more you use it, the more natural it will feel.

Check Yourself

Exercise 1.4.1

How recently has *they* been used as a singular pronoun in English?

- *They* started being used as a singular around 2015.
- *They* has been used as a generic singular for centuries and now it's changing to also be used as a specific singular.
- Never; *they* is always plural.

Answer

"*They* has been used as a generic singular for centuries and now it's changing to also be used as a specific singular."

Hint: Think about how language change is constant and adaptive. Think about how language is used, both spoken and written formats.

Exercise 1.4.2

How did Jack Chambers show that Canadian English changed over time?

- He interviewed speakers who had been born in different decades and compared the differences in their language use.
- He conducted a research project that started in the 1920s and kept gathering data until the 1990s.

Answer

"He interviewed speakers who had been born in different decades and compared the differences in their language use."

Exercise 1.4.3

How can people stop language from deteriorating?

- Publish extensive dictionaries and make them widely available.
- Correct people every time they make a mistake.
- No one can stop language from changing!
- Teach correct grammar in schools.

Answer

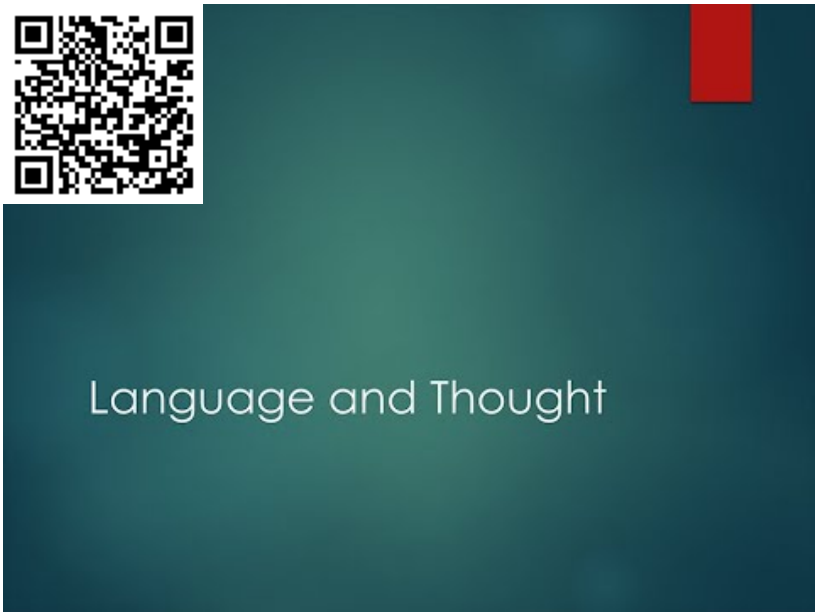
"No one can stop language from changing!"

Hint: Think descriptively, not prescriptively. We will come back to this topic in Chapters 8 and 9.

1.4: Arbitrariness and ongoing changes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.5: Language and thought

Language and Thought, from Sarah Harmon



Video Script

Language and thought. There has been quite a bit of discussion over the years and, especially, it has resurfaced in the last 15 years about whether language and thought are connected. By that I mean whether or not how we speak the language affects how we think, or how we think affects how we speak the language. There has been quite a bit of discussion; it is not a new topic, in fact, this has been openly discussed in academic circles for over 100 years. Certainly with respect to the philosophical side of things, probably…well, how long has philosophy been around? . 🤣 Certainly, this is a discussion that warrants a little bit of light, especially with respect to getting into a mind frame of analyzing language like a linguist would, so let's talk about it.

Frequently, the term the **Sapir-Whorf Hypothesis** is put out as a way to say that language affects how we think. There are some things to keep in mind. First of all, Edward Sapir and Benjamin Lee Whorf never wrote anything together; in fact, I believe Sapir predates Whorf. Also, it's really important to remember that both of them were early anthropologists. 'Early' is the key term there, because even anthropology has moved on significantly from both of their work. It is not something that is espoused necessarily even today in anthropology. But what it really refers to is this concept of **linguistic determinism,** the fact that our language determines how we view the world. That has been absolutely debunked in every way possible, both by linguists and by anthropologists. That is in fact not even exactly what Edward Sapir and Benjamin Lee Whorf wrote about; it is just what people interpreted from their works.

A different version of it, a softer version of it, has come out as **linguistic relativism**. That there might be some connections, that it is potentially the case that how we speak a language and how our language is set up might inform certain perspectives of how we see the world. A linguist will say, probably not. It could be true, but it bottom line is very hard to determine this, to analyze this. That is because we cannot, for example, measure how a person thinks in such detail that we can see the connection between the language and the thought process. We just can't do that. For right now, it is what we affectionately like to call 'armchair philosophy', as in you sit in an armchair and you think deep thoughts about life. It's not actual social science and our colleagues in anthropology would say the same thing.

What we can say?

We can question about a few things. First of all, the work of Dan Sullivan, who promotes this concept of **Thinking for Speaking**, that there is some interesting data that we can collect. For example, there are a number of languages throughout the world, where this concept of past, present, and future (the three verbal tenses) doesn't really exist, but it does exist in a different way. For

example, you may not have ‘I speak’, ‘I spoke’, ‘I will speak’—present past future—you may not have those very three linear breaks between how you say those reactions. But you will convey that same information in some way, shape or form. There's also the concept that in many cultures, you do not refer to things being above or below, up or down, to the left or to the right; you don't think in those terms. You think north, south, east, west. Again, this is something that is interesting and certainly linguistically, as well as anthropologically very interesting. Can we prove that those differentiations in language affect how a person thinks? It may just simply be a case of perspective or focus, but it may not exactly be what we would consider a change in thought process.

The work of John McWhorter really speaks to this. He has written a number of books and has a couple different podcasts on language and specifically really highlights how nothing with respect to language implies actual differences in the psychological linguistic connection. It's just a matter of different perspective that being said. It is interesting to think and perhaps in the future we may get some way of measuring this connection. For right now, though, there really isn't one that we can say shows a different perspective. A lot of it is cultural. That doesn't mean it is not linguistically relevant; it just means it's different and a different element to analyze. It is, overall, still a controversial thing to say, “Well, language has a major impact on thought.” We don't really know that to be true.

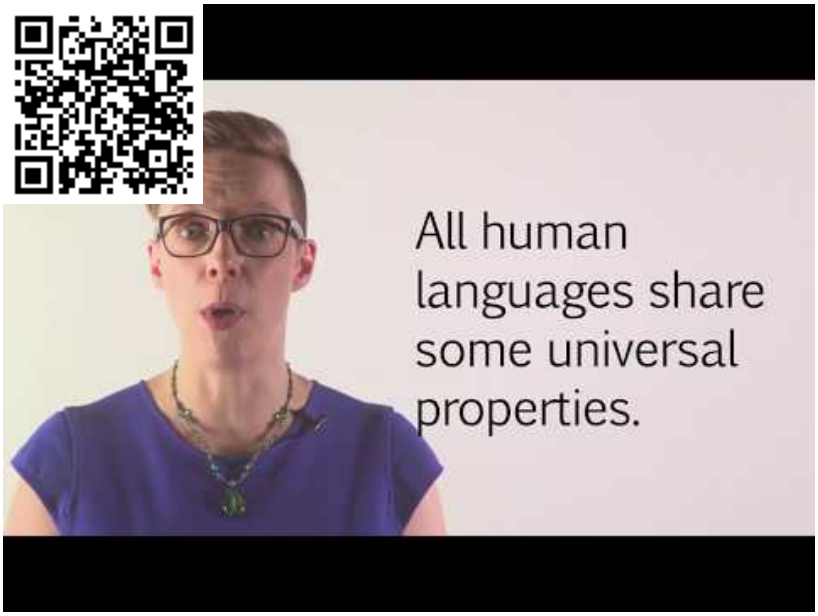
What do we know to be true?

Well, we know that there are some, shall we say, **linguistic universals**, things that we have observed over 100-150 years, give or take, of linguistic analysis. Here's a few things that we know. We know we can prove that there is a difference between **knowledge** or **competence**, on the one hand, and **performance**, on the other. We also know, without question, that human beings understand a language before they're able to produce it. That part we can say, and we can say that that is a psychological linguistic connection that we have to understand the tool before we use the tool. We have to know what a hammer is before we can use it. We have to know what a language does before we can use it, and that is something that we know. As far as other universals, that's the next section.

1.5: Language and thought is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.6: Language universals

1.6.1 From 1.4 Fundamental Properties of Language, from Anderson's *Essentials of Linguistics*



Video Script

Because everybody speaks a language, just about everybody has opinions about language. But there are lots of things that are commonly believed about language that just aren't true.

You might have heard someone say that a given language has no grammar. I've heard people try to argue that Chinese has no grammar, that English has no grammar, that the languages spoken by Indigenous people who live in what is currently Canada have no grammar, even that Swiss German has no grammar.

When people say this, they might mean a few different things. Sometimes they just mean that there's not much variation in the forms of words, which is true of Chinese, but the grammar of Chinese has lots of complexity in its sound system.

But sometimes people who argue that a language has no grammar are actually trying to claim that that language is inferior in some way.

The truth is that **all languages have grammar**. All languages have a sound system, a system for forming words, a way of organizing words into sentences, a systematic way of assigning meanings. Even languages that don't have writing systems or dictionaries or published books of rules still have speakers who understand each other; that means they have a shared system, a shared mental grammar.

When we're investigating mental grammar, it doesn't matter whether a language has a prestigious literature or is spoken by powerful people. Using linguists' techniques for making scientific observations about language, we can study the phonetics, phonology, morphology, syntax and semantics of any language.

Another opinion that you might have heard about language is that some languages are better than others. Maybe you've heard someone say, "Oh, I don't speak real Italian, just a dialect," implying that the dialect is not as good as so-called real Italian. Or maybe you've heard someone say that Québec French is just sloppy; it's not as good as the French they speak in France. Or maybe you've heard someone say that nobody in Newfoundland can speak proper English, or nobody in Texas speaks proper English, or maybe even nobody in North America speaks proper English and the only good English is the Queen's English that they speak in England.

The truth is that **all languages are equally valid**. Just as we said that all languages have grammar, it's also the case that there's no way to say that one grammar is better or worse than another grammar. Remember that linguistics takes a scientific approach to

language, and scientists don't rate or rank the things they study. Ichthyologists don't rank fish to say which species is more correct at being a fish, and astronomers don't argue over which galaxy is more posh. In the same way, linguists don't assign a value to any language or variety or dialect.

It is the case, though, that plenty of people do attribute value to particular dialects or varieties, and sociolinguistic research tells us that there can be negative or positive social consequences for people who speak certain varieties. When people say that British English is better than American English, for example, they're making a social judgment, based on politics, history, economics, or snobbery. But there is no linguistic basis for making that value judgment.

One of the common misconceptions about language arose when scholars first started doing linguistics. At first, they focused on the languages that they knew, which were mostly the languages that were spoken in Europe. The grammars of those languages had a lot in common because they all evolved from a common ancestor, which we now call Proto-Indo-European. When linguists started learning about the languages spoken in other parts of the world, they thought at first that these languages were so unfamiliar, so unusual, so weird, that they speculated that these languages had nothing at all in common with the languages of Europe.

Linguists have now studied enough languages to know that in spite of the many differences between languages, **there are some universal properties that are common to all human languages**. The field of linguistic **typology** studies the properties that languages have in common even across languages that they aren't related to. Some of these universal properties are at the level of phonology, for example, all languages have consonants and vowels. Some of these universals are at the level of morphology and syntax. All languages make a distinction between nouns and verbs. In nearly all languages, the subject of a sentence comes before the verb and before the object of the sentence. We'll discover more of these universals as we proceed through the chapters.

A very common belief that people have about language is something you might have heard from your grandparents or your teachers. Have you heard them say, "Kids these days are ruining English! They should learn to speak properly!" Or if you grew up speaking Mandarin, maybe you heard the same thing, "Those teenagers are ruining Mandarin! They should learn to speak properly!" For as long as there has been language, there have been people complaining that young people are ruining it, and trying to force them to speak in a more old-fashioned way. Some countries like France and Germany even have official institutes that make prescriptive rules about what words and sentence structures are allowed in the language and which ones aren't allowed.

The truth is that **every language changes over time**. Languages are spoken by humans, and as humans grow and change, and as our society changes, our language changes along with it. Some language change is as simple as in the vocabulary of a language: we need to introduce new words to talk about new concepts and new inventions. For example, the verb *google* didn't exist when I was an undergraduate student, but now googling is something I do every day. Language also changes in they we pronounce things and in the way we use words and form sentences. In a later chapter, we'll talk about some of the things that are changing in Canadian English.

Another common belief about language is the idea that you can't learn a language unless someone teaches you the rules, either in a language class or with a textbook or a software package. This might be partially true for learning a language as an adult: it might be hard to do it on your own without a teacher. But think about yourself as a kid. Whatever language you grew up speaking, whether it's English or French or Mandarin or Arabic or Tamil or Serbian, you didn't have to wait until kindergarten to start speaking. You learned the language from infancy by interacting with the people around you who spoke that language. Some of those people around you might have taught you particular words for things, but they probably weren't teaching you, "make the [f] sound by putting your top teeth on your bottom lip" or "make sure you put the subject of the sentence before the verb". And by the time you started school you were perfectly fluent in your language. In some parts of the world, people never go to school and never have any formal instruction, but they still speak their languages fluently.

That's because almost everything we know about our language — our mental grammar — is unconscious knowledge that's acquired implicitly as children. **Much of your knowledge of your mental grammar is not accessible to your conscious awareness**. This is kind of a strange idea: how can you know something if you're not conscious of knowing it? Many things that we know are indeed conscious knowledge. For example, if I asked you, you could explain to me how to get to your house, or what the capital of Canada is, or what the difference is between a cow and a horse. But our mind also has lots of knowledge that is not fully conscious. You probably can't explain very clearly how to control your muscles to climb stairs, or how to recognize the face of someone you know, or how to form complex sentences in your native language, and yet you can do all of these things easily and fluently, and unconsciously. A lot of our job when we study Linguistics is to make explicit the things that you already know implicitly. This is exactly what makes linguistics challenging at first, but it's also what makes it fun!

Check Yourself

? Exercise 1.6.1

It’s important to study Latin because Latin is more logical than other languages.

- True.
- False.

Answer

"False"

The reason: There is no such thing as 'more logical' when it comes to languages. All languages have a 'logic', a system of how sounds are put together, how words are put together, and how phrases are put together.

? Exercise 1.6.2

Spending too much time texting will ruin your ability to write proper English.

- True.
- False.

Answer

"False"

The reason: This is a highly prescriptive perspective, one which linguists don't consider. There is no such mindset. Also: There is significant research to suggest that if we teach standard written and spoken language as a 'dialect' or 'alternative' that should be used in certain situations. More on this in Chapter 8.

? Exercise 1.6.3

The dictionary gives the only correct meaning and pronunciation for words.

- True.
- False.

Answer

"False"

The reason: A dictionary includes the standard definitions and pronunciations, and frequently includes more dialectal and historic uses/pronunciations.

1.6.2: Language universals, from Sarah Harmon



Video Script

So let's talk about these language universals that were hinted at in the previous section. There are certain pieces that we're going to talk about now and then we'll come back to more later on in the course. But to start off, there are certain things that are universal, with respect to language.

All languages have grammar they have some kind of structure, they have some kind of rules as to how to put lexicon or terms together. How to put sentences or phrases together how to make it all makes sense in a given context, there are rules, there are rules as to which sounds can go next to one another and in what order.

All speech communities tweak these rules and structure over time and we do it constantly. Almost always, it is an unconscious thing that we do we just change language we just speak it it comes out in a different way and it either gets continued or it doesn't.

These changes happen all the time, but there are stable patterns for change, and it is interesting to note that certain languages and language families have certain patterns that they always use. So a pattern, that you will see in English might be shared in the other Germanic languages, but it may not be shared with the Slavic languages, even though they are cousins. And they may not be shared with a totally different language let's say teaching a language spoken by one of the Maya communities in Central America. But we can see stable patterns of change across time and across speech communities.

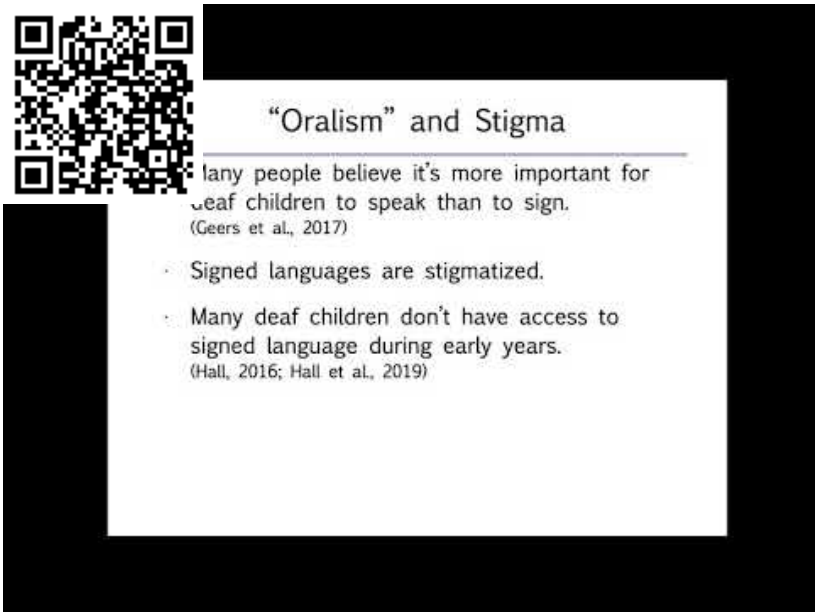
Why do we have this? We'll go into this a little bit more later, but some of it has to do with a term called **Universal Grammar**. That is a term that was coined by Noam Chomsky who was considered the founder of modern linguistics he is part of a group of folks who really started analyzing language. In a more complex way and it is the foundation of that work in the late 50s and early 60s that really continues to motivate us to this day. It's the concept that we have a faculty for language human beings have a faculty for language, we seem to be auto programmed for something now what that's something is. We don't know we still are trying to figure that out, what can we say, well, certain things seem to be in place certain combinations seem to happen frequently other combinations don't seem to happen frequently. We do know this concept of confidence versus performance is the is baked into this, we know that that is a key component to this universal grammar. But we're still trying to figure out quite a bit of it and with each successive year we learn a little bit more so, who knows in 20-30 years we could know significantly more about this universal grammar that we all seem to be, well, born with.

1.6: Language universals is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.4: Fundamental Properties of Language** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

1.7: Signed languages

1.7.1: From 2.10 Classifying Signs, in Anderson's *Essentials of Linguistics*



Video Script

In this chapter we've talked about how the sounds of spoken languages are organized: we classify consonants according to their place & manner of articulation and their voicing, and we classify vowels according to their tongue position and whether the lips are rounded. That's a pretty tidy system for the segments of spoken languages, which are produced with the mouth (and the articulators and the larynx and the lungs), and perceived with the ears. We haven't talked yet about signed languages, which are produced with the hands and arms and (usually) perceived with the eyes. It turns out that, just like consonants and vowels, the signs in signed languages can be classified according to how they're produced, along five parameters. Before we talk about the signs themselves, let's talk a bit more about signed languages in general.

It might be that when you hear the phrase "sign language", you think of American Sign Language, or ASL, which is the signed language used most widely in North America. But ASL is just one of many signed languages in the world. There's BSL, or British Sign Language, and LSQ, Langue des signes québecoise, Auslan, and many others. These languages are not mutually intelligible — in other words, users of BSL don't necessarily understand ASL and vice versa. But all sign languages share some properties with each other, and they also share properties with spoken languages, which we'll examine later in this book.

You should know that not everyone who uses a sign language is deaf — some hearing children acquire sign natively if their parents or other people in their household sign. And many hearing people choose to learn a signed language in addition to their spoken language, the same way people might choose to learn Spanish or Korean. The other piece of the story is that not everyone who's deaf uses a sign language, because of stigma or because of language deprivation. When the word *deaf* is spelled with a lower-case 'd', it's the medical term for people who have little or no hearing. The word *Deaf* with an upper-case 'D' is used by people who participate in Deaf culture. Deaf culture includes using signed languages, and usually does not perceive deafness as an impairment.

This idea that deafness is not a defect can be quite radical in our ableist society where disability is often stigmatized. **Oralism** is the name for the attitude that says that speech is better than sign. Many parents, teachers, and even doctors believe that it's more important for deaf people to learn to speak than to sign. This stigmatizing attitude toward signed languages means that it's quite common for children who are born deaf not to have access to any language during those vital first few years of life. The evidence shows that deaf children who are exposed to sign from an early age have better outcomes than deaf children who only have access to speech, relying on hearing aids and cochlear implants.

The stigma towards signed languages is based on faulty assumptions, from attitudes that believe that signed languages aren't "real" or full languages. But the tools of linguistics reveal that signed languages are just as complex as spoken languages; they have phonetics, phonology, morphology, syntax and semantics. Furthermore, neural imaging has revealed that users of sign languages recruit the same areas of the brain for producing and understanding language as users of spoken languages do.

So since we've paid attention to how linguists classify the sounds of spoken languages, let's look at how the signs of signed languages can be classified. First, I need to point out that when we talk about classifying signs, we're looking at a different level of representation from speech sounds: individual signs correspond to words, not to segments like consonants or vowels. But like consonants and vowels, each individual sign includes multiple **parameters**. The five parameters that make up a sign are location, movement, handshape, orientation, and non-manual markers. Each of these parameters can vary independently of the others. Let's look at them more closely.

Location has to do with where the signer articulates the sign, relative to their body. In ASL, or American Sign Language, almost all signs are articulated above the waist. The two ASL signs for ONION and APPLE differ in their location: the sign for ONION is articulated next to the eyes, while the sign for APPLE is at the mouth.

The **movement** parameter refers to how the hands and fingers move, and what path they take. Compare the ASL signs for CAN and SHOES. In both signs, the hands are in the same shape and the same location, but their movement differs. In the sign for CAN, the fists move downward parallel to each other, while in the sign for SHOES, the signer brings the two fists together at the edges.

The position of the hands and fingers is called **handshape**. When I say the position of the hands and fingers, I don't mean where on the body — that's location — but how is the signer configuring them. Let's look back at the sign for APPLE, which is articulated with a knuckle twisting beside the mouth. Now look at the sign for CANDY. Instead of a knuckle, the index finger is pointing at the side of the mouth. The location and movement are the same as for APPLE, but the handshape is different.

The **orientation** of the hands is also important for ASL signs, that is, the direction that the hands are facing. Compare the two ASL signs for BALANCE and MAYBE. They're very similar, but the orientation of the hands is different in the two signs.

The final parameter that can differ between signs is **non-manual markers**, that is, the parts of the body that aren't the hands. Look here at the sign for LATE, and then look at the sign for NOT YET. What the hands do is pretty similar in these signs, but for NOT YET, the tongue protrudes. If you made the handsign without the tongue movement, you wouldn't have the sign for NOT YET — the non-manual marker is a crucial part of the sign.

These are just a very few examples from ASL, but they illustrate that signed languages are just as complex and just as systematic as spoken languages. Later in the book, we'll talk about the morphology and syntax of signed languages too!

1.7.2: Gestures and Sign Languages, from Sarah Harmon.



Video Script

At this point, it would be remiss of us to not talk about signed languages. As you read previously with the Catherine Anderson piece on sign languages there there's something to them it's not just random gestures that are just put together for no apparent reason. There's a reason to it, there's a systematic component to them and I'm going to go a little more into this. I'm also going to provide a few more clues to understand what a sign language really is.

To start off realistically there's actually four different kinds of signed language, but when I say a 'sign language', you probably think primarily of something called a **primary sign language**. A primary sign language is, for example, American Sign Language, although as we'll talk about as it is not quite as ubiquitous as you think. But it is also not the only kind of signed language. There are in a number of villages, a **village sign language**, which is meant to be a secondary component. This means that the primary language is frequently spoken and then you have certain gestures that must be accompanied by that spoken language. It's not just a matter of talking, you must have certain signs that are shown with your conversation. They happen in a lot of communities around the world. There are also something called **home sign languages**; think baby sign language, the signs that we teach our children, our babies before they are fully lingual. For example, this [gesture] usually means 'more'. We usually have them based off of whatever the local sign language is. The baby sign language that you're familiar with is probably the one that's connected to American Sign Language. The curious one for me is **finger spelling** because finger spelling is the one that we should all learn how to do. You're going to read something as a journal assignment that has to do with a signed name and the difference between a finger spelling of your name and a name sign.

There are some really key facts that you have to remember with respect to sign languages. Just like a spoken language, there is morphology, phonology, syntax, and even a little bit of dialectology. Morphology is how you combine the signs to make a compound word; that one seems pretty straightforward. But phonology? Sound? How does that come into a sign language? Well, it's in the hand movements and how aggressive or subtle you are with the hand movements; that is really the phonology. Also, how you how you move as a whole, this is the syntax. Your telling of the story is not just the actual signs themselves, but how your hands move and even some facial gestures. Then dialectology well, even when you compare the alphabets of different sign languages there's a dialect ology if not a total language difference.

For example, in the top right. Up top this is the American Sign Language alphabet; this is the one that if you took ASL, this is the alphabet that you have. But this one next to it is British, British Sign Language. Notice it is a two-hand system, not a one hand system like we have an American Sign Language. So American and the top left and just to the right of it is British Sign Language. Next to that we have Spanish Sign Language. If you speak Spanish, you'll notice some of those different sounds different letters that we have. The next to that is Irish Sign Language; notice that Irish Sign Language and American Sign Language are both one handed in the alphabet and have a lot of similar in fact almost identical signs. It is very different from British Sign Language. Down on the bottom, this is German Sign Language. The next to that we have Norwegian. And then next to that is Chinese Sign Language, which is based off of Mandarin. You see so many different possibilities, while there are some similarities, even just look at the A, letter A in even just the one-handed sign languages with respect to the alphabets. They're totally different; the A is pretty much the same in American Sign Language, Spanish (and this is Spain). Irish is pretty much the same. German pretty much the same. But notice Mandarin; it's a different sign. Instead of this [gesture], it is this [gesture]. Specifically, these different languages are using different tools.

Speaking of different languages using different tools, let's talk about alternate sign languages. These are frequently used as in lieu of speaking. You might notice some interesting patterns with respect to culture. For example, in a number of the indigenous aboriginal communities of central Australia, if a woman is widowed and in mourning, they are banned from speaking. They use an alternate sign language tied to a village sign language in lieu of speaking they're not allowed to speak, and that is actually true in a number of communities around the world that if you are in mourning, a woman in particular and mourning, especially widowed, you're not allowed to speak, either indefinitely or for a certain period of time, depends on the culture. With respect to the Plains Indians, we are talking mostly various Sioux tribes, and a lot of tribes that are from about Kansas north through Canada. A lot of those areas where you think that they speak similar languages, they actually don't and over the centuries they developed a lingua franca, as it were, but signed instead of spoken. That way they could trade and have discussions and treaties. In Botswana, there is a linguistic community called Ts'ixa. They use a signed language when they are hunting; that makes sense because you don't want to use your voice to alert the animals that you're hunting. But they also use the same signed language when they're giving a narrative or a performance; think hula but with the hands. In various monastic communities, the Benedictine orders in particular are famous for this in Catholic communities, in a number of monastic communities around the world, monks and nuns are not allowed to speak. There's a code of silence. So how do you communicate? You sign; you have a set of symbols that you use and the Cistercian

monks, the Cluniac monks, the Trappist monks all have a different setup of signs. Understand that when we talk about a sign language it's a really rich language. And really it isn't one, but a whole slew of types.

1.7: Signed languages is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.8: Next steps and references

Next steps and references, from Sarah Harmon



Video Script

So where do we go from here? How do we transition to our next several chapters or modules?

Well, this is the layout of the rest of the textbook and the modules in the course will follow along with this. Notice that we're going to go from small to big. So, the small we're going to start with sounds. It can get very complex and it's totally normal to be freaked out; this happens. We're going to start small with the sounds and then go a little bigger and talk about words… and go a little bit bigger and talk about phrases… and then go a little bit bigger and talk about meaning. Phonetics, Phonology, Morphology, Syntax, Semantics and pragmatics. Those are the core areas of linguistics as you saw in an earlier section, and those are really the areas that you need to understand before you start applying to this information to different areas.

After we finish semantics and pragmatics, then we can start talking about speech communities and how different speech communities interact, dialects. That is sociolinguistics. From there we'll talk about historical linguistics, how language changes over time. We'll talk a little bit more about language universals as well. From there, we'll talk about how we acquire or learn a language. It doesn't matter if it's a related language or a completely different language, if it's a language that we learn as a child or a language that we learn as an adult. We'll talk about all of it. Then we'll round out with a little bit on how language processing happens, so how our brain takes language and understands it, as well as produces it. It should be noted that, especially in chapters nine and 10, there are new advances all the time; even as I record, this which is fairly late before the semester starts for 2021, there are changes that are going to be announced. Even what is in this textbook could get modified. There's a lot of what Catherine Anderson says, and a lot of what I say, that is a little bit more generic exactly because of that; we want to give you a good foundation first before we get into the nitty gritty of what actually goes into learning a language, as well as understanding language.

As far as the resources for this chapter, a lot of what I talked about can also be found in just about any intro to linguistics textbook, in their chapter one or their introductory chapter. If you want some more resources, I would look at what both *Introduction to Language* and *Language Files* have, as far as their list of resources. It's really long and I didn't want to put them in the slides here, but you will see them listed out in the actual LibreTexts page.

You can do this. Remember language is difficult, it's complex, it's not simple by any stretch of the imagination. Even when you look at a language and think it's simple, it's not. It is really nuanced and complex, and so to analyze it, we have to take a step back, we have to open our minds, we have to be creatively thinking about any possibility and, most importantly, we need to find patterns, but we need to have fun doing it.

1.8: Next steps and references is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

2: Sounds, Part 1- Phonetics

2.1: Introduction to Phonetics
2.2: Articulators and Airstream Mechanisms
2.3: Voicing, Resonance and Articulation
2.4: IPA and Charts
2.5: Sound Classes and Hierarchy
2.6: All About Vowels
2.7: References

2: Sounds, Part 1- Phonetics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.1: Introduction to Phonetics

First, watch this video of Anne Sullivan describing how she taught Helen Keller how to speak.



Introduction to Phonetics, from Sarah Harmon



Video Script

So in the video above this video script you see a dialogue between Annie Sullivan and us. She's describing what she did to teach Helen Keller, to speak. Remember Helen Keller was both blind and deaf; so a deaf person normally relies on a primary sign language like ASL couldn't do that because Helen was blind. She had to first teach Helen how to finger spell but with the hands, so that you could do [gestures] the different letters on the hand. She could feel them and then she wanted to learn how to speak so Annie Sullivan taught her how to speak. And I think it's really important to understand what Annie taught her is essentially phonetics articulatory phonetics; that's what she taught her how to do.

Why do we need to study phonetics? Well English is a really great example as to why; you can't just rely on how a language is written in order to know how it is pronounced. Those of you who had to learn English as teenagers or adults, you know this very well, so this is a really great poem by T.S. Watts; I love using it.

> *"English--As I Know It" by T.S. Watts*
>
> *I take it you already know*
>
> *of tough and bough and cough and dough*
>
> *Other may stumble you, but not you,*
>
> *on hiccough, thorough, laugh, and through.*
>
> *Well done! And now you wish, perhaps,*
>
> *To learn of less familiar traps?*
>
> *Beware of heard, a dreadful word,*
>
> *That looks like beard and sounds like bird.*
>
> *And dead: it's said like bed, not bead;*
>
> *For goodness' sake, don't call it deed!*
>
> *Watch out for meat and great and threat.*
>
> *(They rhyme with suite and straight and debt.)*
>
> *A moth is not a moth in mother,*
>
> *Nor both in bother, broth in brother.*
>
> *And here is not a match for there*
>
> *Nor dear and fear for bear and pear*
>
> *And then there's dose and rose and lose*
>
> *Just look them up—and goose and choose*
>
> *And cork and work and card and ward*
>
> *And font and front and word and sword*
>
> *And do and go and thwart and cart—*
>
> *Come, come, I've hardly made a start!*
>
> *A dreadful language? Man alive.*
>
> *I'd mastered it when I was five!*

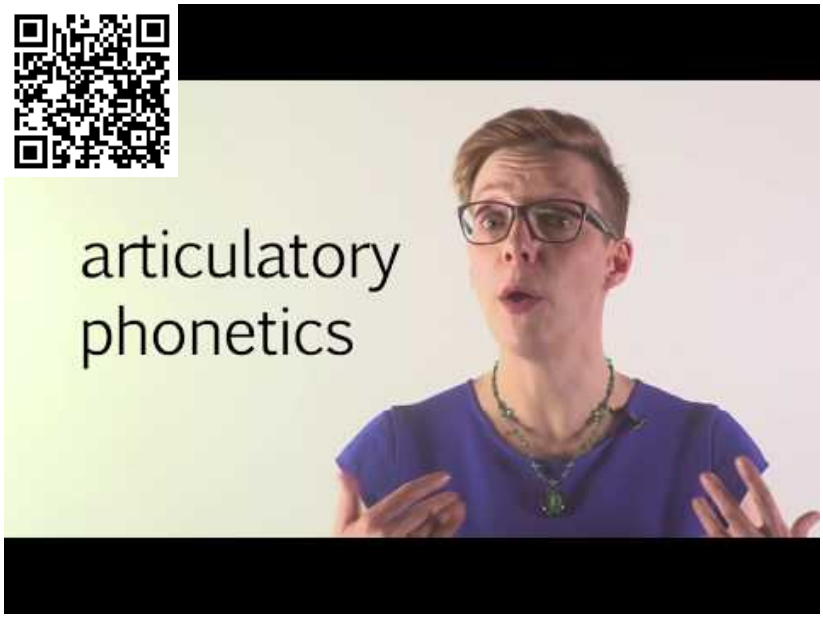
Certainly with respect to language, if we go just strictly by how a language is written, we're not going to be actually describing how it is pronounced. Some languages are written fairly phonetically but even still how you write down the sound is arbitrary.

There's that word again, and we're going to come back to over and over and over again. So, how do we talk about languages sounds? How do we describe them and how do we produce them? That is the subject of phonetics.

2.1: Introduction to Phonetics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2: Articulators and Airstream Mechanisms

How Humans Produce Language, in Anderson's *Essentials of Linguistics*

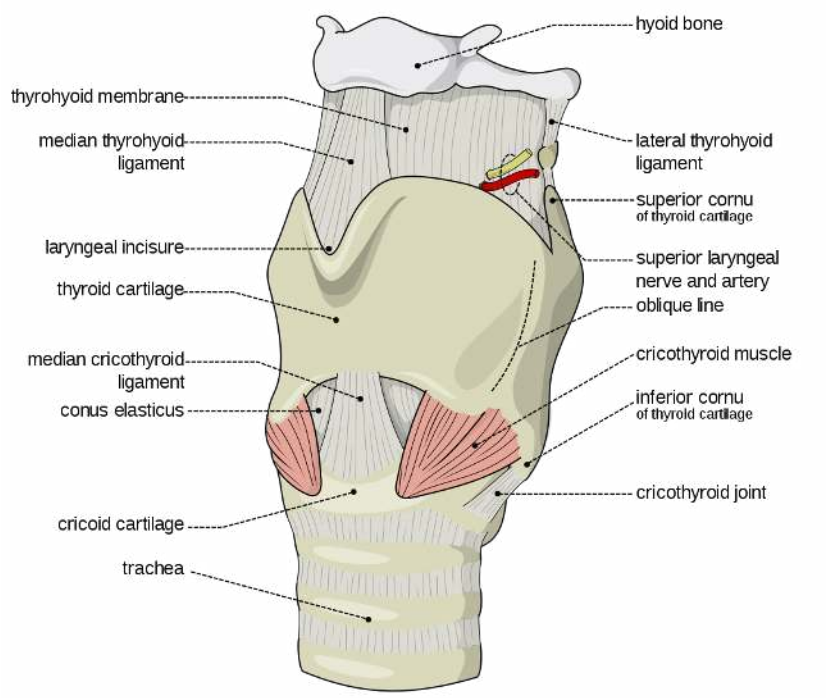


Video Script

The field of phonetics studies the sounds of human speech. When we study speech sounds we can consider them from two angles. **Acoustic phonetics**, in addition to being part of linguistics, is also a branch of physics. It's concerned with the physical, acoustic properties of the sound waves that we produce. We'll talk some about the acoustics of speech sounds, but we're primarily interested in **articulatory phonetics**, that is, how we humans use our bodies to produce speech sounds. Producing speech needs three mechanisms.

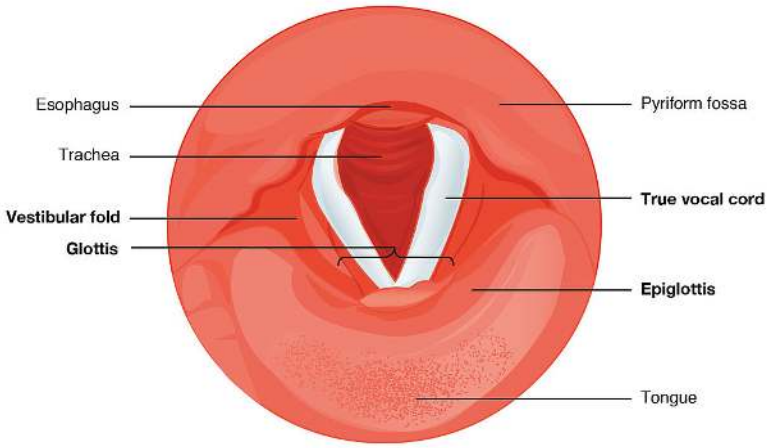
The first is a source of energy. Anything that makes a sound needs a source of energy. For human speech sounds, the air flowing from our lungs provides energy.

The second is a source of the sound: air flowing from the lungs arrives at the larynx. Put your hand on the front of your throat and gently feel the bony part under your skin. That's the front of your **larynx**. It's not actually made of bone; it's cartilage and muscle. This picture shows what the larynx looks like from the front.



By Olek Remesz (wiki-pl: Orem, commons: Orem) [CC BY-SA 2.5-2.0-1.0 (https://creativecommons.org/licenses...sa/2.5-2.0-1.0)], via Wikimedia Commons

This next picture is a view down a person's throat.



By OpenStax College [CC BY 3.0 (http://creativecommons.org/licenses/by/3.0)], via Wikimedia Commons

What you see here is that the opening of the larynx can be covered by two triangle-shaped pieces of skin. These are often called "vocal cords" but they're not really like cords or strings. A better name for them is **vocal folds**.

The opening between the vocal folds is called the **glottis**.

We can control our vocal folds to make a sound. I want you to try this out so take a moment and close your door or make sure there's no one around that you might disturb.

First I want you to say the word "uh-oh". Now say it again, but stop half-way through, "Uh-". When you do that, you've closed your vocal folds by bringing them together. This stops the air flowing through your vocal tract. That little silence in the middle of "uh-oh" is called a glottal stop because the air is stopped completely when the vocal folds close off the glottis.

Now I want you to open your mouth and breathe out quietly, "haaaaaaah". When you do this, your vocal folds are open and the air is passing freely through the glottis.

Now breathe out again and say "aaah", as if the doctor is looking down your throat. To make that "aaaah" sound, you're holding your vocal folds close together and vibrating them rapidly.

When we speak, we make some sounds with vocal folds open, and some with vocal folds vibrating. Put your hand on the front of your larynx again and make a long "SSSSS" sound. Now switch and make a "ZZZZZ" sound. You can feel your larynx vibrate on "ZZZZZ" but not on "SSSSS". That's

because [s] is a **voiceless** sound, made with the vocal folds held open, and [z] is a **voiced** sound, where we vibrate the vocal folds. Do it again and feel the difference between voiced and voiceless.

Now take your hand off your larynx and plug your ears and make the two sounds again with your ears plugged. You can hear the difference between voiceless and voiced sounds inside your head.

I said at the beginning that there are three crucial mechanisms involved in producing speech, and so far we've looked at only two:

- Energy comes from the air supplied by the lungs.
- The vocal folds produce sound at the larynx.
- The sound is then filtered, or shaped, by the **articulators**.

The oral cavity is the space in your mouth. The nasal cavity, obviously, is the space inside and behind your nose. And of course, we use our tongues, lips, teeth and jaws to articulate speech as well. In the next unit, we'll look in more detail at how we use our articulators.

So to sum up, the three mechanisms that we use to produce speech are:

- **respiration** at the lungs,
- **phonation** at the larynx, and
- **articulation** in the mouth.

Check Yourself

? Exercise 2.2.1

What is the voicing of the last sound in the word *soup*?

- Voiced.
- Voiceless.

Answer

"Voiceless"

The reason: [p] is a voiceless sound. When we pronounce it on its own, our vocal cords don't vibrate.

? Exercise 2.2.2

What is the voicing of the last sound in the word *life*?

- Voiced.
- Voiceless.

Answer

"Voiceless"

The reason: The last sound is [f], which is a voiceless sound. When we pronounce it on its own, our vocal cords don't vibrate.

? Exercise 2.2.3

What is the voicing of the last sound in the word *seem*?

- Voiced.
- Voiceless.

Answer

"Voiced"

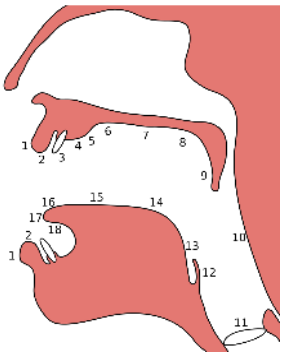
The reason: The last sound is [m], which is a voiced sound. When we pronounce it on its own, our vocal cortds vibrate.

Articulators, in Anderson's *Essentials of Linguistics*



Video Script

We know that humans produce speech by bringing air from the lungs through the larynx, where the vocal folds might or might not vibrate. That airflow is then shaped by the **articulators**.



Created by User:ish shwar (original .png deleted), .svg by Rohieb [GFDL (http://www.gnu.org/copyleft/fdl.html), CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/) via Wikimedia Commons

This image is called a sagittal section. It depicts the inside of your head as if we sliced right between your eyes and down the middle of your nose and mouth. This angle gives us a good view of the parts of the vocal tract that are involved in filtering airflow to produce speech sounds.

Let's start at the front of your mouth, with your **lips**. If you make the sound "aaaaa" then round your lips, the sound of the vowel changes. We can also use our lips to block the flow of air completely, like in the consonants [b] and [p].

We also use our **teeth** to shape airflow. They don't do much on their own, but we can place the tip of the tongue between the teeth, for sounds like [θ] and [ð]. Or we can bring the top teeth down against the bottom lip for [f] and [v].

If you put your finger in your mouth and tap the roof of your mouth, you'll find that it's bony. That is the hard **palate**. English doesn't have very many palatal sounds, but we do raise the tongue towards the palate for the glide [j].

Now from where you have your finger on the roof of your mouth, slide it forward towards your top teeth. Before you get to the teeth, you'll find a ridge, which is called the **alveolar ridge**. If you use the tip of the tongue to block airflow at the alveolar ridge, you get the sounds [t] and [d]. We also produce [l] and [n] at the alveolar ridge, and some people also produce the sounds [s] and [z] with the tongue at the alveolar ridge (though there are other ways of making the [s] sound.)

When we block airflow in the mouth but allow air to circulate through the **nasal cavity**, we get the nasal sounds [m] [n] and [ŋ].

Some languages also have nasal vowels. Make an "aaaaa" vowel again, then make it nasal. [aaaaa] [ããããã]

The articulator that you move to allow air into the nasal cavity is called the **velum**. You might also know it as the soft palate. For sounds made in the mouth, the velum rests against the back of the throat. But we can pull the velum away from the back of the throat and allow air into the nose. We can also block airflow by moving the body of the tongue up against the velum, to make the sounds [k] and [ɡ].

Farther back than the velum are the **uvula** and the **pharynx**, but English doesn't use these articulators in its set of speech sounds.

Every different configuration of the articulators leads to a different acoustic output.

Check Yourself

Exercise 2.2.4

Which articulators are responsible for the first sound in the word *minor*?

- Lips.
- Lips and teeth.
- Tongue and teeth.
- Tongue and alveolar ridge.
- Tongue and palate.
- Tongue and velum.

Answer

"Lips"

The reason: The first sound is [m], which is a bilabial sound. Only the lips are involved in that sound.

Exercise 2.2.5

Which articulators are responsible for the final sound in the word *wit*?

- Lips.
- Lips and teeth.
- Tongue and teeth.
- Tongue and alveolar ridge.
- Tongue and palate.
- Tongue and velum.

Answer

"Lips"

The reason: The first sound is [w], which is a bilabial sound. Only the lips are involved.

Note: It's actually a labio-velar sound, but more on that soon.

Exercise 2.2.6

Which articulators are responsible for the first sound in the word *photography*?

- Lips.
- Lips and teeth.
- Tongue and teeth.
- Tongue and alveolar ridge.
- Tongue and palate.
- Tongue and velum.

Answer

"Lips and teeth"

The reason: The first sound is [f], which is a labiodental sound. Both the bottom lip and the top teeth are involved.

Articulators and Airstream Mechanisms, from Sarah Harmon



Video Script

There's only a little bit more than I'm going to add to Catherine Anderson's discussion on articulators and air stream mechanisms, not much more, because what she puts up there is really good really detailed and I can't do much better so I'll let her take that wheel. There's only a couple of things I want to add.

When we talk about the air stream mechanism, yes, most all sounds that are used for human language are pulmonic, meaning that the air comes from the lungs and then goes out. It is there, and it is pushed out from the lungs all the way. There are, however, some languages that do include sounds that are not pulmonic; they can be either glottalic or velaric. Let me explain what those are.

Glottalic sounds you hear in a number of languages that are indigenous to South America; Quechua is one of them; the early version of it was spoken by the Inca Empire and continues to be spoken to this day. They have, for example, an egressive [t']. So, take a regular [t] sound, and instead of that pulmonic [t] (the is pushed all the way from the lungs), if you are stopping the air also at the glottis and ejecting it out forcefully, that is what a glottalic sound is. It's an egressive, meaning the air flow coming out of the glottis. With the egressive, the air starts from the lungs, but it is stopped also at the glottis and then ejected it out, as I say, the [t']. You can hear like a [t] but it's got more push behind it, and that is the sound. In theory, you could have an ingressive glottalic, but that doesn't really happen.

Velaric sounds, at the velum, meaning the air is stopped at the velum, the back of the mouth, and then pushed out. We have two different kinds: you have ingressive and egressive. With the egressive, air goes out and the famous one is blowing a raspberry. That is a bilabial, ingressive, fricative, velaric. Ingressive velaric, now we're talking about the very famous clicks that we hear in the Bantu languages. The Bantu languages are spoken in central and especially southern Africa. Zulu is, for example, a Bantu language; !Xhosa is a Bantu language; there's a few others. Even when I said that second language name, !Xhosa, and you heard that [click] in front, that's a click. We use clicks all the time, but maybe not to communicate within a human language. Certainly, a lot of us in Europe and the Americas, for example, when we want to call our pet will go right that little kissy sound. That the click you're pulling in air is the second comes from the back of the mouth and it's a bilabial ingressive velaric. Some of us do a different kind of velaric sound when we are disapproving of something. [tsk tsk] The tip of your tongue is up against the back of your teeth and you're sucking in, pulling in from the back of your mouth of the velum. Sometimes, some of these clicks are associated when we're trying to talk or communicate to our pack animals, like horses or oxen or donkeys. Those are all clicks.

When we're talking human language, there are some languages—and the Bantu family is one—that use clicks. It's really cool to see and hear. In the video below this one, you will see a song on YouTube. It's being sung by Miriam Makeba; she was a very famous South African singer. She is among a wide group of black artists from Africa, mostly South Africa. They're all speakers have a Bantu language and they're singing in those languages frequently. A lot of it came out in the jazz movements in the 60s and 70s; Miriam Makeba certainly was part of that. I love her music; her music is amazing and inspiring. Even if I cannot understand it most of the time, it's just the way she conveys sound. But what is really cool is she made a very common Bantu cultural song –I can't remember what exactly it's called in !Xhosa, but it is translated into French and English and etc, as "The Click Song." It's a song about a beetle; she's going to explain more in this video. What is really cool, though, is she sings that in !Xhosa, and then the spelling of the !Xhosa words is there, and under it is the English. In the captioning you can watch and read and listen all at the same time.

Click languages are really awesome; they're amazing and they are unique to the Bantu languages. There are no other language families that are recorded in human history that have the sounds. So where did they come from? That's the mystery and discussion for another time.

"Qongqothwane" by Miriam Makeba, with captions in English and !Xhosa

Unfortunately, the video that I originally had in this text was taken off of YouTube; there are other videos of the same performance, but the captions are not there. Fortunately, I found a different performance of the same song with the captions in both !Xhosa and English. However, the video quality is

quite poor.



2.2: Articulators and Airstream Mechanisms is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.4: How Humans Produce Speech** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **2.7: Articulators** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

2.3: Voicing, Resonance and Articulation

Classifying Consonants, in Anderson's *Essentials of Linguistics*



Video Script

Let's look more closely at the class of sounds we call **consonants**. Remember that consonants have some constriction in the vocal tract that obstructs the airflow, either partially or completely. We can classify consonants according to three pieces of information.

The first piece of information we need to know about a consonant is its **voicing** — is it voiced or voiceless? In the video about how humans produce speech, we felt the difference between voiced and voiceless sounds: for voiced consonants like [z] and [v], the vocal folds vibrate. For voiceless sounds like [s] and [f], the vocal folds are held apart to let air pass through.

The second thing we need to know about consonants is where the obstruction in the vocal tract occurs; we call that the **place of articulation**.

If we obstruct our vocal tract at the lips, like for the sounds [b] and [p], the place of articulation is **bilabial**.

The consonants [f] and [v] are made with the top teeth on the bottom lip, so these are called **labiodental** sounds.

Move your tongue to the ridge above and behind your top teeth and make a [t] or [d]; these are **alveolar** sounds. Many people also make the sound [s] with the tongue at the alveolar ridge. Even though there is quite a bit of variation in how people make the sound [s], it still gets classified as an alveolar sound.

If you're making a [s] and move the tongue farther back, not quite to the soft palate, the sound turns into a [ʃ], which is called **post-alveolar**, meaning it's a little bit behind the alveolar ridge. You also sometimes see [ʃ] and [ʒ] called "alveo-palatal" or "palato-alveolar" sounds because the place of articulation is between the alveolar ridge and the palate.

The only true **palatal** sound that English has is [j].

And if you bring the back of your tongue up against the back of the soft palate, the velum, you produce the **velar** sounds [k] and [g].

Some languages also have **uvular** and **pharyngeal** sounds made even farther back in the throat, but English doesn't have sounds at those places of articulation.

And of course English has a **glottal** fricative made right at the larynx, the sound [h].

In addition to knowing where the vocal tract is obstructed, to classify consonants we also need to know how the vocal tract is obstructed. This is called the **manner of articulation**.

If we obstruct the airflow completely, the sound is called a **stop**. When the airflow is stopped, pressure builds up in the vocal tract and then is released in an burst of air when we release the obstruction. So the other name for stops is **plosives**. English has two bilabial stops, [p] and [b], two alveolar stops, [t] and [d], and two velar stops [k] and [ɡ].

It's also possible to obstruct the airflow in the mouth but allow air to flow through the nasal cavity. English has three **nasal** sounds at those same three places of articulation: the bilabial nasal [m], the alveolar nasal [n], and the velar nasal [ŋ]. Because airflow is blocked in the mouth for these, they are sometimes called nasal stops, in contrast to the plosives which are oral stops.

Instead of blocking airflow completely, it's possible to hold the articulators close together and allow air to flow turbulently through the small space. Sounds with this kind of turbulence are called **fricatives**. English has labiodental fricatives [f] and [v], dental fricatives made with the tongue between the teeth, [θ] and [ð], alveolar fricatives [s] and [z], post-alveolar fricatives [ʃ] and [ʒ], and the glottal fricative [h]. Other languages also have fricatives at other places of articulation.

If you bring your articulators close together but let the air flow smoothly, the resulting sound is called an **approximant**. The glides [j] and [w] are classified as approximants when they behave like consonants. The palatal approximant [j] is made with the tongue towards the palate, and the [w] sound has two places of articulation: the back of the tongue is raised towards the velum and the lips are rounded, so it is called a labial-velar approximant.

The North American English [ɹ] sound is an alveolar approximant with the tongue approaching the alveolar ridge. And if we keep the tongue at the alveolar ridge but allow air to flow along the sides of the tongue, we get the alveolar **lateral approximant** [l], where the word *lateral* means "on the side". The sounds [ɹ] and [l] are also sometimes called "liquids"

If you look at the official IPA chart for consonants, you'll see that it's organized in a very useful way. The places of articulation are listed along the top, and they start at the front of the mouth, at the lips, and move gradually backwards to the glottis. And down the left-hand side are listed the manners of articulation. The top of the chart has the manners with the greatest obstruction of the vocal tract, the stops or plosives, and moves gradually down to get to the approximants, which have the least obstruction and therefore greatest airflow.

In *Essentials of Linguistics*, we concentrate on the sounds of Canadian English, so we don't pay as much attention to sounds with retroflex, uvular, or pharyngeal places of articulation. You'll learn more about these if you go on in linguistics. And you probably noticed that there are some other manners of articulation that we haven't yet talked about.

A **trill** involves bringing the articulators together and vibrating them rapidly. North American English doesn't have any trills, but Scottish English often has a trilled [r]. You also hear trills in Spanish, French and Italian.

A **flap** (or tap) is a very short sound that is a bit like a stop because it has a complete obstruction of the vocal tract, but the obstruction is so short that air pressure doesn't build up. Most people aren't aware of the flap but it's actually quite common in Canadian English. You can hear it in the middle of these words *metal* and *medal*. Notice that even though they're spelled with "t" and "d", they sound exactly the same when we pronounce them in ordinary speech. If you're trying hard to be extra clear, you might say [mɛtəl] or [mɛdəl], but ordinarily, that "t" or "d" in the middle of the word just becomes an alveolar flap, where the tongue taps very briefly at the alveolar ridge but doesn't allow air pressure to build up. You can also hear a flap in the middle of words like *middle*, *water*, *bottle*, *kidding*, *needle*. The symbol for the alveolar flap [ɾ] looks a bit like the letter "r" but it represents that flap sound.

When we're talking about English sounds, we also need to mention **affricates**. If you start to say the word *cheese*, you'll notice that your tongue is in the position to make a [t] sound. But instead of releasing that alveolar stop completely, like you would in the word *tease*, you release it only partially and turn it into a fricative, [tʃ]. Same thing for the word *jam*: you start off the sound with the stop [d], and then release the stop but still keep the articulators close together to make a fricative [dʒ]. Affricates aren't listed on the IPA chart because they're a double articulation, a combination of a stop followed by a fricative. English has only the two affricates, [tʃ] and [dʒ], but German has a bilabial affricate [pf] and many Slavic languages have the affricates [ts] and [dz].

To sum up, all consonants involve some **obstruction** in the vocal tract. We classify consonants according to **three pieces of information**:

- the voicing: is it voiced or voiceless,
- the place of articulation: where is the vocal tract obstructed, and
- the manner of articulation: how is the vocal tract obstructed.

These three pieces of information make up the **articulatory description** for each speech sound, so we can talk about the voiceless labiodental fricative [f] or the voiced velar stop [ɡ], and so on.

Check Yourself

Exercise 2.3.1

What is the articulatory description for the consonant sound represented by the IPA symbol [p]?

- Voiced velar nasal.
- Voiced alveolar approximant.
- Voiceless bilabial stop.

Answer

"Voiceless bilabial stop"

Hint: Look at an IPA chart and the information is there. Also, pronounce that sound on its own, and think about what your articulators are doing.

- The vocal cords don't vibrate, therefore it's voiceless;
- The only articulators involved are your lips; and,
- The air is stopped before it's released all at once.

Exercise 2.3.2

What is the articulatory description for the consonant sound represented by the IPA symbol [ð]?

- Voiceless alveolar fricative.
- Voiced dental fricative.
- Voiced alveolar nasal.

Answer

"Voiced dental fricative"

Hint: Look at an IPA chart and the information is there. Also, pronounce that sound on its own, and think about what your articulators are doing.

- The vocal cords vibrate, therefore it's voiced;
- The articulators involved are your teeth, with the tip of your tongue touching or between your teeth; and,
- The air is released in a continuous stream, with 'friction'.

Exercise 2.3.3

What is the articulatory description for the consonant sound represented by the IPA symbol [ʃ]?

- Voiced velar approximant.
- Voiceless post-alveolar fricative.
- Voiceless labiodental fricative.

Answer

"Voiceless post-alveolar fricative"

Hint: Look at an IPA chart and the information is there. Also, pronounce that sound on its own, and think about what your articulators are doing.

- The vocal cords don't vibrate, therefore it's voiceless;
- The articulator involved is your alveolar ridge, or right behind it, and your tongue is very close to that region; and,
- The air is released in a continuous stream, with 'friction'.



Video Script

Remember that the difference between consonants and vowels is that consonants have some obstruction in the vocal tract, whereas, for **vowels**, the vocal tract is open and unobstructed, which makes vowel sounds quite sonorous. We can move **the body of the tongue** up and down in the mouth and move it closer to the back or front of the mouth. We can also round our **lips** to make the vocal tract even longer.

Take a look at the IPA chart for vowels. Instead of a nice rectangle, it's shaped like a **trapezoid**. That's because the chart is meant to correspond in a very direct way with the shape of the mouth and the position of the tongue in the mouth. We classify vowels according to four pieces of information:

The high/mid/low distinction has to do with how high the tongue is in the mouth. Say this list of words:

beet, bit, bait, bet, bat

Now do the same thing, but leave off the "b" and the "t" and just say the vowels. You can feel that your tongue is at the front of your mouth and is moving from high in the mouth for [i] to fairly low in the mouth for [æ].

We can do the same thing at the back of the mouth. Say the words *boot, boat.*

Now do it again with just the vowels, [u] [o]. Your lips are rounded for both of them, but the tongue is higher for [u] than it is for [o]. The lowest vowel at the back of the mouth is [ɑ]. We don't round our lips for [ɑ], and we often drop the jaw to move the tongue low and back.

We also classify vowels according to whether the lips are rounded or unrounded. In Canadian English, there are only four vowels that have lip rounding, and they're all made with the tongue at the back of the mouth:

[u] as in *boot*

[ʊ] as in *book*

[o] as in *boat*

and [ɔ] as in *bore*

The final piece of information that we use to classify vowels is a little trickier to explain. English makes a distinction between tense and lax vowels, which is a distinction that a lot of other languages don't have. Tense vowels are made with greater tension in the muscles of the vocal tract than lax vowels. To feel this difference, say the two words *sheep* and *ship*. And now make just the vowel sounds, [i], [ɪ]. The [i] sound in *sheep* and the [ɪ] sound in *ship* are both produced with the tongue high and front, and without lips

rounded. But for [i], the muscles are more tense than for [ɪ]. The same is true for the vowels in *late* and *let*, [e] and [ɛ]. And also for the vowels in *food* and *foot*, [u] and [ʊ]

It can be hard to feel the physical difference between tense and lax vowels, but the distinction is actually an important one in the mental grammar of English. When we observe single-syllable words, we see a clear pattern in one-syllable words that don't end with a consonant. There are lots of monosyllabic words with tense vowels as their nucleus, like

day, they, weigh

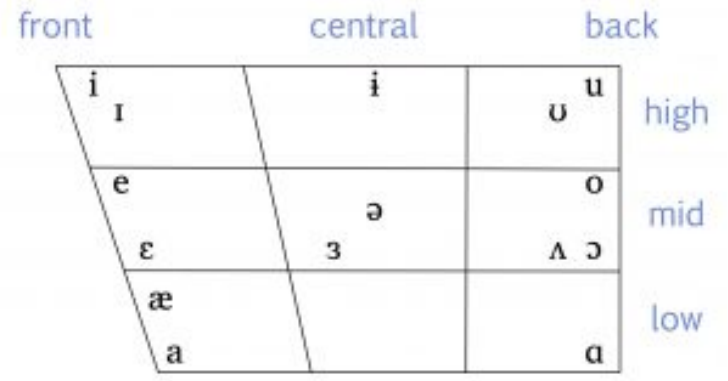
free, brie, she, tea

do, blue, through, screw

no, toe, blow

But there are no monosyllabic words without a final consonant that have a lax vowel as their nucleus. And if we were to try to make up a new English word, we couldn't do so. We couldn't create a new invention and name it a [vɛ] or a [flɪ] or a [mʊ]. These words just can't exist in English. So the tense/lax distinction is an example of one of those bits of unconscious knowledge we have about our language — even though we're not consciously aware of which vowels are tense and which ones are lax, our mental grammar still includes this powerful principle that governs how we use our language.

Here's one more useful hint about tense and lax vowels. When you're looking at the IPA chart, notice that the symbols for the tense vowels are the ones that look like English letters, while the symbols for the lax vowels are a little more unfamiliar. That can help you remember which is which!



So far, all the vowels we've been talking about are **simple vowels**, where the shape of the articulation stays fairly constant throughout the vowel. In the next unit, we'll talk about vowels whose shape changes. For simple vowels, linguists pay attention to four pieces of information:

- tongue height,
- tongue backness,
- lip rounding, and
- tenseness.

Check Yourself

? Exercise 2.3.4

What is the articulatory description for the vowel sound represented by the IPA symbol [i]?

- High front unrounded tense vowel.
- Mid central unrounded lax reduced vowel.
- High front unrounded lax vowel.
- High back rounded tense vowel.
- High back unrounded lax vowel.

Answer

"High front unrounded tense vowel"

Hint: Look at an IPA chart and the information is there. Also, pronounce that sound on its own, and think about what your articulators are doing.

- The tongue is high in the mouth, and in the front part of the mouth;
- The lips are unrounded; and,
- The tongue is tense, or straight and rigid.

? Exercise 2.3.5

What is the articulatory description for the vowel sound represented by the IPA symbol [ɛ]?

- High back rounded tense vowel.
- Mid front unrounded lax vowel.
- Mid back rounded tense vowel.
- High front unrounded tense vowel.
- Mid central unrounded lax vowel.

Answer

"Mid front unrounded lax vowel"

Hint: Look at an IPA chart and the information is there. Also, pronounce that sound on its own, and think about what your articulators are doing.

- The tongue is of middle height in the mouth, and in the front part of the mouth;
- The lips are unrounded; and,
- The tongue is lax, or relaxed.

? Exercise 2.3.6

What is the articulatory description for the vowel sound represented by the IPA symbol [ɑ]?

- High front unrounded tense vowel.
- Mid back unrounded lax vowel.
- Low back unrounded lax vowel.
- Mid back rounded tense vowel.
- Low back unrounded tense vowel.

Answer

"Low back unrounded tense vowel"

Hint: Look at an IPA chart and the information is there. Also, pronounce that sound on its own, and think about what your articulators are doing.

- The tongue is of low height in the mouth, and in the back part of the mouth;
- The lips are unrounded; and,
- The tongue is tense, or straight and rigid.

2.3: Voicing, Resonance and Articulation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.11: Classifying Consonants** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **2.12: Classifying Vowels** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

2.4: IPA and Charts

Describing Speech Sounds: the IPA, from Anderson's *Essentials of Linguistics*



Video Script

In the first part of this book, we're concentrating on the sounds of human speech. You might have already noticed that there's a challenge to talking about speech sounds — English spelling is notoriously messy.

Take a look at these words:

say, weigh, they, rain, flame, lei, café, toupee, ballet

All of them contain the same vowel sound, [e], but the sound is spelled with nine different combinations of letters. Some of them are more common ways than others of spelling the sound [e], but even if we take away the ones that English borrowed from other languages, that still leaves five different ways of spelling one sound. One of the problems is that English has only five letter characters that represent vowels, but more than a dozen different vowel sounds. But it's not just the vowels that are the problem.

English has the opposite problem as well. Take a look at these words:

cough, tough, bough, through, though

Here we've got a sequence of four letters that appear in the same order in the same position in each word, but that sequence of letters is pronounced in five different ways in English. Not only can a single sound be represented by very many different spellings, but even a single spelling is not consistent with the sounds that it represents.

Even one letter can be pronounced in a whole lot of different ways. Look at:

cake, century, ocean, and cello

The letter "c" represents four quite different sounds. Clearly, English spelling is a mess. There are a lot of reasons for why that might be.

The area where English first evolved was first inhabited by people who spoke early forms of Germanic and Celtic dialects. But then Normans invaded and brought all kinds of French and Latin words with their spellings. When the technology to print books was invented, there was influence from Dutch. So even the earliest form of English was influenced by many different languages.

Modern English also borrows words from lots of languages. When we borrow words like *cappuccino* or *champagne*, we adapt the pronunciation to fit into English but we often retain the spelling from the original language.

Another factor is that the English spelling system was standardized hundreds of years ago when it became possible to print books. A lot of our standard spellings became consistent when the Authorized Version of the Bible was published in the year 1611. Spelling hasn't actually changed much since 1611, but English pronunciation sure has, so the way we produce the sounds of English has diverged from how we write the language.

Furthermore, English is spoken all over the world, with many different regional varieties. British English sounds quite different from Canadian English, which is different from Australian English, and Indian English is quite different again, even though all of these varieties are spelled in nearly the same way.

There's even variation within each speaker of English, depending on the context: the way you speak is going to be different depending on if you're hanging out with your friends or interviewing for a job or talking on the phone to your grandmother.

The important thing to remember for our purposes is that everyone who knows a language can speak and understand it, and children learn to speak and understand spoken language automatically. So in linguistics, we say that **speaking and listening are the primary linguistic skills**. Not all languages have writing systems, and not everyone who speaks a language can read or write it, so those skills are secondary.

So here's the problem: as linguists, we're primarily interested in speech and listening, but our English writing system is notoriously bad at representing speech sounds accurately. We need some way to be able to refer to particular speech sounds, not to English letters. Fortunately, linguists have developed a useful tool for doing exactly that. It's called the **International Phonetic Alphabet**, or **IPA**. The first version of the IPA was created over 100 years ago, in 1888, and it's been revised many times over the years. The last revision was fairly recent, in 2015. The most useful thing about the IPA is that, unlike English spelling, there's no ambiguity about which sound a given symbol refers to. Each symbol represents only one sound, and each sound maps onto only one symbol. Linguists use the IPA to transcribe speech sounds from all languages.

When we use this phonetic alphabet, we're not writing in the normal sense, we're putting down a visual representation of sounds, so we call it **phonetic transcription**. That phonetic transcription gives us a written record of the sounds of spoken language. Here are just a few transcriptions of simple words so you can begin to see how the IPA works.

- *snake* [snek]
- *sugar*[ʃʊgəɹ]
- *cake*[kek]
- *cell*[sɛl]
- *sell* [sɛl]

Notice that some of the IPA symbols look like English letters, and some of them are probably unfamiliar to you. Since some of the IPA symbols look a lot like letters, how can you know if you're looking at a spelled word or at a phonetic transcription? The notation gives us a clue: the transcriptions all have **square brackets** around them. Whenever we transcribe speech sounds, we use square brackets to indicate that we're not using ordinary spelling.

You can learn the IPA symbols for representing the sounds of Canadian English in the next unit. For now, I want you to notice the one-to-one correspondence between sounds and symbols. Look at those first two words: *snake* and *sugar.* In English spelling, they both begin with the letter "s". But in speaking, they begin with two quite different sounds. This IPA symbol [s] always represents the [s] sound, never any other sound, even if those other sounds might be spelled with the letter "s". The word *sugar* is spelled with the letter "s" but it doesn't begin with the [s] sound so we use a different symbol to transcribe it.

So, **one IPA symbol always makes the same sound**.

Likewise, **one sound is always represented by the same IPA symbol**.

Look at the word *cake*. It's spelled with "c" at the beginning and "k-e" at the end, but both those spellings make the sound [k] so in its transcription, it begins and ends with the symbol for the [k] sound. Likewise, look at those two different words *cell* and *sell*. They're spelled differently, and we know that they have different meanings, but they're both pronounced the same way, so they're transcribed using the same IPA symbols.

The reason the IPA is so useful is that it's unambiguous: each symbol always represents exactly one sound, and each sound is always represented by exactly one symbol. In the next unit, you'll start to learn the individual IPA symbols that correspond to the sounds of Canadian English.

Check Yourself

? Exercise 2.4.1

The vowel sounds in the words *neat* and *spread* are both spelled “ea”. Do the vowels in the two words sound the same as each other or different?

- Same.
- Different.

Answer

"Different"

The reason: When you pronounce the vowel sound in *neat*, your tongue is high in your mouth, and is represented by [i]. When you pronounce the vowel sound in *spread* is in the middle of your mouth, height-wise, and is represented by [ɛ].

? Exercise 2.4.2

Are the final sounds in the words *face* and *mess* the same as each other or different?

- Same.
- Different.

Answer

"Same"

The reason: The sound at the end of both of those words is an s-like sound, and is represented by [s].

? Exercise 2.4.3

Are the first sounds in the two words *gym* and *gum* the same as each other or different?

- Same.
- Different.

Answer

"Different"

The reason: The first sound of *gym* is a voiced, post-alveolar africate, and is represented by [dʒ]. The first sound of *gum* is a voice, velar stop, and is represented by [g].

IPA Symbols and Speech Sounds, in Anderson's *Essentials of Linguistics*

The following tables give you some sample words so you can start to learn which IPA symbols correspond to which speech sounds. In these tables, the portion of the English word that makes the relevant sound is indicated in boldface type, but remember that English spelling is not always consistent, and your pronunciation of a word might be a little different from someone else’s. These examples are drawn from the pronunciation of mainstream Canadian English. To hear an audio-recording of the sound for each IPA symbol, consult the consonant, vowel, and diphthong charts available here.

The sounds are categorized here according to how they’re produced. You’ll learn more about these categories in units 2.6, 2.7 and 3.2.

Stops

[p]	**p**each, a**pp**le, ca**p**
[b]	**b**ill, a**b**ove, ri**b**

[t]	**t**all, in**t**ernal, ligh**t**
[d]	**d**ill, a**d**ore, ki**d**
[k]	**c**ave, ti**ck**et, lu**ck**
[ɡ]	**g**ive, ba**gg**age, di**g**

Fricatives

[f]	**ph**one, ra**ff**le, lea**f**
[v]	**v**ideo, li**v**ely, lo**v**e
[θ]	**th**in, au**th**or, ba**th**
[ð]	**th**ere, lea**th**er, brea**th**e
[s]	**c**elery, pa**ss**ing, bu**s**
[z]	**z**ebra, depo**s**it, shoe**s**
[ʃ]	**sh**ell, o**c**ean, ru**sh**
[ʒ]	**g**enre, mea**s**ure, rou**ge**
[h]	**h**ill, a**h**ead

Affricates

[tʃ]	**ch**ip, a**ch**ieve, di**tch**
[dʒ]	**j**ump, a**dj**oin, bri**dge**

Nasals

[m]	**m**ill, ha**mm**er, broo**m**
[n]	**n**ickel, su**nn**y, spoo**n**
[ŋ]	si**ng**er, wro**ng**

Approximants

[l]	**l**amb, si**ll**y, fa**ll**
[ɹ]	**r**obot, fu**rr**y, sta**r**
[j]	**y**ellow, ro**y**al
[w]	**w**inter, flo**w**er

Flap

[ɾ]	bu**tt**er, pe**d**al (*only between vowels when the second syllable is unstressed*)

Front Vowels

[i]	s**ee**, n**ea**t, p**ie**ce
[ɪ]	p**i**n, b**i**t, l**i**ck
[e]	s**ay**, pl**a**ce, r**ai**n (*in spoken Canadian English*, [e] *becomes* [eɪ])

[ɛ]	t**e**n, s**ai**d, br**ea**d
[æ]	m**a**d, c**a**t, f**a**n
[a]	f**a**r, st**a**rt

Back Vowels

[u]	p**oo**l, bl**ue**
[ʊ]	l**oo**k, g**oo**d, b**u**sh
[o]	thr**ow**, h**o**le, t**oe** (*in spoken Canadian English*, [o] *becomes* [oʊ])
[ʌ]	b**u**s, m**u**d, l**u**nch
[ɔ]	st**o**re, m**o**re, c**o**rn
[ɑ]	d**o**g, b**a**ll, f**a**ther

Central Vowels

[ə]	b**e**lieve, cinn**a**mon, s**u**rround (*in an unstressed syllable*)
[ɨ]	ros**e**s, want**e**d (*in an unstressed syllable that is a suffix*)
[ɚ]	weath**er**, edit**or** (*in an unstressed syllable with an r-quality*)
[ɝ]	b**i**rd, f**ur** (*in a stressed syllable with an r-quality*)

Diphthongs

[aɪ]	fl**y**, l**ie**, sm**i**le
[aʊ]	n**ow**, fr**ow**n, l**ou**d
[ɔɪ]	b**oy**, sp**oi**l, n**oi**se
[ju]	c**ue**, f**ew**

You can also use the following charts:

Mainstream American English Consonants

Figure 1 shows the IPA symbols for the consonants in Mainstream American English (as well as Canadian English):

	bilabial	labio-dental	dental	alveolar	post-alveolar	palatal	velar	glottal
stop	p b			t d			k g	ʔ
nasal	m			n			ŋ	
flap				ɾ				
fricative		f v	θ ð	s z	ʃ ʒ			h
approximant				ɹ		j		
lateral approximant				l				

Figure 2.4.1: Canadian English and Mainstream American English consonant chart, using IPA symbols. (Copyright; author via source)

Remember that this table lists the manner of articulation from top to bottom in order of how obstructed the vocal tract is: the greatest obstruction is for stop consonants, and the least obstruction is for approximants. And from left to right, the table depicts a place of articulation, starting at the front of the mouth with the bilabial sounds and moving all the way back to the glottis. When there are two symbols given in one cell, the one on the left is voiceless and the one on the right is voiced.

There are three speech sounds that are part of Canadian English that don't fit neatly into this table. Remember that affricates have a two-part manner of articulation: they begin with a complete obstruction of the vocal tract, but that obstruction is released only partially. You can think of an affricate like a stop combined with a fricative, and the symbols that we use to transcribe them reflect that. Because they have a two-part manner of articulation, the affricates [tʃ] and [dʒ] don't fit into the consonant chart. The other sound that doesn't fit on the chart is the approximant [w]. It has two places of articulation: the lips are rounded, and the body of the tongue moves towards the velum. So the IPA's name for the segment [w] is a **labial-velar approximant**.

Mainstream American English and Canadian English Vowels

Figure 2 shows the IPA symbols for the vowels of Canadian English:

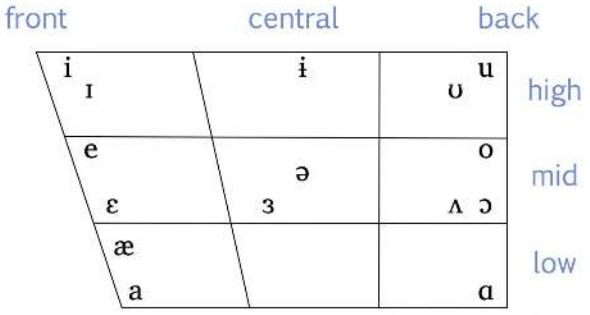


Figure 2.4.2: Canadian English vowel chart, using IPA Symbols. (Copyright; author via source)

Mainstream American English has the exact same vowels, with the exception of [ɜ] and [ɑ].

Remember that the vowel trapezoid is meant to correspond to the position of the tongue in the mouth.

Broad and Narrow Transcription, in Anderson's *Essentials of Linguistics*

Learning to use the IPA to transcribe speech can be very challenging, for many reasons. One reason we've already talked about is the challenge of ignoring what we know about how a word is spelled to pay attention to how the word is spoken. Another challenge is simply remembering which symbols correspond to which sounds. The tables in Units 2.4 and 3.2 may seem quite daunting, but the more you practice, the better you'll get at remembering the IPA symbols.

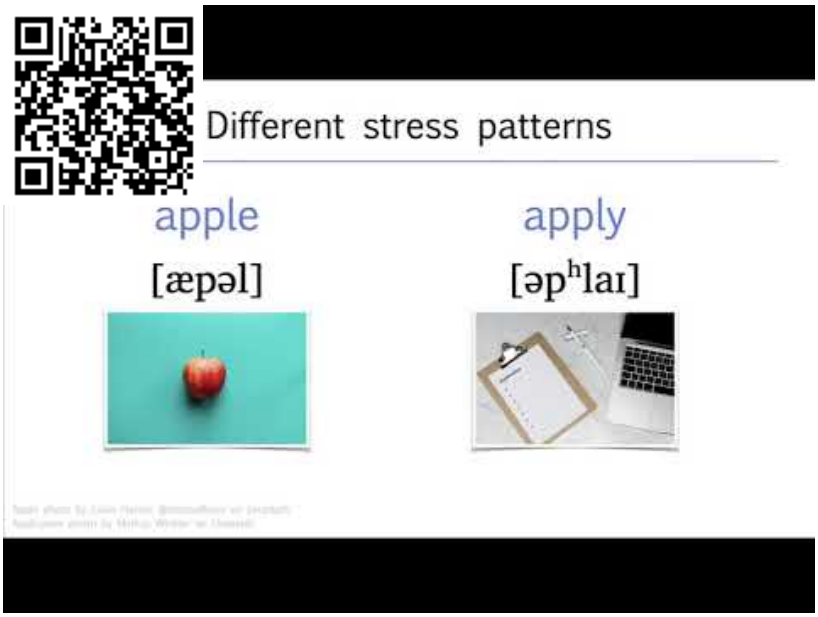
A challenge that many beginner linguists face is deciding exactly how much detail to include in their IPA transcriptions. For example, if you know that Canadian English speakers tend to diphthongize the mid-tense vowels [e] and [o] in words like *say* and *show*, should you transcribe them as the diphthongs [eɪ] and [oʊ]? And the segment [p] in the word *apple* doesn't sound quite like the [p] in *pear*; how should one indicate that? Does the word *manager* really begin with the same syllable that the word *human* ends with?

Part of learning to transcribe involves making a decision about exactly how much detail to include in your transcription. If your transcription includes enough information to identify the place and manner of articulation of consonants, the voicing of stops and fricatives, and the tongue and lip position for vowels, this is usually enough information for someone reading your transcription to be able to recognize the words you've transcribed. A transcription at this level is called a **broad transcription**.

But it's possible to include a great deal more detail in your transcription, to more accurately represent the particulars of accent and dialect and the variations in certain segments. A transcription that includes a lot of phonetic detail is called a **narrow transcription**. The rest of this chapter discusses the most salient details that would be included in a narrow transcription of the most widespread variety of Canadian English.

Transcribing Vowels in Canadian English, in Anderson's *Essentials of Linguistics*

(Note: This is an excellent explanation of how to transcribe vowels, and is meant to give you an idea of what to do. The next subsection below will also have how to focus on Mainstream American English, as well as something closer to what we use in California.)



Video Script

Canadian English and American English have a lot in common, but some of the more noticeable differences between them are in how vowels get pronounced. US textbooks often classify this symbol that looks like a lower-case [a] as a low back vowel, but you probably noticed in an earlier section that this book puts [a] in the low front position and lists this character, known as the script [ɑ], in the low back position. What's the difference between them? To figure this out, let's look at a pair of words.

How do you pronounce these two words?

I'm a fairly typical speaker of middle-class, middle-aged Canadian English, and for me, these two words are homophones: they sound exactly the same, whether I'm saying, "*I caught the ball*" or "*I slept on a cot*". Caught/cot. So my dialect has what Linguistics calls the "caught/cot merger"". This is pretty typical of Canadian English, but not of many other varieties of English.

In many varieties of US English, the past tense of the verb *catch* is pronounced with the low-back [ɑ] like mine, but the noun uses the low front [a], [kʰat]. I like to think of this low front vowel as the Chicago White Sox vowel, because it's very typical of Chicago English, where the south-side baseball team is known as the [saks].

And in many of the varieties of English spoken in the UK, the past tense of *catch* has a mid-back rounded vowel, [kʰɔt], while the noun version has the low back vowel [kʰɑt].

Since this book concentrates on Canadian English, I'm going to suggest the following convention. When transcribing Canadian English, use the low back vowel that we represent with the script a [ɑ] for the simple vowel in words like *father, box, and log*.

The low-front [a] doesn't usually appear in Canadian English as a simple vowel. It shows up in the major diphthongs, [aɪ] like in *fly*, and [aʊ] as in *brown*, and it turns up before an [ɹ] as in *car* or *farther*.

And this mid-back rounded vowel also doesn't show up as a simple vowel in Canadian English. It appears in the diphthong [ɔɪ] like in *coin*, and before [ɹ] like *fork* or *short*.

Note that this convention that I'm suggesting does over-simplify the variation that exists in the real world, but that's ok at this introductory level.

With the low vowels taken care of, let's talk about the mid and central vowels. Think about these two words, *funny* and *phonetics*. They start with almost the same syllable, don't they? But there's one important difference. In *funny*, the stress is on the first syllable, while in *phonetics*, the second syllable is stressed. This has consequences for the vowel in the nucleus of each syllable.

Remember that we talked about syllable stress in an earlier section. Stressed syllables are more prominent than unstressed syllables. They're louder, longer, and higher in pitch. So stressed syllables get pronounced with full vowels. In the word *funny*, the first syllable has this mid-back unrounded vowel that we represent with the wedge symbol [ʌ]. But when a syllable is unstressed, it gets reduced — it's shorter and quieter. That means that speakers pronounce the vowel with the mid-central reduced vowel that has the funny name **schwa** [ə] and looks like an upside-down e.

You can hear this stress difference between the "uh" sound in *bun* and the first syllable of *banana*, and between the first syllables of *apple* and *apply*. So being able to recognize stress is important in producing an accurate transcription.

Speaking of unstressed syllables, in an earlier section we talked about what happens when a syllable is so reduced that the vowel nucleus disappears entirely and the sonorant consonant from the coda becomes a vowel. This is another oversimplification, but at this intro level let's treat a syllable with a reduced nucleus as equivalent to one that has a syllabic consonant as its nucleus. So we'll consider these transcriptions on the left as equivalent to the ones on the right.

Many words in English end with an unstressed *-er* syllable, so [ɹ] is maybe the sonorant that becomes syllabic the most frequently. It just so happens that the IPA also has a way of transcribing a vowel that takes on a **rhotic**, or r-like quality. So in addition to considering the schwa-r transcription as equivalent to the syllabic-r, we'll also consider the rhotic-schwa transcription to be equivalent. Again, this is an oversimplification, but it makes sense for when you're first learning to do IPA transcription.

Now, you already know that schwa is a reduced vowel that only appears in unstressed syllables, so what about when a word has that "er" sound in a stressed syllable, like in *bird*? We'll use a different symbol to represent the stressed one: this is the mid-central, unreduced vowel. Notice that looks a lot like the symbol for the mid-front [ɛ] vowel, but it faces the opposite direction!

The only place this vowel shows up in Canadian English is when it's rhotic, that is, in a stressed syllable with an [ɹ] in the coda. You can hear the difference if you compare the word *bird* with the second, unstressed syllable in *amber*.

There's one more central vowel we need to pay attention to, and that's the high central vowel that looks like a little crossed-out lower-case letter i. Again, it has a very predictable distribution in Canadian English. We'll use it only in unstressed syllables where the syllable is a suffix. For example, the final syllable in *heated* and *excited* is a past-tense suffix, and the final syllable in *horses* and *quizzes* is a plural suffix, so we'll transcribe them all with this high central vowel.

That's a lot of details! But paying attention to these subtle differences will develop your phonetic listening skills, and your transcription skills.

Check Yourself

Exercise 2.4.4

In Mainstream American English, what vowel is in the nucleus of the first syllable of the word *delay*?

- [e]
- [ɜ]
- [ɨ]
- [ə]

Answer

"[ə]"

Hint: When you pronounce that vowel, your tongue is in the middle of the mouth, somewhat towards the back.

Exercise 2.4.5

In Mainstream American English, what vowel is in the nucleus of the last syllable of the word *ended*?

- [ɛ]
- [ə]
- [ɨ]
- [ɜ]

Answer

"[ɨ] for some; [ə] for others"

Hint: When you pronounce that vowel, your tongue is in the central part of your mouth. For many American dialects, it's in the same place as the previous example; for others, especially in the Western U.S., the vowel is a bit higher.

Exercise 2.4.6

In Mainstream American English, what vowel is in the nucleus of the first syllable of the word *dusty*?

- [ə]
- [ʊ]
- [ʌ]
- [u]

Answer

"[ʌ]"

Hint: When you pronounce that vowel, your tongue is towards the back of the mouth, and about midway up in height. Your tongue is also relaxed

IPA and Charts, from Sarah Harmon



Video Script

Catherine Anderson does an amazing job in describing why we have IPA, the International Phonetic Alphabet, in the first place and describing the sounds. Of course, she's Canadian and so the charts that she used are all Canadian Standard English. I figured it would be important to bring in the Mainstream American English, just so that you have a little bit of context.

	Bilabial	Labiodental	Dental	Alveolar	Post-Alveolar	Palatal	Velar	Glottal
Stop	p b			t d			k g	ʔ
Nasal	m			n		(ɲ)	ŋ	
Fricative		f v	ϴ ð	s z	ʃ ʒ		(x)	h
Affricate						ʧ ʤ		
Lateral				l			L	
Rhotic				ɾ (r) ɹ				
Glide, Semi-cons.	ʍ w					j	ʍ w	

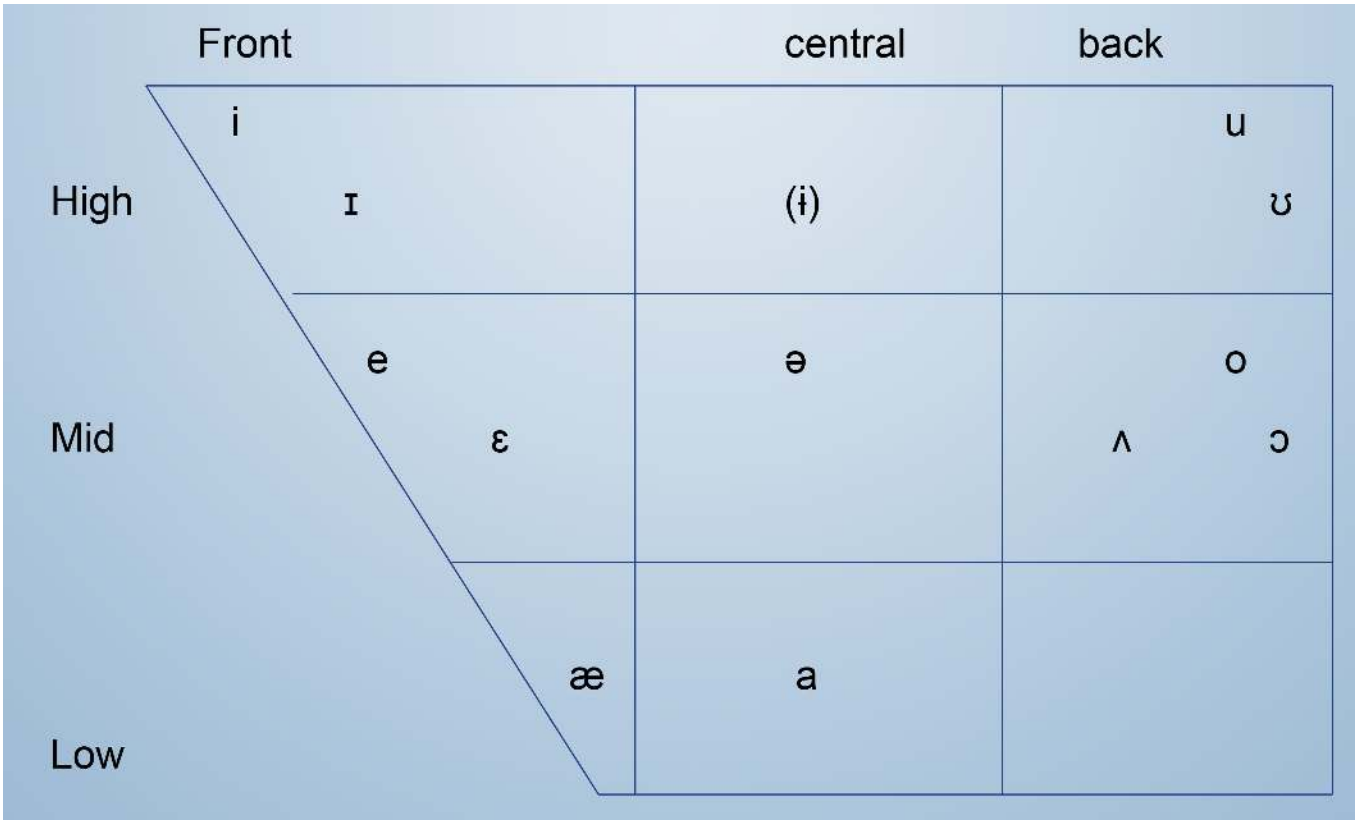
One thing to note is that for the consonants are going to be identical; Canadian English, American English, it doesn't matter. The constants are exactly the same. What is interesting is that I would argue in Mainstream American English—or Standard American English, if you wish, but that phrasing is going out of style and for a number of reasons—has added a few other sounds. It's exactly because of its proximity to Spanish; we have so many native Spanish speakers here in the United States. In major population areas

like the Bay Area, and most of California, you can argue certainly the Phoenix area, and most of Arizona, New Mexico, Texas, and throughout the entirety of the US, there are certain sounds that are starting to migrate in as official sounds of English. For example, that piece of flatbread that you have with Mexican food, you don't call it a 'tortila'; you call it a ‘tortilla’. Now, you may not go full Spanish pronunciation of 'tortilla', but you certainly don't say 'tortila'. A really common chili pepper that is used in a lot of cooking, including Mexican cooking, is not a 'jalapano'. Nobody calls it a 'jalapano'. You at the very least say ‘jalapeño’ and you do that [ɲ] to go with it. Sometimes you do a velar [x] to go with it, because in Spanish, of course, it is a 'jalapeño'.

There are certain sounds that have started to creep into Mainstream American English, and those are the ones you see in parentheses. You have the palatal [ɲ]; for example, I teach not at Canada College but at Cañada College, and anybody who's in the San Francisco Bay Area, especially in San Mateo County, you know that is Cañada College; you don't call it Canada College. That, for example, that have that we here in jalapeño in the Spanish pronunciation of that term, some people do actually use that velar fricative of instead of the glottal [h] sound. The trill [r] that you hear in Spanish, and in so many other languages, is also starting to creep in; it's not entirely they're not as entrenched just say that palatal nasal but certainly the rest of them are there.

As far as whether you call these **glides, semi consonants,** or **approximates**, they're all kind of the same thing. If you get into hardcore phonetics, then you start understanding the differences between those terms, but for here I will call them a glide or an approximate, and she tends to call them an approximate. Sometimes you will see them as semi consonants. They're all interchangeable terms as far as you're concerned. Of course, both lateral and rhotic approximates tend to be separated out on most IPA charts. I don't think she did it in hers, but I am doing it in mine, so that you can see.

In English, whether it's Canadian English American English or just about any other English, the consonants are pretty much the same. What is different is the vowels. So, you had the Canadian English vowel chart above in the Catherine Anderson work, this is the Mainstream American English sound chart for the vowels.



The barred 'i' [ɨ] that central high vowel that it is in parentheses, because not all dialects have it, although it is becoming more common west of the Mississippi River. But pretty much all of these sounds are in existence in some way, shape or form in Mainstream American English and in most American English dialects, though there is some difference, and we'll come back to that when we talk about sociolinguistics and different dialects.

What you can see here is that there is a slew of English vowel sounds. Its enormous. Most languages don't have that many sounds; they might have typically five to seven. With English, it depends on what dialect you're talking about; English can have anywhere from seven to 14 different vowel sounds. Those are just the single vowels, not even counting the diphthongs, triphthongs, quadrathongs; those would be the drawls that you see and hear.

This is a real issue when we're talking about English. For those of you for whom English is not your native language and is still a language you are acquiring, this no doubt is the hardest part about the English sound system. And if American English was not the first dialect of English that you came across, it gets complicated even more. Those would be the big differences between the various dialects of English. We'll come back to this more later when we get to sociolinguistics.

2.4: IPA and Charts is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.8: Describing Speech Sounds- the IPA** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **3.2: IPA for Canadian English** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

2.5: Sound Classes and Hierarchy

2.5.1 From 2.5 Sonority, Consonants, and Vowels, in Anderson's *Essentials of Linguistics*



Video Script

Remember that there are three steps involved in producing speech sounds. The process starts with respiration as air flows up from the lungs. Phonation occurs at the larynx, where the vocal folds may or may not vibrate to produce voicing, and then we use our mouth, jaw, lips, teeth and tongue to shape the sound, which is called articulation.

In phonetics, we classify sounds according to how they're produced, and also according to the acoustic properties of the sounds. The primary acoustic property that we're interested in is called **sonority**. Sonority has to do with the amount of **acoustic energy** that a sound has. A simple example of this is that a loud sound is more sonorous and a quiet sound is less sonorous. But sonority is not just about loudness. Sounds that are made with lots of airflow from the lungs, and with vocal folds vibrating, are sonorous sounds. Sounds that have less airflow or don't have voicing from the vocal folds have less sonority. Those two pieces of information, sonority and articulation, allow us to group sounds into three broad categories

We produce **vowels** with the vocal tract quite open and usually with our vocal folds vibrating so vowels have a lot of acoustic energy: they're sonorous. Vowel sounds can go on for a long time: if you're singing, when you hold the note, you hold it on the vowels. Make some vowel sounds and notice how you can hold them for a long time: "aaaaa iiiii uuuuu".

The sounds that we call **consonants** are ones where we use our articulators to obstruct the vocal tract, either partially or completely. Because the vocal tract is somewhat obstructed, less air flows from the lungs, so these sounds have less energy, they're less sonorous, and they're usually shorter than vowels. Consonant sounds can be voiced or voiceless.

There's also an intermediate category called **glides** that have some of the properties of vowels and some of the consonants. The vocal tract is unobstructed for glides, like for vowels, but they are shorter and less sonorous than vowels. We'll learn more about glides when we take a closer look at vowels.

This acoustic notion of sonority plays a role in every language of the world because spoken words are organized around the property of sonority. Every single spoken word is made up of one or more syllables. You probably know that a **syllable** is like a beat in the rhythm of the word, so you know that *ball* has one syllable, *basket* has two syllables, and *bicycle* has three.

But what is a syllable, in phonetic terms? **A syllable is a peak of sonority that is surrounded by less sonorous sounds**. What that means is that a syllable is made up of a vowel, or some other very sonorous sound, with some sounds before it and after it that are less sonorous, usually glides and consonants. The most sonorous sound, the peak of sonority, is called the **nucleus** of a syllable.

Looking back at those words, we can see that the word *ball* contains the sonorous vowel sound [ɑ], with two less-sonorous consonants, [b] and [l] on each side of it. Likewise, *basket* has two vowel sounds [æ] and [ɪ], with the consonants [b] before the first syllable, [sk] between the two vowels, and [t] after the second vowel. Can you figure out what the vowel and consonant sounds are in the word *bicycle*? Remember that written letters don't necessarily map directly onto speech sounds!

Check Yourself

? Exercise 2.5.1

What kind of sound is the first sound in the word *early*?

- Vowel.
- Glide.
- Voiced consonant.
- Voiceless consonant.

Answer

"Vowel"

Hint: There is zero touching or approximation of the tongue to any articulator.

? Exercise 2.5.2

What kind of sound is the first sound in the word *junior*?

- Vowel.
- Glide.
- Voiced consonant.
- Voiceless consonant.

Answer

"Voiced consonant"

Hint: When you pronounce that sound on its own, your vocal cords vibrate, and your tongue is touching certain articulators.

? Exercise 2.5.3

What kind of sound is the first sound in the word *winter*?

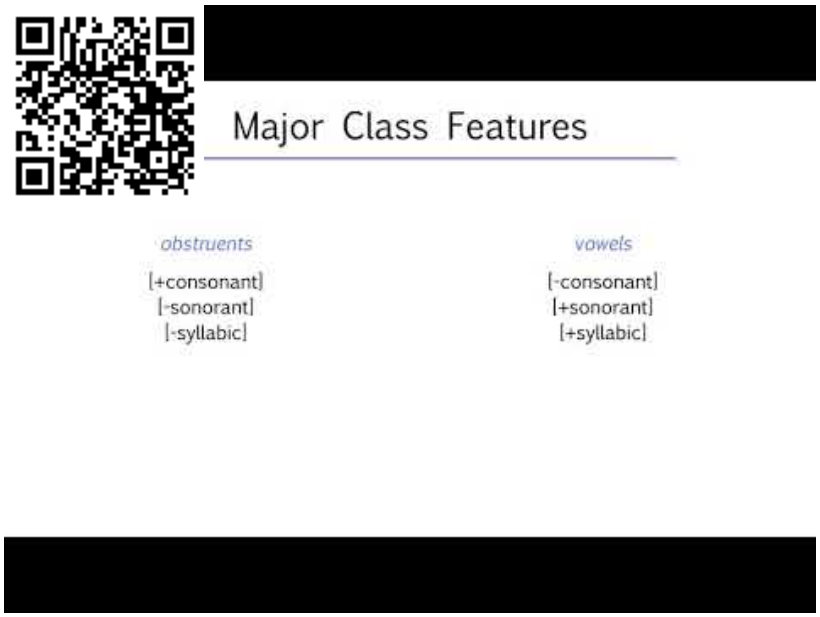
- Vowel.
- Glide.
- Voiced consonant.
- Voiceless consonant.

Answer

"Glide"

Hint: When you pronounce that sound on its own, your tongue glides around in your mouth.

2.5.2: From 4.4 Natural Classes, in Anderson's *Essentials of Linguistics*



Video Script

We saw in the last unit that a **natural class** is a group of speech segments that have some features in common. Because they share features, they also tend to behave similarly to each other in the grammar of a language. We also learned how to use the notation of a **feature matrix** to describe the features that the members of a natural class have in common. Remember that the more features we list, the smaller the natural class gets, and the fewer features we list, the larger the class we're describing.

The two largest natural classes can be described with just one feature! All segments that have an obstruction in the vocal tract are [+consonant], and the vowels, which have no obstruction in the vocal tract, are [-consonant]. Notice that the feature *consonant* is binary, meaning it has exactly two values: plus or minus. This means that we don't need a separate feature to label the sounds that are vowels: using the binary feature notation we can just say that consonants are [+consonant] and all vowels are [-consonant].

Now, why did I say that most consonants are [+consonant]? Could there be consonants that aren't totally consonants? Think back to when we first started learning to categorize sounds. Remember we said the glides /j/ and /w/ have very little obstruction in the vocal tract. So even though /j/ and /w/ often behave like consonants, they get labelled as [-consonant] in the feature system. This is how we indicate that also they share some properties with vowels, namely, the property of having the vocal tract unobstructed.

The next feature that groups segments into the major classes has to do with sonority, or acoustic energy. Because of the relatively open vocal tract, all vowels are [+sonorant]. But not all consonants are [-sonorant]. Approximants also have a relatively open vocal tract, so all glides and liquids are also [+sonorant]. And when air circulates through the nasal cavity, that also creates a lot of acoustic energy so all the nasal consonants are also [+sonorant]. All other sounds, that is, stops, fricatives, affricates, and so on, are minus sonorant because they have lower sonority. Informally, we call the non-sonorant class of sounds **obstruents**, but we don't need a separate feature label for obstruents because of how the binary feature system works: we just identify them as [-sonorant].

The third major feature groups sounds according to whether they can be the nucleus of a syllable or not. There's a lot of overlap between sounds that are [+sonorant] and [+syllabic], but the two classes aren't totally identical. I'm sure you can already guess that all vowels have the feature [+syllabic]. By default, all consonants are [-syllabic], but remember that some consonants can sometimes serve as the nucleus of a syllable — do you remember which ones? The liquids and nasals become [+syllabic] only when they're the nucleus of a syllable.

These three features are called the **major class features** because they allow us to group segments into these broad categories. At one end of the spectrum, we have the vowels, with no obstruction in the vocal tract, so they're sonorant, and they can serve as the nucleus of a syllable. At the other end, we have the obstruents: these are the consonants that have low sonority and can't be the nucleus of a syllable. And then in between, we have the sonorant consonants, which are usually [-syllabic] but can become [+syllabic], so if we don't list that feature, we indicate that it can be either plus or minus. And then there are the glides, which are almost like vowels except they're [-syllabic]. There are a few details to notice here: the nasals and liquids are consonants that are

sonorant, and that can become syllabic. The glides are sonorants, but they don't serve as the nucleus of a syllable, and they don't even count as consonants!

The next set of features we should consider are those that describe manners of articulation. The *continuant* feature has to do with whether air is allowed to flow in the oral cavity (that is, in the mouth). So the main job the *continuant* feature does is to distinguish between the stops, which are [-continuant] because airflow is blocked, and all other sounds. Notice that the nasals count as [-continuant] along with the plosive stops because they involve a complete closure in the mouth.

Speaking of *nasals*, the feature [+nasal] labels nasal sounds. /m n ŋ/ are obviously [+nasal], and in English, all other sounds are [-nasal] by default. But sometimes assimilation leads a vowel to take on a nasal feature. In the word *pants*, for example, the [æ] vowel becomes nasalized in anticipation of the following nasal sound, which we would label as a change from [-nasal] to [+nasal].

The *voice* feature is pretty obvious: if the vocal folds are vibrating, then a segment is [+voice]. So all the voiced consonants are [+voice] and the voiceless consonants are [-voice]. In English, the vocal folds vibrate for vowels, so the vowels are also [+voice].

Now let's look at the features that describe which articulators are active in producing a sound. Notice that the notation for the place features is different: these ones are not binary features so there's no plus or minus, and they're written in all capital letters to distinguish them from the binary features. The idea is that there's no known language that makes meaningful phonological distinctions on the basis of what a sound's place of articulation is NOT, so we don't need a minus-value for these features. The feature [LABIAL] identifies any sound that involves the lips. [CORONAL] classifies sounds that are produced using the tip of the tongue, and [DORSAL] classifies sounds that are produced with the body and back of the tongue.

Let's start by looking at labial sounds. If a sound doesn't involve the lips at all, then we don't even list [LABIAL] in its feature matrix. But if it does involve the lips, then we specify whether the lips are [+round], like the rounded vowels and the glide [w], or [-round], like for the bilabial and labio-dental consonants.

We label sounds that are made with the tip of the tongue with the feature [CORONAL], and then we can classify them further according to where the tip is. The word *anterior* is just a fancy word for "front". The dental and alveolar sounds are made with the tip of the tongue towards the front of the mouth and are [+anterior]. Post-alveolar sounds have the tip of the tongue pointing farther back so they're [-anterior]. There's another distinction for the coronal sounds. We use the label [+strident] for the sounds that are acoustically noisy and sound like hissing, that is, [s z ʃ ʒ tʃ dʒ]. The other coronals are [-strident]. You'll see why this distinction matters in one of the exercises at the end of this chapter.

Sounds that are made not with the tip of the tongue but with the tongue body have the feature [DORSAL]. Only a few consonants in English have the feature [DORSAL]: the velar sounds [k ɡ ŋ] and the glides [j w]. And because the body of the tongue is the primary articulator that we use to make different vowel sounds, all vowels have the feature [DORSAL].

Once we've identified a sound as a dorsal sound, then we use binary features to specify the position of the tongue body. We specify vowel height with the features [±high] and [±low]. All high sounds are [-low], and all low sounds are [-high], but there are also vowels that are both [-low] and [-high]: the mid vowels! We identify the front or back position of the tongue with [±back], but we don't need to include a [±front] feature because these central vowels pattern with the natural class of back vowels.

There's one other feature we need for categorizing the vowels of English, and that's the tense/lax distinction. We label the tense vowels as [+tense] and the lax vowels as [-tense].

There's a full chart of features for the segments of English presented below, but don't be intimidated by it. A feature matrix is just a more organized way of presenting the information that you already learned about phonetic segments. If you start to think of segments in terms of natural classes and what features the natural class has, you'll start to get the hang of it.

Attributes of various consonant sounds

	p	b	t	d	k	ɡ	f	v	s	z	θ	ð	ʃ	ʒ	m	n	ŋ	l	ɹ	j	w
Major Class Features																					
[consonant]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–

	p	b	t	d	k	g	f	v	s	z	θ	ð	ʃ	ʒ	m	n	ŋ	l	ɹ	j	w
[sonorant]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+
[syllabic]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Manner Features																					
[nasal]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	–	–	–	–
[continuant]	–	–	–	–	–	–	+	+	+	+	+	+	+	+	–	–	–	+	+	+	+
[lateral]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–
[voice]	–	+	–	+	–	+	–	+	–	+	–	+	–	+	+	+	+	+	+	+	+
Place Features																					
LABIAL	✓	✓					✓	✓							✓						✓
[round]	–	–					–	–							–						+
CORONAL			✓	✓					✓	✓	✓	✓	✓	✓		✓		✓	✓		
[anterior]			+	+					+	+	+	+	–	–		+		+	+		
[strident]			–	–					+	+	–	–	+	+		–		–	–		
DORSAL					✓	✓											✓			✓	✓
[high]					+	+											+			+	+
[back]					+	+											+			–	+

Attributes of various vowel sounds

	i	ɪ	e	ɛ	æ	a	u	ʊ	o	ʌ	ɔ	ɑ
Major Class Features												
[consonant]	–	–	–	–	–	–	–	–	–	–	–	–
[sonorant]	+	+	+	+	+	+	+	+	+	+	+	+
[syllabic]	+	+	+	+	+	+	+	+	+	+	+	+
Manner Features												
[continuant]	+	+	+	+	+	+	+	+	+	+	+	+
[voice]	+	+	+	+	+	+	+	+	+	+	+	+
Place Features												
LABIAL							✓	✓	✓		✓	
[round]							+	+	+		+	
DORSAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
[high]	+	+	–	–	–	–	+	+	–	–	–	–
[low]	–	–	–	–	+	+	–	–	–	–	–	+
[back]	–	–	–	–	–	–	+	+	+	+	+	+
[tense]	+	–	+	–	–	+	+	–	+	–	–	+

Check Yourself

? Exercise 2.5.4

In the set of segments listed below, which segment must be excluded to make the remaining segments constitute a natural class?

- [i]
- [e]
- [æ]
- [o]
- [u]

Answer

"[æ]"

Hint: All the other sounds are either high or mid in height.

Exercise 2.5.5

In the following set of segments, which segment must be excluded to make the remaining segments constitute a natural class?

- [p].
- [f].
- [t].
- [k]

Answer

"[f]"

Hint: All the other sounds are stops.

Exercise 2.5.6

This set of segments constitutes a natural class: [i ɛ æ]. Which segment could you add to the set while still preserving the natural class?

- [ɪ].
- [ɔ].
- [ə].

Answer

"[ɪ]"

Hint: [i ɛ æ] are all front vowels.

2.5.3 Sound Classes and Hierarchy, from Sarah Harmon



Video Script

With respect to sound classes and hierarchy, Catherine Anderson again does a really great job of explaining pretty much everything there's only one tiny piece that I'm going to add, and this is something we're going to come back to when we get to sociolinguistics and historical linguistics and language acquisition.

She talked about different **sound classes**. I'll give you a few more when we talk about labels or coronals interiors or sibilants. These all have common features; these sounds have common features:

- Labials are produced using the lips in some way
- Sibilants are always 's'-like sounds.
- Coronals use the corona of the tongue, so that's crown part of your tongue
- Interiors are any sound produced from the alveolar ridge forward.
- And posteriors, by the way, would be everything from the palatal region back

There's lots of ways to think about classification of sounds, and this will come into play when we talk about how these sounds combined in different languages.

The only other thing to bring up has to do with **sound hierarchy**. Again, we'll come back to this later, but there is a hierarchy of sounds meaning from the most occlusion to the least occlusion. **Occlusion** means stuff getting in the way. Think about all that has to be done when you have to make a voiceless stop like a [t, p, k]. With those sounds, you have to stop your glottis from moving, and you are stopping the airflow before releasing it all. That takes a lot of effort on the part of your articulators. Because of that, they tend to be the 'hard sound', as it were. As we go down that list you get softer and softer as to how sounds are produced.

This comes to play in a few ways and we'll talk about them pretty soon when we get to phonology, and then more when we talk about how phonology gets applied to the everyday world. Suffice it to say that this sound hierarchy might explain, for example, if you come from a language that has a lot of fricatives and you're trying to learn a language that has a lot of stops, why that might be a little more difficult. Overall, this is something that will definitely come into play as we continue down our linguistic path.

2.5: Sound Classes and Hierarchy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.10: Sonority, Consonants, and Vowels** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **4.5: Natural Classes** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

2.6: All About Vowels

2.6.1 From 2.8 Diphthongs, in Anderson's *Essentials of Linguistics*



Video Script

The last unit talked about simple vowels, where the tongue position stays pretty constant throughout the duration of the vowel. In addition to simple vowels, many languages include **diphthongs**, where we move our articulators while producing the vowel. This gives the sound a different a different shape at the end from how it begins. The word *diphthong* comes from the Greek word for "two sounds".

There are **three major diphthongs** in English that have quite a noticeable change in the quality of the vowel sound.

Say these English words out loud: *fly, tie, ride, smile*. Now make the vowel sound [aɪ] again but hold it at the beginning [aaa]. The first part of the sound is the low front [a], but then the tongue moves up quickly at the end of the sound, ending it [ɪ]. So the [aɪ] sound is a diphthong, and it gets transcribed with two consecutive symbols:[aɪ].

In the words *now, loud, brown*, the tongue again starts low and front [a], and then it moves high and to the back of the mouth, and the lips get rounded too! The second part of this diphthongs is but the high back rounded [ʊ]. The [aʊ] diphthong is transcribed like this: [aʊ].

The third major diphthong in English occurs in words like *toy, boil, coin*. It starts with the tongue at the back of the mouth and lips rounded [ɔ], then moves to the front with lips unrounded. It is transcribed like this: [ɔɪ].

Some linguists also consider the vowel sound in *cue* and *few* to be a diphthong. In this case, the vowel sound starts with the glide [j] and then moves into the vowel [u].

In addition to these major English diphthongs, speakers of Canadian English also have a tendency to turn the mid-tense vowels into diphthongs.

For example, let's look at the pair of vowels [e] and [ɛ] from the words *gate* and *get*. They're both mid, front, unrounded vowels, but [e] is tense – it's made with greater tension in the muscles of the vocal tract than [ɛ]. Canadian English speakers pronounce the lax vowel in *get* as a simple vowel [ɡɛt], but for the tense vowel, we tend to move the tongue up at the end: [ɡeɪt]. We do it so systematically that it's very hard for us to hear it, but it's always there.

We do the analogous thing for the mid-back vowel [o] like in *show* and *toe*: at the end of the [o] vowel, the tongue moves up a little bit so we produce the vowel as [oʊ]. Notice that the lips are rounded for both parts of this diphthong.

To sum up, **a diphthong is a vowel sound that involves movement of the tongue from one position to another**. Nearly all dialects of English include the three major diphthongs [aɪ] , [aʊ] , and [ɔɪ]. These ones are called the major diphthongs because they involve large movements of the tongue.

In Canadian English, speakers also regularly produce diphthongs for the tense vowels, [eɪ] and [oʊ], but not all English dialects do this. Some linguists consider these ones to be **minor diphthongs**.

Check Yourself

Exercise 2.6.1

What is the diphthong sound in the word in the word *proud*?

- [ɔɪ]
- [oʊ]
- [aʊ]
- [aɪ]
- [eɪ]

Answer

"[aʊ]" (also known as [aw])

Hint: If you break down the vowel sounds in that diphthong, the first one is [a], and the second one is [ʊ]. It may sound more like the glide [w] for many Mainstream American English speakers.

Exercise 2.6.2

What is the diphthong sound in the word in the word *rain*?

- [ɔɪ]
- [eɪ]
- [oʊ]
- [aɪ]
- [aʊ]

Answer

"[eɪ]" (also known as [ej])

Hint: If you break down the vowel sounds in that diphthong, the first one is [3], and the second one is [ɪ]. It may sound more like the glide [j] for many Mainstream American English speakers.

Exercise 2.6.3

What is the diphthong sound in the word in the word *sigh*?

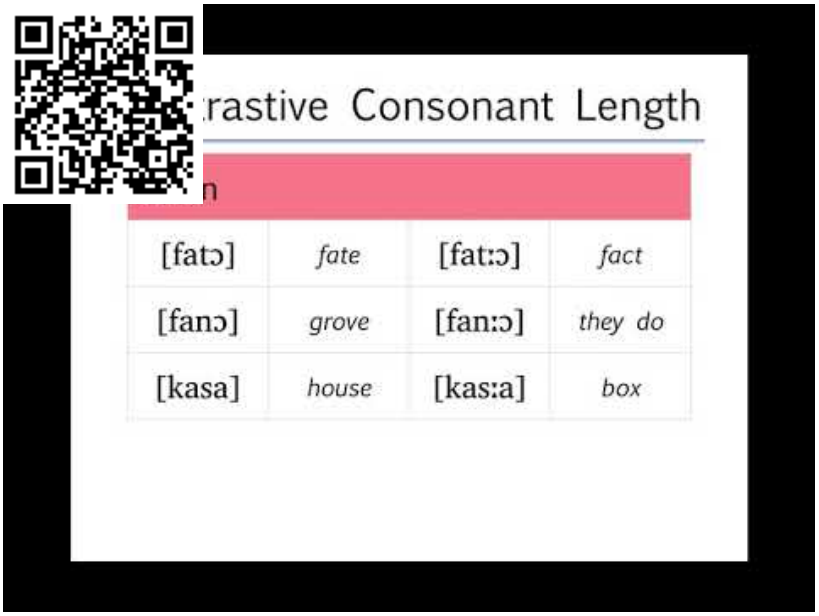
- [aɪ]
- [aʊ]
- [oʊ]
- [eɪ]
- [ɔɪ]

Answer

"[aɪ]" (also known as [aj]

Hint: If you break down the vowel sounds in that diphthong, the first one is [a], and the second one is [ɪ]. It may sound more like the glide [j] for many Mainstream American English speakers.

2.6.2: From 3.3 Stress and Suprasegmentals, in Anderson's *Essentials of Linguistics*



Video Script

So far all the sounds we've been considering are segments: the individual speech sounds that we represent with IPA symbols. But when we speak, we also include sounds that are above or beyond the level of the segments. This sound information is called **prosody**, or **suprasegmental** information, and it makes up the rhythm, timing, meter, and stress of the words and sentences that we speak. The primary pieces of suprasegmental information are the **pitch** of sounds, the **loudness**, and the **length**.

The **pitch** of a sound is how high or low it is. We produce high pitched sounds when our vocal folds have a high-frequency vibration, and when our vocal folds vibrate more slowly, the resulting sound is lower in pitch.

Some languages use pitch information to signal changes in word meaning. If a language uses pitch this way, the pitch information is called **tone**. These example words are from Yoruba, a language spoken in Nigeria. If you look just at the segmental level, these words all seem to be transcribed the same. But speakers of Yoruba vary their pitch when they speak these words so that the meaning of the word changes depending on whether the second syllable has a high tone, a mid-tone, or a low tone. Probably the best-known tone language is Mandarin, which has five different tones. Looking at these five words, you can see that they contain the same segments, but it's the tones that distinguish their meaning.

Languages also use pitch in another way, not to change word meaning, but to signal information at the level of the discourse, or to signal a speaker's emotion or attitude. When pitch is used this way, it's called **intonation** rather than tone. English uses pitch for intonation — let's look at some examples.

Sam got an A in Calculus.
Sam got an A in Calculus!
Sam got an A in Calculus?
Sam? got an A? in Calculus?

All of these sentences contain the same words (and the same segments) but if we vary the intonation, we convey something different about the speaker's attitude towards the sentence's meaning. Notice that we sometimes use punctuation in our writing to give some clues about a sentence's prosody.

Another component of suprasegmental information is the **length** of sounds. Some sounds are longer than others. Listen carefully to these two words in English. *beat, bead.* The vowel sound in both words is the high front tense vowel [i]. But in *bead*, the vowel is a little longer. This is a predictable process in English — vowels get longer when there's a voiced sound in the coda of the syllable. The diacritic to indicate that a segment is long looks a bit like a colon [iː].

So a sound can change in length as the result of a predictable articulatory process, or, like intonation, length can signal discourse-level information about an utterance. Consider the difference between, *That test was easy*, and, *That test was eeeeeeeeeeeeeaaasyyyyyyyy*. Some languages use length **contrastively**, that is, to change the meaning of a word. In these words in Yapese, a language of the Western Pacific region, you can see that making a vowel long leads to a completely different word with a new meaning. In these words from Italian, consonant length can change the meaning of a word, so *fato* means *fate*, but *fatto* means *fact*.

In English, pitch, loudness and length also contribute to the stress pattern in words. English words that are longer than one syllable usually alternate between stressed and unstressed syllables. Stressed syllables are more prominent than unstressed syllable, and what makes them prominent is that they're louder, longer, and higher in pitch than unstressed syllables. Here are some examples.

The words *happy*, *music*, *sweater* have primary stress on the first syllable, while the words *beside*, *around*, *descend* are stressed on the second syllable. If you're having a hard time hearing the stress difference, try humming the words to hear the difference in pitch. Stress on the first syllable sounds like this [humming] and stress on the second syllable sounds like this [humming].

Being able to identify stressed syllables is important when we're learning to do phonetic transcription, because in English, stressed syllables usually get pronounced with a full vowel, while the vowel in unstressed syllables gets **reduced**. What does it mean to be reduced? That short mid-central vowel that has the name schwa and the symbol [ə] like an upside-down "e" is the most neutral vowel in English. So the "uh" sound in the first syllable of *banana* gets transcribed with a schwa because it's unstressed, but the "uh" in *bunny* gets a full vowel because it's in a stressed syllable. We'll see later in this chapter that stress makes a difference to alveolar stops and to aspirated consonants as well!

To sum up, suprasegmental information, also known as prosody, is that sound information that's above the level of the segment. It consists of pitch, loudness, and length. Many languages use prosody to provide discourse-level information, and some languages also use prosody to change word meanings.

Check Yourself

? Exercise 2.6.4

Young children's voices are usually recognizably different from adult's voices. Which factor is likeliest to be different between children's speech and adults' speech?

- Word length.
- Tone.
- Pitch.

Answer

"Pitch"

Hint: Think about when children speak, and how the pitch of their speech is less varied than what adults produce.

? Exercise 2.6.5

In English, *yes-no* questions often conclude with rising pitch, whereas *wh*-questions often have a falling pitch on the final words. Is this pitch difference a difference in tone or in intonation?

- Tone.
- Intonation.
- Pitch.

Answer

"Intonation"

Hint: "Intonation" is more association with syntactical elements, like question formation. Tone and pitch are more lexical, typically. Also: English is not a tone or pitch language.

? Exercise 2.6.6

English uses pitch as one factor in syllable stress. There are many English pairs of words like *record* (noun) and *record* (verb), which are spelled the same but differ in their stress patterns. Which of the following is true for this pair of words?

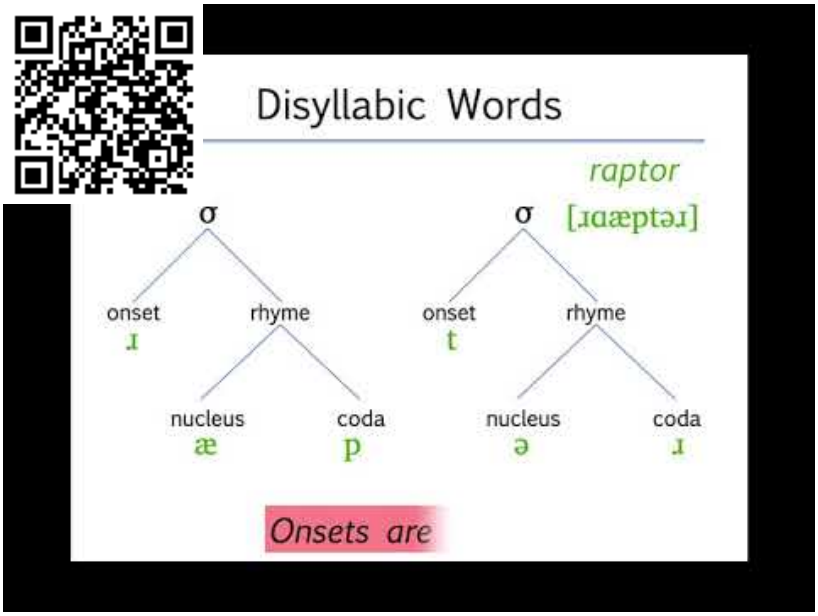
- The first syllable has higher pitch than the second in the noun *record*.
- The second syllable has higher pitch than the first in the noun *record*.

Answer

"The first syllable has higher pitch than the second in the noun *record*."

Hint: Think about which syllable is a bit higher or louder when you say the noun *record*.

2.6.3: From 3.4 Syllable Structure, in Anderson's *Essentials of Learning*



Video Script

In a previous unit we saw that a syllable is a peak of sonority surrounded by less sonorous sounds. We know that sonority is acoustic energy, and now that we understand how speech is produced, we know that the most sonorous sounds, the ones that have the most acoustic energy, are the sounds that are produced with the vocal tract unobstructed. The most sonorous sounds are vowels. Consonants, on the other hand, have an obstruction in the vocal tract so they're less sonorous. So we might also think of a syllable as a vowel surrounded by some consonants. That's a good beginning definition, but it's a little more complex than that, as we'll see in this unit and the next. Our mental grammar doesn't just organize words into syllables, but it also structures what's inside a syllable. Let's take a look. The name for the most sonorous part of a syllable is the **nucleus**. In a typical syllable, the nucleus will be a vowel, produced with an unobstructed vocal tract. The segments that come before the nucleus are called the **onset**, and if there are any segments after the nucleus they're called the **coda**. The nucleus and coda together form a unit that we call the **rhyme**, and linguists like to use the Greek letter sigma (σ) to label the entire syllable.

Let's look at how this works in some English words. When we say a word is "monosyllabic" that just means that it has one syllable. We'll start with a nice simple word like *big* [bɪg]. The nucleus is the most sonorous part, so in this word, the vowel [ɪ] is the nucleus. The consonant that comes after the vowel nucleus [g] is the coda, and the consonant that comes before [b] is the onset. The only part of a syllable that always has to be there is the nucleus. Some syllables have an onset but no coda, like the word *day* [deɪ], and some syllables have a coda but not onset, like the word *eat* [it]. And the occasional syllable has neither an onset nor a coda, just a nucleus, like the word *I* [aɪ]!

What about a single-syllable word that has more consonants in it? Let's look at *screens*. Again, the vowel [i] is the nucleus of this syllable, and the consonants [nz] that come after the nucleus form the coda. There are three consonants [skɹ] before the nucleus, and they form the onset. When there's a group of consonants in the onset or coda we call them a **cluster**.

Monosyllabic words are pretty straightforward. How does it work with words that have more than one syllable, like *raptor?* It's got two syllables, so it has two nuclei [æ] [ə]. The consonant at the beginning of the word [ɹ] is the onset of the first syllable, and the consonant at the end of the word [ɹ] is obviously the coda of the second syllable. What about these two consonants in the middle? In the word *raptor*, the [p] is the coda of the first syllable and [t] is the onset of the second syllable, but there are other logical possibilities. We could just as easily say that the first syllable has a coda cluster [pt], or that the second syllable has an onset cluster [pt]. How does the mental grammar organize consonants in the middle of a multi-syllabic word?

Well, it's not random, and the mental grammar doesn't just try to distribute consonants evenly. There's a systematic principle that operates in the mental grammar, which is that **onsets are greedy**. To see what that means, let's look at a word that has a bunch of consonants in the middle, like *emblem*. There are three consonants [mbl] in the middle of this word, so there are four logical possibilities for how they could be organized. It could be that all the consonants go in the onset of the second syllable. It could be that they all go in the coda of the first syllable, or they could be divided up between the coda of the first and the onset of the second, with a couple of possible permutations. What does the mental grammar do with these consonants?

The principle that onsets are greedy means that an onset will take as many consonants as it can. So this first option here has the greediest onset: it has the greatest number of consonants in an onset position. But it looks pretty weird, doesn't it, to have a syllable start with [mbl]? A greedy onset takes as many consonants as it can within the grammar of that language. It's a principle of English grammar that words don't begin with a cluster like [mbl], and neither do syllables. Of these four options, the one that has the greediest onset that is possible within English is this one: the [m] is the coda of the first syllable, and the consonant cluster [bl] is the onset of the second syllable.

Let's look at one more example to illustrate this idea that onsets are greedy. Consider the word *ugly*. The two vowels [ʌ] [i] form the two nuclei of the syllables; there's no onset for the first syllable, and no coda for the second syllable. So there are three logical possibilities for these middle consonants [ɡl] — they could both be the coda; they could both be the onset; or they could split the difference. Which does the mental grammar do? The onset is greedy, so it wants to take as many consonants as it can. We know that [ɡl] is a possible onset in English, because there are lots of words that start with [ɡl], like *glue*, *glass*, *glamour*. So because [ɡl] is a possible, grammatical onset cluster in English, the onset of the second syllable takes all of it, and leaves no consonants in the coda of the first syllable.

Let's sum up. Syllables are units within words, and they also have an inner structure of their own. Every syllable has a nucleus, which is the most sonorous part of the syllable: a vowel or another sonorous sound. If there are consonants, which are less sonorous, they make up the onset and coda of the syllable. And in the middle of a word, onsets are greedy: they'll take as many consonants as they can, within the constraints of the grammar of the language.

Check Yourself

? Exercise 2.6.7

In the following transcriptions, a dot [.] represents a potential syllable boundary. Which one shows the syllable boundary in the correct location for the word *dispute*?

- [dɪsp . jut]
- [dɪ . spjut]
- [dɪspj . ut]
- [dɪs . pjut]

Answer

"[dɪs . pjut]"

Hint: Think about where a 'natural break' is in that word, and that is where the syllable boundary will be. It is connected to the morphology--more on that in the next chapter.

? Exercise 2.6.8

In the following transcriptions, a dot [.] represents a potential syllable boundary. Which one shows the syllable boundary in the correct location for the word *melting*?

- [mɛ . ltɪŋ]
- [mɛl . tɪŋ]
- [mɛlt . ɪŋ]

Answer

"[mɛlt . ɪŋ]"

Hint: Think about where a 'natural break' is in that word, and that is where the syllable boundary will be. It is connected to the morphology--more on that in the next chapter.

? Exercise 2.6.9

In the following transcriptions, a dot [.] represents a potential syllable boundary. Which one shows the syllable boundary in the correct location for the word *access*?

- [æ . ksɛs]
- [æk . sɛs]
- [æks . ɛs]

Answer

"[æk . sɛs]"

Hint: Think about where a 'natural break' is in that word, and that is where the syllable boundary will be. It is connected to the morphology--more on that in the next chapter.

2.6.4: From 3.5 Syllabic Consonants, in Anderson's *Essentials of Linguistics*



Video Script

Do you remember our definition of a **syllable** from a couple of units ago? We said that a syllable has a **nucleus: the peak of sonority**, which is surrounded by less sonorous sounds. We already know that vowels are the most sonorous sounds, so most syllables have a vowel as the nucleus. We know that glides are also fairly sonorous, but they're too short to serve as the nucleus of a syllable. Thinking about all the consonant sounds we know, some of them are more sonorous than others. Stops are not very sonorous because they have so little airflow because the vocal tract is completely obstructed. And fricatives also aren't very sonorous because of the obstruction in the vocal tract. But **nasal** consonants are quite sonorous because the airflow resonates through the nasal cavity even when the oral cavity is stopped. And the **liquids**, **[l]** and **[ɹ]**, are also quite sonorous because air is allowed to flow around the tongue.

These sonorous consonants can sometimes serve as the nucleus of a syllable in their own right. In other words, there are some syllables that don't have a vowel at all, just a sonorous consonant. Let's look at some examples.

In the word *rhythm*, the second syllable is unstressed, and it's pretty short. Most of the time, in ordinary rapid speech, that second syllable doesn't have a vowel in it at all. Our articulators go right from the [ð] sound at the end of the first syllable into the [m] sound. The [m] itself becomes the nucleus of the syllable. It is said to be a syllabic consonant, and we use a special notation to transcribe it: [ɹɪðm̩]. Look at that little vertical line below the [m] symbol — that's called a **diacritic**. Diacritics are special additional notations we add to IPA symbols to give extra information about the sounds. That vertical line is the diacritic for a **syllabic consonant**.

Here's an example of a liquid consonant becoming syllabic. When we speak the word *funnel*, we don't produce a vowel in the second, unstressed syllable. Instead, we pronounce the [l] as a syllabic [l̩], so that it is the nucleus of the syllable. The notation is the same, with the diacritic for the syllabic [l̩]: [fʌnl̩].

Check Yourself

Exercise 2.6.10

The video indicated that the word *funnel* can be transcribed to indicate that the second syllable consists of a syllabic [l̩]. The word *elbow* is also spelled with the letters 'el'. Say the two words to yourself several times. Which is the correct transcription for *elbow*?

- [l̩boʊ].
- [ɛlboʊ].

Answer

"[ɛlboʊ]"

Hint: There is a true vowel sound before the [l], so it's not a syllabic sound.

Exercise 2.6.11

The words *human* and *manager* both contain a syllable that is spelled with the letters 'man'. In which word does that syllable contain a syllabic [n̩]?

- Human.
- Manager.

Answer

"Human"

Hint: The stressed syllable in *human* is the first one, making the second syllable unstressed with a sonorant as the nucleus, [n]. The same is not true for *manager*, as the [n] is an onset of a syllable.

? Exercise 2.6.12

In the word *umbrella*, is the [m] syllabic?

- Yes.
- No.

Answer

"No"

Hint: There is a true vowel sound before the [m], so it's not a syllabic sound.

2.6.5: All About Vowels, from Sarah Harmon



Video Script

Catherine Anderson's explanation of how to break down the syllable really mirrors what I do with mine, so realistically I'm not going to say much more. The only thing that is important to understand is that different languages will have different constraints as to what can be an onset, what can be a coda, and even what can be a nucleus. Certainly, the vowel is always a candidate as far as the nucleus is concerned, but having syllabic consonants will also factor in, depending on the language, maybe even the dialect.

What is also important to bring up his nasalization and lengthening of vowels because different languages do different things. For example, **lengthening of vowels** means literally you were keeping it for an extra half to full beat; if you are a musician, you know what that means. Think of how, when you say 'sat' versus 'sad'. When you think of 'sat' versus 'sad', really the only difference is on the vowel, whether that is a long or short vowel: 'sat', very short; 'sad', very long. Same thing with 'bit' versus 'bid'. 'Bit' versus 'bid' is a difference of length. In English there are other reasons why those vowels are long. It is not a phonemic issue, meaning is not something that is going to matter if we say 'sat' versus 'saat'. They still both mean the same thing. But in other languages it does, and Breton is a really great example. It's a Celtic language spoken in Brittany—that part of northern France. You can see that there are plenty of cases where vowel length is a factor, that everything is exactly the same, except for the vowel is longer in one lexicon versus the other. The difference between the term for 'white' and the term for 'bees' is just how long that vowel is. The same thing is true with the term for 'every' or 'all' versus the term for 'owl'. Breton does this, even though it has borrowed a significant number of terms from French, in particular—the word for 'every' or 'all' you could probably see that if you know, French or any other Romance language. But this lengthening of vowels is a very Celtic thing to do. Most Celtic languages do have some kind of vowel lengthening phenomenon going on, so this is pretty cool; we'll come back to more cases like this pretty soon.

Nasalization is again something that we see in French, we see in Breton, we see in Portuguese, and we see it in a number of languages. The difference between an oral vowel and a nasal vowel means that those are two totally different terms. The word for ‘handsome’ in French versus ‘good’, the only difference is that vowel, whether it is nationalized or not: [bo] versus [bõ]. The word for ‘low’ versus the word for a wedding announcement: [ba], [bã]. I do my nasalization a little bit stronger because I speak some Portuguese, but French, it's right there and Breton, we see the same thing. Breton, by the way, probably got this from French.

Both of these phenomena of lengthening and nasalization of vowels, we will see more of as we go through a lot of different phonological patterns, as well as when we start talking about the history of these dialects and these languages.

2.6: All About Vowels is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.9: Suprasegmentals** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **3.3: Syllabic Consonants** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

2.7: References

For this chapter, it's recommended to look at the list of references that are in *Introduction to Language*, Chapter 5, "Phonetics," by Rodman, Fromkin, and Hyams. Additionally, there are a long list of references in *Language Files* on this topic in a wide number of chapters.

2.7: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

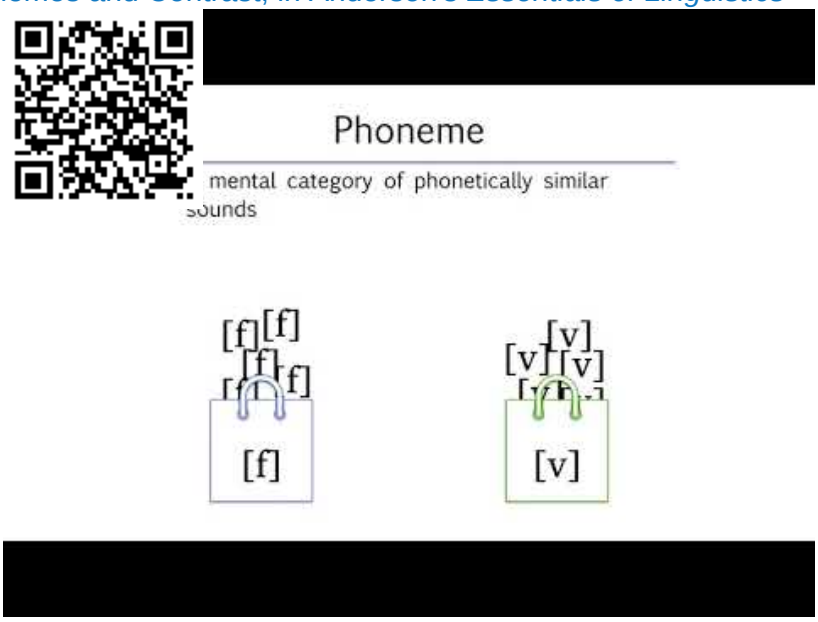
3: Sounds, Part 2- Phonology

3.1: Comparing Sounds and Distribution
3.2: Assimilation and Dissimilation
3.3: Other Phonological Rules
3.4: Analyzing Phonological Data
3.5: References

3: Sounds, Part 2- Phonology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.1: Comparing Sounds and Distribution

3.1.1 From 4.1 Phonemes and Contrast, in Anderson's *Essentials of Linguistics*



Video Script

In the last couple of chapters, we've seen lots of ways that sounds can differ from each other: they can vary in voicing, in place and manner of articulation, in pitch or length. Within the mental grammar of each language, some of these variations and meaningful and some are not. Each language organizes these meaningful variations in different ways. Let's look at some examples.

In the English word *please*, I could pronounce it with an ordinary voiced [l]: [pʰliz] it would be a little unnatural but it's possible. Or, because of perseveratory assimilation, I could devoice that [l] and pronounce it [pʰl̥iz]. We've got two slightly different sounds here: both are alveolar lateral approximants, but one is voiced and one is voiceless. But if I pronounce the word [pʰliz] or [pʰl̥iz], it means the same thing. The voicing difference in this environment is not meaningful in English and most people never notice if the [l] is voiced or not.

In the words *van* and *fan*, each word begins with a labio-dental fricative. In *van*, the fricative is voiced and in *fan* it's voiceless. In this case, the difference in voicing is meaningful: it leads to an entirely different word, and all fluent speakers notice this difference! Within the mental grammar of English speakers, the difference between voiced and voiceless sounds is meaningful in some environments but not in others.

Here's another example. I could pronounce the word *free* with the ordinary high front tense vowel [i]. Or I could make the vowel extra long, *freeeee*. (Notice that we indicate a long sound with this diacritic [iː] that looks a bit like a colon.) But this difference is not meaningful: In English, both [fri] and [friː] are the same word. In Italian, a length difference is meaningful. The word *fato* means "fate". But if I take that alveolar stop and make it long, the word *fatto* means a "fact". The difference in the length of the stop makes [fatɔ] and [fatːɔ] two different words. (N.B., In the video there's an error in how these two words are transcribed; it should be with the [a] vowel, not the [æ] vowel.)

So here's the pattern that we're observing. Sounds can vary; they can be different from each other. Some variation is meaningful within the grammar of a given language, and some variation is not.

Until now, we've been concentrating on phonetics: how sounds are made and what they sound like. We're now starting to think about **phonology**, which looks at how sounds are organized within the mental grammar of each language: which phonetic differences are meaningful, which are predictable, which ones are possible and which ones are impossible within each language. The core principle in phonology is the idea of **contrast**. Say we have two sounds that are different from each other. If the difference between those two sounds leads to a difference in meaning in a given language, then we say that those two sounds contrast in that language.

So for example, the difference between *fan* and *van* is a phonetic difference in voicing. That phonetic difference leads to a substantial difference in meaning in English, so we say that /f/ and /v/ are **contrastive** in English. And if two sounds are contrastive in a given language, then those two sounds are considered two different phonemes in that language.

So here's a new term in linguistics. What is a **phoneme**? A phoneme is something that exists in your mind. It's a mental category, into which your mind groups sounds that are phonetically similar and gives them all the same label. That mental category contains memories of every time you've heard a given sound and labelled it as a member of that category. You could think of a phoneme like a shopping bag in your mind. Every time you hear the segment [f], your mental grammar categorizes it by putting it in bag labelled /f/. /v/ contrasts with /f/ — it's a different phoneme, so every time you hear that [v], your mind puts it in a different bag, one labelled /v/.

If we look inside that shopping bag, inside the mental category, we might find some phonetic **variation**. But if the variation is not meaningful, not contrastive, our mental grammar does not treat those different segments as different phonemes. In English, we have a phonemic category for /l/, so whenever we hear the segment [l] we store it in our memory as that phoneme. But voiceless [l̥] is not contrastive: it doesn't change the meaning of a word, so when we hear voiceless [l̥] we also put it in the same category in our mind. And when we hear a syllabic [l̩], that's not contrastive either, so we put that in the same category. All of those [l]s are a little different from each other, phonetically, but those phonetic differences are not contrastive because they don't lead to a change in meaning, so all of those [l]s are members of a single phoneme category in English.

Now, as a linguist, I can tell you that voiceless [f] and voiced [v] are two different phonemes in English, while voiceless [l̥] and voiced [l] are both different members of the same phoneme category in English. But as part of your developing skills in linguistics, you want to be able to figure these things out for yourself. Our question now is, how can we tell if two phonetically different sounds are phonemically contrastive? What evidence would we need? Remember that mental grammar is in the mind — we can't observe it directly. So what evidence would we want to observe in the language that will allow us to draw conclusions about the mental grammar?

If we observe that a difference between two sounds — a phonetic difference — also leads to a difference in meaning, then we can conclude that the phonetic difference is also a phonemic difference in that language. So our question really is, how do we find differences in meaning?

What we do is look for a **minimal pair**. We want to find two words that are identical in every way except for the two segments that we're considering. So the two words are minimally different: the only phonetic difference between them is the difference that we're interested in. If we can find such a pair, where the minimal phonetic difference leads to a difference in meaning, it's contrastive, then we can conclude that the phonetic difference between them is a phonemic difference.

We've already seen one example of a minimal pair: *fan* and *van* are identical in every way except for the first segment. The phonetic difference between [f] and [v] is contrastive; it changes the meaning of the word, so we conclude that /f/ and /v/ are two different phonemes. Can you think of other minimal pairs that give evidence for the phonemic contrast between /f/ and /v/? Take a minute, pause the video, and try to think of some.

Here are some more minimal pairs that I thought of for /f/ and /v/: *vine* and *fine*, *veal* and *feel*. Minimal pairs don't have to have the segments that we're considering at the beginning of the word. Here are some pairs that contrast at the end of the word: *have* and *half*, *serve* and *surf*. Or the contrast can occur in the middle of the word, like in *reviews* and *refuse*. What's important is that the two words are minimally different: they are the same in all their segments except for the two that we're considering. And it's also important to notice that the minimal difference is in the IPA transcription of the word, not in its spelling.

So we've got plenty of evidence from all these minimal pairs in English that the phonetic difference between /f/ and /v/ leads to a meaning difference in English, so we can conclude that, in English, /f/ and /v/ are two different phonemes.

Check Yourself

? Exercise 3.1.1

Are the phonetically different segments [m] and [n] phonemically contrastive in English?

- Yes.
- No.

Answer

"Yes"

Hint: There are minimal pairs where the only difference is the nasal: *mine* [maɪn] and *nine* [naɪn].

? Exercise 3.1.2

Are the phonetically different segments [p] and [pʰ] phonemically contrastive in English?

- Yes.
- No.

Answer

"No"

Hint: [pʰ] is a variant of [p]

? Exercise 3.1.3

Do the words *sight* and *site* form a minimal pair in English?

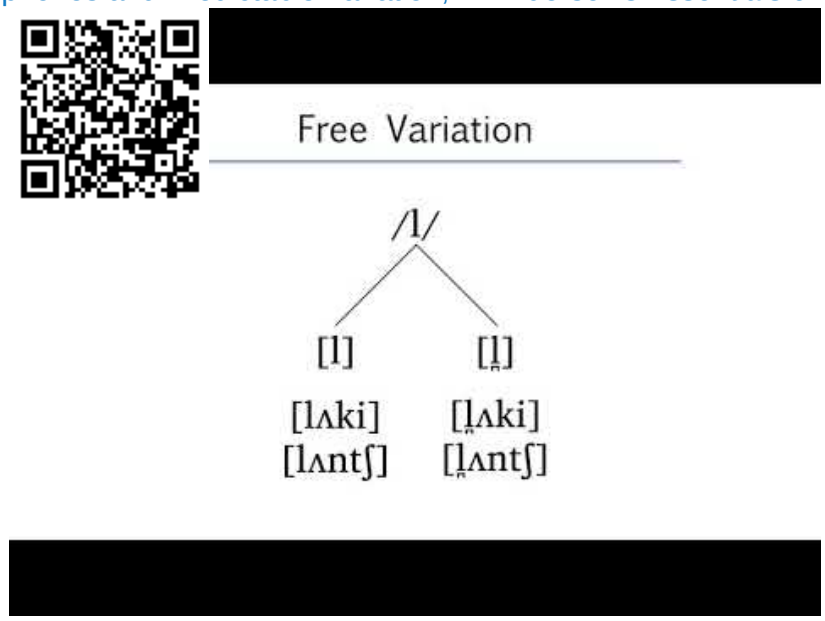
- Yes.
- No.

Answer

"No"

Hint: They have different meanings, and they have the exact same IPA transcriptions: [saɪt].

3.1.2: From 4.2 Allophones and Predictable Variation, in Anderson's *Essentials of Linguistics*



Video Script

In our last unit, we learned about the notion of a **phoneme**. Remember that a phoneme is something that exists in your mind: it's like a shopping bag in which your mind stores memories of examples of phonetically similar sounds that are all members of one category. Not all the sounds that you store in one phoneme category have to be identical; in fact, your mental category has room for a lot of **variation**. Any variants that are not contrastive, that don't lead to a meaning change, are members of that same phoneme category and are called **allophones**.

We've already seen some examples of allophones of English phonemes as we've been learning to transcribe sounds. We know that the alveolar lateral approximant [l] has a voiceless variant [l̥] and a syllabic variant [l̩], but our minds categorize all of them as members of the same phoneme. This shopping-bag metaphor is going to get a little unwieldy, so let's look at another notation that we can use to represent this phoneme category.

We say that /l/ is the label for the phoneme category itself, it's the most general form of the phoneme. Notice that instead of using square brackets, for the symbol that represents the whole category we use slashes. In any given word, the phoneme /l/ might get spoken as any one of its allophones, each of which gets represented in square brackets. But where does each allophone appear? Which allophones do we use in which words? One of the big things that phonology is concerned with is the **distribution** of allophones: that is, what phonetic environments each allophone appears in. The distribution of allophones is a key part of the mental grammar of each language — it's something that all speakers know unconsciously.

Some allophones appear in **free variation**, which means that it's pretty much random which variant appears in any environment. But most allophones are entirely **predictable**: linguists say that allophonic variation is **phonetically conditioned** because it depends on what other sounds are nearby within the word.

Let's start by looking at free variation because it's the simpler case. Take our phoneme /l/, as in the words *lucky* and *lunch*. Most of the time you pronounce these words with a plain old ordinary voiced alveolar lateral approximant. But sometimes you might be speaking extra clearly — maybe you're trying to talk to a relative who's hard of hearing, or maybe you're concentrating on teaching some speech sounds to a language learner. So instead of making the /l/ sound at the alveolar ridge, you stick your tongue right out between your teeth and say *lucky* or *lunch*. Now you're making a dental [l̪], not an alveolar [l], but it's still a member of the phoneme category for /l/ — it doesn't change the meaning of the word so this phonetic difference is not contrastive. It's just free variation within the category.

But most allophonic variation is predictable: different allophones show up in different **environments**. Let's look at a few words. If we look at this set of words: *plow*, *clap*, *clear*, *play*, we can see that whenever /l/ follows a [p] or [k], it is **devoiced**. But now look at this other set of words *(blue, gleam, leaf, fall, silly*), when /l/ appears in any other environment, like following a voiced stop, or at the beginning of a word, at the end of a word, or in the middle of a word, it's the ordinary [l]. If we looked at a whole lot more words and recorded a lot of English speakers, we'd find that whenever /l/ is in a consonant cluster following a voiceless aspirated stop, it also becomes voiceless, but when /l/ is in other environments, it stays voiced. We never find voiceless [l̥] in other environments, and we almost never find voiced [l] following a voiceless stop. That pattern is called **complementary distribution**.

That's an important phrase, and it's going to come up a lot in the next few units. It means that there's no overlap in where we find the allophones: We see voiceless [l̥] following voiceless stops, but never anywhere else, and we never see voiced [l] in that environment. Likewise, we see voiced [l] in lots of different environments, but we never see voiceless [l̥] in any of those places. When we see complementary distribution, that's good evidence that the two segments we're considering are allophones of one phoneme. Can you think of any other examples of English phonetic segments that are in complementary distribution? Think about what happens when you're transcribing voiceless stops.

So let's sum up. If we have two phonetic segments that are related but different from each other, and we find some **minimal pairs** to show that this phonetic difference is **contrastive**, then we conclude that those two segments are **two different phonemes**.

And if we have two phonetic segments that are related but different, and they're not contrastive, then we look to see what the **distribution** of these segments is, that is, what **environments** we see them in. If they're not contrastive and they're in **complementary distribution**, then we conclude that they're **allophones of the same phoneme**.

Check Yourself

? Exercise 3.1.4

Remember that in English, voiceless stops are aspirated at the beginning of a word and the beginning of a stressed syllable, but never in the middle of a word nor at the end of a word. Which term best describes this pattern?

- Phonemic contrast.
- Minimal pair.
- Complementary distribution.

Answer

"Complementary distribution"

Hint: This describes a scenario where a sound only occurs in a specific environment, and that if you swapped it with the other 'option' or 'variant' it wouldn't change the meaning.

? Exercise 3.1.5

The symbol [ɬ] represents a velarized [l]. Looking at the following set of transcribed English words, what can you conclude about [l] and [ɬ] in English?

leaf [lif]

fall [fɑɬ]

luck [lʌk]

spill [spɪɬ]

lemon [lɛmən]

wolf [wʊɬf]

- [l] and [ɬ] are phonemically contrastive in English.
- [l] and [ɬ] are in complementary distribution in English.

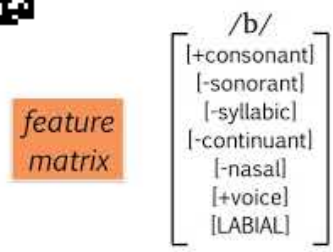
Answer

"[l] and [ɬ] are in complementary distribution in English."

Hint: This describes a scenario where a sound only occurs in a specific environment, and that if you swapped it with the other 'option' or 'variant' it wouldn't change the meaning.

3.1.3: From 4.3 Phonetic Segments and Features, in Anderson's *Essentials of Linguistics*

[redacted]



[redacted]

Video Script

In our thinking about speech sounds so far, we've focused almost entirely on **segments**. Segments are the individual speech sounds, each of which gets transcribed with an individual symbol in the IPA. We've seen that any given segment can influence the segments that come before and after it, through coarticulation and other articulatory processes. And we've also seen that segments can be grouped together into **syllables**, which we look at in more detail in another unit. Within the grammar of any language, two different segments might contrast with each other or might not.

So we've been talking as if segments are the smallest unit in speech, but in fact, each speech segment is made up of smaller components called **features**. Each feature is an element of a sound that we can control independently. To see how features work, let's look at a couple of examples. We can describe the segment [b], for example, as being made up of this set of features. First, [b] is a consonant (meaning it has some obstruction in the vocal tract), so it gets the feature *consonant* indicated with a plus sign to show that the consonant feature is present. Looking at the next feature, sonorant, notice that it's indicated with a minus sign, meaning that [b] is not a sonorant. The feature *sonorant*, of course, has to do with sonority. We know that stops have very low sonority because the vocal tract is completely closed for stops, so stops are all coded as [-sonorant]. The next feature, *syllabic*, tells us whether a given segment is the nucleus of a syllable or not. Remember that the most common segments that serve as the nucleus of a syllable are vowels, but stops certainly cannot be the nucleus, so [b] gets labelled as [-syllabic]. These first three features, consonant, sonorant, and syllabic allow us to group all speech segments into the major classes of consonants, vowels, and glides. We'll see how in a couple of minutes.

This next set of features has to do with the manner of articulation. The feature *continuant* tells us how long a sound goes on. Stops are very short sounds; they last for only a brief moment, so [b] gets a minus sign for *continuant*. We also know that [b] is not made by passing air through the nasal cavity, so it also gets a minus sign for the feature *nasal*. And [b] is a voiced sound, made with vocal folds vibrating, so it is [+voice].

The last feature we list for [b] is [LABIAL] because it's made with the lips. (Stay tuned for an explanation of why some features are listed in lower-case and some in upper-case.)

This whole list of features is called a **feature matrix**; it's the list of the individual features that describe the segment [b], in quite a lot of detail! Because features are at the phonetic level of representation, we use square brackets when we list them. You often see a feature matrix listed with a large pair of square brackets, like this, but we'll just use individual square brackets on each feature.

Now I want you to notice something. If we take this whole feature matrix and change the value of just one feature, changing the feature *voice* from plus to minus, now we're describing a different segment, [p]: [p] has every feature in common with [b] except for voicing. Likewise, if we take the feature matrix for [b] and change the value of the feature *continuant* from minus to plus, now we're describing the segment [v], which has all the same features as [b] except that it can continue for a long time because it's a fricative. Or if we take the feature matrix for [b] and change the feature *nasal* from minus to plus, this has the effect of also changing the *sonorant* feature to plus as well, because circulating air through the nasal cavity adds sonority. Now, this feature matrix describes the properties of the segment [m].

So each feature is something that we can control independently of the others with our articulators. And changing just one feature is enough to change the properties of a segment. That change might lead to a phonemic contrast within the mental grammar of a language, or it might just result in an allophone of the same phoneme.

It turns out that segments that have a lot of features in common tend to behave the same way within the mental grammar of a language. And we can use these features to group segments into natural classes that capture some of these similarities in their behaviour.

Let's look again at the feature matrix for [b]. If we take away the feature that describes its place of articulation, we end up with a smaller list of features. This smaller list describes not just a single segment, but a class of segments: all the voiced stops. By not mentioning the place feature, we've allowed this matrix to include segments from any place of articulation, as long as they share all these other features. These three segments have all these features in common: they're a natural class. If we remove another feature, the voicing feature, the natural class gets bigger: now we've got a feature matrix that describes all the stops in English, including those that are [+voice] and those that are [-voice]. So you can see that this system of features is very powerful for describing classes of segments that have things in common. We'll learn more about natural classes in the next unit.

Check Yourself

? Exercise 3.1.6

Which feature distinguishes the segments [w] and [o]?

- [syllabic].
- [consonant].
- [LABIAL].

Answer

"[consonant]"

Hint: One is on the consonant chart, the other is on the vowel chart.

? Exercise 3.1.7

Which feature distinguishes the segments [p] and [f]?

- [LABIAL].
- [sonorant].
- [continuant].

Answer

"[continuant]"

Hint: [p] is a stop, meaning that the airflow is stopped. [f] is a fricative, meaning that the airflow is continuous.

? Exercise 3.1.8

Which feature distinguishes the segments [p] and [b]?

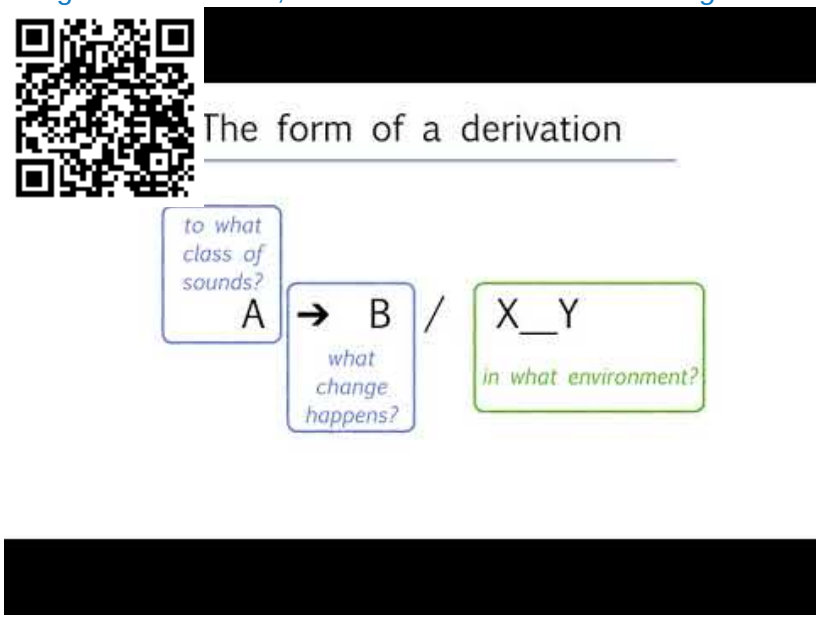
- [voice].
- [LABIAL].
- [continuant].

Answer

"[voice]"

Hint: These sounds are the same, just that one has the vocal cords moving/vibrating.

3.1.4: From 4.5 Phonological Derivations, in Anderson's *Essentials of Linguistics*



Video Script

Earlier in this chapter, we talked about the difference between **phonemic** and **phonetic** representations. Remember that when we talk about a phonemic representation, we're referring to how a sound or a word is represented in our mind. At the phonemic level, the mind stores segmental information, but not details about allophonic variation. But the phonetic representation is how we

actually speak words, and because of coarticulation and various articulatory processes, when we speak a given phoneme, it gets produced as the particular allophone that's conditioned by the surrounding environment. For example, the word *clean* is represented phonemically like this /klin/ in our minds but when we speak, these particular phonetic details [k^hl̥in] are part of what we say. We speak in allophones but we hear in phonemes.

The systematic, predictable relationship between the phonemic and phonetic representations is part of the mental grammar of every fluent speaker of a language. Phonologists have developed a notation for depicting this relationship, which is sometimes known as a **derivation** or a **rule**. Remember of course that when we talk about rules in linguistics, we don't mean those prescriptive rules that your high school English teacher wanted you to follow. We mean the principles that our mental grammar uses to link the underlying phonemic representation to the surface form. Our mental grammar keeps track of every predictable phonetic change that happens to a given natural class of sounds in a given phonetic environment.

Let's think about that now familiar process of liquid devoicing. We've seen lots of English examples like *clean* where the voiced [l] becomes voiceless following the voiceless [k^h] because of perseveratory assimilation. In fact, we've seen enough data from English to observe that this doesn't just happen to one segment; it happens to the natural class of liquids in the environment of another natural class: voiceless stops.

The way that we write a derivation takes a particular form that looks like this. This notation is read as "A becomes B in the environment between X and Y". The left side represents the phonetic change that happens: a particular phoneme or natural class of phonemes becomes a given allophone or undergoes a change to one or more features. The right-hand side shows the phonetic environment that the change occurs in.

So how would we use this notation to represent the predictable process of liquid devoicing? Let's start by describing the pattern in words. The change happens to the liquids [l] and [ɹ]. What happens to them is that they go from voiced to voiceless. And where it happens is following voiceless stops. So now let's describe the pattern using a feature matrix.

We start with the feature matrix for the liquids. They're consonants that are sonorant but not nasal and by default, they're [+voice]. The change that happens is that their voice feature goes from plus to minus. The other features stay the same so we don't list them in the feature matrix that describes the change.

Now we have to say where this change happens. The big slash just means "in the environment" and we know that this change happens following something, so we put the horizontal line that indicates the location of the change following the feature matrix that represents the environment and then we fill in the details of the particular environment. Voiceless stops are consonants that are [-continuant] and [-voice].

We could say that [l] becomes voiceless [l] in the environment following [p] or [k] and [ɹ] becomes voiceless [ɹ] in the environment following [p] or [t] or [k] but using feature matrices captures the broader generalization that this allophonic variation happens to an entire natural class in the environment of another natural class.

Check Yourself

? Exercise 3.1.9

Which phonological rule accurately represents the process, "vowels become nasalized before a nasal consonant"?

- [-consonant] [−sonorant] [+syllabic] → [+nasal] / ___ [+consonant] [+sonorant] [+nasal]
- [+vowel] [+sonorant] [−nasal] → [+nasal] / [+nasal] [+continuant] ___

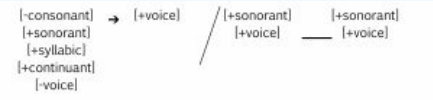
Answer

[-consonant] [−sonorant] [+syllabic] → [+nasal] / ___ [+consonant] [+sonorant] [+nasal]

Hint: Look for the attribute [+nasal] in the right-most portion of the equation.

Exercise 3.1.10

Which sentence accurately describes the process depicted in this phonological rule?



- Voiceless fricatives become voiced between voiced sonorants.
- Voiceless stop consonants become voiced before or after a voiced sonorant.

Answer

"Voiceless fricatives become voiced between voiced sonorants."

Hint: Look for the attribute [+continuant] in the left-most portion of the equation.

Exercise 3.1.11

Which sentence accurately describes the process depicted in this phonological rule?



- The segment [ə] is deleted in the environment of a strident consonant.
- The segment [ə] is epenthesized following a strident consonant.

Answer

"The segment [ə] is epenthesized following a strident consonant."

Hint: The left-most element of the equation is a 'null', meaning that you're starting with nothing in that position.

3.1: Comparing Sounds and Distribution is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.1: Phonemes and Contrast** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **4.2: Allophones and Predictable Variation** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **4.3: Phonetic Segments and Features** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **4.7: Phonological Derivations** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

3.2: Assimilation and Dissimilation

3.2.1 From 3.7 Articulatory Processes: Assimilation, in Anderson's *Essentials of Linguistics*



Video Script

So far, we've been talking about individual speech sounds as if they're all separate from each other. But we know of course that we don't articulate individual segments when we speak – we don't produce the word *book* as [b] – [ʊ] – [k]. When we're speaking, our articulators are always moving – they're moving away from from the position for the sound they just made, and preparing to make the sound that's coming up. You can feel this really easily by saying a couple of words. I want you to prepare to say this word, but don't actually say it: just put your mouth in the position to say the word *key.* Pay attention to how you're holding your mouth. What do you notice? Now get ready to say this word, but don't actually say it, just freeze in the position: *cool*. What position are your articulators in? Both *key* and *cool* start with the voiceless velar stop [k^h], so if we articulated speech segments individually, we'd expect our mouths to be in the same position for both words. But the vowels in each word are quite different: [i] in *key* is high and front, and [u] in *cool* is high, back and rounded. So when we produce that [k] sound, our mouths are already preparing for the next vowel. This is called **coarticulation**: the articulation of every speech sound is shaped by the sounds that come before and after it. When we're doing detailed, narrow phonetic transcription, we can include details about coarticulation and other articulatory processes.

Probably the most common articulatory process is **assimilation**. You can guess from its name that it involves sounds becoming more similar to each other. Sounds often become more similar to what's coming up in the word. Here's an example; say the words *cat* and *can*. They both have the vowel as the nucleus, but for *can*, when we produce that [æ] we're already anticipating the upcoming nasal so we've already got the velum lowered to allow air into the nasal cavity. So the vowel gets nasalized too — it gets assimilated to the following nasal. We transcribe a nasal vowel with the diacritic for nasalization, like this:[æ̃]. Because this nasalization is in anticipation of an upcoming nasal consonant, we call this process **anticipatory assimilation**: the vowel is becoming more similar to the sound that follows it. In some books, you might see this called **regressive assimilation**, since the nasal property of the [n] is moving backwards or regressing onto the vowel.

Assimilation can go in the other direction too: sometimes the properties of one speech segment persevere into the next segment. Say these two words out loud: *bleed*, *please*. The two [l] sounds in these two words are a little different from each other. For *bleed*, the vocal folds are vibrating for the voiced [b] and they keep vibrating to produce the voiced [l]. We know that [l] is usually voiced so there's nothing remarkable about that. But for *please*, the vocal folds are held apart for the voiceless [p^h]. We start making the [l] before the vocal folds start to vibrate, so the [l] becomes voiceless in this context. We say that the [l] following a voiceless stop is **devoiced**, and it gets transcribed with the diacritic for voicelessness, like this: [l̥]. In this case, the voiceless property of the [p] is persevering; it's sticking around to have an influence on the [l], so we call it **perseveratory assimilation**. You might also see this called **progressive assimilation** because the voicelessness of the first sound progresses, or moves forward, onto the following

sound. One thing to note about the diacritic for voicelessness: it only gets used when a sound that is ordinarily voiced becomes voiceless in one of these articulatory processes. An [l] is usually voiced, so if it gets devoiced it gets the diacritic. But a sound like [h] or [s] is already voiceless, so it wouldn't make any sense to transcribe it with the diacritic.

So **assimilation** can be **anticipatory**, where a speech sound is influenced in anticipation of the sound that's about to be spoken after it, or **perseveratory**, where a sound is influenced by properties persevering, or lingering, from the sound that was just spoken.

Check Yourself

? Exercise 3.2.1

What articulatory process is at work when the word *bank* is pronounced as [bæŋk]?

- Assimilation (Anticipatory / Regressive).
- Assimilation (Perseveratory / Progressive).

Answer

"Assimilation (Anticipatory / Regressive)"

Hint: The sound that changes is before the 'influencer'.

? Exercise 3.2.2

What articulatory process is at work when a child pronounces the word *yellow* as [lɛloʊ]?

- Assimilation (Anticipatory / Regressive).
- Assimilation (Perseveratory / Progressive).

Answer

"Assimilation (Anticipatory / Regressive)."

Hint: The sound that changes is before the 'influencer'.

? Exercise 3.2.3

What articulatory process is at work when the word *cream* is pronounced as [k^hɻiim]?

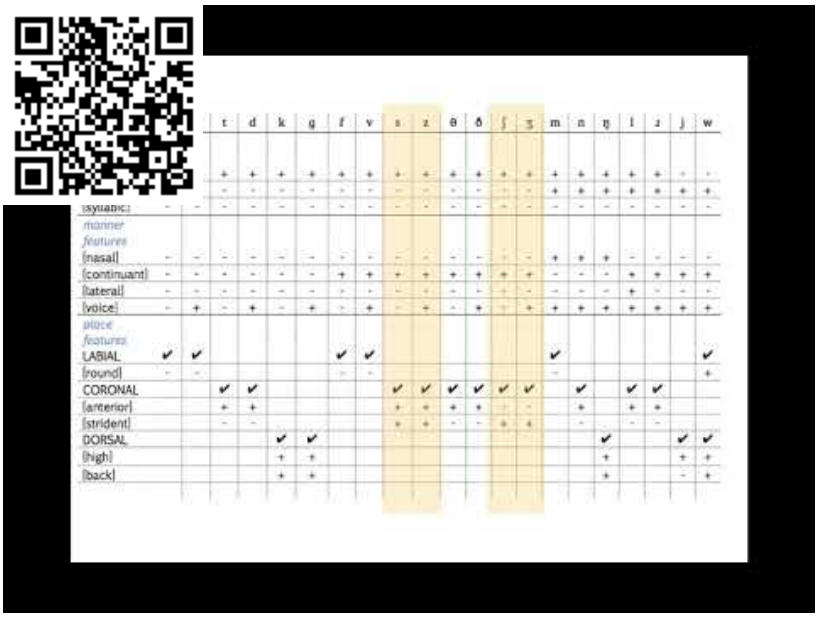
- Assimilation (Anticipatory / Regressive).
- Assimilation (Perseveratory / Progressive).

Answer

"Assimilation (Perseveratory / Progressive)."

Hint: The sound that changes is after the 'influencer'.

3.2.2: From 4.6 Phonological Derivations in Everyday Speech, in Anderson's *Essentials of Linguistics*



Video Script

The last unit showed us how we can use the formal notation of a derivation (or rule) to represent what's happening in the mental grammar of a speaker as they use their language. In this unit we'll look at some of the processes that we use really frequently when we speak English, and how they can be represented with a phonological rule.

Let's start by talking about plurals in English. When we were learning to transcribe, we noticed that the common English plural suffix, which is usually spelled with the letter "s", gets transcribed in three different ways. This is sometimes called the "cats, dogs, horses" phenomenon because *cats* ends with a voiceless fricative [s], *dogs* ends with the voiced fricative [z], and *horses* with a whole syllable [ɨz]. Here are some other words with each of these different plural forms.

> *cups*, *peacocks*, *myths*, *cliffs* all take the voiceless [s]
> *bees*, *fans*, *pencils*, *leaves* all take the voiced [z]
> and *edges*, *mazes*, *dishes*, *beaches* take the [ɨz] form

We don't have to look too hard to figure out that words that end in a voiceless consonant take the voiceless plurals [s] while the voiced [z] is for words that end in a voiced segment. But why is there this third form of the plural. Why does that high central vowel get **epenthesized**, and where does it happen? Look down this list of words and you'll see that they all end in fricatives, [s], [z], [ʃ] or [ʒ]. But it's not all fricatives, as we can see from *myths, cliffs, leaves*. Looking at the feature chart, we see that it's a particular class of CORONAL fricatives — the ones that are [+strident]. We can describe this process in words by saying that the English plural suffix [z] gets an extra vowel [ɨ] following a strident consonant. How can we represent that with a rule?

Well, we start by thinking about the change that happens. In this case, a vowel is getting epenthesized. We've been describing phonetic changes as something becoming something else, but epenthesis is really a case of nothing become something. So we represent it this way: this zero with the diagonal line through it means "nothing", and the something that gets inserted is the high central vowel [ɨ]. And what's the environment where it happens? Following a strident consonant, but not just any time there's a strident. Our mental grammar doesn't go around sticking extra vowels into every word with a strident in it. It happens specifically when we're sticking a [z] at the end of a word. Notice that this correctly predicts that we'll also get that extra vowel when we add the simple present suffix to a verb, so *breathes* just gets [z] for simple present in *she breathes* but *reaches*, where the verb *reach* ends with the affricate [tʃ] gets the epenthesized vowel: *reaches.*

So the idea we're working with here is that every single fluent speaker of English, every time they speak the plural form of a word that ends with [s] [z] [ʃ] [ʒ] [tʃ] or [dʒ], their mental grammar automatically applies this rule, and it happens so regularly and so rapidly that most of us aren't even aware of it.

Now let's look at that common process of **flapping**. When we were learning to do phonetic transcription, we learned that a word that's spelled with a "d" or a "t" in a particular environment usually gets pronounced with a flap [ɾ] instead of a stop, for people who speak varieties of Canadian and US English, and also for most speakers of Australian English. Here are some examples:

water, ladder, total, model, bottom, modem

Looking at this set of words, we can see a pretty clear pattern, which we can describe in words this way: [t] and [d] become the flap [ɾ] between vowels in the onset of an unstressed syllable. (It's actually a little more complex than that, but linguists are still arguing about what the exact environment is, so this is close enough for our purposes.) We can describe this process with a phonological derivation something like this:

The class of sounds that the change happens to is the alveolar stops [d] and [t]. So these are consonants that are [-sonorant], which excludes the nasals and liquids, and [-continuant], which excludes fricatives. And they're the ones made at the tip of the tongue, that is, the coronals. These sounds become the flap, in the environment between vowels, when the second vowel is unstressed. So that's a lot of fancy notation to describe a process that your mental grammar does rapidly, unconsciously, hundreds of times a day.

So we've looked at a couple of examples of phonological derivations that represent allophonic and allomorphic variation in our everyday speech. (Glance ahead a couple chapters to learn what "allomorphic" means!) Linguists use this formal notation to represent them, but remember that these are unconscious processes in our mental grammar that operate hundreds of times a day without us even noticing.

Check Yourself

? Exercise 3.2.4

Which of the following correctly illustrates the environment "in the onset of a syllable"?

- #__
- σ__
- __#
- __σ

Answer

"σ__"

Hint: The symbol σ refers to a syllable boundary.

? Exercise 3.2.5

Which of the following correctly illustrates the environment "between vowels"?

- #V__
- __V#
- V__V
- [DORSAL]__[DORSAL]

Answer

"V__V"

Hint: The symbol V refers to any vowel.

3.2.3 Assimilation and Dissimilation, from Sarah Harmon



Video Script

As Catherine Anderson stated, assimilation is a really common device that pretty much every language seems to have. It is true to say that pretty much every language has at least one assimilation rule. Dissimilation is also pretty common, although not as common as assimilation. I like to use some concrete examples, just to showcase a little bit as to what happens.

For assimilation, I like to use this example from Spanish. The reason is simple: most of you either are native speakers of Spanish or you've come across Spanish in some way, shape or form. Many of you have studied it. But I bet you didn't realize any of you that there really aren't two or three nasal phonemes in Spanish. There's only one, and it changes place depending on the consonant that comes afterwards. No matter how it is written, that nasal is pronounced in the same place of articulation as the consonant that comes afterwards. In Spanish, this is not said, [enfasis], because that wouldn't make any sense. That nasal has to be in a labiodental position your teeth and your lips have to touch, because that's where the following consonant is also articulated. So it's [eɱfasis]. Notice that my top teeth are touching my bottom lips. [eɱfasis]

In the case of [imposible], that is straightforward because of how it is written and most speakers who are literate in a Latinate alphabet are going to recognize the sound. But it is truly just the case that that nasal sound is before a bilingual consonant. [imposible]

The following is [iɲṭroduksjon]. That first nasal is said as a dental because the following consonant is a dental consonant; the [ṭ] and [ḍ] sounds in Spanish or not like an English, as tongue is hitting the back of the teeth. [iɲṭroduksjon]. That final [n] is left as an alveolar nasal.

In the case of that vest kind of like thing that comes out of the Andes region of South America, in English we call it a *poncho* and that nasal is pretty much an alveolar nasal. But in Spanish that's not the case. In Spanish, it's going to conform; that [tʃ] is post-alveolar or palatal, depending on your perception of it, and the nasal is also going to be post-alveolar nasal [poɲtʃo].

Finally, the term *Inca* in English, that [n] also kind of skewed to the back to the velar region, and in Spanish, much more [iŋka]. This is a really great example of articulation that gets assimilated, the place of articulation in this case. Assimilation rules are super common.

But dissimilation rules do also exist, and in this case, I'm going to turn to a specific dialect of Spanish—Caribbean Spanish—in a phenomenon that only happens there, meaning it doesn't happen in any other dialects of Spanish.

If you know Spanish, you know that the diminutive the suffix that we put on a word to make it small or cute. You know that that in Spanish is frequently either *-ito* or *-ico*, depending on your dialect; most use *-ito*, some use *-ico*. Some use both. In the Caribbean it's not just using both but using them interchangeably. It all depends on what the final consonant is in the root word. For example, in the word *bota*, which means 'boot', the final content is a [t̪] sound so that's alveolar/dental. If you are going to make it small and cute like a baby booty, then you cannot use [botita] as you would in most other dialects of Spanish; you have to have a different

sound, so you have to use [botika]. [botika] A [boka] in Spanish, is a ‘mouth’ and notice that the final consonant is the [k] sound. Well, if I want to say a ‘small mouth’, a ‘cute mouth’, I can't say [bokika], because you can't have the same sound twice. you have to say [bokita]. The default is the -ito, -ita version. Notice that if you want to make something really, really, really short in Spanish —and Spanish speakers, you know this—we just stack the diminutive; we keep adding them on. If you're telling somebody, “Just a minute,” “Momento.” Or you can make it short, “momentito.” That's saying, “Okay, just a minute.” You can make it, “Wait, wait, wait a second.” That phrasing in Spanish, usually you would just say, “Momentitito,” and you can put 2, 3, 4 of the diminutives on. But in the Caribbean, while you are doing that you have to switch out—meaning that you have to make sure that you dissimulate between the different versions. So, 'momento', that last consonant is the [t] sound, and that means the next that first diminutive has to be the *-ico*. But because that ends in a [k], if you're going to stick on another one, it has to be the *-ito*, and if you put on a third one after that it would have to be *-ico*: *momento*, *momentico*, *momentiquito*, *momentiquitiquito*. it's like a series of Russian nesting dolls.

3.2: Assimilation and Dissimilation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.5: Articulatory Processes- Assimilation** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

3.3: Other Phonological Rules

Aspirated Stops in English, in Anderson's *Essentials of Linguistics*



Video Script

We know now that we can use the IPA to transcribe speech sounds, and that our transcription can be either broad or narrow. When we make a **narrow transcription**, we're including as much detail as possible about how speakers produce sounds, which often means including diacritics. To give an accurate narrow transcription of Canadian English, we would have to include a property that is part of nearly every variety of English – **aspiration** on voiceless stops.

To illustrate what aspiration is, I'm going to ask you to say a silly sentence: *The spy wanted to buy a blueberry pie*.

Now say it again, and hold your hand in front of your mouth. *The spy wanted to buy a blueberry pie*.

Did you feel any differences between the words *spy*, *buy* and *pie*? For native speakers of English, the word *pie* is produced with a little puff of air as the [p] is released. That puff of air is called **aspiration**. English speakers systematically produce aspiration on voiceless stops at the beginning of a stressed syllable, but not on voiced stops. To understand why we have to think about voicing and about the manner of articulation.

Remember that voiced sounds are produced by vibrating the vocal folds, whereas voiceless sounds have the vocal folds held open so air can pass freely between them. Remember also that producing a stop involves closing off the vocal tract completely for a moment, then releasing the obstruction and allowing air to flow freely again.

Think about the voiced stop at the beginning of the word *buy*. The lips are closed – that's the stop closure – and the vocal folds start vibrating for the voiced [b]. Then the lips open and the stop is released, and the vocal folds keep vibrating for the diphthong [aɪ].

But in the word *pie*, things work differently. The lips are closed for the bilabial stop. But because [p] is a voiceless stop, the vocal folds are not vibrating. We open the lips to release the stop, but 30 or 40 milliseconds pass before we start vibrating the vocal folds. That 30-40 milliseconds between when the stop closure is released and the voicing begins is called the **voice onset time** or **VOT**. In English, voiceless stops in certain positions have a VOT of 30-40 milliseconds, so we say that they're aspirated. But voiced stops have a much shorter VOT, of about 0-10 milliseconds. In other words, the vocal folds start vibrating at almost exactly the same time as the stop closure is released, so voiced stops in English are unaspirated. The diacritic to indicate aspiration on a stop is a little superscript h, like so: [p^h, t^h, k^h].

But to make matters even more complicated, it's not all voiceless stops that get aspirated in English – only voiceless stops at the beginning of a **stressed syllable**. In words like *appear* and *attack*, the voiceless stop isn't the first sound in the word, but it comes at the beginning of a stressed syllable so it gets aspirated. [əp^hiɹ] [ət^hæk]

But in the words *apple* and *nickel*, the voiceless stop comes after a stressed syllable and before an unstressed syllable, so it doesn't get aspirated. [æpəl] [nɪkəl]

We don't aspirate voiceless stops at the ends of words, like in *brick*. [bɹɪk]

And we don't aspirate voiceless stops following an [s], even if they're at the beginning of a stressed syllable:

Aspiration of voiceless stops is something that native speakers do so regularly and so automatically that it's very hard for us to perceive it because it's just always there. To convince you, I'm going to record someone saying this sentence and show you the waveforms. This program is known as a waveform editor. And here's Kendrick's voice saying that sentence.

> *The spy wanted to buy a blueberry pie.*

Here's the waveform: this is a visual representation of the sound waves that Kendrick just produced. See that I can select certain parts of the sentence and play them back. *spy, buy, pie*

Look first at *buy* – you can see that there's very a brief silence: that's where Kendrick's lips were closed for the bilabial stop. Then when he releases his lips the waveform gets nice and big for the sonorous vowel [aɪ].

Look over here at *pie*. You see the same silence where the lips are closed, and the same big waveform for the vowel [aɪ] but before the vowel, there's this noisy burst of turbulence – that's the aspiration.

And now look at *spy*. We see the turbulence at the beginning for the fricative [s], followed by the silence while the lips are closed and the nice sonorous vowel. But there's no burst of noise following the release of the lips because the [p] in *spy* was not aspirated. In fact, if I select just the *-py* portion of *spy*, what does it sound like? To a native speaker of English, this part sounds like *buy*, because the [p] is unaspirated.

When you're transcribing words with the voiceless stops [p t k], your challenge will be to figure out if the stops are aspirated or unaspirated, so you can indicate the aspiration in your narrow transcription. In most varieties of English, aspiration happens in these predictable environments.

- Voiceless stops are aspirated at the beginning of a word, and at the beginning of a stressed syllable.
- Voiceless stops are unaspirated at the beginning of an unstressed syllable. They're also unaspirated in any other position, like at the end of a syllable or the end of a word.
- And even if a syllable is stressed, a voiceless stop is unaspirated if it follows [s].
- In English, voiced stops are never aspirated. They're always unaspirated.

One thing that I want you to remember is that this pattern of aspiration is particular to the grammar of English, but stops behave differently in other languages. In French and Spanish, for example, voiceless stops are almost always unaspirated. And some languages, like Thai, actually have a three-way distinction between voiced, unaspirated voiceless, and aspirated voiceless stops.

Check Yourself

? Exercise 3.3.1

The following city names all contain the letter 't' within the word. In which of them is the letter 't' pronounced as [t^h]?

- Victoria.
- Ottawa.
- Edmonton.

Answer

"Victoria"

Hint: [t^h] has to occur at the beginning of a stressed syllable.

? Exercise 3.3.2

The following words all contain the segment /k/. In which of them is it pronounced as the allophone [k^h]?

- Actually.

- Accomplish.
- Attic.

Answer

"Accomplish"

Hint: [k^h] has to occur at the beginning of a stressed syllable.

? Exercise 3.3.3

The following words all contain the segment /p/. In which of them is it pronounced as the allophone [p^h]?

- Appearance.
- Zipper.
- Application.

Answer

"Appearance"

Hint: [p^h] has to occur at the beginning of a stressed syllable.

Other Articulatory Processes, in Anderson's *Essentials of Linguistics*



Video Script

In our last unit, we talked about **assimilation**, when speech segments become more similar to nearby sounds because of coarticulation. There are other **articulatory processes** that shape the words that we say. Some of these processes occur simply as a result of speaking quickly and naturally. Some of them make speech more clear for a listener. Some of them happen over time within a dialect, as speakers start unconsciously changing the way they produce sounds.

While we were learning to do IPA transcription we talked about **vowel reduction**. It's a very common process in rapid, natural speech. In English, the vowel in an unstressed syllable often gets reduced to the mid-central vowel schwa [ə]. This happens in lots of words. For example, we don't usually pronounce this word *electric* as [ilɛktɹɪk]. Instead, because the first syllable is unstressed, the vowel gets reduced, and we say [əlɛktɹɪk]. Likewise, this word *today* doesn't get pronounced as [tudeɪ]. The vowel in the first, unstressed syllable gets reduced and we say [tədeɪ].

In fact, sometimes an unstressed vowel gets reduced so much that it disappears altogether! This process is called, obviously, **deletion**. In some varieties of English, reduced vowels are systematically deleted in certain predictable environments, like in *police*

or *garage*. Deletion can also occur within consonant clusters. It's pretty common for speakers to delete the first [ɹ] in *surprise* or the [d] in *Wednesday*. Deletion also happens when we borrow words from other languages. For example, take the Greek word *pteron*, which means "wing". When we borrow this word and incorporate it into *helicopter*, we pronounce both the [p] and the [t]. But when it comes at the beginning of a borrowed word, like *pterodactyl*, we just delete the [p] altogether, since English doesn't allow two stops in a syllable onset.

Sometimes when we're speaking, extra segments find their way into our words, as a result of coarticulation. Can you guess what word I'm saying? [pʰɹɪnts] Was it *prince* or *prints*? Only one of them is spelled with a "t", but we pronounce them both the same way. In *prince*, an alveolar stop appears between the alveolar nasal and the alveolar fricative. The articulatory process that inserts an extra sound is called **epenthesis**. In English, this tends to happen between nasals and stops or between nasals and fricatives. Another example is in the word *something*, where we often epenthesize a little bilabial stop [p] between the bilabial nasal [m] and the voiceless fricative [θ]. Or when George W. Bush famously pronounced the word *nuclear* as [nukjəlɚ], he was epenthesizing a [j] between the [k] and [l].

Some articulatory processes result from speech errors. Some of these errors are characteristic of children's speech, and some of them just occur in everyday rapid speech. Children's speech often includes the process of **metathesis**, exchanging the position of speech segments. When my niece was little, she used to pronounce the word *hospital* as [hɑstɪbəl], exchanging the positions of the two stops. Metathesis can also happen when we borrow words from another language. When English speakers want to buy a burrito from the restaurant chain called *Chipotle*, we often metathesize the [t] and [l] and say, [tʃəpolti], because the "tl" sequence is rare in English.

Many of these articulatory processes are frequent and systematic in natural speech. In the next chapter, we'll see that they play an important role in our mental grammar.

Check Yourself

? Exercise 3.3.4

What articulatory process is at work when the word *idea* is pronounced as [aɪdijɚ]?

- Reduction.
- Deletion.
- Metathesis.
- Epenthesis.

Answer

"Epenthesis."

Hint: There is an extra sound produced or added to the sequence.

? Exercise 3.3.5

What articulatory process is at work when the word *gorilla* is pronounced as [ɡɹɪlʌ]?

- Reduction.
- Deletion.
- Metathesis.
- Epenthesis.

Answer

"Deletion."

Hint: There is a vowel that was deleted.

? Exercise 3.3.6

What articulatory process is at work when the word *you* is pronounced as [jə]?

- Reduction.
- Deletion.
- Metathesis.
- Epenthesis.

Answer

"Reduction."

Hint: The diphthong was reduced to a single vowel sound.

Other Phonological Rules, from Sarah Harmon



Video Script

Now that we've talked about assimilation and dissimilation, let's talk about other types of phonological rules. To be clear, they're actually quite a few, but we're only going to focus on a certain number of them.

To start off, we're going to do a kind of easy one, in the sense that these are both types of assimilation rules. For example, **neutralization**. Those of you who are experienced English speakers, you do not actually say 'rider' and 'rider'; you do not pronounce that clearly the end in the middle of those lexicon. Rather, you say both of them exactly the same [ɹajɾəɹ]. And notice that you do the same with 'later' and 'ladder'; you frequently just say [læɾəɹ]. What happened to that [t] and that [d]? Well, they neutralized, and specifically in English it happens in the middle of the word when there's two vowels on either side of that stop. They're neutralizing to something very similar but lower in the sound hierarchy. Think about it: the highest position on that sound hierarchy, was a voiceless stop. A voiced stop was pretty high up as well, and are exactly the same sound; it's just a matter of whether your vocal cords are moving or not. They have to neutralize to something that's very similar, just lower on the sound hierarchy. A tap is always going to be lower; it's not quite as an as occlusive, it doesn't stop the air quite as much. It's also in the same place of articulation; it's also an alveolar sound, just like [t] and [d]. They neutralize that position. This is possibly one of the most common assimilation rules.

The other most common assimilation rule is something called **palatalization** and it's exactly what you think it is. The sound is moving to the palatal region of the mouth. Experienced English speakers, ‘I bet you' is not how we usually say that is it; we usually

say 'I betcha'. So what's going on? That [t] in 'bet' is getting pulled from the alveolar region to the palatal region because of the [j] that's coming next to it. That's a palatal glide or palatal approximate. So, because of that we're going to pull that sound. Why does it happen? Well, there's a lot of theories, although none are really proven. Suffice it to say, think about the natural resting place of your tongue when you're not talking, you're not eating, you're not doing anything with your tongue. Where is your tongue? It's in the middle of your mouth, near your palate. That could be the reason that palatalization is the most used assimilation rule there is.

Let's talk about a few others.

While **feature subtraction** is fairly rare, **feature addition** is not; it happens quite often. When we talked about aspiration in English, how [p^h, t^h, k^h] are aspirated and there's an extra puff of air when you are putting them at the beginning of a stress syllable. That (feature addition) is something fairly common to do, to add an aspiration, add nasalization, add some kind of feature.

What is also really common, though, is not just adding a feature but adding a segment (**segment addition**), an extra sound or two. I'm going to go back to Spanish because most all of you have some experience with a native Spanish speaker; maybe you yourself are native Spanish speaker. You know that in Spanish, for anything having to do with writing the verb root is 'scribir'. 'Inscribir', to inscribe; 'transcribir', to transcribe; 'suscribir', to subscribe. Notice we use them in English as well it's a cognate. But the verb meaning to write is not 'scribir'; it is 'escribir'. In Spanish, you have to stick an [e] in front of a certain combination; that combination specifically is an [s] and a consonant. You cannot start a word with an [s] sound and a consonant; it just can't happen in Spanish, so the rule is you **epenthesize**, you add in that in. By the way, French used to have this rule, too, but then the [s] dropped. Notice the verb in French for writing is also 'écrire', it's also a cognate, but then 'to write' is not 'escrire'; it's a 'écrire'. It's just that the S deleted along the way.

Speaking of French, we cannot talk about French without talking about **apocope**. Apocope is another word for segment subtraction, so getting rid of something. If you have studied French or you are a native speaker French, you know that you drop consonants at the end of a word regularly. In fact, it's one of the jokes of learning French is to look at the how the word is spelled and then not pronounce half the consonants. There's some truth to that joke and it's because of apocope. 'Petite' is the term for small, like petite in English. This is the word for 'rose', you probably guessed that this is the word for a 'sheep'. But how you pronounce the word for small is going to change, depending on whether the word that comes afterwards starts with a consonant or not. If it does, you delete that so you do not say [pɛtit ʀoz], you say [pɛti ʀoz]. You just don't, say that second [t]. Notice that with 'agneau' you do: [pɛtit aɲø]. Apocope is an interesting phenomenon, and if epenthesis is common, apocope is almost as common.

Finally, let's look at **metastasis**. Sometimes is called **permutation**; metastasis is the more common term. It's when you reorder sounds. [lehit] in Hebrew is the reflexive pronoun it means 'self'. If you want to say 'to arrange yourself' or 'to use yourself' or 'to apologize to yourself', instead of [lehit] always being the form, it's going to change if the verb it goes with starts with a sibilant, an [s] like sound. Instead of [lehit sader] you switch the [s] and the [t]: [lehis tader]. Instead of [lehit ʃsameʃ] you switch the [s] and the [t]: [lehiʃ tameʃ]. Instead of saying [lehit tsadek], you're going to switch the [t] and the affricate [ts]: [lehts tadek]. Pretty cool right? Lest you think that this doesn't exist in languages like English…it does. For example, the verb 'to ask' well that's Standard American English or Mainstream American English. If you go to African American English, as well as Appalachian English and a number of other dialects, it is [æks]. You might think that [æsk] is the original, but it's not. In Old English, the verb was [aksjan], so clearly in certain dialects of English, including Queen's English and Standard British English some time ago, there was a metastasis between that [k] sound and the [s] sound so that it is [æsk] versus [æks]. However, many dialects continue with the older pronunciation.

Finally, let's talk about **sequential constraints**. sequential constraints are when a specific language has rules as to what combinations are allowed, and what are not allowed. For example, English cannot start a stressed syllable with a consonant that is not plosive, not aspirated and not voiceless; that refers to the aspiration rule we talked about earlier. You cannot say, for example, [to^{w}me^{j}to^w], without pronouncing that aspiration. It doesn't make sense, and it sounds weird. Likewise, we talked about the

epenthesis rule in Spanish, that you cannot start a word with an [s] and a consonant; it has to have an [e] in front of it. For those who have studied German, you know that it cannot have a word end with a voiced consonant; it's always going to devoice, no matter how it's spelled. There are always **accidental gaps**, meaning possible words that follow the sequential constraints of phonemes for that language but that just don't exist. They haven't been used yet in the lexicon. That doesn't mean they won't get used in the future, and in many cases, this is where marketing gurus have fun. They use those accidental gaps to think of different names for different types of products.

But what this all points to is the fact that we have some kind of **universal grammar**. That grammar is not just how we put words together or how we put phrases together, it also includes how we put sounds together and that this template, this early set of rules gets massaged and added to as we learn our individual native language or native languages.

3.3: Other Phonological Rules is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.4: Aspirated Stops in English** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **3.6: Other Articulatory Processes** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

3.4: Analyzing Phonological Data

3.4.1 Analyzing Phonological Data, from Sarah Harmon



Video Script

Now that we have learned all about phonetics and phonology, let's talk about how we analyze phonological data. To do that, I'm going to go back to the same rules that we found when we talked about complimentary and contrastive distribution.

When we look at a data set, we use this setup. We first look, are there any minimal pairs? Then we look at the distribution of those two sounds: whether or not they have the exact same environment, or if they have different environments. If they have different environments, then we know they're in complementary distribution and we start looking for the phoneme—whichever one is the most general of the sounds. But, if they have overlapping distribution or contrasting distribution, then we know that they're different phonemes entirely

Let's take a phonological data set and let's analyze it together so that you know how to get this done.

What you see here is a data set for German, and specifically we're going to analyze these two sounds. [x] is representative and IPA of the voiceless velar fricative. It's the sound that we also have in Spanish, with respect to the 'j' sound; in German, that's not the letter it represents, but you get the idea it's that same sound. German also has a palatal version of that same sound, [ç]. We're going to compare and contrast [x] versus [ç]. We're going to notice whether or not they are in complementary distribution or contrastive distribution. Are they phonemes in and of themselves? Are they both allophones of the same phoneme? Remember 'allophone' just means 'variant'.

The first thing you have down here are the questions; this mimics the process that I was just describing. First thing first, are there any minimal pairs? Do you see any terms over here in German written in IPA that are exactly the same over here in the palatal side? Are they're exactly the same, and the only difference is that [x] versus [ç], depending on the same place? Take a second pause the recording now.

You should not have found any minimal pairs there are; in fact, none, and so we know that there are none.

The next thing we have to do we have to decide, are they in complementary distribution or are they in contrastive distribution? In order to do that we need to look at the exact environment. What is the sound that comes immediately before and what is the sound that immediately comes afterwards? In order to do that, we literally go word by word, term by term. We look at the sound immediately before and the sound immediately after and we document it.

he word for 'eight' [axt]. The sound that comes immediately before is [a]. We write an underscore to say that's where our analysis sound is. Then the sound that comes immediately after is the [t]. And that's all we need, we do not need to write anything else.

he word for 'book' [bux]. The sound that comes immediately before is [u] so we write that then we have an underscore and then, in this case. It is at the end of the word, so in order to mark that we use the **word boundary**. In non-linguistic terms, that's a hashtag, that's a number sign, or it's a pound sign; it has a lot of roles, but in linguistics we call that a word boundary.

Notice that we're staying within the same sound; we're not crossing back and forth across the data set, and that is typical this is how we start building our pattern.

he next word for 'hole' [lɔx] Okay, we have that backwards 'c' that back vowel. For some of you, you may have to use your character map. If you have downloaded an IPA font map, you will have access to the sounds and you can scroll down. In my case, there's the sound right there, so I'm going to select it and I'm going to copy it. I have that back [ɔ]. Underscore. Again that [x] sound is at the end of the word, so I'm going to mark it as such. Yes, I am using square brackets, because I know these are allophones; I don't know which one is the phone yet.

- The word for 'high'. [ho:x] Okay, so we have a long [o:]. Underscore, and then set the end of the word again so market accordingly.
- The word for 'flight' [fluxt] . Okay, sound before is [u]. Then we have the underscore and then we have [t].
- The word for a 'brook' like a stream, [bax] . You have an [a] and you have an underscore and a word boundary.
- Then the word for cake. [ku:xən]. What comes before that? Long [u:]. What comes afterwards? Schwa [ə]. Again, I'm going to have to go find that so pull up my character map. And I copy it and paste it.

Okay, so those are the environments that we see. What do we notice? We noticed that before we have a variety of vowels: [a, u, o]. We noticed that it is after those vowels. And then we noticed that after that sound we could have a [t], we could have a word boundary we could have a [ə], we could have a lot of different possible combinations. Just by looking at this, it is hard to tell what the pattern is. We need to look at the other side

Now let's look at the environments for [ç] the palatal.

- For 'I', [ɪç]. First things first notice each I'm going to have to use my character map to find that capital 'i', that [ɪ]. Notice that is before it, and then what comes after it a word boundary.
- Okay, the word for 'real', [ɛçt]. I'm going to need to go find that vowel, which I can do. Then that comes before; what comes after is a [t].

- The word for 'sick' [zi:ç]. Again, what do we see before [i:], what do we see afterwards, a word boundary.

- The word for 'smile'. [lɛçəln]. We have two vowels. I'm going to copy both.

- The word for to 'smell' [ri:çən]. Okay long [i:], underscore.

- The word for 'to fence' [fɛçtən]. That epsilon [ɛ] and then [t].

- And the word for 'kitchen' [kyçə]. That long but front high but rounded vowel [y] and then schwa [ə].

So what do we see for environments? For the vowels that go in front they're all vowels. Notice: [ɪ, y, i, ɛ]. There's something interesting about those; they share a feature. Afterwards, we see word boundary; we see [t]; we see [ə]; we see lots of things there's nothing there that shares anything. But, we do see something here: all of those happen to be front vowels. Look at your IPA charts and you'll notice that all of those vowels are in the front of the trapezoid. So, they are front vowels. Are they in complementary distribution, where we see one, we don't see the other? None of these environments for [x], the velar, are exactly the same as the ones for [ç], the palatal. And the same is true in reverse; none of the palatal environments, are the same as the ones for the velar. We're comparing columns. So, are they in complimentary distribution? Yes, they are. They have different environments.

What is the phoneme and spell out the allophones. That's really simple; remember the phoneme is going in slashes. They are our allophones: we know that one is [x] and the other one is [ç]. Okay. But which one has the more varied environments? Think about it: we were able to say that the palatal has all one type of environment—it is a very similar environment and they have something in common—but we could not do that for any of the sounds afterwards. Nor can we say anything about those environments for the velar. Therefore, if we have a really specific environment here for the palatal, that tells us that [ç], the palatal, is the variant, it is not the phoneme. That's because it has a really specific occurrence after front vowels. Everything else is going to be the velar [x]. We're going to want to say that /x/, the velar, is the phoneme.

Now, this might have been really difficult; I can almost guarantee it was difficult. In walking you through—we'll practice this in class, and for those not in my class and are using this text, you can practice on your own with your instructors—this should help walk you through how to analyze a basic phonological problem. We are not going to go any deeper than this, at least not right now. In a different course, definitely.

3.4: Analyzing Phonological Data is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.5: References

In addition to Rodman, Fromkin and Hyams' *Introduction to Language* and OSU's *Language Files*, here are a couple of other resources that are particularly useful:

- For more general information on Phonetics and Phonology, check out the seminal work by Peter Ladefoged and Keithe Johnson, *A Course in Phonetics* (2011/Boston: Wadsworth/Cengage Learning). It's still the gold standard for learning about phonetics and many phonological processes.

3.5: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

4: Words- Morphology

1: Back to the Arbitrary
2: Morphological Definitions
3: Morphemes
4: Affixation and Other Morphological Processes
5: Morpho-phonology
6: Parts of Speech and Word Formation
7: Creation of New Lexicon
8: Morpho-syntax
9: Analyzing Morphological Data
10: References

4: Words- Morphology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1: Back to the Arbitrary

4.1.1 Back to the Arbitrary, from Sarah Harmon



Video Script

We're going to talk about how these little puzzle pieces of language, these morphemes, combined together in a given language and how they work in a language. First, we have to take a minute to revisit that one hallmark of language that I brought up earlier that I said was probably the most important piece to understand what, at least with respect to language, and that has to do with arbitrariness. Let's go back to the arbitrary, and let's talk about what this means really.

When I say arbitrariness, I'm referring to the **Principle of the Arbitrary**; we'll get more into this when we get to the chapter on Meaning, specifically Semantics. Just to give you a little taste of what is to come, the Principle of the Arbitrary was set up by Ferdinand de Saussure about 120 years ago. He was a Swiss mathematician who focused on logic and philosophy; of course, logic, philosophy, and mathematics all are tied together. He is considered the true founder of what we now call linguistics; Noam Chomsky is the modern renovation, shall we say, but even he said, if it weren't for Saussure and his initial statements and theories about language, we would not have linguistics at all. That is true; his work inspired generations of linguists. Those of us that have read his work understand the importance of it, especially given that he was working just simply with math. He was looking at language as if it were well a math problem. I'll get more into him later, but really quickly, let's focus on the Principle of the Arbitrary. It states that there is no particular reason why a specific sound is associated with a specific meaning; in this case, we're going to focus on morphemes.

That means a few different things. For example, the fact that we even have synonyms—two terms that mean pretty much the same thing, if not in fact two terms that mean exactly the same thing—that is an example of the Principle of the Arbitrary. For example, in most dialects of North American English *woodchuck* and *groundhog* are the exact same animal, there is no difference in some dialects there is a preference for one term versus the other but in most dialects of English certainly North American English, there is no difference between those two terms.

How about that piece of furniture that is in most North American living rooms or sitting rooms or family rooms, and it's kind of long and pretty comfy; we frequently lay out on it, and probably take a nap. The fact that you have two, or even sometimes three different terms for that same piece of furniture, is arbitrary. Why should there even be two, let alone three? The fact that most dialects of English, you can use two of those interchangeably—certainly out here in California. In fact, most of the Western U.S.,

sofa and *couch* are 100% interchangeable; there is no difference. For many in the eastern half of North America, chances are you can use *davenport* along with either *sofa* or *couch*. Two, if not all three of those terms are in your native dialect of English.

How about the fact that those three most important things to human life—air, water, food—are completely different in every single language. Even in related languages, they get pronounced differently, even if they come from the same root. The fact that, for example, I have here water in English, the fact that in Spanish and Portuguese it is *agua*, depending on your dialect depends on what flavor of vowel that might have too, and the fact that their sibling language, French, has the same term, but it's not because it's pronounced differently and is spelled differently. Even though it comes from the same root, you do not say *agua* the same way that you say *eau*.

Why? Why is that the case? That really comes back to this concept of arbitrariness. There is no particular reason why a given term is linked to a given meaning; it just happens. Certainly, this is true with respect to things that we create, technology (remember, 'technology' is not only smartphones; it's more than that). The fact that we give a name to a given item that we create, yet that name does not necessarily get used in every single culture or speech community that uses that technology, is an example of arbitrariness. What we might call an English a (computer) *mouse*, first of all doesn't look much like a mouse anymore. Second, that term is not necessarily used in any other language beyond English. It could be used as-is, or it could be just tweaked a little bit, but that's arbitrary. It could be a totally different term, and it may not have anything to do with that rodent animal that is being used to describe this thing. All of that is arbitrary. Every single language, every single speech community could call it a totally different thing.

Arbitrariness also connects to this concept whereby we have the same sound, but you referring to multiple entities. For example, in English, that sound combination that gets written in English, as b-a-n-k, *bank*. That could refer to two totally different things. It could refer to a financial institution or it could refer to the side of a river. Those are unrelated in every way and yet the same sound combination, [bæŋk], is referred to two totally different things that is arbitrary. Why should English have that setup?

Just to take this one step further, if I give you a sound combination [ni]. In English that refers to a specific part of the body; in Spanish means 'not' or 'neither'; in French, its sibling language, it refers to a 'nest'. How can the same sound combination in three different languages, two of them very closely related? How can that be referring to three completely different concepts? That is arbitrariness.

Already arbitrariness is one of those concepts that we have to understand. As we analyze a language, the reason is straightforward. When we understand that a given language or even a given dialect has an arbitrary connection between sound and meaning, we then allow ourselves to analyze the language or the phenomenon in an objective way. We are not imposing our own biases; we're not imposing our own language on to that. We don't say that something is weird or odd. We just say that it is arbitrary. "It seems to happen here." "Why?" "Well, sometimes we have an answer, but frequently, we do not."

When we talk about arbitrariness, it's important to bring up that while the vast majority of the aspects of language are in fact arbitrary, there are a couple of exceptions with respect to the principle of the arbitrary. What I refer to as the fact that there are sometimes sound symbolisms. One of them, of course, is **onomatopoeia**, as we talked about in the introduction chapter. Onomatopoeia, of course, is when we say that a given term represents a sound. It could be a sound in nature or a sound that is caused by some process. Either way, it is called an onomatopoeic sound. It is true that there is some symbolism, some connection between the term and the meaning, but it's not exactly the case. Think back to that chart that I showed you where you have all of these different terms for the same types of sounds, but the representation was not the same across the board. In fact, there were a number of differences, even in sibling languages. Sometimes there was a difference in syllables, in the types of sounds being used to represent that other sound, but it really depended on a number of things. I'll go one step further: the fact that the sound of a dog barking in English has at least three different representations is still arbitrary, we can say 'bow wow', we can say 'wuff wuff', and we can say 'arf arf'. All three of those are onomatopoeics, so why should we have three different ones that don't necessarily share

anything in common with respect to their sounds. That is arbitrary, and that is where we see that the Principle of the Arbitrary really does rule everything with respect to language.

There is also a semi universal where the sound [i], that high front vowel. It tends to show up with respect to more themes that are diminutive. A diminutive is a type of morpheme that make things small and cute; think *eenie*, *-ity*, etc. Those kinds of more things in English, if you speak Spanish or French or Portuguese or Italian or any of the other Romance languages tend to be something in the *-ito*, *-ico*, *-ino* variety; there's lots of different versions. It is the case that in a large number of languages, and not just Indo-European languages, but throughout the world, we do see this high vowel get inserted into the morpheme for the diminutive.

Why is that the case? When you think about how you pronounce that sound, it is a high front bow and in the very top corner front corner of your mouth, so the sound is made in a very small place. There's a really small area where that sound is produced. Is that the reason? Maybe, but maybe not. There are plenty of languages that have a diminutive that do not have [i], as a part of that sound combination, in fact, may not even have front vowels at all. It's not a universal; it's a semi universal and, to be honest it's still arbitrary. Why should that sound [i] represents something small? Either way, onomatopoeics and even this diminutive rule, represent a very small fraction of all the possible combinations of sound and meaning, puzzle pieces putting together different words in any given language. Overall, arbitrariness really does rule the day.

1: Back to the Arbitrary is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2: Morphological Definitions

Compound Words, in Anderson's *Essentials of Linguistics*



Video Script

We've seen that English frequently uses affixation to derive new words. Affixation is quite **productive**, meaning that our mental grammar uses the process for many different words, even for new words that come into the language. You've probably generated new words yourself sometimes by adding affixes to existing words.

Another extremely productive derivational process in English is **compounding**. Compounding is different from affixation. In affixation, a bound morpheme is affixed to a base. Compounding derives a new word by joining two morphemes that would each usually be free morphemes.

For example, if I take the free morpheme *green*, an adjective, and combine it with the free morpheme *house*, a noun, I get the new word *greenhouse*. We can tell that this is a new word because its meaning is different from what we would get if we just combined the two words to make a phrase. We could walk down the street describing houses: This is a brown house and this one here is a tall house and here is a red house and here is a greenhouse. But a greenhouse is something different from a house that's green! It's a new word, derived by compounding.



Credit: Joe Shlabotnik on Flickr; Attribution 2.0 Generic (CC BY 2.0)



Image “Greenhouse at Wilson Farm, East Lexington MA“ by John Phelan is licensed under CC BY 3.0.

Another way that words derived by compounding differ from words derived by affixation is that a compound word doesn’t really have a base or root that determines the meaning of the word. Instead, both pieces of a compound make a sizeable contribution to the meaning. For example, *yoga pants* are pants that you wear to do yoga, and *emerald green* is the particular colour of green that emeralds are. So it doesn’t make sense to say that compounds have a root.

On the other hand, there is one part of a compound that has a special role, which we can see if we think about the categories of the words that make up a compound. If you look at these examples,

dry clean
stir fry
outrun
power wash

Each compound is made up of a different category of the word on the left plus a verb on the right. But in each case, the compound word is a verb. Even if both parts of a compound contribute to the meaning of the compound, it’s the **head** of a compound that determines its category. We say that English is a **head-final** language because in English the second part of the compound determines the category of the compound. Some languages are **head-initial**, with the head as the first element in a compound.

In many compounds, the head determines the category and also constrains the meaning of the compound. So *dog food* is a kind of food, not a kind of dog, and *yoga pants* are a kind of pants, not a variety of yoga. Compounds like this, where the meaning relationship between the head and the whole compound is obvious, are called **endocentric**. But in some compounds, the meaning relationship is not so transparent. For example, a *redhead* is a person, not a kind of head; a *nest egg* is money that you’ve saved, not a kind of egg; a *workout* is not a particular kind of out, and *facebook* is not a book at all! Compounds where the meaning of the head does not predict the meaning of the compound are said to be **exocentric**.

Check Yourself

? Exercise 2.1

In the sentence, “The room contained a bearskin rug,” what kind of compound is *bearskin*?

- Endocentric.
- Exocentric.

Answer

"Endocentric."

Hint: Think about what *bearskin* means, and the parts are represented there.

? Exercise 2.2

In the sentence, “Randy worked as a cowhand on the ranch,” what kind of compound is *cowhand*?

- Endocentric.
- Exocentric.

Answer

"Exocentric."

Hint: Think about what *cowhand* means, and the parts aren't represented there. Cows don't have hands.

? Exercise 2.3

In the sentence, “Hyunji manages a bookshop,”, what kind of compound is *bookshop*?

- Endocentric.
- Exocentric.

Answer

"Endocentric."

> Hint: Think about what *bearskin* means, and the parts are represented there.

Closed Class Categories (Function Words), in Anderson's *Essentials of Linguistics*

In Unit 7.1 we saw that linguists group words into syntactic categories according to how they behave. Content words convey a lot of the meaning of a sentence. But not many sentences would be complete if they contained only nouns, verbs, or adjectives. There are also several smaller categories of words called **closed-class categories** because the language does not usually add new words to these categories. These categories don't have many members, maybe only a few dozen, in contrast with the many thousands of words in the open-class categories. They're the **function words** or **non-lexical categories** that do a lot of grammatical work in a sentence but don't necessarily have obvious semantic content.

The category of **determiners** doesn't have many members but its members occur very frequently in English. The two little words *the* and *a* are the most recognizable members. Determiners most often appear before a noun, as in:

a student

an orange

the snake

the ideas

Any word that can appear in the same position as *the* counts as a determiner, like demonstratives:

those students

these oranges

that snake

this idea

Quantifiers and numerals also behave like determiners:

many students

twelve oranges

most snakes

several ideas

And the words that you might have encountered as "possessive adjectives" or "possessive pronouns" behave like determiners as well:

my sister

your idea

their car

The category of **prepositions** seems to have slightly more obvious semantic content than most other closed classes. Prepositions often represent relationships in space and time. They also have consistent syntactic distribution, usually appearing with a noun phrase immediately following them:

on the table

in the basket

around the block

through the centuries

near campus

after class

A very small category of words that does an important job are the **conjunctions**. There are only three conjunctions, *and*, *or*, *but*. The job that conjunctions do is to conjoin two words or phrases that belong to the same category:

oranges and lemons

brushed her teeth and went to bed

strong and fast

soup or salad

singing or dancing

hated her roommate but loved her roommate's sister

small but mighty

You might have learned that words like *because* and *although* are a type of conjunction, but they don't behave like *and*, *or*, *but*. Their behaviour is more similar to a category of words we label as **complementizers**. Complementizers are function words that introduce a clause, which is a sentence embedded inside a larger sentence:

Sam told us that she loved baseball.

She hoped that the Blue Jays would win the World Series.

Leilani wondered whether it would rain that afternoon.

She asked her roommate if she had heard the forecast.

The roommate checked the forecast because she wanted to go for a run.

She decided to go running although a storm was forecast.

Mel washed the dishes while the cupcakes were in the oven.

Morphological Definitions, from Sarah Harmon



Video Script

I'm going to add a little bit more with respect to the morphological definitions that Catherine Anderson has posted above. This just gives a little more clarity when we're talking about what these definitions are, as we analyze morphology.

In the previous section use the term **morpheme**. What is a morpheme? It is the smallest unit in a language that has meaning there's minimal matching of sound and meaning. That's a very fancy definition so let's break that down a little bit more. If I have the term *preschool*, it is a lexicon; it is a minimal free form. In this case, *preschool* has two different morphemes, two different puzzle pieces that combined to create that lexicon. There is the morpheme *school*, and that is the root, and there is the morpheme *pre-*, a prefix that is a type of affix. They combine their two puzzle pieces to building blocks to combine to make the lexicon *preschool*.

You'll notice I keep saying **lexicon**; I'm not using the term 'word' and there's a good reason, as we saw in Chapter 1. 'Lexicon' is just a little bit more precise. We often use 'term' or 'terminology' as the same concept; 'lexicon' is used in linguistic circles. The term 'word' is not exactly precise enough, because in a given language, a word could have just a couple different morphemes put together and is an entire lexicon of a specific part of speech—it's all a noun or a verb or an adjective, etc. The problem is that in many languages, that is not the case that a word is frequently multiple pieces put together multiple parts of speech, all combined; we'll see examples of that later on. We use the term 'lexicon' to denote that this is a unit, a free-standing unit doesn't have to be combined with anything else; maybe it is, but it doesn't have to be it can stand alone. A lexicon can have only one single morpheme, or it can have multiple morphemes.

Remember that arbitrariness is always going to be a factor, that we are combining morphemes in a certain way, and that is language specific. Why a given language combines morphemes in one way versus another, that's arbitrary. Most of you have experienced in learning a second language, whether it was English or another language and, in many cases they had very different rules as far as how to put lexicon together, how to put morphemes together, and how to put sentences together. All of those rules are arbitrary; those no real reason why North American English does one thing, but South Asian English does a different thing. There's no reason why English does one thing, but Japanese does a different thing, and Kikuyu does a third thing, and Cherokee does a fourth thing. It simply is the way the languages work; they're put together in different ways.

One more term that I think will be very crucial, not just for morphology but as we go further, and especially in semantics and pragmatics when we get to Meaning, and that is a **mental lexicon**. I think of a mental lexicon as a type of mental dictionary, but it does so much more than a dictionary does. A dictionary tells you the definition, it might tell you the part of speech, it may give you the pronunciation, but it really doesn't give you much more. A mental lexicon gives you all of that, plus the slew of grammatical functions, a very large corpus of synonyms, antonyms, maybe relational memory and meaning. We'll come back to this when we get to Meaning and, specifically, having to refer to an ontology but more on that late. For now, as we go through morphology, just remember that these are mental lexicon are our storage units, or a box of knowledge with respect to a given language or even a given dialect. It includes all the possible entries of lexicon and the relationships.

2: Morphological Definitions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **7.4: Compound Words** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

3: Morphemes

4.3.1 From 6.1 Words and Morphemes, in Anderson's *Essentials of Linguistics*

What's a word? It seems almost silly to ask such a simple question, but if you think about it, the question doesn't have an obvious answer. A famous linguist named Ferdinand de Saussure said that a word is like a coin because it has two sides to it that can never be separated. One side of this metaphorical coin is the **form** of a word: the sounds (or letters) that combine to make the spoken or written word. The other side of the coin is the **meaning** of the word: the image or concept we have in our mind when we use the word. So a word is something that links a given form with a given meaning.

Linguists have also noticed that words behave in a way that other elements of mental grammar don't because words are **free**. What does it mean for a word to be free? One observation that leads us to say that words are free is that they can appear in **isolation**, on their own. In ordinary conversation, we don't often utter just a single word, but there are plenty of contexts in which a single word is indeed an entire utterance. Here are some examples:

> What are you doing? Cooking.
>
> What are you cooking? Soup.
>
> How does it taste? Delicious.
>
> Can I have some? No.

Each of those single words is perfectly grammatical standing in isolation as the answer to a question.

Another reason we say that words are free is that they're **moveable**: they can occupy a whole variety of different positions in a sentence. Look at these examples:

> Penny is making soup.
>
> Soup is delicious.
>
> I love to eat soup when it's cold outside.

The word *soup* can appear as the last word in a sentence, as the first word, or in the middle of a sentence. It's free to be moved around.

The other important observation we can make about words is that they're **inseparable**: We can't break them up by putting other pieces inside them. For example, in the sentence,

> Penny cooked some carrots.

The word *carrot* has a bit of information added to the end of it to show that there's more than one carrot. But that bit of information can't go just anywhere: it can't interrupt the word *carrot*:

> *Penny cooked some car-s-rot.

This might seem like a trivial observation – of course, you can't break words up into bits! – but if we look at a word that's a little more complex than *carrots* we see that it's an important insight. What about:

> Penny bought two vegetable peelers.

That's fine, but it's totally impossible to say:

> *Penny bought two vegetables peeler.

even though she probably uses the peeler to peel multiple vegetables. It's not that a plural *-s* can't go on the end of the word *vegetable*; it's that the word *vegetable peeler* is a single word (even though we spell it with a space between the two parts of it). And because it's a single word, it's inseparable, so we can't add anything else into the middle of it.

So we've seen that **a word is a free form that has a meaning**. But you've probably already noticed that there are other forms that have meaning and some of them seem to be smaller than whole words. **A morpheme is the smallest form that has meaning**. Some morphemes are **free**: they can appear in isolation. (This means that some words are also morphemes.) But some morphemes can only ever appear when they're attached to something else; these are called **bound** morphemes.

Let's go back to that simple sentence,

Penny cooked some carrots.

It's quite straightforward to say that this sentence has four words in it. We can make the observations we just discussed above to check for isolation, moveability, and inseparability to provide evidence that each of *Penny*, *cooked*, *some*, and *carrots* is a word. But there are more than four units of meaning in the sentence.

Penny cook-ed some carrot-s.

The word *cooked* is made up of the word *cook* plus another small form that tells us that the cooking happened in the past. And the word *carrots* is made up of *carrot* plus a bit that tells us that there's more than one carrot.

That little bit that's spelled *–ed* (and pronounced a few different ways depending on the environment) has a consistent meaning in English: past tense. We can easily think of several other examples where that form has that meaning, like *walked*, *baked*, *cleaned*, *kicked*, *kissed*. This *–ed* unit appears consistently in this form and consistently has this meaning, but it never appears in isolation: it's always attached at the end of a word. It's a **bound morpheme**. For example, if someone tells you, "I need you to walk the dog," it's not grammatical to answer "-ed" to indicate that you already walked the dog.

Likewise, the bit that's spelled *–s* or *–es* (and pronounced a few different ways) has a consistent meaning in many different words, like *carrots*, *bananas*, *books*, *skates*, *cars*, *dishes*, and many others. Like *–ed*, it is not free: it can't appear in isolation. It's a **bound morpheme** too.

If a word is made up of just one morpheme, like *banana*, *swim*, *hungry*, then we say that it's **morphologically simple**, or **monomorphemic**.

But many words have more than one morpheme in them: they're **morphologically complex** or **polymorphemic**. In English, polymorphemic words are usually made up of a **root** plus one or more **affixes**. The root morpheme is the single morpheme that determines the core meaning of the word. In most cases in English, the root is a morpheme that could be free. The affixes are bound morphemes. English has affixes that attach to the end of a root; these are called **suffixes**, like in *book<u>s</u>*, *teach<u>ing</u>*, *happi<u>er</u>*, *hope<u>ful</u>*, *sing<u>er</u>*. And English also has affixes that attach to the beginning of a word, called **prefixes**, like in *<u>un</u>zip*, *<u>re</u>heat*, *<u>dis</u>agree*, *<u>im</u>possible*.

Some languages have bound morphemes that go into the middle of a word; these are called **infixes**. Here are some examples from Tagalog (a language with about 24 million speakers, most of them in the Philippines).

[takbuh]	*run*	[tumakbuh]	*ran*
[lakad]	*walk*	[lumakad]	*walked*
[bili]	*buy*	[bumili]	*bought*
[kain]	*eat*	[kumain]	*ate*

It might seem like the existence of infixes is a problem for our claim above that words are inseparable. But languages that allow infixation do so in a systematic way — the infix can't be dropped just anywhere in the word. In Tagalog, the position of the infix depends on the organization of the syllables in the word.

4.3.2: From 6.2 Allomorphs, in Anderson's *Essentials of Linguistics*

The previous unit showed us that a **morpheme** is the smallest unit that pairs a consistent form with a consistent meaning. But when we say that the form of a morpheme is consistent, there's still some room for variability in the form. Think back to what you know about phonology and remember that a given phoneme can show up as different allophones depending on the surrounding environment. Morphemes work the same way: a given morpheme might have more than one **allomorph**. Allomorphs are forms that are related to each other but slightly different, depending on the surrounding environment.

A simple example is the English word *a*. It means something like "one of something, but not any particular one", like in these examples:

a book
a skirt
a friend
a phone call

But if the word following *a* begins with a vowel and not a consonant, then the word *a* changes its form:

an apple
an ice cream cone
an iguana
an idea

The two forms *a* and *an* are slightly different in their form, but they clearly both have the same meaning. And each one shows up in a different predictable environment: *a* before words that start with consonants and *an* before words that begin with vowels.

Another example of allomorphy in English is in the plural morpheme. In written English, the form of the plural morpheme is spelled *-s*, as in:

carrots
books
hats
friends
apples
iguanas

But it's spelled *–es* in words like:

churches
bushes
quizzes

And in fact, even in the cases where it's spelled *-s*, it's pronounced as [s] for words that end in a voiceless segment (*carrots*, *books*, *cliffs*) and as [z] for words that end in voiced sounds (*worms*, *dogs*, *birds*). So it's got two written forms (*-s* and *-es*) and three spoken forms ([s], [z], [ɨz]), but a consistent meaning of "more than one". Each form is an allomorph of the plural morpheme. Can you figure out what the relevant environment is that predicts which allomorph appears where?

4.3.3: From 6.3 Inflectional Morphology, in Anderson's *Essentials of Linguistics*



Video Script

We saw in our last units that words can be made up of morphemes, which are the smallest linguistic unit that links form with meaning. Morphemes can do a couple of quite different jobs in a word.

Inflectional morphemes are morphemes that add grammatical information to a word. When a word is inflected, it still retains its core meaning, and its category stays the same. We've actually already talked about several different inflectional morphemes:

The **number** on a noun is inflectional morphology. For most English nouns the inflectional morpheme for the **plural** is an *–s* or *–es* (e.g., *books*, *cars*, *dishes*) that gets added to the singular form of the noun, but there are also a few words with irregular plural morphemes. Some languages also have a special morpheme for the **dual** number, to indicate exactly two of something. Here's an example from Manam, one of the many languages spoken in Papua New Guinea. You can see that there's a morpheme on the noun woman that indicates dual, for exactly two women, and a different morpheme for plural, that is, more than two women.

Manam (Papua New Guinea)		
/áine ŋara/	*that woman*	singular
/áine ŋaradiaru/	*those two women*	dual
/áine ŋaradi/	*those women*	plural

The **tense** on a verb is also inflectional morphology. For many English verbs, the past tense is spelled with an *–ed*, (*walked*, *cooked*, *climbed*) but there are also many English verbs where the tense inflection is indicated with a change in the vowel of the verb (*sang*, *wrote*, *ate*). English does not have a bound morpheme that indicates future tense, but many languages do.

Another kind of inflectional morphology is **agreement** on verbs. If you've learned French or Spanish or Italian, you know that the suffix at the end of a verb changes depending on who the subject of the verb is. That's agreement inflection. Here are some examples from French. You can see that there's a different morpheme on the end of each verb depending on who's doing the singing.

French		
1st	je chante	*I sing*
2d	tu chantes	*you sing*
3d	elle chante	*she sings*
1st	nous chantons	*we sing*
2d	vous chantez	*you (pl.) sing*
3d	elles chantent	*they sing*

And in some languages, the morphology on a noun changes depending on the noun's role in a sentence; this is called **case** inflection. Take a look at these two sentences in German: The first one, *Der Junge sieht Sofia*, means that, "The boy sees Sofia". Look at the form of the phrase, *the boy*, "*der Junge*". Now, look at this other sentence, *Sofia sieht den Jungen*, which means that "Sofia sees the boy". In the first sentence, the boy is doing the seeing, but in the second, the boy is getting seen, and the word for boy, *Junge* has a different morpheme on it to indicate its different role in the sentence. That's an example of **case** morphology, which is another kind of inflection.

German	
Der Junge sieht Sofia.	*The boy sees Sofia.*
Sofia sieht den Jungen	*Sofia sees the boy.*

Check Yourself

? Exercise 3.1

What type of grammatical information does the inflectional affix in the word *speeches* communicate?

- Number.
- Tense.
- Subject agreement.
- Case.

Answer

"Number."

Hint: If you take off the *-es*, it makes the base noun singular.

? Exercise 3.2

What type of grammatical information does the inflectional affix in the word *climbed* communicate?

- Number.
- Tense.
- Subject agreement.
- Case.

Answer

"Tense."

Hint: If you take off the *-ed*, it makes the base verb present tense (in a way; there is another process going on).

? Exercise 3.3

What type of grammatical information does the inflectional difference between *he* and *him* indicate?

- Number.
- Tense.
- Subject agreement.
- Case.

Answer

"Case."

Hint: These pronouns are used in different roles in the sentence.

4.3.4: From 6.4 Derivational Morphology, in Anderson's *Essentials of Linguistics*



Video Script

The last unit talked about inflection, which is one of the jobs that morphology can do. The other big job that morphemes have is a **derivation**. The derivation is the process of creating a new word. The new, derived word is related to the original word, but it has some new component of meaning to it, and often it belongs to a new category.

One of the most common ways that English derives new words is by affixing a derivational morpheme to a **base**. For example, if we start with a verb that describes an action, like *teach* and we add the morpheme *–er*, we derive a morphologically complex noun, *teacher*, that refers to the person who does the action of teaching. That same *-er* morpheme does the same job in *singer*, *dancer*, *baker*, and *writer*.

Verb	**Suffix**	**Noun**
teach	-er	teacher
sing	-er	singer
dance	-er	dancer
bake	-er	baker
write	-er	writer

If we start with an adjective like *happy* and add the suffix *–ness*, we derive the noun that refers to the state of being that adjective, *happiness*.

Verb	**Suffix**	**Noun**
teach	-er	teacher
sing	-er	singer
dance	-er	dancer
bake	-er	baker
write	-er	writer

Adding the suffix *–ful* to a noun derives an adjective, like *hopeful*.

Noun	**Suffix**	**Adjective**
hope	-ful	hopeful
joy	-ful	joyful
care	-ful	careful
dread	-ful	dreadful

Adding the suffix*–ize* to an adjective like *final* derives a verb like *finalize*.

Adjective	**Suffix**	**Verb**
final	-ize	finalize
modern	-ize	modernize
social	-ize	socialize
public	-ize	publicize

Notice that each of the morphologically complex derived words is related in meaning to the base, but it has a new meaning of its own. English also derives new words by prefixing, and while adding a derivational prefix does lead to a new word with a new

meaning, it often doesn't lead to a category change.

Prefix	Verb	Verb
re-	write	rewrite
re-	read	reread
re-	examine	reexamine
re-	assess	reassess

Each instance of derivation creates a new word, and that new word could then serve as the **base** for another instance of derivation, so it's possible to have words that are quite complex morphologically.

For example, say you have a machine that you use to compute things; you might call it a *computer* (*compute* + *-er*).Then if people start using that machine to perform a task, you could say that they're going to *computerize* (*computer* + *-ize*) that task. Perhaps the *computerization* (*computerize* + *-ation*) of that task makes it much more efficient. You can see how many words have many steps in their derivations.

An interesting thing to note is that once a base has been inflected, then it can no longer go through any derivations. We can inflect the word *computer* so that we can talk about plural computers, but then we can't do derivation on the plural form (**computers-ize*). Likewise, we can add tense inflection to the verb *computerize* and talk about how yesterday we *computerized* something, but then we can't take that inflected form and use it as the base for a new derivation (**computerized-ation*). Inflection always occurs as the last step in word formation.

? Exercise 3.4

Which of the following best describes the derivation of the word *assignment*?

- Noun + *–ment* ➔ Verb.
- Adjective + *–ment* ➔ Noun.
- Verb + *–ment* ➔ Noun.
- Verb + *–ment* ➔ Verb.

Answer

"Verb + *–ment* ➔ Noun."

Hint: Think about the role of *assign* and the role of *assignment*.

? Exercise 3.5

Which of the following best describes the derivation of the word *skillful*?

- Adjective + *–ful* ➔ Verb.
- Adjective + *–ful* ➔ Adjective.
- Verb + *–ful* ➔ Noun.
- Noun + *–ful* ➔ Adjective.

Answer

"Noun + *–ful* ➔ Adjective."

Hint: Think about the role of *skill* and the role of *skillful*.

Exercise 3.6

Which of the following best describes the derivation of the word *simplify*?

- Verb + *–ify* ➔ Verb.
- Adjective + *–ify* ➔ Verb.
- Noun + *–ify* ➔ Adjective.
- Adjective + *–ify* ➔ Noun.

Answer

"Adjective + *–ify* ➔ Verb."

Hint: Think about the role of *simple* and the role of *simplify*.

4.3.5: From 6.5 Inflectional Morphology of Some Indigenous Languages, in Anderson's *Essentials of Linguistics*

Talking about morphology when your primary language is English is sometimes a little disappointing because English does not have very much inflectional morphology. Many other languages do much more interesting jobs with inflectional morphology. Many of the Indigenous Languages spoken by the First Peoples of what is currently Canada have rich morphological systems that communicate a great deal of information.

Number in Inuktitut

Inuktitut is one of the dialects spoken by the Inuit people who live in the Arctic region. There is a good deal of dialect variation across the Inuit languages. Inuktitut is the variety that is the official language of the territory of Nunavut, and has about 40,000 speakers.

All languages make a distinction between singular and plural nouns, but some languages, like Inuktitut, also use inflectional morphology to indicate **dual** number when there are exactly two of something, as in the following examples:

matu	*door*
matu**uk**	*doors (two)*
matu**it**	*doors (three or more)*

nuvuja	*cloud*
nuvuja**ak**	*clouds (two)*
nuvuja**it**	*clouds (three or more)*

qarasaujaq	*computer*
qarasauj**aak**	*computers (two)*
qarasauja**it**	*computers (three or more)*

Subject-Verb Agreement in Inuktitut

The three-way distinction between singular, dual, and plural in Inuktitut applies not only to nouns but also to verbs that agree with their noun subjects:

nirijunga	*I eat*
nirijuguk	*the two of us eat*
nirijugut	*we (three or more) eat*

nirijutit	*you (one of you) eat*
nirijusik	*you two eat*
nirijusi	*you (three or more) eat*

nirijuq	*he or she eats*
nirijuuk	*the two of them eat*
nirijut	*they (three or more) eat*

Animacy in Cree

The Cree languages are the Indigenous languages that have the greatest number of speakers, about 80,000 according to Statistics Canada's 2016 Census.

You might know a language that categorizes nouns according to their gender, like French, which makes a distinction between masculine and feminine nouns, adjectives, and determiners. Of course, grammatical gender has a quite arbitrary relationship to concepts of social and biological gender. Other languages categorize nouns along different criteria. Cree distinguishes words along a dimension called **animacy**. The animacy distinction is approximately related to whether something is alive or not, but the categories for animate vs. inanimate things are somewhat arbitrary, just the like the categories for masculine vs. feminine things in languages that mark grammatical gender. The animacy of a noun affects which demonstrative determiners may be used with it, the form of the plural morphology, and the morphology of the verb that agrees with it.

In Plains Cree (Nêhiyawêwin), the noun *atim* (dog) is animate, while *astotin* (hat) is inanimate. The sentences below shows how the noun's animacy affects the other words in the sentence.

	animate	inanimate
singular	atim*dog*	astotin*hat*
plural	atim**wak***dogs*	astotin**a***hats*
singular demonstrative	**awa** atim*this dog*	**ôma** astotin*this hat*
plural demonstrative	**ôki** atim**wak***these dogs*	**ôhi** astotin**a***these hats*
singular transitive verb	niwâpam**âw** atim.*I see a dog.*	niwâpa**htên** astotin*I see a hat.*
plural transitive verb	niwâpamâw**ak** atim**wak**.*I see dogs.*	niwâpa**htên** astotin**a***I see hats.*

Pronouns

All languages make at least a three-way distinction among pronouns — the first person (*I/me* in English) is the person talking; second person (*you*) is the person being addressed, and third person (*she, he, they, it*, etc.) is anybody or anything else. Some languages make even more distinctions in pronouns.

In Ojibwe (Anishnaabemowin), which has about 20,000 speakers, there are two pronouns for the first-person plural. The pronoun *niinwi* refers to the speaker plus other people but not the person being addressed (that is, "we but not you"). This is known as the exclusive we. The pronoun for inclusive we ("all of us including you") is *kiinwi*. The distinction between inclusive and exclusive *we* is sometimes referred to as **clusivity**.

Cree also makes an inclusive/exclusive distinction in the first-person plural. The inclusive form is *niyanân* and the exclusive form is *kiyânaw*.

Subject-Verb Agreement in Cree

In the third person, Cree makes a distinction between **proximate** and **obviative** third person. You might think of this distinction as something similar to the near/far distinction between *this* and *that* in English, where *this* is used for something that is closer to the speaker and *that* is for something farther away. But, like in English, the proximate/obviative distinction is not just about physical distance; it can also allude to distance in time, or within a conversation, to someone that has been mentioned recently (proximate)

versus someone that is being mentioned for the first time (obviative). The distinction is marked on the verbal morphology, as illustrated below:

proximate	Regina wîkiwak. *They live in Regina.*	kiskinwahamâkosiwak. *They are in school.*
obviative	Regina wîkiyiwa. *Their friend/someone else lives in Regina.*	kiskinwahamâkosiyiwa. *Their friend/someone else is in school.*

These few examples illustrate that the rich morphological systems of these languages can communicate a great deal of information efficiently.

3: Morphemes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.3: Inflectional Morphology** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **6.4: Derivational Morphology** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

4: Affixation and Other Morphological Processes

Affixation and Other Morphological Processes, from Sarah Harmon



Video Script

I wanted to take a moment to talk about some specific affixation and other morphological processes. I want to not just focus on English; I want to show you what happens in so many other languages. Some of these languages will look familiar to you, while others will be very different and new. That's the point: I want to showcase to you the variety that we have within language.

When we start off with morphological processes, we always start with affixation. You probably have heard at some point of the terms prefix and suffix. A **prefix** goes before the root; a **suffix** goes after the root. They're very common, not just in English, but in all Indo-European languages; we love both prefixation and suffixation. We tend to use prefixation for derivational morphology; suffixation can be used both for derivational morphology and inflectional morphology, not just an English, but in most Indo-European languages.

I'm showcasing here an example of suffixation from Spanish because I think it would be interesting for many of you to understand exactly what happens in Spanish, with respect to verbs. Most of you either are fluent in Spanish natively or learned, or you have learned about Spanish, so this should give you a new vision, as it were, with respect to verbs in Spanish.

Realistically, in Spanish, and all the Romance languages, we have a root, followed by a series of suffixes that indicate what that verb is going through, the various types of inflection. When we talk about inflection, especially in the Romance languages, but even in other Indo-European languages, there is a big combination of roles into one affix. In this case, I have the verb *tener* that means 'to have', as in to possess something, to own something. You see that *ten-* is the root and for many, many forms of the verb it just stays as *ten-*.

(Of course, if you know, Spanish, you know there is a stem change involved, but that's for a Spanish class, that's not morphology. Right now, let's just focus on the fact that you have a root, ten-.)

You have a specific suffix that goes afterwards, and that suffix can tell you a lot of different information. If it ends in [-er]—Spanish is written fairly phonetically so I'm leaving it here in how it is written, because Spanish is written basically in IPA—that -er suffix [-eɾ] tells you is that it is the infinitive, the base form of the verb. Think of it as the form of the verb that you would look up in the dictionary;in English, that would be 'to have'. If we swap that *-er* for *-emos*, now we have conjugated that verb, but in this specific way. As we can see in the middle three examples, *-mos* shows up in all three of them. We notice that we have first person plural as an indicator for all three of them. First person is the speaker; second person is the audience; third person is somebody not involved in the conversation. Plural, so 'we'. 'We' is first person plural, in English, *nosotros* is the form in Spanish, and the *-mos* at the end of that inflection tells you it is first person plural. That vowel or lead-in to the *-mos* tells you so much more. It can tell you the tense, the aspect, and the mood Those are three elements of deixis; we'll come back to that when we go to pragmatics in the

chapter on Meaning. But for now, know that *-emos*, that *-e-* refers to both present tense and indicative, meaning something that stated declared. *-íamos*, that *-ía-* is still indicative, but it is past tense and durative aspect, meaning ongoing. And the *-dre-* in the *-dremos* is future indicative.

Those of you who speak Spanish especially native speakers of Spanish this is breaking down those puzzle pieces and understanding what they refer to. Think about, Spanish speakers, what *tener*, *tenemos*, *teníamos*, *tendremos* mean. They have multiple connotations, but each one of them have different roles, then *tener* is the infinitive 'to have'. *Tenemos* is' we have right now', as a declarative statement. *Teníamos* would be 'we used to have' or something along those lines. *Tendremos*, 'we will have', or 'we might have in the future'. Those are all different versions, and that inflection carries with it multiple levels of meaning. Then, just to show you a little bit more, if we want to do a past participle so a form of the verb that gets used a lot of compound forms it's the *-ido*, which tells you it is a past participle. I bet you didn't realize those inflections, those suffixes in Spanish carried so much meaning. This is just Spanish, and just one verb in Spanish. We see this throughout all human languages; if you have inflection, it carries a lot of meaning and, in this case, it's a suffix.

But there's two more types of affixation that we do not see in most Indo-European languages, although there is a little caveat to one of them. We see we have infixation and we have circumfixation. Let's start with **infixation**; as you would expect infixation means in the root. I'm going to show you two different examples; one is a more traditional way of thinking about infixation. This is Bontoc, which is one of the Filipino languages; the Filipino languages are part of a larger family called Australio-Pacific. This is a huge language family that covers the languages all the way out to Rapa Nui and Polynesia (think Hawaii, Fiji, Tonga, Samoa, even, New Zealand and the Maori language) and then covers all the languages throughout the Pacific Islands, through the northern coast of Australia, Papua New Guinea, Indonesia, all of that area. They are the poster children for infixation; Indo-European languages love suffixation for inflection, but Australio-Pacific languages love infixation. If you speak Tagalog, Ilokano, or any other Filipino language, as well as the other Australio-Pacific languages, you know about this.

For example, you want to take an adjective and you want to make it a verb, like 'to be…' that adjective to be strong, to be read what you do is you infix *-um-* right after the first consonant: *fikas* 'strong' becomes *fumikas* 'to be strong'; *kilad* 'red', the color red becomes *kumilad* 'to be read'. This is just one example; all of the Australio-Pacific languages love infixation, especially for any kind of inflection or derivation.

There's a second kind that you only see in the Semitic languages; I have Hebrew here, but Arabic, Amharic and all of the Semitic languages, which is a branch of the Afro Asiatic language family, they all have some kind of continental root. The vowels are the different inflections and derivation of morphine this is radically different from just about anything else that we see in any other language family, it makes the Semitic languages stand out in a very unique way. For example, *ktb*, if you will, but I'm going to use the pronunciation, the IPA. [ktb] is the continental root for 'to write' in Hebrew; I believe it's also the same in Arabic and it might even be in Amharic. But we're going to stick to Hebrew. Depending on which vowels you put in and where you put them, that depends on how you conjugate the verb 'to write'. For example, 'katab' with two 2 [a] thrown in there, that is just a declarative statement, 'write'. Compare that with [kutib], sticking an [u] and [i] in between those consonants, that means 'have been written' like 'the letter has been written', a perfect passive combination. If you use [a] and [u] and change where they go [aktub] this is 'be writing' like 'my mother is writing a letter'. If you change that [a] to an [u], and the [u] to an [a] so flip the vowels same placement, that is the passive, something 'is being written', so 'the letter is being written'. The vowels that you choose and where you choose to place them drastically change the meaning of what you're trying to say.

This is arbitrariness. Why the Semitic languages do this and seemingly no other language that we know of does is very peculiar and interesting and amazing.

The last type of application is called **circumfixation** and it means just that the inflection goes around the root. While it is not a very common way of doing inflection, it is something we see in the other Germanic languages—not English, although we used to, and that's a story for historical linguistics. German, Dutch, the Scandinavian languages, Icelandic, all of them, they all do some kind of circumfixation. Frequently, this is to form the past participle of the verb. To conjugate the verb to that 'have loved', 'have kissed', etc. Notice that it is not just the [ge-] in the front, the prefix, but also [-t] at the end, the suffix. It's a prefix and suffix combined; that is circumfixation.

A couple more processes that we should bring up right now. **Umlaut** or **morpheme-internal change**; usually, this is a vowel, which is why umlaut is the term we frequently use. English certainly is an example of this, although it is not a productive element; it is something that used to be rife in Old English. The difference between *man* and *men*, *goose* and *geese*, *hold* and *held*. Certain plural formations and certain past tense formations show umlaut. Those vowel changes, that is what umlaut is. You may have heard the term 'umlaut' to refer to those two dots that go over a vowel (e.g. ü, ä), especially in Germanic languages; in fact umlaut is a

German word, and that is why it tends to be referred to with respect to vowels that change. In Modern English, certainly for the last several hundred years, this is unpredictable; it is a relic, as it were. When we get to historical linguistics we'll talk more about this relic. In some cases, there are changes, but in most cases, they just have **fossilized**—they've stayed the same.

Suppletion is the last one. **Suppletion** is when we have what we call a radical change of the morpheme. Another term for suppletion is **merger** and that really does describe what happens when you have two or more forms that merged together and form a single form. For example, if you want to impress some people and talk about language in a party setting (and who doesn't!), bring up this little factoid: if a language has the verb 'to be', it will be the most irregular form in the entire language, and if it has the verb 'to go' and pretty much every language does, it will be the second most irregular form in the entire language. That is because over the history of human language, there have frequently been multiple versions of 'to be', either as an actual verb on its own or idiomatic versions of it. These multiple forms get suppleted or merged together over time. If you've ever wondered why 'to be' has a present conjugation as *I am*, *you are*, *he* or *she is*, with the past tense conjugation is *I was*, *you were*, and wondering how those forms match, let alone how they match 'to be'? That's suppletion; that's merger. The forms merged together over time at some point in the history of English; there were separate verbs and they merge together. The same with 'to go' that has a past tense form of 'went'; again, there used to be two different verbs meaning ‘to go’ in Old English and they merged together. We'll come back to this when we get to historical linguistics. In fact, affixation as a whole is a huge area of historical linguistics that yields so much fruit with how a language has changed over time and how maybe it might change in the future.

4: Affixation and Other Morphological Processes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5: Morpho-phonology

Morpho-phonology, from Sarah Harmon



Video Script

There's a phrase that we say with respect to language, that language does not occur in a vacuum. What that means is that language doesn't work just on its own. Any given area of language doesn't just change or stay the same on its own. Every aspect of language is related to every other aspect of language. When we talk about morphology, we can't just talk about lexicon and morphemes; we have to also include other aspects of language. The next several sections in this chapter are going to cover combinations of morphology and some other aspect of language. In this case, we're going to focus on sound so morph-phonology.

Just to give you a little hint, when we were talking about phonological rules in the previous chapter, some of those rules did not have anything to do with morphology at all. Think about, for example, the aspiration rule in English, when consonants are in the front of a stressed syllable, they get aspirated—that little puff of air. That has nothing to do with the morphology—it has only to do with phonology—so that is a **morphologically-irrelevant sound change**.

That being said, in phonology we have seen various types of **morphologically-relevant rules**:

- Whether it was the Hebrew reflective pronoun [lehit] that swaps places if the next verb it goes with starts with a sibilant;
- Whether it's Spanish and its epenthesis rule, the fact that an [s] + consonant cannot start off a word or lexicon; you have to epenthesize, you have to add that [e] in front of it;
- Whether it’s Classical Latin with respect to some of its lengthening;

these rules have to do with morphology in some way.

Let me show you something that probably hits a little closer to home. Look at these pluralized nouns. I've written them, both in the English writing system, as well as in IPA.

- buts,
- buds,
- buses,
- bushes,
- batches,
- badges,
- buys,
- bins,
- bills.

Notice that we technically write the plural of these nouns, the plural affixes, in one of two ways, either as an *-s* or as an *-es*. You were probably taught it has something to do with that sound that goes before. However, look at the IPA we have [-s], we also have [-z], and we have [-ɨz]; for us in western North American English, we tend to see [-ɨz] but in many cases, you also can say [-əz], so it's either a schwa [ə] or a barred 'i', that central high vowel [ɨ]. Either way, there's a vowel there in between that [-z] and the previous syllable. We don't actually have two; we have three forms of that plural suffix, the plural inflection. Can you spot the difference? Can you tell when you use each one of those forms? I'll give you a minute; pause here for a second.

Okay, are you ready to see the answer?

The rule is if the last constant of the root is a sibilant—an 's' like sound—then you have a dissimilation rule and you use [-ɨz] or [-əz]. If you have a voiceless consonant as that last consonant sound, then you have [-s], an assimilation rule. If you have any voiced constant or vowel or semi consonant/glide, then you have [-z] as the rule. That is it; it depends on the last sound of that root. Is it a sibilant? Is it voiceless? Or, is it voiced? That is an example of a morph-phonologic rule.

We talked about phonemes in an earlier chapter, which went in slashes and were the default form, and then allophones, which went in square brackets and went to describe what we actually do what we actually pronounce. The phoneme is always one of those variants and then we have multiple versions. We can do the same with morphemes: the **morpheme** is going slashes and then how they are realized how they are pronounced are the **allomorphs** and they go in square brackets. We will work with this more towards the end of the chapter when we go to analyze morphological data sets.

5: Morpho-phonology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6: Parts of Speech and Word Formation

4.6.1 From 7.1 Nouns, Verbs and Adjectives: Open Class Categories, in Anderson's *Essentials of Linguistics*

In Linguistics, we observe how parts of language behave. When we find a set of words that all behave similarly, we can group them into a category, specifically, into a syntactic category. You might have learned about some of these categories as “parts of speech”. This unit gives an overview of the behaviour of the biggest categories.

You’ve probably learned that **nouns** are words that describe a person, place or thing. But when we’re studying morphology and syntax, we categorize words according to their behaviour, not according to their meaning.There are two elements to a word’s behaviour:

- What inflectional morphemes does the word take?
- What is the word’s syntactic distribution? In other words, what position does it occupy in a sentence?

What behaviour can we observe that allows us to categorize words as nouns? Looking at the inflectional morphology, we observe that most nouns in English have a singular and a plural form:

singular	plural
tree	trees
book	books
song	songs
idea	ideas
goal	goals

English uses a plural morpheme on a noun to indicate that there is more than one of something. But there is a **subcategory** of nouns that don’t have plural forms. **Mass nouns** like *rice*, *water*, *money*, *oxygen* refer to things that aren’t really countable, so the nouns don’t get pluralized. Nouns that refer to abstract things (such as *justice*, *beauty*, *happiness*) behave like mass nouns too. If they don’t have plural forms, why do we group them into the larger category of nouns? It’s because their syntactic distribution behaves like that of count nouns. Most English nouns, singular, plural, or mass, can appear in a phrase following the word *the*:

the tree, the trees
the book, the books
the song, the songs
the idea, the ideas
the goal, the goals
the rice
the money
the beauty (e.g., the beauty of the scenery)
the happiness (e.g., the happiness of the children)

In their syntactic distribution, **pronouns** (*I*, *me*, *you*, *we*, *us*, *they*, *them*, *he*, *him*, *she*, *her*, *it*) do the job that noun phrases do. A pronoun rarely appears with *the*, but it can replace an entire noun phrase:

The woman read the book.
*The she read the book.
She read **it**.

In *Essentials of Linguistics*, we'll group pronouns into the larger category of nouns, remembering that they're a special case.

Verbs behave differently to nouns. Morphologically, verbs have a past tense form and a progressive form. For a few verbs, the past tense form is spelled or pronounced the same as the bare form.

bare form	past tense form	progressive form
sing	sang	singing
think	thought	thinking
stay	stayed	staying
bake	baked	baking
remember	remembered	remembering
read [ɹid]	read [ɹɛd]	reading
set	set	setting

Adjectives appear in a couple of predictable positions. One is between the word *the* and a noun:

the red car

the clever students

the unusual song

the delicious meal

The other is following any of the forms of the verb *be*:

That car is red.

The students are clever.

The song is unusual.

The meal was delicious.

Many adjectives can be intensified with the words *very* or *more*:

very clever

more unusual

very delicious

And some adjectives (but not all) have comparative and superlative forms:

red – redder – reddest

smart – smarter – smartest

tall – taller – tallest

tasty – tastier – tastiest

The behaviour of adverbs is a little more difficult to observe. Unlike adjectives, adverbs don't have comparative or superlative forms, but like adjectives, they can be intensified with *very* or *more*:

very quickly

very cleverly

more importantly

The above examples illustrate that many adverbs are derived by affixing *-ly* to an adjective, but there are also many adverbs that are not derived this way, and there are also some common English words that have the *-ly* affix that aren't adverbs but adjectives, like

friendly, *lonely*, *lovely*, so the affix is not a reliable clue. The syntactic distribution of adverbs is also a little slippery. Adverbs can precede or follow verbs (or verb phrases; see Unit 8.5) to provide information about the verb:

The children sang *beautifully*.

The students complained *loudly* about the pop quiz.

They had *just* arrived when the fire alarm rang.

Samira tripped and *nearly* broke her wrist.

The visitors will arrive *tomorrow*.

And adverbs can precede adjectives or other adverbs to provide information about the adjective/adverb:

This meal is *surprisingly* tasty.

An *extremely* expensive car drove by.

The children finished their homework *remarkably* quickly.

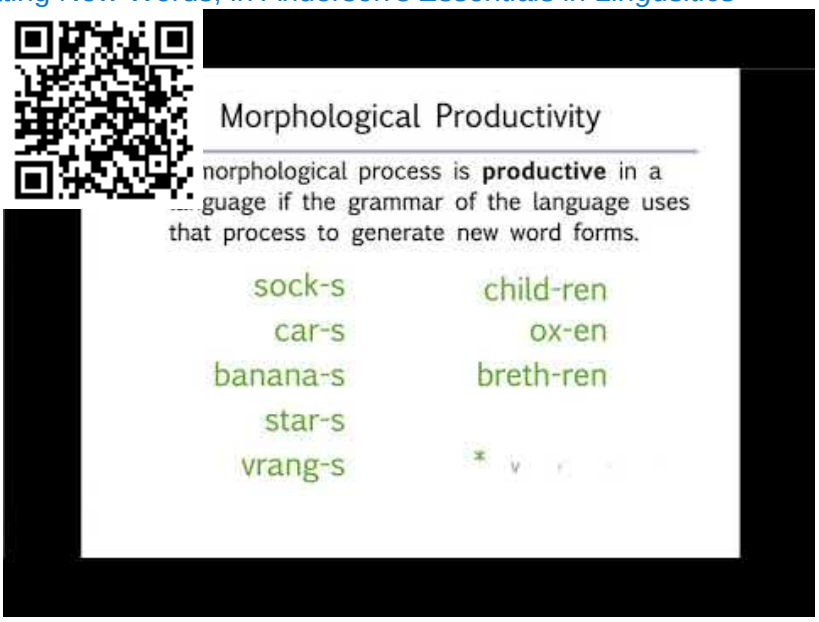
Because their behaviour is more variable than that of words in the other open-class categories, adverbs can be a challenge to identify. In the rest of this book, we'll label adverbs as "A", the same label that we use for adjectives.

The three syntactic categories of nouns, verbs and adjectives, are called **open-class categories**. The categories are considered open because when new words get added to the language, they are almost always in one of these three categories — the categories are open to new members. These categories are sometimes also called **lexical categories** or **content words** because these categories are the ones that do most of the lexical semantic work in a sentence: they convey most of the meaning of a sentence. The semantic content of the words from other categories (like *the*, *of*, *in*, *that*, etc.) is not as obvious as the semantics of the words from lexical categories.

6: Parts of Speech and Word Formation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7: Creation of New Lexicon

4.7.1 From 6.6 Creating New Words, in Anderson's *Essentials in Lingusitics*



Video Script

Back in Chapter 1 we learned that mental grammar is **generative**, that is, it allows users to create, or generate, brand new words and sentences that have never been spoken before. And in fact, one of the fastest ways that languages change, and the easiest way to observe, is by new words entering the language.

There are all kinds of different ways that new words can make their way into a language. It's possible to **coin** a new word, that is, to create a completely new form that hasn't existed before. So I made up this form *vrang*; I don't know what it means because I just made it up. But that was pretty hard to do — any new form I tried to make up turned out to have some obscure definition. So brand new coinages are possible, but they don't actually happen very often.

One way that English gets a lot of new words is by borrowing them from other languages. For example, the Welsh word *hiraeth* means longing or yearning. It's become common enough for English-speakers to use this Welsh word that in 2020, the Oxford English Dictionary added it. You can probably think of many other common English words that started out as borrowings from other languages and became deeply embedded in the English lexicon, like *anime*, from Japanese, *limousine* from French, and *boomerang* from Australian Indigenous languages.

Of course, one of the most obvious ways to derive a new words is with an affix. You might recognize the suffix *–ology*, which usually means "the study of". So mythology involves studying myths, criminology is the study of criminality, and epidemiology is the study of epidemics. The Oxford English Dictionary recently added *garbageology*, the study of a society or community by investigating what people consider to be garbage.

In English, affixation is one of the most **productive** ways to derive new words: No matter what the word is, you can almost always add an affix to derive a new, related word from it. Some other new affixed words that have found their way into the dictionary are *enoughness*, *farmette* (a small farm), and *unfathom*.

Another extremely productive way of deriving new words in English is by **compounding**, that is, by taking two existing words, both of which are free morphemes, and sticking them together. For example, the year 2020 saw the words *plant-based*, *jerkweed*, and *delete key* added to the dictionary. You can learn more about compounds in Chapter 7.

So we can say that **productivity** is a property of morphological processes in the grammar of a language. A given process is productive if it's one that the language uses a lot, and uses to generate new forms. For example, in English the plural morpheme spelled *–s* is extremely common, and we see it on words like *socks*, *cars*, *bananas*, *stars*, and thousands of others. In contrast, a plural affix *–en* is very rare in English: we see it on the plural forms *children*, *oxen*, and the very old-fashioned word *brethren*, but

pretty much nowhere else. And if we coin a new word, like *vrang*, and then decide we have more than one vrang, the plural we use is going to be *vrangs*, not *vrangen*.

If you look through the lists of new words that get added to dictionaries each year, you'll see that besides affixation and compounding, there are other morphological processes that occur in English. Here are some of them.

One thing that English does a lot is take a word from one syntactic category and just move it to another category with a new meaning. For example, the old meaning of *ghost* is the noun meaning, and then there's the newer verb meaning, where if you ghost someone you just stop replying to their messages and kind of disappear from their life. Not very nice! Likewise, *catfish* and *sundown* have newer, verb meanings that are different from their original compound noun meanings.

Acronyms pretty frequently make their way into English and some of them stick around, especially in typed form online, like a link that's not-safe-for-work, the classic *LOL*, and of course, "too long ,didn't read".

Clipping happens when we take a long word and just clip part of it off. Usually the meaning doesn't change, but often the clipped form becomes much more frequent then the long form. Does anyone even know that *fax* is shortened from facsimile? And certainly no-one talks about electronic mail anymore.

A few years ago clipping had a brief moment in the way some young people talked, so you might have an outfit that's *totes adorbs*, or a *relaish* that's not serious, just *cazh*. This trend seems to have lost its popularity, the way language fads often do.

The word-formation process that I've left for last is my favourite because I find a lot of them so funny. That's the **blend**, or **portmanteau**, the process whereby two words are kind of jammed together, but not in a compound. Instead some parts of the two words overlap with each other, like when *spoon* and *fork* combine to make *spork*. The best blends, the ones that stick around in the language and become permanent, seem to share a syllable like the second syllable in both *hungry* and *angry*, or at least share some segments and the rhythmic pattern, like *athleisure*. And then there are some that just seem to be trying too hard, *peanutritious*, *Christmasketball*, and (shudder) *covidpreneur*. I'm no prescriptivist, but I hope these words die a quick death.

All these words are examples of the **generativity** of grammar. Languages are constantly adding new words, using the **productive** morphological processes that are part of the grammar. Pay attention to the new words you discover as you read and listen, and see if you can figure out how they're formed.

Check Yourself

? Exercise 7.1

In 2020 the Merriam-Webster Dictionary added the word *contactless* to its word list, referring to, for example, a way of paying with a credit card without contacting the credit machine. How was *contactless* derived?

- affixation
- loanword
- zero-derivation
- clipping
- coinage
- compounding
- blend/portmanteau
- acronym

Answer

"affixation"

The reason: An affix was added to an already existing word.

? Exercise 7.2

In 2020 the Merriam-Webster Dictionary added the word *zonkey* to refer to the offspring of a zebra and a donkey. How was *zonkey* derived?

- affixation

- loanword
- zero-derivation
- clipping
- coinage
- compounding
- blend/portmanteau
- acronym

Answer

"blend/portmanteau"

The reason: This is a mix of *zebra* and *donkey.*

? Exercise 7.3

In 2020 the Merriam-Webster Dictionary added the word *WFH* to its word list, referring to the situation of people working from home. How was *WFH* derived?

- affixation
- loanword
- zero-derivation
- clipping
- coinage
- compounding
- blend/portmanteau
- acronym

Answer

"acronym"

The reason: It is a string of letters that correspond to the initials of each of the lexicon in the phrase.

4.7.2 Creation of New Lexicon, from Sarah Harmon



Video Script

Let's take a minute to walk through some more derivational as well as creational morphology. This is when we create new words, either out of the blue or on their own, as well as when we derive or create from a pre-existing lexicon. When talking about the creation of new terms or new lexicons, let's focus on a few of these. Catherine Anderson has already talked about some and I'm going to give you a few more.

Clipping, acronyms, and blending. Anderson more or less described those, and these are terms that you probably already know. Certainly, **acronyms** are very ubiquitous, at least among those of us that use text as a means of communication. **Blending** is something you're very familiar with, and in fact English is a great example of blending. We do this pretty regularly and we're one of the few languages that do this regularly and pervasively. In fact, we've been doing blending for a few hundred years; this is not a new phenomenon. **Clipping** is something that most languages do, and especially if we're talking about really technical terms, or if we are talking about high level language, or really educated speech. Think about an *exam*, *math*, and a *dorm*; if you have experience on a four-year university, you know what a dorm is. Even *prof*, although that one is used as much anymore. How many of you actually say *examination* anymore? or *dormitory*? Even the term *mathematics* is not very common certainly for average speakers of American English. Most of the time they're very specific terms.

A few more that I want to bring up and we'll come back to these when we get to historical linguistics because these are very common when we talk about. The history of the language and how it incorporates or creates new terms. **Borrowing** is a really common strategy that all speech communities use. We all borrow terms from other languages, especially if it's referring to technology, flora or fauna that is not native to us. *Coup d'etat* was a term that was created by a French philosopher to describe a certain political movement, and that term has been used throughout the world, regardless of what language we're talking about. A *coyote* is a term that I love bringing up; in fact, any of these terms that we borrow or that you see in Spanish that ending *-ate* or *-ote* up, they...have a very wonderful history in and of themselves. We have them in English like 'coyote', 'chocolate', 'tomato'; they were borrowed in from Spanish but Spanish actually borrowed the term from Nahuatl, which is the language that the Aztecs spoke. In many cases the Nahuatl term is borrowed from a different Mesoamerican language. it is a real historical domino effect and it shows the beauty of borrowing across cultures and across speech communities. Sometimes we borrow a term but make it more like the sounds in our own language; for example, [kojote] sounds a little Spanish so [kaɪjoʷtej] is what we say in English. In the case of a *tsunami*, we're pretty close. We say a tsunami; we just clip that [t]. In Japanese, you actually really shorten that 'u' sound almost make it voiceless: [tsu̥nami]. Sometimes, you borrow that and other times you don't.

Coinage is just creating the term on its own. These three lexicon, by the way, *nerd*, *geek*, *chuck*, have an interesting history of their own.

Name generalizations are a way for marketing gurus to have their day because we take that term and apply it to everything in that category. For example, if you work in an office environment, especially up until about 10 years ago, you used to not *photocopy* something; you used to *Xerox* something. How many of you actually say *facial tissue* or *tissue* when you go and grab that thing out of the box? Almost always you say a *Kleenex*, regardless of what brand that you're using. Maybe not so much here in the United States, but in many other English-speaking places, when you go to vacuum something you're actually *Hoovering*. Even **eponyms** are a part of this. We've taken the name, not from a brand name but from a place. That specific pastry that many of us have joined the mornings that comes originally from Denmark is a *Danish*. Yes, that lovely snack that we have, or even a lunch item that we have, with meat between two bread pieces of bread? Yes, it's from the Earl of sandwich you've heard that story before and those would be two types of eponyms.

Derivation is also really a key piece here, or sometimes it's called **misanalysis** (I use actually as misanalysis. This is when we analyze something incorrectly sometimes; they are native terms or lexicon and sometimes they are items that we have borrowed in. The examples here for **subtractive derivation** or frequently what we call **backformation**. We speakers misanalyze it; we think it's really a derivation of something else when it's not and then we create a whole new history for those. I'll come back to this more when we get to historical linguistics because this is always an historical process and always this happens not within one generation, but over several generations. Although I will say the one exception to that is the second one: *hamburger*. A *Hamburger steak* was brought in, I believe, at the end of the 19th century into English speaking cultures, most probably connected to the Victorian court and Britain. Within two generations, we get this concept of it not being a *Hamburg steak*, but a *hamburger* and then within two more generations of that, *burger* is the lexicon that can stand alone, and it can be modified by any number of combinations most of them being compounds like *tofu burger*, *veggie burger*, *turkey burger*, *Spam burger*, you get the idea. We can also have an **additive** situation, where we are adding on pieces, either morphemes or compounding various lexicon. The one I love to bring up, especially as a native Californian is this one, the second one. If you're in Southern California, you know where the La Brea Tar Pits are and

you know what they are; they are quite literally tar pits. Here's the fun bit and the part that, unless you speak Spanish, you probably didn't know: *la brea* in Spanish is 'the tar' or 'the tar pits'; it means both. Quite literally, if we're translating everything out, we're saying: 'the tar pits tar pits'. We're really saying that. We are thinking of *la brea*, not as a Spanish term, meaning 'the tar pits'; we're thinking of it as a place name that we just borrowed in.

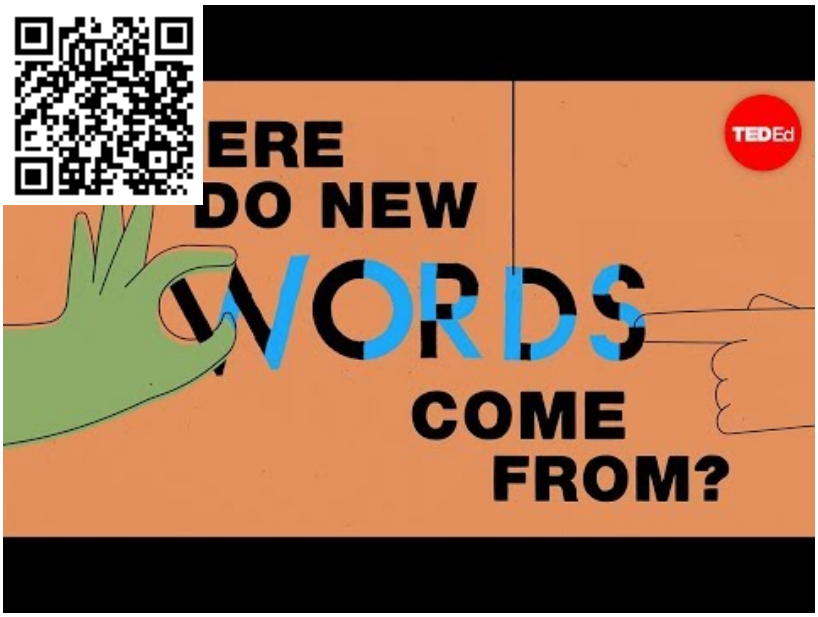
A couple more to go through, and these are really interesting. I'm showing you not just English, but many other languages, although English certainly has its place in these types of processes. Let's talk about **composition** first. This is where Germanic languages shine, including English. We compound everything; think about these lexicon in English—specifically, when they are combined they mean one thing, while when they're separated out they mean something different. *Girlfriend*/*boyfriend*: if we mean a specific loved one, then they're combined, but if we mean a friend who happens to have a specific gender identity, then we split them out. *Girlfriend*, *lifeguard chair*, *air conditioner*, *looking glass*, *textbook*, *aircraft carrier*, *bookcase*, *tennis racket*. I could go on and on; Germanic languages love composition. In fact, look up the word for 'research' in German; do it on your own and there's the perfect example of composition.

Reduplication is also a really important piece, maybe not so much in English and another Indo-European languages, but certainly throughout the Australio-Pacific languages, many of the Niger-Congo (languages in Africa) and throughout the world. Reduplication is just what you think; it is you are duplicating part or all of the lexicon. I'm going to give you an example of **total reduplication**, meaning all of the lexicon is being reduplicated, as well as an example of **partial reduplication**, when only part of it is being duplicated. Indonesian pluralization is total reduplication, so you have a term for 'house', a term for 'mother' and a term for 'a fly' like an insect: *rumah*, *ibu*, *lalat*. If you want to pluralize them, you repeat them: *rumahrumah*, *ibuibu*, *lalatlalat*. Just all combined together in the in a full stand-alone term, a lexicon. Here's an example of Tagalog, where you have partially reduplication. In the Australio-Pacific languages tends to be very common as a type of inflection. In this case, you have the base form, the infinitive form of the verb, and then you have the future form; the future form is made by repeating the first syllable. *Bili* is the term for 'to buy'; *bibili* would be 'will buy'; *kaim*, *kakaim*; *pasok*, *papasok*, you get the idea.

The last two processes will also revisit when we come to historical linguistics because, again, these are processes that tend to happen over time: extension and narrowing, and then functional shift. **Extension** and **narrowing** is increase and decrease; extension like 'holiday' or 'business', narrowing like 'hound' or 'doctor'. I'll explain these more when we get to historical linguistics, but suffice it to say that, over time, meaning changes. **Functional shift** also happens over time, where you have terms that change roles in the sentence and can have more than one role. For example, laugh used to only be a verb and then, when we wanted to describe it as a noun, we just use the same term; we didn't change it in any way. The same thing happened with process, position, contrast, and so many more. This is just an example of English, but other languages do this, too. It is interesting to note that, at least with respect to English, sometimes we can change a little bit of the pronunciation as well. For example, frequently the noun version will have the stressed vowel as the first syllable, but the verb version will be stress on the last syllable. This is something we'll come back to both extension/narrowing and functional shift when we get to historical linguistics.

4.7.3 More on the Creation of New Words

Watch this TedEd video in which Marcel Danesi explains how new words are derived or created. (The video is captioned.)



7: Creation of New Lexicon is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8: Morpho-syntax

Morpho-syntax, from Sarah Harmon



Video Script

One really interesting way that we incorporate morphology into our languages is having it combined with syntax. As the last real content area of morphology, as we ultimately will lead to syntax, let's talk about how these to combine and then combine them really just amazing ways.

One way that we can analyze languages and talk about how these morphemes come together to create lexicon, and then larger sentences and phrases. One way we can analyze this is whether a language is more analytical or more synthetic. When we talk about an **analytical language**, we're talking about something that's **isolating;** that's frequently the term we use. What we say is that there's very few, if any, affixes. When they do get thrown in, they are not that common and they don't give as much information; most of the time a truly analytical language has no affixation whatsoever. Individual lexicon are put together and are analyzed in the way they combine. It is interesting to note that most analytical languages rely heavily on word order; that's something we'll get to again when we talk about syntax. I have Mandarin here; Mandarin is one of the Sino-Tibetan languages and the truth is all of the Sino-Tibetan languages are heavily analytical. They have no inflection whatsoever, at least not in the way that we look at it in the rest of the world. They do not put a piece onto a morpheme and combine them to create larger lexicon. They just have individual one-syllable, and occasionally two-syllable, lexicon that get combined in certain ways and that, together, depending on the word order, convey a certain meaning. This is Mandarin, but you can say the same for Shanghainese, Cantonese, Myanmar or Burmese, Tibetan. All of these languages are spoken in China and Southeast Asia and are connected to that Sino-Tibetan family, the same can also be said for a number of the Southeast Asian languages in Vietnam, Laos, Cambodia, Thailand, Indonesia and others. I will not butcher Mandarin; my pronunciation is terrible, but what you can see here is the following. You have a morpheme that means ‘first person’, a second morpheme that means ‘plural’, and so combined when they are together, they mean first person plural, 'we'. Notice that they are not combined in a traditional sense. They are individual lexicon; they stand alone. You have a verb here, and then you have a noun so 'we play piano', 'we are playing piano'. Really, you can't tell tense in this case either, because you have no tense marker; you have nothing that says *yesterday*, *today*, *tomorrow*. Compare that to be where you have the same combination, but you do have a past tense marker. That is how tense is marked—not really marked as a morpheme that is attached to the verb, rather it is a stand-alone. It is this if we are saying 'we play piano yesterday'. For an English speaker or a speaker of most other languages, that would sound really odd; most languages have a tendency to mark the verb in some way to let you know that it has a tense marker or something else. Analytical/isolating languages do not do this. By the way, notice, for example, the combo for 'we', the first person plural, and notice that when it's an object, it has the exact same form; there's no difference. Mandarin, like most analytical languages, heavily lives on word order, it has to be in a certain order for anyone to understand it.

Compare this to Hungarian. Hungarian is a Finno-Ugric language; it's one of my favorites to use, not just because we have a significant Hungarian population here in the United States, especially here in California, but also because we like to think of all the European languages as being similar. In fact, there are some that are very different. Finno-Ugric languages have some similarities with Indo-European languages, but they are quite different in many ways. Hungarian Estonian and Finnish are the three main Finno-Ugric languages, but there are more. Hungarian is a **synthetic language**, meaning that you have combinations of affixes and roots, and they are combined to create larger morphemes. If you come if you look at the terms for 'the man'. You see as a subject and you see it as an object; that little [-t] at the end of *embert* is telling you that it's an object, not a subject. The same is true if you compare the term for ‘the dog’ as a subject versus ‘the dog’ as an object; the object has that [t] at the end. When we look at a possessive, 'our house' versus 'your house', notice that has is the root for 'house' and depending on the suffix depends on whose house it is; that's how you modify it. This includes what we like to think of as a prepositional phrase—in Hungarian, it's a post-positional phrase (more on that in syntax)—it is part of the inflection of the lexicon.

Synthetic languages and analytical languages are on a spectrum; there isn't really a ‘super synthetic’ language, although we'll get to that concept in a minute. Most languages are on the spectrum; English is a great example of being smack in the middle. In many ways, it's quite analytic; we don't have that many inflectional morphemes. However, we have plenty of affixes; most of them are derivations. We certainly use them in combination with a root to convey information, so we do have synthetic elements as well. We're kind of in the middle of that spectrum: more analytical languages on one side, more synthetic languages on the other.

The other way to think about how more fields combine in a given language is to talk about them as being agglutinating, functional, or polysynthetic. Let me explain what those terms are and individually give you some examples.

Agglutinating languages put morphemes together loosely; you can see the edges of the puzzle pieces as you fit them together. Hungarian is a great example; you can pick out the different morphemes. Another example is Swahili. I love using Swahili and the other Niger-Congo languages because it's really straightforward to pick out their different morphemes. You can frequently see them pretty clearly; you just need to take a minute to look at the pieces. This is the verb 'to read' in present tense, in past tense, and then future tense. You have first person singular, so 'I', second person singular, 'you', and third person singular and, in this case, masculine, so 'he'. Notice that the first person marker is *ni-*, and you can see it horizontally across compared to *u-* as the second person singular marker and *a-* the third person singular masculine marker. You have a morpheme that gives you the tense, and then you have the root of the verb. You can see those puzzle pieces pretty clearly. When you learn an agglutinating language. I'm not going to say it's easier to learn the morphology, but certainly there's a piece that's a little easier, in that you're able to see the pieces.

Compared to a language like English or Spanish, which is fusional, and it's a different story. **Fusional languages** combine roles in a given suffix or affix. I say suffix because most functional languages tend to have a lot of suffixation. I earlier showed you Spanish verb *tener*, or 'to have', and it's multiple conjugations Well, this is *hablar*, 'to speak'. Those of you who have experience with Spanish, you know this verb well. This is just another little bit of a breakdown of those suffixes that give us those inflections and, as you, as you can see, they have multiple roles. That [-o] at the end of [ablo] tells you first person singular present indicative; I actually forgot to put the mood on. This slide gives you four pieces of information in one syllable, one sound even. The same is true for that [a] and that [e], as well as that [n]; technically it's the vowel that goes with it, but that changes as well. When you learn a fusional language, one of the challenges is remembering that that one little syllable or one little inflection carries so much information, especially if the information is very different than what you do in your native language. There's a real juxtaposition as to trying to decipher what that sound means. Those of you who have had to learn Spanish, you know that one of the other pieces that's really important is the stress on the syllable. If you are stressing the root of the verb or if you're stressing the affix and that can drastically change, meaning that, for example, [ablo] as you see it written here, means 'I speak now' in the present tense indicative course. However, if I say [abló] and stress the last syllable, that is third person, singular, past tense and now we have to add an aspect; we have to say punctual as well. The difference between ‘I speak (right now)’, and ‘he spoke (yesterday)’ is all in where you put that stress syllable. Fusional languages have affixes that include multiple meanings and even multiple roles. English [-s] as an inflection is a great example.

We have agglutinating, and we have fusional, and those two are on a spectrum, but way the end of the spectrum is really a third one and it's called **polysynthetic**. It combines that synthetic (as in synthetic and analytic language) and fusional or agglutinating, that depends on the language. It combines these aspects in really interesting ways: highly complex lexicon that have multiple affixes and stems. We're not just talking about a language that has a stem and a couple different aspects of stuck on somewhere. We're talking about languages in which a given lexicon, a stand-alone word has an entire phrase, or even sentence, involved. This is Sora, a language spoken in southern India, it is a Dravidian language, it is not Indo-European. Sora is a really great example of polysynthetic languages in action. You see this sentence here, 'he is catching fish', but notice that there are only two lexicon: the

lexicon for 'he', so the subject, and then the lexicon for the entire verb phrase. All of it in one lexicon: the morpheme for 'catch', the morpheme for 'fish', the morpheme for 'not past' (what we would consider maybe just present but, it's just not-past), and the morpheme for 'do', as in 'this is an action and an active sentence'. It's not just having a whole verb phrase in a single stand-alone lexicon; you can have an entire sentence in a single standalone lexicon: 'catch-tiger-non past-do-first person agent', which is saying the first person is this person doing the action, so 'I will catch a tiger', all in one lexicon. Polysynthetic languages are completely fascinating. The fact that you can pack in so much information into one lexicon, suddenly learning a language like English or Spanish or Japanese or Russian isn't so difficult.

8: Morpho-syntax is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9: Analyzing Morphological Data

4.9.1 Analyzing Morphological Data, from Sarah Harmon



Video Script

When we go to analyze morphological data, the process is going to be really similar to what we did for phonological data. This is for a good reason; in many ways we're trying to see those puzzle pieces. In order to see those puzzle pieces, we have to compare and we have to contrast. We're going to take some morphological data and I'm going to walk you through how to solve some puzzles. When we go to do this, take note of the environment in which every morpheme exists, how it is used, if there are changes, or if there are patterns that we can pick up. Take note of the meaning—what we term as **glosses**, the direct translation—and look at their usages. Notice that we're going to use languages that are very different than English. Let's take a very straightforward morphological data set and let's analyze and pick out some puzzle pieces.

This is Isthmus Zapotec. It is a Mayan language spoken in South-Central Mexico, it is spoken kind of down near Chiapas, Michoacán, and Southern Mexico. This is a data set in which you have three different columns. The leftmost column is the bare noun, so the straightforward noun with no inflection. The middle column has a genitive inflection, meaning possession. Think of the apostrophe-s that we have an English (e.g. Mariela***'s*** book), that is a genitive; it means possessive or possession. We have a genitive here, that is third person singular and, in this case, masculine, 'his'. We also have a second person singular genitive. You'll notice in the glosses we have the bare noun, and then we have 'his...' whatever the noun is, and then we have 'your...' whatever the noun is. We're being asked to find three puzzle pieces: the genitive marker or possessive marker, the second person and third person singular markers, the difference between 'his' and 'your'. We have these written out in international phonetic alphabet (IPA): [palu], [ku:ba], [tapa], [geta], [bere], and [doʔ]. We then have third person singular second person singular forms.

How do we analyze this? First thing you do look for patterns. This is the bare noun. Can you see it in [spalube] and [sku:babe]? Can you see the bare noun? I bet you can and I bet specifically you can see [palu]...[palu]...[palu]. Do the same for the other nouns. [ku:be]...[ku:be]. [tapa]...[tapa]...[tapa]. [geta]... and notice there's also a devoicing thing going on but [g] and [k] are the exact same sound, so we know that's the same morpheme; ...[keta]...[keta]. Same thing here: [bere]...[b] and [p] are same sound just voicing difference so [pere]...[pere]. [doʔo]...same thing, [t] and [d], right? [toʔ].

Those are the bare nouns. Now, do we see something a puzzle piece, that is, in common with all of these third person singular genitives? The prefix, right? That [s-] at the beginning, do we see this also here? Yes, we do. It's entirely probable. For example,

that [s-] is in common, and we take note of that.

Now let's look at what comes after. Do we see the same suffix in all of the third person singular forms? We do; let's take note of that. Is it a different suffix in the second person singular? Yes, it is, but it's the same form in this entire column; let's take note of that.

Now we get to see what we can see. Let's look at that data. We see a prefix that's the same in both of those columns and we see a suffix that's different. We know that all of these in the second and third column, are all genitive in some way. That pretty much gives us our clues; that tells us that, for example, the genitive marker in general is [s-]. We write the hyphen after it so that we notify to everybody that it is a prefix. It goes before the root; where that hyphen is, that is where we would put the root. We can say that the different person markers, third person versus second person, they are suffixes. Third person is that middle column, so that is [-be]; second person is that third column, and that is [-lu]. Again, we are using the hyphen to say where the root would go.

This is a fairly straightforward process but we'll be using this process whenever we analyze anything having to do with morphology.

9: Analyzing Morphological Data is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10: References

For this section, most of the resources are from *Introduction to Language* by Rodman, Fromkin, and Hyams (2018) and *Language Files* (2019). Their list of resources for Morphology will have some solid resources if you wish to look deeper.

10: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

5: Phrases- Syntax

5.1: Syntax Terminology
5.2: Word Order and Lexical Categories
5.3: Phrase Structure Rules, X-Bar Theory, and Constituency
5.4: Analyzing a PS Tree
5.5: Grammatical Dependencies
5.6: Grammaticality
5.7: References

5: Phrases- Syntax is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.1: Syntax Terminology

Syntax Terminology, from Sarah Harmon



Video Script

Just a quick piece about syntactic terminology. Mostly because it's going to sound very familiar compared to what we talked about in morphology.

When we talk about syntax, we talk about lexical categories and about functional categories. When we talk about these things, we're talking about their role with respect to the overall structure of a given phrase or sentence. You'll notice that this pretty closely mirrors what we see when we talk about morphology. We talk about lexical vs functional lexicon, and open versus closed class lexicon. There is one exception: you'll notice that P (Preposition) is now a lexical category, not a functional category. That's what happens with prepositions or adpositions because it can be either before or after the noun phrase.

The reality is prepositions/adpositions are tweeners. They are somewhat lexical syntactically because they can have a phrase; they frequently are part of the phrase, and they are the head of that phrase. That being said, morphologically they are closed class and tend to act a little more functional. We can't make new ones easily; we don't tend to have multiple versions of them; they tend to pretty much stay as is and they don't tend to be borrowed from one language to another. Just know that prepositions have a dual role.

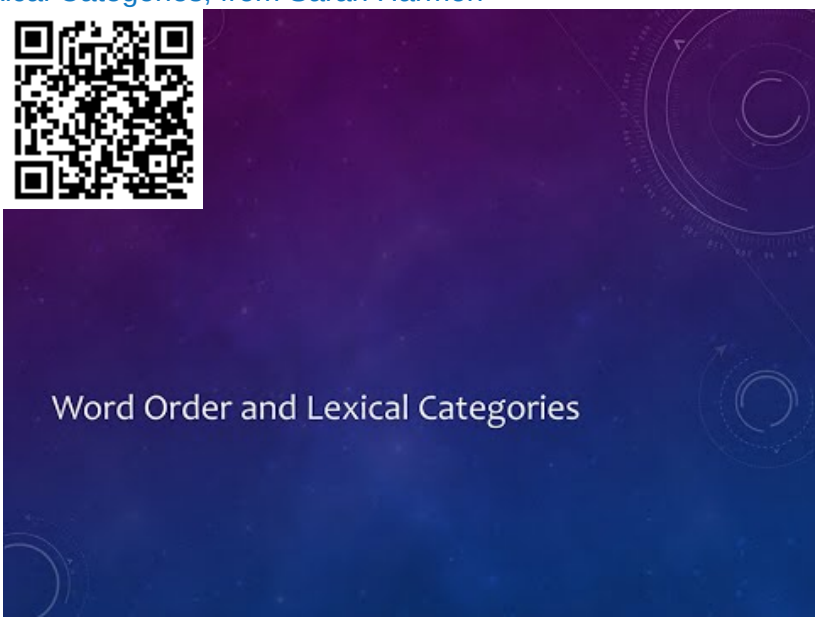
Everything else is fairly self-explanatory. The **lexical categories** all can be heads of their own phrases, so nouns, verbs, adjectives, adverbs and prepositions. **Functional categories** are only going to be complements; they're going to support a lexical item in some way, shape or form; these are determiners, auxiliaries, pro forms (pronouns and others), conjunctions and complementizers.

In the next section, we're going to talk about word order, and then after that we're going to start talking about actual structure. It is going to be important to remember that there is this big distinction between a lexical category and a functional category and we're going to see them both used in a variety of ways.

5.1: Syntax Terminology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.2: Word Order and Lexical Categories

Word Order and Lexical Categories, from Sarah Harmon



Video Script

Let's talk word order and a little bit more about lexical categories. For the most part, this is not something that Catherine Anderson brings up, but I think it's something that's important to talk about, especially when we compare English with just about any other language.

I found this little comic a while back, and if you are a Star Wars fan, and certainly those who have been watching—whether it's the movies, or the Mandalorian or anything else—you know about how Yoda has a different way of talking. Certainly, if Luke Skywalker says, “Why do you talk backwards all the time?” Luke is showing his perspective; Yoda also has his own perspective: “Backwards talking, I am not. English centric view of grammatical structure, you have.” Not only is it perfect Yoda-ese, as it were, but it also showcases why it's really important to take that step back as linguists, to be objective and just report on and describe.

First of all, let's talk about word order. It's a really important thing; every language has some kind of default word order. In some cases, that's the only possible word order; in other languages, you might be able to play with it a little bit.

When we talk word order, we are focusing on specifically the order of the subject, the object and the verb. We frequently just use those initials: S, O, V. There's a really curious fact about human languages, not just the ones that are being spoken now but anything that we have recorded and deciphered in the human existence.

- 35% of the world languages are SVO languages; that means that their canonical word order their default word order is subject verb object. English certainly falls in this category pretty much every Indo-European language is part of this category, with a couple of exceptions that will get to. Swahili, Thai, Hausa, which is spoken in western Africa, there's so many examples of this.
- 19% of the world languages are VSO languages, and that means the verb is first and then the subject, and then the object. Irish, in fact all of the Celtic languages, are VSO languages, so a little tweak on that VSO. Classical Arabic is, although modern Arabic is not always. Tagalog, as well as all the other Filipino languages and most into most Australio-Pacific languages, are VSO.
- By far and away the most common word order is SOV, subject object verb. Turkish, Japanese, Persian or Farsi (both terms are equal). Farsi, by the way, is an Indo-European language; Hindi tends to be SOV as well.

It's really interesting to note that some 96% of the world's languages actually have the subject before the object in some way. That's an important thing that will come back to when we talk about typology and historical linguistics, but suffice it to say that humans tend to like their subjects before their objects. That makes Yoda-ese a little unique, by the way.

Just as a comparison, let me show you a little bit about English versus Japanese. We'll come back to this when we talk about topology, when we talk about how different word orders affect other elements. You will notice, no doubt that English is an SVO language, that means that we have our subject and then our verb and then our object. SVO languages tend to have prepositions, which is what we do. We tend to be head first, which means that, in a given phrase the head, the crucial piece of that phrase, has to be the first thing there and, if not the very first thing one of the very first things. We tend to rely on word order for a structure; there can be cases of SVO languages that you some sort of case marking, but by and large, not as much.

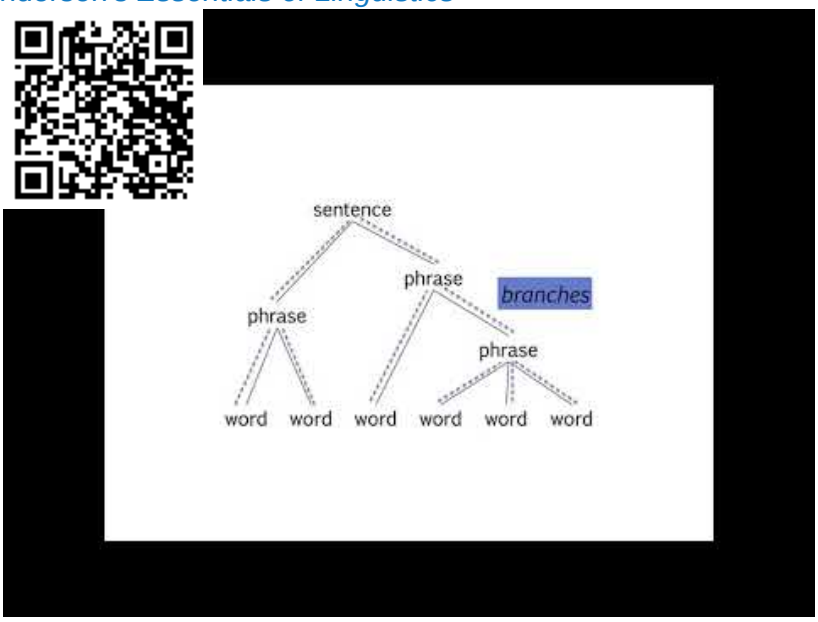
Japanese is a very traditional SOV language; again, that means the subject, and then the object and then the verb is the overriding structure of a sentence. They tend to be head last, meaning the verb is at the end of the verb phrase. The preposition is not a preposition but a post position because it's at the end of the phrase. An interesting note is that most SOV language is also mark case. Just to remind you, case is a type of inflection that tells you who or what the subject is, who or what the direct object is, who the indirect object is, and the like. Japanese has case marking, as does German and the other Germanic languages, Russian and the other Slavic languages. Latin was a case marketing language and there's so many more examples

As we walk through a little bit more about lexical categories and everything else, just keep in mind that every language has its set of rules, which leads us to the next section, how do you put a phrase together.

5.2: Word Order and Lexical Categories is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.3: Phrase Structure Rules, X-Bar Theory, and Constituency

Tree Diagrams, in Anderson's *Essentials of Linguistics*



Video Script

We're about to start looking into how sentences are organized in our mental grammar. Before we do that, we need to be familiar with a particular kind of notation called a **tree diagram**. We'll see that, within each sentence, words are grouped into phrases. Phrases can be grouped together to form other phrases, and to form sentences. We use tree diagrams to depict this organization. They're called tree diagrams because they have lots of **branches**: each of these little lines that join things in the diagram is a branch. Within a tree diagram, we can talk about the relationships between different parts of the tree.

Every place where branches join together is called a **node**. Each node corresponds to a set of words that act together as a unit called a **constituent**, which we'll talk about later in this chapter.

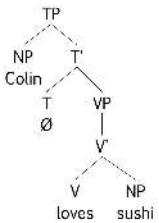
Each branch connects one node to another. The higher node is called the **parent** and the lower one is the **child**. A parent can have more than one child, but each child has only one parent. And, as you might expect, if two child nodes have the same parent, then we say that they're **siblings** to each other. (Just so you know, most linguistics textbooks call these nodes "mother, daughter and sister" nodes, but we're using non-gendered terms in this book.)

If a node has no children, we call it a **terminal node**.

Having this vocabulary for tree diagrams will allow us to talk about the syntactic relationships between the parts of sentences in our mental grammar.

Check Yourself

In the following tree diagram:



? Exercise 5.3.1

What is the structural relationship between V *loves* and NP *sushi*?

- V and NP are sisters.
- V is the mother of NP.
- V is the daughter of NP.
- V and NP are not related in any of these three ways.

Answer

"V and NP are sisters."

The reason: They are on the same level, both coming off of a mother node, V'.

? Exercise 5.3.2

What is the structural relationship between NP *Colin* and V *loves*?

- Deg and V are sisters.
- Deg is the mother of V.
- Deg is the daughter of V.
- NP and V are not related in any of these three ways.

Answer

"NP and V are not related in any of these three ways."

The reason: There are too many steps/nodes separating them.

? Exercise 5.3.3

Which node is the sister of NP *Colin*?

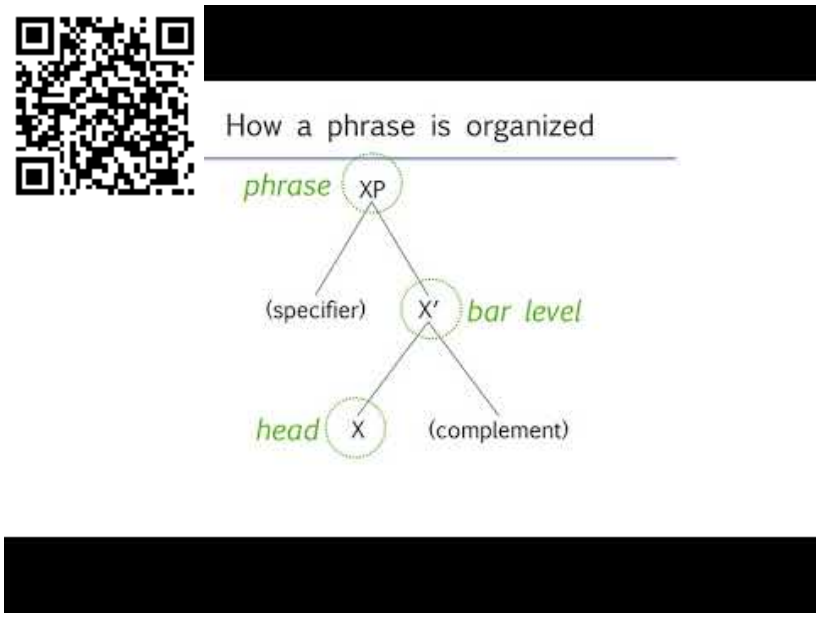
- TP.
- T'.
- T.

Answer

"T'."

The reason: They are on the same level, both coming off of a mother node, TP.

X-Bar Phrase Structure, in Anderson's *Essentials of Linguistics*



Video Script

We're starting to look at how our minds organize sentences. We'll see that within each sentence, our mental grammar groups words together into phrases and phrases into sentences. We saw in the last unit how we can use tree diagrams to show these relationships between words, phrases and sentences.

The theory of syntax that we're working within this class is called **X-bar theory**. X-bar theory makes the claim that every single phrase in every single sentence in the mental grammar of every single human language, has the same core organization. Here's a tree diagram that shows us that basic organization. Let's look at it more closely. According to x-bar theory, every phrase has a **head**. The head is the terminal node of the phrase. It's the node that has no daughters. Whatever category the head is determines the category of the phrase. So if the head is a Noun, then our phrase is a Noun Phrase, abbreviated NP. If the head is a verb (V) then the phrase is a verb phrase (VP). And likewise, if the head is a preposition (P), then the phrase is a preposition phrase (PP), and Adjective Phrases (AP) have Adjectives as their heads.

So the bottom-most level of this structure is called the head level, and the top level is called the phrase level. What about the middle level of the structure? Syntacticians love to give funny names to parts of the mental grammar, and this middle level of a phrase structure is called the bar level; that's where the theory gets its name: X-bar theory.

So if every phrase in every sentence in every language has this structure, then it must be the case that every phrase has a head. But you'll notice in this diagram that these other two pieces, the **specifier** and the **complement**, which we haven't talked about yet, are in parentheses. That's to show that they're optional — they might not necessarily be in every phrase. If they're optional, that means that it should be possible to have a phrase that consists of just a single head — and if we observe some grammaticality judgments, we can think of phrases and even whole sentences that seem to contain a head and nothing else. We could have a noun phrase that consists of a single noun — *Coffee?* or *Spiderman!* We could have verb phrase that has nothing in it but a verb, like *Stop!* or *Run!* Or an adjective phrase might consist of only a single adjective, like *Nice…* or *Excellent!*

But X-bar theory proposes that phrases can have more in them than just ahead. A phrase might optionally have another phrase inside it in a position that is sister to the head and daughter to the bar level. If there's a phrase in that position, it's called the complement. The most common kinds of head-complement relationship we see are a verb taking an object or a preposition taking an object. Let's look at some examples. Here we've got a verb phrase, with the verb *drank* as its head. That head has the noun phrase *coffee* as its sister. The NP *coffee* is sister to the verb head and daughter of the V-bar node so it is a complement of the verb.

Here's another example that has the same structure, but a different category. The head of this phrase is the preposition *near*, so the phrase is a preposition phrase. The complement of the preposition is the noun phrase *campus* and the whole phrase is *near campus*. Try to think of some other examples of verbs and prepositions that take noun phrases as their complements.

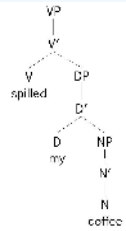
The other common place we see a head-complement relationship is between a determiner and a noun. In phrases like *my sister*, *those shoes*, and *the weather*, the determiner is a head that takes an NP complement.

X-bar theory also proposes that phrase can have a specifier. A specifier is a phrase that is sister to the bar-level and daughter to the phrase level. The most common job for specifiers is as the subjects of sentences, so we'll look at those in another unit.

Check Yourself

Exercise 5.3.4

In this tree diagram, what position does the DP *my coffee* occupy?



- Head.
- Specifier.
- Complement.

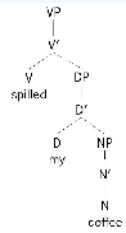
Answer

"Complement."

The reason: It's the complement of the V, both under the mother node of V'.

Exercise 5.3.5

In this tree diagram, what position does the D *my* occupy?



- Head.
- Specifier.
- Complement.

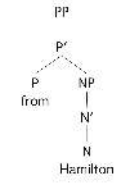
Answer

"Head."

The reason: It's the D in the DP.

Exercise 5.3.6

In this tree diagram, what position does the NP *Hamilton* occupy?



- Head.
- Specifier.
- Complement.

Answer

"Complement."

The reason: It's the necessary element and complement of a P.

Constituents, in Anderson's *Essentials of Linguistics*

We've started to use tree diagrams to represent how phrases are organized in our mental grammar. And we're using the tree diagram notation to represent every single phrase as having X-bar structure. But so far I've just asked you to believe me about X-bar structure: I've told you that this is what the theory claims, but we haven't yet talked about any evidence that our mental grammar really is organized into phrases that have X-bar structure. This unit shows some of the linguistic evidence that phrases have some reality in the mental grammar.

When we draw a tree diagram, we're making a claim about how a sentence or phrase is organized in our mind. Every time we draw two or more branches coming together at a node, we're making the claim that the node corresponds to a unit. In other words, all the daughters of that node behave together as a unit. Some of these nodes are at the phrase level, and some of them are at the bar-level. The more generic term for a group of words that act together to form a unit is a **constituent**.

So what's our evidence that constituents exist in our minds? Within a given sentence, how can we tell if a given string of words acts as a unit? Here again is where we rely on observing our grammaticality judgments, using a few simple tools.

Replacement Test

Here's a simple sentence:

> The students saw their friends after class.

Let's consider the string of words *their friends*. Because you've already started to practice drawing trees, you probably have an instinct that this is a noun phrase. But if you're going to claim that it's a constituent, it would be nice to have some evidence for that claim. One piece of evidence is that we can replace this set of words. Take the pronoun *them* and replace the string of words we're investigating:

> The students saw their friends after class.
>
> The students saw them after class.

Then we ask ourselves whether the resulting sentence is grammatical. Replacing *their friends* with *them* does indeed leave us with a grammatical sentence, which is one piece of evidence that *their friends* is a constituent.

Let's test another chunk of this sentence. Let's try the string of words *after class*. If we replace that set of words with the word *then*:

> The students saw their friends after class.
>
> The students saw their friends then.

And when we observe our grammaticality judgment, it turns out that this replacement is also grammatical. That's some evidence that words *after class* behave together as a constituent in this sentence.

We can do the same thing with the string *the students*. Replace that string with the pronoun *they*:

> The students saw their friends after class.
>
> They saw their friends after class.

And observe our grammaticality judgment, and we find evidence that the students is a constituent as well.

What happens if we try to replace a string of words that isn't a constituent?

> The students saw their friends after class.
>
> *The they friends after class.

*The did friends after class.

*The then friends after class.

*The them friends after class.

We can try lots of replacements, but when we ask ourselves whether the result is grammatical, the answer is No. There doesn't seem to be anything that can replace the string of words *students saw their*. The fact that nothing can replace that string of words suggests that *students saw their* is not a constituent in this sentence.

At this point, you're probably wondering how you know what you can use as a replacement. Here are some handy tips:

- Noun Phrases can be replaced with Pronouns (*it*, *them*, *they*).
- Verb Phrases can be replaced with *do* or *do so* (or *did*, *does*, *doing*).
- Some Preposition Phrases (but not all) can be replaced with *then* or *there*.
- Adjective Phrases can be replaced with something that you know to be an adjective, such as *happy*.

Let's see how this replacement tool works for a verb phrase. We'll go back to our sentence and look for the verb, *saw*. Let's test this set of words: *saw their friends*. Since *saw* is the past tense of *see*, we'll try replacing it with *did*, the past tense of *do*, and observe our grammaticality judgment.

The students saw their friends after class.
The students did after class.

This replacement is grammatical, so that provides us with some evidence that the set of words *saw their friends* is indeed a constituent.

You can use this evidence as you're drawing trees. If you can't quite figure out which groups of words go together into certain phrases, you can try replacing different chunks of the sentence. The parts that allow themselves to be replaced, that is, the parts that can be replaced and still leave a grammatical sentence are constituents, and those parts will be joined under one node.

You can also use this evidence when you're trying to figure out what category a certain phrase is: If you can replace it with a pronoun, then you've got a noun phrase and you can look for the noun as the head. If you can replace it with *do* or *do so*, then you've got a verb phrase which will have a verb as its head. *Then* and *there* are a little less reliable because they sometimes replace PPs or APs, but you'll be able to tell the difference between prepositions and adjectives because prepositions usually have complements and adjectives don't.

Movement Test

Replacement is not the only tool we have for checking if a set of words is a constituent. Some constituents can be moved to somewhere else in the sentence without changing its meaning or its grammaticality. Preposition Phrases are especially good at being moved. Look at this sentence:

Nimra bought a top from that strange little shop.

Let's start by targeting the last string of words by moving it to the beginning. Move the string of words then ask yourself whether the resulting sentence is grammatical.

Nimra bought a top from that strange little shop.

From that strange little shop Nimra bought a top.

Yes, it is. Standing here in isolation, the sentence might sound a little unnatural, but we can imagine a context where it would be fine, such as, "At the department store, she bought socks, at the pharmacy she bought some toothpaste, and at that strange little shop, she bought a top." On the other hand, if we target a smaller string of words:

Nimra bought a top from that strange little shop.

*From that strange, Nimra bought a top little shop.

If we try to move that string to the beginning of the sentence, the result is a total disaster. The fact that the resulting sentence is totally ungrammatical gives us evidence that the string of words *from that strange* is not a constituent in this sentence.

Cleft Test

There's a version of the movement tool that can be useful for other kinds of phrases. It's called Clefting. A cleft is a kind of sentence that has the form:

It was ____ that …

To use the cleft test, we take the string of words that we're investigating and put it after the words *It was*, then leave the remaining parts of the sentence to follow the word *that*. Let's try it for the phrases we've already shown to be constituents.

Nimra bought a top from that strange little shop.

It was from that strange little shop that Nimra bought a top.

The students saw their friends after class.

It was their friends that the students saw after class.

It was after class that the students saw their friends.

And let's try the cleft test on another new sentence.

Rhea's sister baked these delicious cookies.

It was these delicious cookies that Rhea's sister baked.

It was Rhea's sister that baked these delicious cookies.

The cleft test shows us that the string of words *these delicious cookies* are a constituent, and that the words *Rhea's sister* are a constituent. But look what happens if we apply the cleft test to another string of words:

Rhea's sister baked these delicious cookies.

*It was sister baked that Rhea's these delicious cookies.

Rhea's sister baked these delicious cookies.

*It was these delicious that Rhea's sister baked cookies.

Rhea's sister baked these delicious cookies.

*It was cookies that Rhea's sister baked these delicious.

All of these applications of the cleft test result in totally ungrammatical sentences, which gives us evidence that those underlined strings of words are not constituents in this sentence.

Answers to Questions

If a string of words is a constituent, it's usually grammatical for it to stand alone as the answer to a question based on the sentence.

Rhea's sister baked these delicious cookies.

What did Rhea's sister bake? These delicious cookies.

Who baked these delicious cookies? Rhea's sister.

The answer-to-questions test can also help us identify a verb phrase using *do*-replacement:

Who baked these delicious cookies? Rhea's sister did.

Notice that in the answer, "Rhea's sister did", the word *did* automatically replaces the verb phrase *baked these delicious cookies*.

Again, if a string of words is not a constituent, then it is unlikely to be grammatical as the answer to a question. In fact, it's difficult to even form the right kind of question:

What did Rhea's sister bake cookies? *these delicious

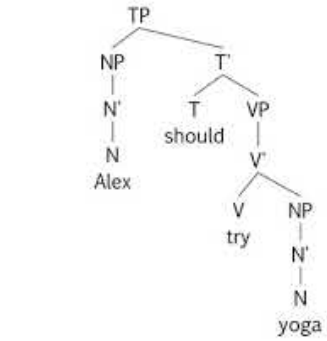
Who of Rhea's these delicious cookies? *sister baked

Remember that tree diagrams are a notation that linguists use to depict how phrases and sentences are organized in our mental grammar. We can't observe mental grammar, so observing how words behave is how we make inferences about the mental grammar. These four tests are tools that we have for observing how words behave in sentences. If we discover a string of words

that passes these tests, then we know that the phrase is a constituent, and therefore there should be one node that is the mother to that entire string of words in our tree diagram.

Not every constituent will pass every test, but if you've found that it passes two of the four tests, then you can be confident that the string is actually a constituent. When you're drawing trees, use these tests as a check every time you draw a mother node.

Sentences are Phrases, from Anderson's *Essentials of Linguistics*



Video Script

We've been developing a model for how the mind generates sentences. Our idea is that we draw words, morphemes, and morphosyntactic features from the mental lexicon, and then the operation **MERGE** organizes all these things and into an x-bar structure. And the theory makes the quite powerful claim that every phrase in every sentence is organized in an x-bar structure, with a head, a bar-level and a phrase level.

So it's fairly easy to understand that verb phrases have verb heads, and preposition phrases have preposition heads. But what kind of phrases are sentences? If we look carefully, we can observe that every sentence includes one and only one **tense** feature, which occupies a head position that we label as T, for Tense. And of course, where there is a head there are also the bar level and the phrase level. So a sentence is a T-phrase.

Let's see how this works in a fairly simple sentence. "*Alex should try yoga*." We already know how to depict verb phrases. The verb *try* is a transitive verb, and the NP *yoga* is its complement. The modal auxiliary *should* occupies the T head position. A T head always takes a VP as its complement. So every sentence that we see from here on will have a T head with a VP complement.

Whenever there's a head, there's also a bar-level and a phrase level. And the last phrase we have left in the sentence is the noun phrase *Alex*. It's the subject of the sentence. We said that specifiers are kind of special, and the subject is the most special of all. The Specifier of TP is the position for the phrase, usually a noun phrase, that's the subject of the sentence. Subjects go in SpecTP.

To sum that all up, every sentence is a T-phrase. The T-head of the T-phrase takes a VP as its complement. And the specifier of TP is a noun phrase, and the name for the noun phrase that occupies that position is the subject of the sentence.

What kinds of things can occupy the T-head position? The example that we just saw had a modal auxiliary in the T-head position. But not all sentences have modals in them. If the sentence does not have a modal auxiliary, then the T-head position will be occupied by a morphosyntactic tense feature. And when we're looking at English, there are only two tense features, [-past] and [+past].

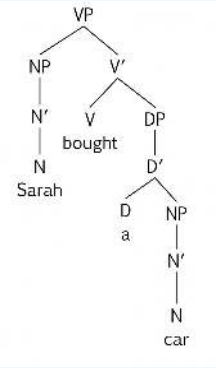
Let's see how this works in another simple sentence, very similar to the last one: *Alex loves yoga*. Remember that the theory claims that when we draw a verb in its tensed form from the mental lexicon, we also bring along the tense feature as well. *Loves* is in the present tense, so it has the [-past] feature.

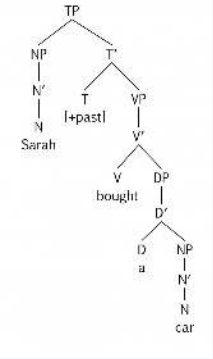
The structure of this sentence is the same as the previous one we looked at. The only difference is that, instead of a modal auxiliary in the T position, we have that morphosyntactic tense feature. We put this feature in brackets to indicate that it's present in our mental grammar but it doesn't actually get pronounced when we say the sentence out loud. You could think of it as present in the underlying form, but not in the surface form. The job of this feature in the T-head position is to make sure that its complement has the correct form. Since this feature is [-past], it wants to make sure that the verb in its VP complement is in the [-past] (that is, the present-tense) form. If there were some non-tensed form of the verb there, like *Alex loving yoga* or *Alex love yoga*, the sentence would be ungrammatical.

Check Yourself

Exercise 5.3.7

Which of the following is the correct representation for the sentence, *Sara bought a car*?





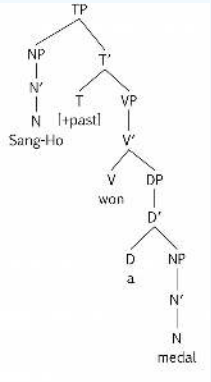
Answer

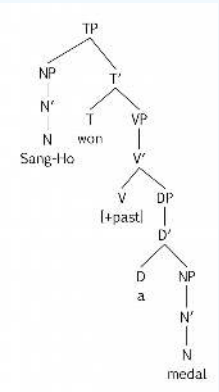
The second option, the one headed by a TP node.

The reason: All sentences are Tense Phrases, or TP's.

Exercise 5.3.8

Which of the following is the correct representation for the sentence, *Sang-Ho won a medal*?





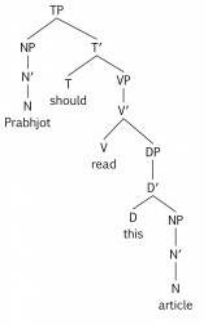
Answer

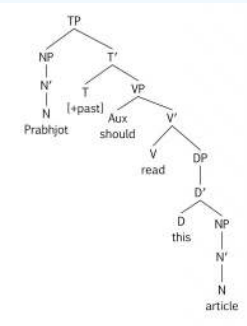
The first one, with the Tense node as a sister node to the VP node.

The reason: Tense nodes are always sister nodes to VPs.

? Exercise 5.3.9

3. Which of the following is the correct representation for the sentence, *Prabhjot should read this article*?





Answer

The second one, with *should* as an auxiliary.

The reason: *Should* is an auxiliary, not a tense.

English Verb Forms, in Anderson's *Essentials of Linguistics*

Every English verb has five different forms, but only two of the forms have a tense feature. The tensed forms are indicated with a morphosyntactic feature, either [+past] or [-past]. The form that a verb in the V-head position takes depends on what tense feature is in the T-head position, among other things.

bare	(non-tensed)	eat	walk	sing	take
[-past]	(tensed)	eats	walks	sings	takes
[+past]	(tensed)	ate	walked	sang	took
past participle	(non-tensed)	eaten	walked	sung	taken
present participle	(non-tensed)	eating	walking	singing	taking

But there are some quirks of English that can make things confusing:

Is it bare or [-past]?

For just about every verb, the [-past] form is recognizable in the 3rd-person-singular form (*she eats/walks/sings/takes*). The 1st & 2nd-person forms (*I eat/walk/sing/take* and *You eat/walk/sing/take*) look just like the bare form (*eat/walk/sing/take*).

If you're looking at a verb and can't tell if it's in the bare form or the [-past] form, give it a 3rd-person subject and then look for the –*s* morpheme:

I want to visit Saskatoon. (bare or [-past]? Can't tell: they're ambiguous)

She wants to visit Saskatoon. (*wants* is [-past], *visit* is bare)

Is it [+past] or past participle?

For many English verbs, the [+past] form (*She bought a donut*) and the past participle form (*She has bought a donut*) are the same.

If you're looking at a verb and can't tell if it's in the [+past] form or the past participle form, try replacing it with the verb *eat*:

She bought a donut after she had walked the dog. ([+past] or past participle? Can't tell: they're ambiguous)

She ate a donut after she had eaten the dog. (Silly, but grammatical)

The form *ate* is [+past], while *eaten* is past participle, so we can conclude that, in that sentence, *bought* is [+past] while *walked* is the past participle)

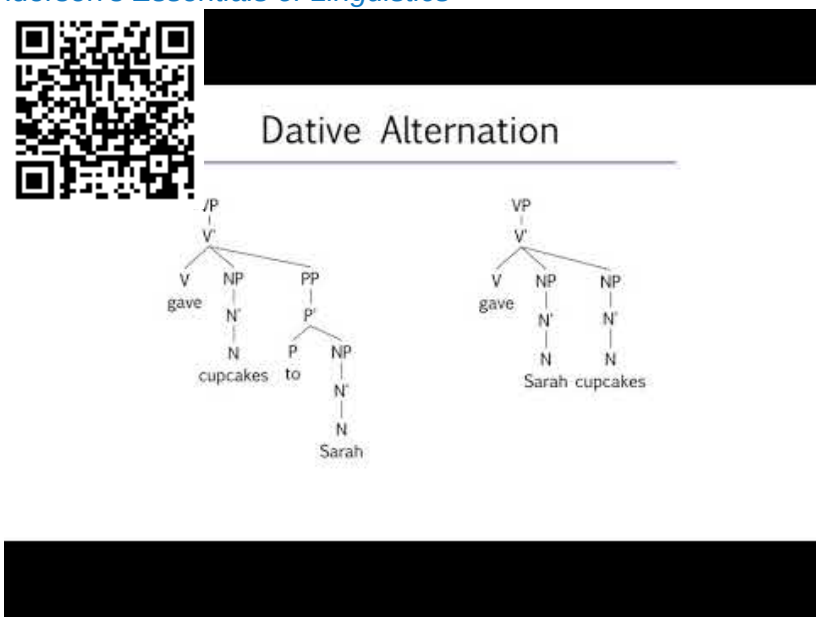
What about auxiliaries?

The modal auxiliaries never change their form: they occupy the T-head position in their own right.

The non-modal auxiliaries, like main verbs, change their form depending on what tense feature is in the T-head position, among other things.

bare	(non-tensed)	be	have	do
[-past]	(tensed)	am/are/is	has	does
[+past]	(tensed)	was/were	had	did
past participle	(non-tensed)	been	had	done
present participle	(non-tensed)	being	having	doing

Subcategories, in Anderson's *Essentials of Linguistics*



Video Script

Let's consider a simple sentence, *Jamie might bake cupcakes.*This is a perfectly grammatical English sentence, and we can account for it all using x-bar structure. If we change the verb *bake* to the verb *eat*, our sentence is still grammatical, *Jamie might eat cupcakes.* And that makes sense of what we know about how categories work — we group verbs together into the verb category because they behave the same way.

But what about these sentences?

> *Jamie might arrive cupcakes.*
>
> *Jamie might hope cupcakes.*

Are these grammatical? My mental grammar doesn't generate these, and I bet yours doesn't either. And their ungrammaticality isn't just a matter of them not making semantic sense, either. Since the verb *arrive* often has something to do with a location, we could try changing *cupcakes* to *Toronto*, but the sentence is still ungrammatical: the grammar of English does not generate the sentence, *Jamie might arrive Toronto.* But why aren't these sentences grammatical? There's no doubt that *arrive* and *hope* are verbs, and they seem to fit into the same x-bar structure that was grammatical for lots of other sentences. Why doesn't our mental grammar generate the sentence *Jamie might hope cupcakes?*

It's something to do with the verbs themselves. Some heads are picky about the kinds of complements they're willing to take. And this is especially true for verbs. Within the large category of verbs, we can group verbs further into **subcategories** according to the kinds of complements they take. For each head, the mental lexicon stores not just syntactic category information, but also

subcategory information. The subcategory information tells us what kinds of complements each head will accept. So let's look at a few verb subcategories.

Transitive Verbs have one complement, a noun phrase, so they have this basic structure. The verb *baked* is transitive when it has an NP or DP complement like *cupcakes*. Here are some other transitive structures: *drank coffee*, *likes Linguistics*, *needs money*, *speaks Mandarin*.

When there is a noun phrase in the complement of a verb, we call it the **direct object**. And the direct object NP or DP doesn't have to be a single word. It could be a fairly complex phrase itself. As long as it's a noun phrase and it's the complement of a verb head, we call it the direct object, and the verb is a transitive verb.

Intransitive verbs have no complement at all. These are verbs that describe an action or state that involves just a single participant, like s*neezed* or *arrived* or *dances* or s*lept.*

There's a small set of verbs that are called **ditransitives**. They're a little special because they have two complements, but for them to count as ditransitives, they have a special kind of behaviour, called the **dative alternation**. The best example of a ditransitive verb is the verb *give*. Take a look at this structure and notice that the V-head gave has two sisters — two complements — the NP *cupcakes* and the PP *to Sarah*. But this verb give has another possible grammatical structure that means exactly the same thing. In this alternate structure, the verb has two NP complements. The NP *Sarah*, which was the complement to the preposition in the other structure, is now the first complement, and *cupcakes* has become the second complement.

The fact that our mental grammar generates both these structures for this verb and its complements is called an **alternation**. There are other alternations in our mental lexicon, but this particular one is called the dative alternation, which comes from the Latin word for *give*. Most of the verbs that allow the dative alternation are verbs that have a meaning that's related to giving. *Send* is another example:

> *She sent a letter to her grandmother. // She sent her grandmother a letter.*

Or to *hand* someone something:

> *She handed a coffee to her friend. // She handed her friend a coffee.*

The last subcategory of verbs to talk about is another small one, but it's an interesting subcategory. Some verbs take complements that are entire sentences. Each of these verbs, *hope*, *doubt*, *wonder, ask*, has a complement that could stand alone as a sentence:

> *Ann hopes that the Leafs will win.*
>
> *Bev doubts that the Leafs can win.*
>
> *Carla wondered if she should cancel her season's tickets.*
>
> *Divya asked whether Eva liked hockey.*

Each of these sentences, or clauses, is embedded inside the larger sentence. And each one is introduced by a word from the category of **complementizers**. The words *that*, *if*, and *whether* are called complementizers because they introduce complement clauses. Let's look at the structure of one of these sentences.

First, the embedded clause, which could stand as a sentence in its own right — it has a tense feature in the T-head position. The complement to the T-head is, as always, a VP. In this clause, the verb is intransitive so it has no complements, and the entire phrase is made up of the word *win*. This clause has a subject, a DP in the SpecTP position, *the Leafs*.

So this whole TP could be a sentence in its own right, but we know that in this case, it's embedded inside a larger sentence — it's the complement to the verb *hopes*. And often when a clause is in complement position, it gets introduced by a complementizer, which is a head of its own that we label as C. Notice that because the complementizer *that* is a C-head, there is also a C-bar and CP level as well.

Now from here on it's quite simple. This whole CP is simply the complement of the verb *hopes*, so it's sister to the V-head and they're both daughters of V-bar. And then this matrix clause has its own T-head, T-bar and TP levels, and the subject NP in SpecTP is *Ann*.

OK, let's recap. We've seen now that in addition to category information, the mental lexicon includes subcategory information for some heads. Verbs belonging to different subcategories are choosy about the form their complement takes. This means that it would be possible for a given sentence to be ungrammatical even if it has an x-bar structure if the complement is the wrong kind for that subcategory of head. And we've looked at four different verb subcategories:

- transitive verbs have one NP or DP as their complement
- intransitive verbs have no complements
- ditransitive verbs have two complements that can alternate position in the dative alternation
- and there is a set of verbs that take clauses as their complements

? Exercise 5.3.10

What is the subcategory of the bolded verb in this sentence?

*The soccer players **kicked** the ball.*

- Verb with a complement clause
- Ditransitive
- Intransitive
- Transitive

Answer

"Transitive"

The reason: The verb *kicked* requires a NP clause as a complement.

? Exercise 5.3.11

What is the subcategory of the bolded verb in this sentence?

*Many birds **fly** over Ontario each fall.*

- Verb with a complement clause
- Ditransitive
- Intransitive
- Transitive

Answer

"Intransitive"

The reason: The verb *fly* cannot have an NP as a complement.

? Exercise 5.3.12

What is the subcategory of the bolded verb in this sentence?

*This game **teaches** children the alphabet.*

- Verb with a complement clause
- Ditransitive
- Intransitive
- Transitive

Answer

"Ditransitive"

The reason: The verb *teaches* requires both an NP complement and a PP complement.

Grammatical Roles, in Anderson's *Essentials of Linguistics*

We use **grammatical role** labels to identify the syntactic position of Noun Phrases or Determiner Phrases within each clause. It's vital to remember that grammatical role labels are defined strictly according to syntactic positions, not according to the meaning of a noun phrase or its semantic relationship to the verb. We'll come back to this idea in the next chapter.

The **subject** is the NP or DP that appears in the Specifier of TP. The underlined phrases in the following sentences are all subjects of their respective clauses.

Zora kicked the soccer ball.

Yasmin guessed that Xavier would probably be late.

That tall woman nearly knocked me over.

The view from the top floor is quite impressive.

Understanding Calculus takes a lot of work.

The town where I was born recently elected a new mayor.

The **direct object** is an NP or DP that is the complement to a Verb head. Each of the following underlined phrases is a direct object.

Zora kicked the soccer ball.

Xavier's lateness annoyed Yasmin.

William convinced Veronica that class was cancelled.

Ursula asked the fellow who works at Tim Horton's what time the store closed.

Stefanie bought a gift certificate for $100 for her mother.

If we refer to an NP or DP simply as *the object*, by default we mean the direct object, not the indirect object (see below).

If a Verb head takes a complement that is some category other than an NP or DP, then that complement phrase does not count as a direct object. The phrases following the verbs in these sentences are NOT direct objects, even though they are complements to V-head, because they are not NPs/DPs.

Yasmin guessed Xavier would be late.

Rana seemed unhappy.

The parcel was on the porch.

We can identify two additional grammatical roles for NPs/DPs, according to the syntactic positions they occupy. An **indirect object** only appears with a ditransitive verb. It is the NP or DP that alternates between being the complement of a P-head and the complement of a V-head, for a verb that allows the dative alternation. The underlined phrases below are all indirect objects:

Stefanie bought a gift certificate *for* her mother.

Stefanie bought her mother a gift certificate.

Quinn texted directions to the party *to* her friends.

Quinn texted her friends directions to the party.

Preeti sent a bouquet of flowers *to* her aunt.

Preeti sent her aunt a bouquet of flowers.

If a verb does not allow the dative alternation, then it does not have an indirect object.

If an NP or DP appears as the complement to a preposition, but does not an alternate position to become the complement of a verb in the dative alternation, then it is not an indirect object, but an **oblique**. Oblique is the catch-all label for all other positions that NPs or DPs can occupy in a sentence. The underlined phrases below are all obliques:

Oscar bought a bicycle from that store on Locke Street.

Norma left her business card on the table.

Massimo watched a documentary about antibiotics.

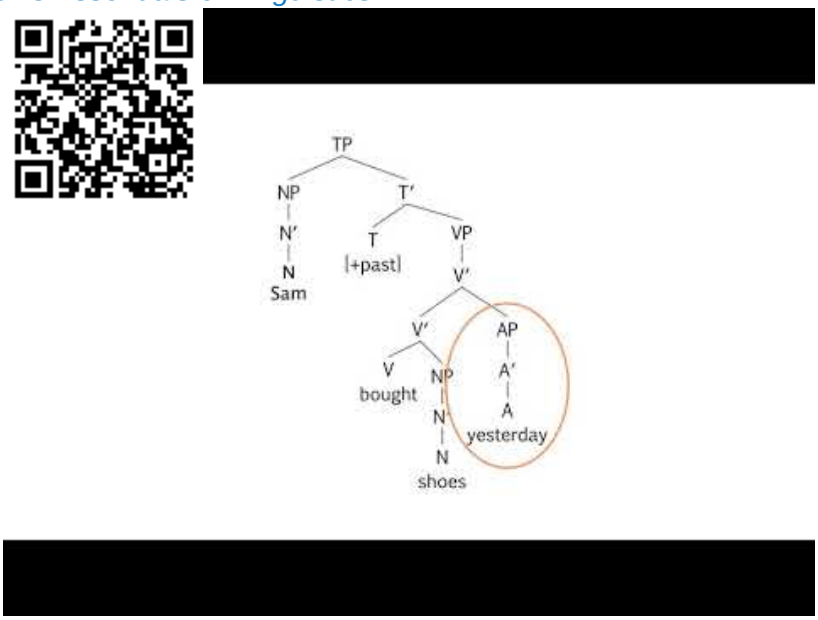
These four labels: subject, direct object, indirect object, and oblique, describe Noun Phrases or Determiner Phrases only in terms of what position they occupy in a clause. Look at the following two sentences:

A famous food critic reviewed this restaurant.

This restaurant was reviewed by a famous food critic.

Even though the person who does the reviewing is the same person (the famous food critic) in both sentences, the DP a *famous food critic* is not the subject of both sentences, only of the first. The subject of the second sentence is the DP *this restaurant*. Grammatical role labels describe the syntactic position of noun phrases.

Adjuncts, in Anderson's *Essentials of Linguistics*



Video Script

We've been working at representing how phrases and sentences are organized in the mental grammar, and to do that we've been using x-bar theory, which claims that every phrase in every sentence in every language of the world is organized into an x-bar structure. An x-bar structure has a head, a bar-level and a phrase level. It might, optionally, have a complement phrase as the sister to the head and daughter to the bar-level. It might, optionally, have a specifier as sister to the bar level and daughter to the phrase level.

But as you've been drawing trees and thinking about how sentences are organized in your mental grammar, you might have encountered some kinds of sentences that don't seem to fit into an x-bar structure. The X-bar structures that we've looked at so far have left out one element. The additional level of structure that we need is called an **adjunct**, and here's what it looks like. What structural relationships do you notice? The adjunct is sister to the bar-level, but here's something we haven't seen before: it's also daughter to a bar-level. This is an instance of **recursion**. A recursive structure is a structure that contains another structure inside it that has the same type as itself. Some linguists argue that recursion is a fundamental property of all human languages and that it's one of the things that makes human languages different from all other species' communication systems. Try to think of some other examples of recursive structures that we've already seen.

So if adjunction is recursive, you've probably already figured out that it can happen over and over again within a single phrase. Every time we add an adjunct as sister to the bar level, we add another bar-level as its mother, to which we could add another adjunct as its sister, which adds another bar level as its mother, and so on. So an x-bar phrase can have 0 or 1 complements, and it can have either 0 or 1 specifies, but because adjuncts are recursive, it could theoretically have an infinite number of adjuncts.

We've got a pretty clear idea of what complements do: they complete the meaning of a head. And the specifier position is a special position for subjects. So why do we need adjuncts? Well, like complements and specifiers, adjuncts are optional: a phrase might have one or more, or it might have no adjuncts. Adjuncts often add extra information that's not totally necessary for the meaning of the sentence, the kind of information that's often contained in APs or PPs, like where an event happened or how it happened. Because adjuncts are optional, they can often be moved or even removed altogether without changing the grammaticality of the sentence. And many adjuncts can appear on either side of their x-bar sister, whereas in English, complements pretty much always come after their sisters and specifiers come before. Let's look at some examples of adjunct phrases.

In this sentence, *Sam bought shoes yesterday*, the phrase *yesterday* is giving us extra information about when the buying happened, so it's adjoined within the verb phrase headed by *bought*.

In this next one, *Sam bought new shoes yesterday*, the adjective *new* is giving us extra information about the noun, *shoes*, so it's adjoined within the noun phrase that has *shoes* as its head. Notice that this AP is still in adjunct position: it's sister to N' and daughter to N', but it happens to come before its sister instead of after.

And look at all the adjuncts in this one: *Ted snored loudly for several hours at night*. We know that the verb *snore* is an intransitive verb: it doesn't take anything as its complement, so the head has no sister. But then there are three separate adjunct phrases, each of which gives us extra information about the snoring: the AP *loudly* is sister to V' and daughter to another V' node.The PP *for several hours* is sister to V' and daughter of another V' node. And the PP *at night* is sister to V' and daughter to another V' node. I've drawn these phrases with triangles; that's just a shorthand that indicates that we're not depicting the full inner structure of these phrases.

One piece of evidence that these phrases are adjuncts and not specifiers is that we can rearrange them in the sentence without changing the meaning or the grammaticality of the sentence.

> *Ted snored loudly for several hours at night.*
>
> *Ted loudly snored at night for several hours*
>
> *Ted snored for several hours at night, loudly.*

If any of these phrases were complements, we wouldn't be able to move them around, because a complement is always sister to the head, so it has to be right beside the head. But because adjunction is recursive, an adjunct always introduces another bar-level that can accommodate another adjunct.

So from now on, when you're drawing trees, you don't just have to decide whether each phrase goes in a complement or specifier position, but you also have to consider whether it might be an adjunct.

Check Yourself

? Exercise 5.3.13

Is the underlined phrase an adjunct or a complement?

Sam ran the Around-the-Bay race.

- Adjunct.
- Complement.

Answer

"Complement"

The reason: The verb *ran* can have an NP as a complement and fills in needed information.

? Exercise 5.3.14

Is the underlined phrase an adjunct or a complement?

Sam ran this morning.

- Adjunct.
- Complement.

Answer

"Adjunct"

The reason: The phrase *this morning* is not needed information to complete *ran*; it's additional

Exercise 5.3.15

Is the underlined phrase an adjunct or a complement?

The baby slept through the night.

- Adjunct.
- Complement.

Answer

"Adjunct"

The reason: The phrase *this morning* is not needed information to complete *ran*; it's additional

Phrase Structure Rules, X-Bar Theory, and Constituency, from Sarah Harmon



Video Script

Anderson does an excellent job of walking you through a phrase structure tree. I'm just going to add a little bit more.

If you want a basic set of phrase structure rules for English—and it should be noted, this is most all dialects of English, not all but most all—this is one set that you can use.

S → NP VP

S → CP VP

CP → (Comp) S

NP → (Det) (AdjP) N’ (CP) (PP) (CoordP)

N’ → (AdjP) N’ (PP)

N’ → N

NP → NP’s N’

NP → Pro

NP → [proper name]

VP → V’ (CP) (NP) (AdjP) (PP) (CoordP) (AdvP)

V’ → V

V’ → Aux V’

AdjP → Adj (N’) (CoordP)

AdvP → Adv (PP) (CoordP)

PP → P NP (CoordP)

CoordP → Coord XP

Figure 5.3.1: English Phrase Structure Rules. (Copyright; Sarah Harmon)

Catherine has a slightly modified version of this but either set will work just fine. What's important to note is that for every type of phrase, there is a head and a complement. The **head** is the crucial piece of that phrase, and then the **complement** is the supporting cast, as it were. For example, the head of a noun phrase is always going to be a noun.

All of these little codes, as it were, are a nice shorthand for trying to decipher what role that piece has. Instead of thinking of individual morphemes and lexicons and puzzle pieces, think bigger. Think big steel beams that are being used to set up the structure of a building; that's basically what phrase structure rules are.

As Anderson talks about, we're going to be using X-Bar Theory. It has been around for quite some time, about 40 years. It is still the prevailing method of doing basic syntactic analysis; we'll talk more about that soon enough. She does an excellent job of explaining it, it is really a great tool, not just to analyze the syntax. Morphology as well; you can use this type of structure this branching tree structure to analyze how more things are put together in a language, to create lexicon. Then you can use it to create the structure of a sentence or phrase in a given language. It has used been used for pretty much every language that is in existence; there's only a very few that it doesn't work for.

Catherine talks quite a bit about constituency, so I’m going to leave that there, but I do want to bring up recursion. It is one of the great ways to analyze a specific phrase structure of a specific language: to look at what has allowed with respect to recursion. **Recursion** actually refers to that hallmark of language called **duality**, that we have a limited number of resources, but we can make an infinite number of statements or sentences in a given language. For example, if we scroll back to our phrase structure rules, this more or less covers every single possible sentence and phrase in English. You pretty much can blueprint English based off of these rules. Yet, these few rules explain pretty much every possible statement that you can make in English; that's recursion. It also shows up in the fact that you can have nesting of different types of phrases. For example, a noun phrase can have a propositional phrase as a complement, so ‘the cat in the house’. ‘In the house’ is modifying or describing ‘cat’. Notice that a prepositional phrase can have a noun phrase as one of its components; in fact, it has to have a noun phrase. ‘In the House’, there's your prepositional

phrase, and then that noun phrase itself can get further modified by another propositional phrase: ‘the cat in the house down the hill on the river next to the forest’ and on and on and on. You can keep building quite a number of these recursive statements.

That's possible. Is it probable? Well, no, it's not. There is a tendency to kind of keep a limit as to how many recursive elements you put in. If you think about it, that makes sense; if I start to describe where I live, ‘I live in a house on a street in the town by the river next to the forest through the gap down the gulch…’. After a minute, you stop being able to keep track of where my house is. There seems to be a limit, three to four phrases are the rough limit. That being said, could we build a structure as tall and as far as we can see? Sure. Do we do that? No, we tend to keep a limit to our height of building; we also tend to put a limit on how many recursive elements we tend to include.

5.3: Phrase Structure Rules, X-Bar Theory, and Constituency is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.4: Tree Diagrams** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **8.5: X-bar Phrase Structure** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **8.7: Sentences are Phrases** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **8.13: Adjuncts** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

5.4: Analyzing a PS Tree

5.4.1 Analyzing a Phrase Structure Tree, from Sarah Harmon



Video Script

Catherine showed you a number of different types of trees, how to break a phrase down into various types of trees. I thought I would take a minute to do a version of that, as a way to show you how to analyze a specific phrase. I will warn you that when we linguists get a little funky and decide to create new sentences to showcase to people; we tend to use a theme. My theme is *Archer*, the TV show, so the character names and some of the descriptions that you will see refer to *Archer*. I'm a bit of a fan, but you can use them as examples in any way, shape or form.

Let's take a basic noun phrase-verb phrase-noun phrase kind of sentence. By that I mean, something that doesn't have too many moving parts, just to get you started. *Sterling sees Lana*. I'm going to use sentence, instead of TP, but same idea. You always know that your top most node, you start from up here and then you work your way down. It's always important to include intermediary steps. The reason is, we want to be able to showcase all of the possible structure, just like when you're building a sizable building and not just like a little box. You want to be able to not just put up the four corners of the walls; you need to be able to show the intermediary steps in between. Same thing with phrase structure trees. By the way, I am following a very common, simplified version of X-Bar Theory that shows proper names just directly coming off the NP node. That's just to make your life a little easier; if you're in an upper division syntax class, things get a little more complicated. For this level, I won't do that; I will keep it simple.

Two more sentences, where you have noun phrases and where the noun is modified. In this case *car* has a determiner with it an article; *waffles* has an adjective with it. Notice, too, the different structures. We do want to keep everything on the line we do not want to have all sorts of crazy lines when you do a phrase structure tree. In fact, when I was taught to do phrase structure trees, I was taught to do them on graphing paper, so that every line had its place, and you could see relationships very clearly. I've mimic that in the slides so that you get the idea.

This is an example of a propositional phrase down here. And an example of a coordinate structure; coordinates mean that you're using some kind of conjunction, like *and*, *but*, *or*; frequently it's *and*. They have to be the same type of phrase. *Pam* and *Cheryl* are both proper nouns, so we can do this, we cannot say *Cyril works with Pam and quietly*. That isn't going to work, and that sentence probably made no sense to you—it should not have—because the Phrase Structure Rules say that if you're going to coordinate, it has to be the same kind of phrase in each part.

This is an example of a noun phrase that has both an adjective that describes it, as well as a prepositional phrase. This gives you a longer type of phrase that you can see splayed out here.

I wanted to also showcase a type of sentence where you have an auxiliary verb. This explains why we even have this V-bar level to begin with. We want to express that this auxiliary has the tense, aspect, mood, all the other inflections, but the main verb is what is

being modified by the noun phrase, in this case. We do this to show the relationship between each of these phrases each of these pieces.

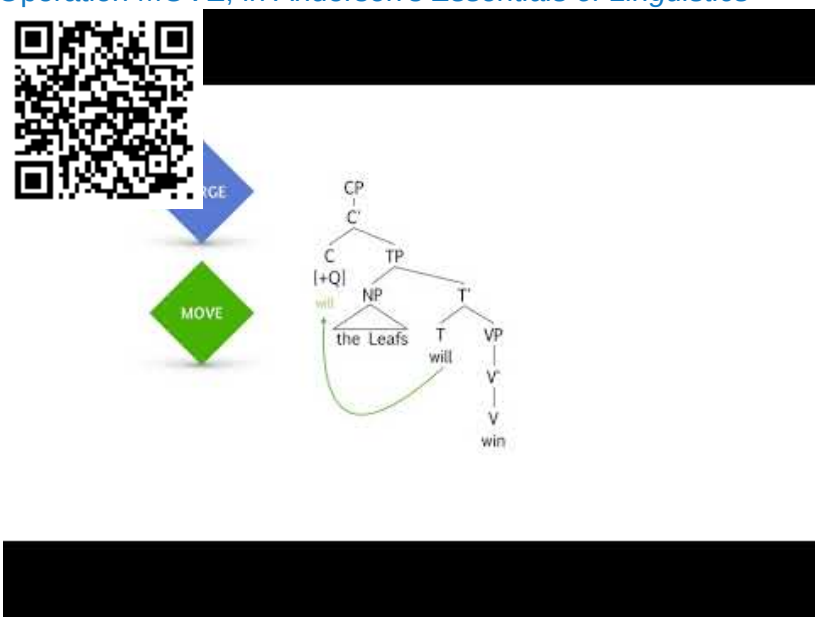
I wanted to showcase a little bit with a pronoun here. And one example of an embedded sentence, when you have a smaller clause, as it is usually called. In a larger clause, you have the different setup. Here you have a sentence down here or a TP—Tense Phrase, if you wish—and then it's embedded within that verb phrase. This is something that in colloquy will get worked on a little bit more.

Know in your quiz for the section, you will not have to do any phrase structure trees. However, you will have to start understanding a little bit about this relationship between heads and complements, phrases and everything between. It is wise to work on that a little bit. we'll play with this more later.

5.4: Analyzing a PS Tree is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.5: Grammatical Dependencies

5.5.1 From 8.9 The Operation MOVE, in Anderson's *Essentials of Linguistics*



Video Script

We've developed a model of the mental grammar that says that we draw words, morphemes, and morphosyntactic features from our mental lexicon, and then the operation MERGE combines all these items into a grammatical x-bar structure. In this unit, we're going to look at another operation of the mental grammar, called **MOVE**, which takes a part of a sentence and moves it somewhere else in the sentence. But why would we want to do that? Let's look at some examples to figure it out.

Let's take our now-familiar sentence, *Ann hopes that the Leafs will win*. *Hope* is one of those verbs that takes a whole clause as its complement, and that clause is introduced by the complementizer *that* in C-head position. But what if Ann isn't hoping, what if she's asking whether there's any possibility of her hopes being realized? If she's asking instead of hoping, the complementizer *that* doesn't work as well here. It's grammatical if we use the complementizer *if*. I'm going to suggest that the reason the complementizer *that* doesn't fit very well in this position is because the verb head *ask* doesn't just subcategorize for a complement clause — it's even pickier than that! The verb *ask* subcategorizes for a question clause. And the difference between a regular complement clause and a question clause is that a question clause has a [+Q] feature in the C-head position.

So we're seeing another instance of a morphosyntactic feature occupying a head position. The strange thing about features is that they don't get pronounced. We know that morphemes and words link up a form (either spoken or written) with a meaning. But features have meaning and don't have any form of their own. The tense feature does its job by making sure that its complement has the right form. And one way that the [+Q] feature does its job is by making sure that the thing in the C-head has the right form. So *if* is ok when the C-head has a [+Q] feature in it, but *that* is not so good.

There's another way that the [+Q] feature can make itself noticed, though. In English, if we want to ask a question, we don't do it like this: *If the Leafs will win?* How do we form that question when it's not embedded inside another clause? Of course, it's *Will the Leafs win?* How did that modal *will* get from its position between the subject and the main verb up to the beginning of the sentence? Here's where the operation MOVE comes in.

The theory claims that the operation MERGE generates this structure for the question. This looks just like the declarative sentence, *The Leafs will win*, except that it has a [+Q] feature in the C-head position. That [+Q] feature needs people to know that it's there; it's not like the null complementizer that doesn't mind being silent: it changes the meaning of the sentence so it wants to be pronounced. For questions that are main clauses themselves, it's not grammatical in English to just stick a question complementizer into the C-head position. So instead, the operation MOVE comes along, picks up the modal from the T-head position, and moves it up to the C-head position.

But when the modal moves from T up to C, the T position doesn't disappear from the tree. Instead, the thing that moved leaves behind a little footprint or shadow of where it used to be. In the theoretical literature, it's called a "deleted copy" or a "trace" — the idea is that it's still there in the T-head position in our mind, even though we pronounce it up in the C-head position.

Notice that now we've got two levels of representation in our syntax. MERGE generates the underlying form of a sentence, what we'll call the **Deep Structure**. And then MOVE comes along and, well, moves things, to give us the **Surface Structure**. A lot of the sentences that we've been looking at so far have the same deep structure and surface structure, but when we look at questions, we can see that a sentence can have some systematic differences in how it's represented at the two levels.

One thing I want to point out is that when we observe sentences, and when we make grammaticality judgments to observe whether a sentence is grammatical or not, what we're observing is always the surface structure. The Surface Structure is the form of the sentence that we speak: it's the form that's out there in the world. We can't ever observe the Deep Structure; that's the form of the sentence that exists in our minds. But some of the things that we observe about Surface Structures allow us to conclude that the Deep Structure does exist in our mind and that it's related in a systematic and predictable way to the Surface Structure. We can represent the relationship between Deep and Surface Structure using a tree diagram.

Just to make this idea about levels of representation clearer, let's look again at these sentences. The idea is that when these two sentences, one a question and one a declarative sentence, are generated by MERGE, they have almost the exam same Deep Structure. The only difference between their Deep Structures is that the question sentence has a [+Q] feature in C and the declarative doesn't. The declarative sentence is pretty much ok, so it doesn't need the MOVE operation to do anything, so its Surface Structure is the same as its Deep Structure. But the question sentence needs to get its [+Q] feature expressed, so MOVE comes along and forms a Surface Structure that's different from its Deep Structure. So the two sentences that were almost the same in their Deep Structures have one crucial difference between them in their Surface Structures.

We saw that when a sentence has a modal in the T-head position, that modal can move up to the C-head position to support a [+Q] feature. What happens when we want to form a question from a sentence that doesn't have a modal in the T-head position? What about a sentence like this one? *Samira phoned*.

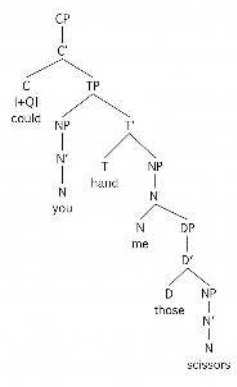
We know that the T-head node has a tense feature in it. In a declarative sentence, that tense feature makes sure that the V-head in its complement has the right feature — in this case, the past tense form *phoned*. But we certainly don't form a question from this sentence by asking, "*Phoned Samira?*" What happens in this kind of sentence is that we bring the auxiliary *do* into the T-head position. Because in this case the tense feature is [+past], *do* becomes its past-tense form *did*. Now that the past tense feature is on the auxiliary, it won't be on the lexical verb, so the verb *phoned* in V-head just takes its bare form *phone*. And then to get the [+Q] feature supported, MOVE takes *did* from the T-head position and moves it up to the C-head position, leading to the Surface Structure, "*Did Samira phone?*"

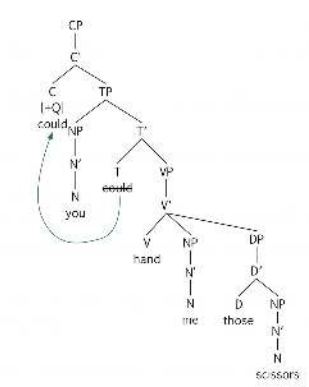
Your challenge now is to think about how our mental grammar forms *yes-no* questions for sentences that have non-modal auxiliaries in them. What do the Deep Structure and Surface Structure look like for each one? Try to figure it out!

Check Yourself

? Exercise 5.5.1

Which tree diagram correctly represents the question, "*Could you hand me those scissors*?"





https://socialsci.libretexts.org/@go/page/112738

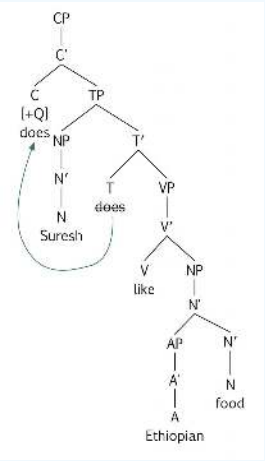
Answer

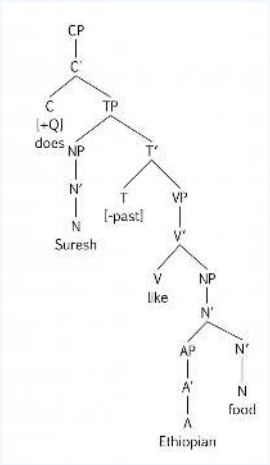
The second one, with the movement of *could* from the T node to the C node.

The reason: This is part of forming a question, where we move the conjugated verb to the front of the phrase.

? Exercise 5.5.2

Which tree diagram correctly represents the question, "*Does Suresh like Ethiopian food?*"





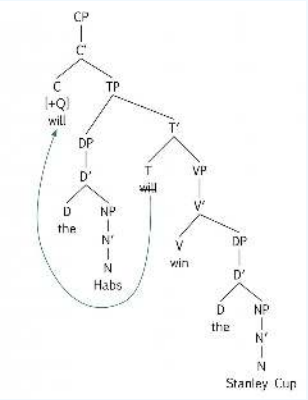
Answer

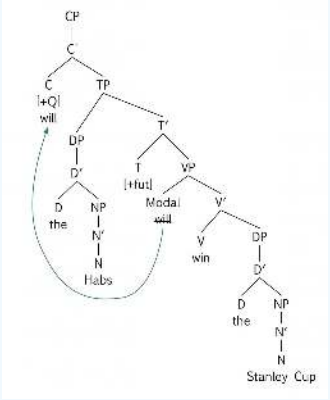
The first one, with the movement of *does* from the T node to the C node.

The reason: This is part of forming a question, where we move the conjugated verb to the front of the phrase.

? Exercise 5.5.3

Which tree diagram correctly represents the question, "*Will the Habs win the Stanley Cup?*"



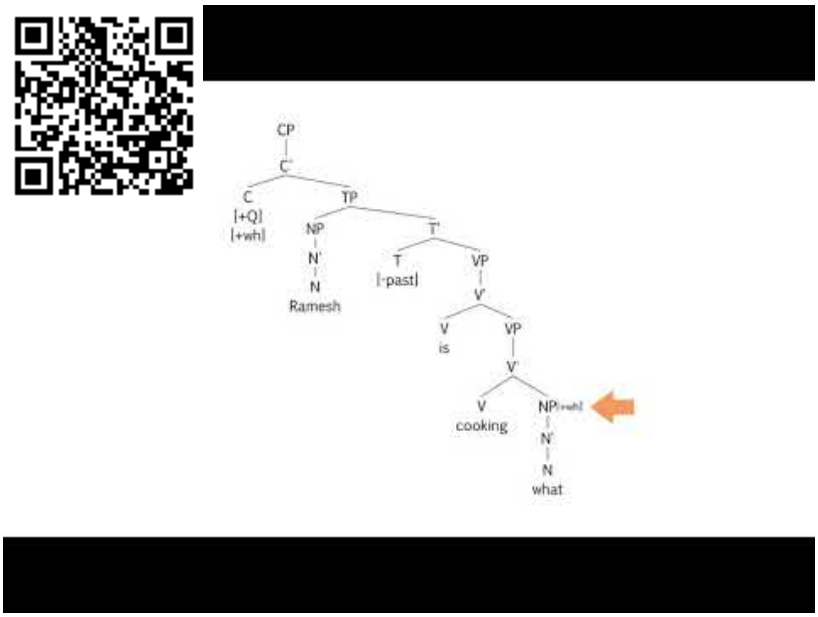


Answer

The first one, with the movement of *will* from the T node to the C node.

The reason: This is part of forming a question, where we move the conjugated verb to the front of the phrase.

5.5.2: From 8.10 Wh-Movement, in Anderson's *Essentials of Linguistics*



Video Script

In the last unit, we extended our model of the syntax component of the mental grammar. The idea is that we draw words, morphemes and features from our mental lexicon, and the operation MERGE combines them using X-bar principles. The structure that MERGE generates is called a Deep Structure; it's the underlying form of a sentence that we hold in our minds, but it's not always exactly like the form that we speak out loud. For some sentences, a second operation, MOVE, takes some elements from the Deep Structure and moves them to another position in the Surface Structure.

The MOVE operation that we've seen so far takes a head and moves it to another head position, leaving a trace behind. For obvious reasons, this kind of movement operation is called *head movement*, and the primary job that head movement does is to form *yes-no* questions. For this question, *Has Faiza eaten lunch yet?*, there's only a small set of possible answers that are grammatical: *yes*, *no*, *I don't know*, *maybe*.

But of course, *yes-no* questions aren't the only kinds of questions we ask in English. Take a look at these questions:

What is Ramesh cooking?

Who is he cooking for?

When is Leela arriving?

Where did he buy the ingredients?

Why is he making samosas?

How do they taste?

These questions don't take *Yes* or *No* as their answers, they take phrases.

What is Ramesh cooking? Samosas.

Who is he cooking for? Leela.

When is Leela arriving? In an hour.

Where did he buy the ingredients? At the store.

Why is he making samosas? Because Leela loves them.

How do they taste? Delicious!

We call these *wh*-questions because most of the English words that we use in these questions are spelled with "wh". Notice that *how* counts as a *wh*-word even though it doesn't have the letter "w" in it.

What we're observing here is the Surface Structures of these *wh*-questions. What might their Deep Structures be? To figure that out, let's think about what the corresponding declarative sentence would be: imagine the situation where we're answering each question with a full sentence.

Ramesh is cooking samosas.

He is cooking for Leela.

He bought the ingredients at the store.

Our theory claims that in the Deep Structure of a *wh*-question, the *wh*-phrase is generated in the position that it would occupy if it were the answer to the question. In other words, the Deep Structure of the sentence, *What is Ramesh cooking* is really, *Ramesh is cooking what*. That *wh*-word *what* is a pronoun that stands in for the Noun Phrase that refers to what he really is cooking. In the situation where we're asking this question, we don't know that the answer to the question is *samosas*, so we need a question word that lets us refer to the samosas without knowing their identity. The *wh*-word *what* does that job and the idea is that the structural relationship between the verb *cooking* and the pronoun *what* is the same as the structural relationship between *cooking* and *samosas*.

Let's see how it looks in a tree diagram. Here's the declarative sentence, *Ramesh is cooking samosas*. The samosas are the direct object: they're the NP in the complement of the verb *cooking*. And in this declarative sentence, the Surface Structure and the Deep Structure are the same. But if we didn't know what Ramesh was cooking, and we wanted to ask the question, MERGE generates a slightly different Deep Structure, like this.

Notice that the C-head contains a [+Q] feature because we're going to be asking a question, and a [+wh] feature because the question is going to be a *wh*-question. Also notice that this *wh*-phrase, *what*, has a *wh*-feature on it too.

So do we form the surface structure by moving what up into the C-head position? There are two reasons that's not going to work. The first is obvious: we know, from observing our own grammaticality judgments, that, *What Ramesh is cooking* is not the surface form of our sentence! And the second reason is for the sake of the theory: C is a head position, so it would mess up the consistency of our theory if we allowed a phrase to move into a head position. Besides, we're going to need that C head position for something else very soon.

The landing site for *wh*-movement is a position we haven't yet used, it's the **specifier of CP**, sister to C-bar and daughter to CP. The idea is that moving a *wh*-phrase into the specifier of CP supports the *wh*-feature in C.

Now, of course, this still isn't the right Surface Structure for this sentence. What still needs to happen? We still have this [+Q] feature in C-head that needs to get supported as well, so in this *wh*-question, in addition to the *wh*-movement of the *wh*-phrase up to the Specifier of CP, we also have head-movement of the auxiliary from V to T to C. And that leads to the grammatical Surface Structure, *what is Ramesh cooking?*

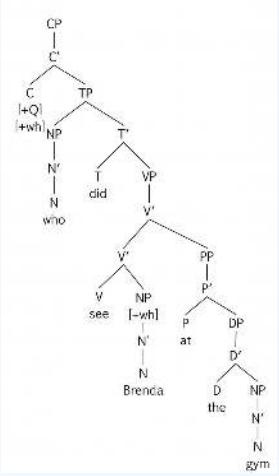
So we've now seen two different ways that the MOVE operation works. For head movement, it's a head that moves, and it moves to another head position. For *wh*-movement, it's a whole phrase that moves, and it ends up in the Specifier of CP. So when you're depicting movement, always check that you've got the right kind of node moving to the right position.

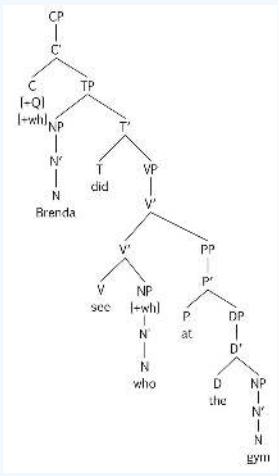
Now it's time for you to practice. Look back at these other *wh*-questions we generated. These are the Surface Structures of these questions. Try drawing trees to represent the Deep Structures of these sentences, and then draw the movement operations that generate the Surface Structure.

Check Yourself

? Exercise 5.5.4

Which tree diagram correctly represents the Deep Structure for the question, "*Who did Brenda see at the gym?*"





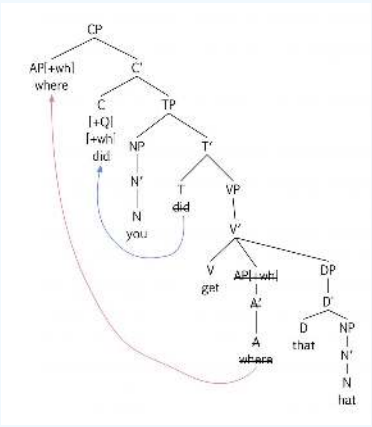
Answer

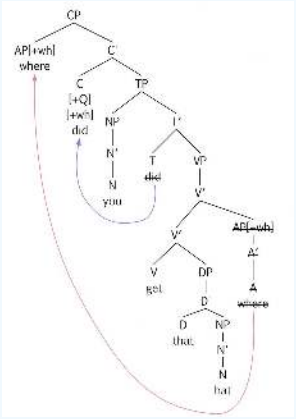
The second one, where *who* comes out of the complement NP node to the verb *see*

The reason: The question is about the direct object, not the subject, of the verb.

? Exercise 5.5.5

Which tree diagram correctly represents the Surface Structure for the question, "*Where did you get that hat?*"





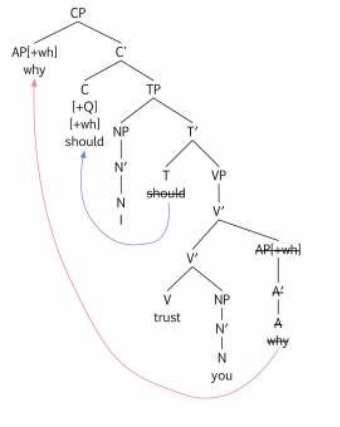
Answer

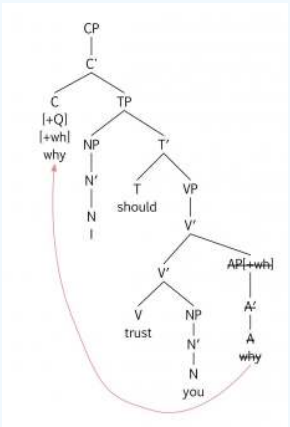
The second one, with *where* coming out of the adjunct node off of the V'.

The reason: That information is an adjunct, not a complement, of the verb *get*.

? Exercise 5.5.6

Which tree diagram correctly represents the Surface Structure for the question, "*Why should I trust you?*"





Answer

The first one, with *should* coming out of the T node to the C node, and with *why* coming out of the adjunct node of the V'.

The reason: Both *should* and *why* move.

5.5.3: From 8.11 Do-Support, in Anderson's *Essentials of Linguistics*

When we form a question that includes a modal auxiliary, the modal moves from the T-head position to the C-head position:

Could you ~~could~~ read this for me?

When a sentence contains no auxiliary, only a lexical verb, it appears that the lexical verb cannot move out of its V-head position. So we've proposed that the auxiliary *do* enters at the T-head position and then moves up to the C-head position:

***Did** you ~~did~~ see her tattoo?*

**Saw you ~~saw~~ her tattoo?*

One piece of evidence that *do* enters the sentence at the T-head position and moves to C-head, rather than just entering at the C-head position, is that it bears the tense feature of the sentence: if the tense is [+past], then we observe the form *did*, but if the tense feature is [-past], we observe the form *do* or *does*. Another piece of evidence is that whatever verb follows the inserted *do* is in its bare form, not its [-past] or [+past] forms:

***Did** she talk to Darren?*

**Did she talks to Darren?*

**Did she talked to Darren?*

***Does** she speak Italian?*

**Does she speaks Italian?*

**Does she spoke Italian?*

Both of these observations suggest that the inserted *do* gets its tense morphology from the tense feature (either [+past] or [-past]) in the T-head.

Negation

English also uses *do*-support to form negated sentences, which follow the same pattern: sentences with modals don't need *do*, but sentences with lexical verbs and no auxiliaries do need *do*:

I could not believe that rumour.

I **did not could believe that rumour.*

**She speaks not Italian.*

*She **does** not speak Italian.*

If we accept that *not* is in a fixed position between T-head and its VP-complement, then the distribution of *do* makes sense. Just like in questions, the evidence suggests that lexical verbs cannot move out of their V-head position up to the T-head position.

Non-modal auxiliaries

This pattern of how *do* behaves in questions and negative sentences gives us a clue about how the other non-modal auxiliaries, *have* and *be*, behave.

Notice that the verb *be* can always move up to C-head in questions, both when it's a genuine auxiliary:

Are you ~~are~~ going to the concert?

Was she ~~was~~ joking about that?

And when it's the only verb in the sentence:

Are you ~~are~~ serious?

Is this ~~is~~ the place?

Likewise, *be* appears before *not* both when it's an auxiliary and when it's the only verb:

You are not ~~are~~ going to the concert.

She was not ~~was~~ joking about that.

You are not ~~are~~ serious.

This is not ~~is~~ the place.

But *have* seems to have two different patterns of behaviour. When it is a genuine auxiliary, it behaves like *be*. It can move up to C-head in questions:

Have they ~~have~~ moved to Texas already?

Had she ~~had~~ already heard the news?

And appears before *not* in negated sentences:

They have not ~~have~~ moved to Texas already.

She had not ~~had~~ already heard the news.

But when *have* is the only verb in the sentence, it behaves like a lexical verb. It can't move up to C-head and can't appear before *not*.

**Has she ~~has~~ five sisters?*

**Have you ~~have~~ a headache?*

**She has not ~~has~~ five sisters.*

**You have not ~~have~~ a headache.*

Instead, when have is behaving like a lexical verb, it needs *do-support*.

Does she have five sisters?

Do you have a headache?

She **does** not have five sisters.

You **do** not have a headache.

The Deep Structures

From all of this evidence, we can conclude that MERGE treats these three kinds of heads differently:

Modals are generated in T-head, from where they can move to C-head if necessary to support a [+Q] feature.

Be is generated in V-head, but moves up to T-head (to the left of *not*) and from there up to C-head if necessary

Have is generated in V-head, and can move up to T-head and from there up to C- head only if it is an auxiliary (that is, only if it has a VP complement). But if it is the only verb in the sentence (and has no VP complement), then it behaves like a lexical verb.

Lexical verbs in English are generated in V-head and cannot move to T-head or C-head.

5.5.4 Grammatical Dependencies, from Sarah Harmon



Video Script

Catherine did an excellent job and describing all of those transformational rules that explain why things move around the way they, like for a passive sentence or a negation or a question or anything like that. I'm going to also add on a little bit more, and that has to do with some other ways that we can show dependency within the structure of any given language.

I'm going to start with something that is fairly straightforward in most other languages, just not so much in English. It's the concept of subject verb agreement. Native English speakers, if you learned a different language, with a few exceptions, this was something that really bugged you. This was an element that was difficult for you to pick up on. The reason why is because English does not have much in the way of inflection on verbs and, in particular, we do not have much with respect to subject verb agreement. It just isn't there. However, when we learn a language that does have it, it throws us for a loop. Sometimes **Subject Verb Agreement** is called **Linear Agreement**. it is a rule that states that there needs to be some kind of inflection on the verb to show some kind of match with the subject in some way, shape or form. Every language does things differently; in the case of English, we have a tendency, like I said, to have a little bit of inflection. For example, take P and Q. *Krieger sees the box*. We know this is correct, that this is the third person singular of the verb and that the subject is also third person singular. Compare that to Q: *Lana and Sterling see the box*. Notice that the inflection is not on the verb and that's because it is not a third person singular form.

I'll give you R and S, which show you pretty much the same thing in the case of *is* versus *are*, *has* versus *have*. English is not always the best example of this rule, but it is a little bit of as an example. Certainly, if you learn any other into European language —most every Finno-Ugric language, pretty much everything outside of the Sino-Tibetan languages and the various languages of Southeast Asia—chances are everything else will have some kind of linear agreement.

There's a second agreement rule that is very important to bring up and it's this one, **Structure Dependent Agreement Rule**. It basically states that a verb agrees in person and number with the subject of the sentence. The subject of the sentence is defined as the NP immediately dominated by the S node or the TP node, whichever version you want. Why is it that important? It explains why we have this linear agreement in the first place. We need to be able to tie in the morphology and the syntax together. Not only is it enough to in most languages have some kind of marker of some kind on the subject, you needed to agree in some way with the verb.

Let me show you some examples. I'm going to show you some examples of the same sentence in English, Spanish, Italian, Japanese and Hungarian. I picked those languages for a reason. Spanish because most all of you have some familiarity with it. Italian because, while it is sibling language of Spanish, the inflection is not always the same on the verb. I picked Japanese because

it is an SOV language; it is definitely different, but you will see and hopefully observe a pattern with these endings and with the verb. Finally, I gave you Hungarian. Hungarian, as a Finno-Ugric language, is very interesting for a few reasons, and we'll come to those in a second.

If you want to say *I watch the movie*, in Spanish and Italian you do not have to use the pronoun for 'I', the first person singular pronoun, because it is baked into the verb. That verb form can only go with first person singular subjects. This means that you can leave off that *yo* or *io*, depending on which one you're talking about. You can leave it off because *miro* and *vedo*, respectively, tell you everything you need to know about who the subject is. But you do not have that with either Hungarian or Japanese. In the case of Japanese, the verb does not show much with respect to subject agreement. Because of that, you have to have the pronoun for 'I', and in this case the pronoun for 'I', *watashi*, has a case marker on it to show you that is also this subject So *watashi-wa*. In Hungarian, you have word order, and you have a verb that does not necessarily tell you who the subject is. Therefore, you do have to have the pronoun, which is marked for being the subject.

This is 'you-singular', and 'we', so first person plural, in the same exact sentences. The only difference is subject. so now, I want you to look at *miro* versus *miras* versus *miramos*. I want you to look at *vedo*, *vedi*, *vediamo*. I want you to look at *miru*, *miru*, *miru*—notice that this verb does not change it all, which is why we have the various forms in Japanese for the subject. But notice that Hungarian *nézem*, *nézed*, *nézzük* does change, but you still need the pronoun to go with it because it doesn't quite give you enough information.

Every language is going to play with this. For example, in Spanish and Italian, while you do not need the subject pronouns, that's why they're in parentheses, because the verb tells you everything. You can add it, but when you add the pronoun, you are being emphatic, or you are usually making some kind of contrast between the subject and someone else. You can say, for example in Italian *io vedo il film*, but you're really stressing that *io*, that 'I' and really what you're doing is you're saying, '*I* watch the film; not them, not her, not you. I do'. But it's not commonly done. The third person singular and plural is a different story, because you need to know who the subject is. But the first and second person in most Romance languages, you don't need the subject pronoun; it's there it's already a part of the verb. On the other hand, you do not have that luxury in Japanese and you do not have that luxury of Hungarian because of other structural pieces that are there. That is to say, the structure dependent agreement rule states that the language needs to make some kind of 'agreement handshake'. If it's not there, then you need to make it clear in some other way; either you need a case marker, or you need some other element in the language to show you who the subject is. Sometimes it's word order and sometimes it's just something else.

5.5: Grammatical Dependencies is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.14: Move** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **8.1: Wh-Movement** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

5.6: Grammaticality

Grammaticality, from Sarah Harmon



Video Script

The last thing to talk about with respect to syntax or structure is grammaticality. This is to be a fascinating area, I find, because it really underlies this element of language that no one piece stands alone. We have already talked about how morphology and syntax will blend, and we have talked about how morphology and phonology blend. Grammaticality is how syntax and semantics blend, and that's an important aspect with respect to linguistics, especially given what Noam Chomsky has said in the past.

Let me explain. When Noam Chomsky first came up with what we now call to formative syntax or generative theory, there's a few names for it. When he first came up with that, his initial goal was to show that syntax and semantics are completely different things. In doing that, he established what grammaticality is. **Grammaticality** focuses on only the syntax and not the semantics. Grammaticality is how we know a sentence to work with respect to structure. Going back to that construction metaphor, if something is grammatical, then the beams are straight and true, everything is plumb, and the structure will stand. It doesn't tell you anything about the colors being used or the shape of the little spout that's coming off of some random wall. It is only talking about the internal structure: the studs, the roof, you get the point.

Chomsky's example sentence was this: *Colorless green ideas sleep furiously*. Now, when I say that sentence, you probably have a massive question mark on your face, and rightly so, because that sentence does not make sense. But it is a grammatical sentence, and that was Chomsky's point: that grammaticality is different than meaning. Meaning is a different aspect; certainly, there's nothing about that statement that makes sense semantically.

· If something is colorless it can't be green, that's just an oxymoron.

· Ideas can't have color anyway; they're abstract, they're not concrete.

· Ideas also cannot sleep; it's just impossible.

· Even if they could, they cannot sleep furiously, because it is impossible to sleep furiously. *Furiously* would imply emotion, where sleep is usually devoid of it.

So, it does not make sense semantically. But it does grammatically. It follows the phrase structure rules perfectly. Not only that, this is what helps to explain, in some ways, why we can get various combinations but they don't make sense. They make sense structurally, but they don't necessarily make sense semantically.

All of you have had experience learning a second language in some way, shape or form. One of the frustrating pieces of learning a new language is learning how the structure and the meaning combine. When we start to learn grammaticality, it starts to make some

sense. But until we're fluent, we will always produce sentences that maybe resemble a little bit of our *colorless green ideas asleep furiously*, in that they make sense structurally, you just wouldn't put those concepts together.

So then, what makes a sentence ungrammatical? It means that you have not followed the phrase structure rules. We always mark that in syntax with an asterisk. So, for example, those same words, the same lexicon, if I mess up the order and say *Furious sleep ideas green colorless*. That isn't ungrammatical sentence. How do we know? It breaks a whole bunch of phrase structure rules.

· The first one is that it has the verb phrase before the subject noun phrase; that's not going to work in English.

· There is an adverb out of place, although we have talked before about where that could happen so that one maybe is a little questionable, but certainly.

· It has the adjective after the noun not before it, which in English is not possible.

Grammaticality helps us in a number of ways, specifically with respect to understanding how the brain processes language. We're going to come back to that towards the end of the course. Suffice it to say that it helps us to understand the subconscious knowledge of a native speaker. You can add to that somebody who becomes a near-native speaker, somebody who has learned a language with such fluency that they may not say everything 100% perfectly, but they're very close. It tells us that this structure is part universal grammar and part acquisition. As children or adults, when we learn a language, we are creating these phrase structure rules, as it were. This is not a conscious thought; this is subconscious. We start kind of puzzling things together. But there has to be some sort of blueprint, if you were, to kind of explain what's going on.

With that, we will leave the world of syntax, and the next chapter enter the world of semantics, pragmatics, meaning.

5.6: Grammaticality is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.7: References

As has been mentioned in earlier chapters, both *Introduction to Language* (Rodman, Fromkin, and Hyams) and *Language Files* (Ohio State University Press) has a full list of references that are excellent.

There is one other book that I used: Thomas Payne's *Describing Morphosyntax: A Guide for Field Linguists* (1997; Cambridge University Press), which has more explanations on morphosyntax and syntax that are a bit more straight forward.

5.7: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

6: Meaning- Semantics and Pragmatics

6.1: Arbitrariness and Compositionality
6.2: Lexical Semantics
6.3: Phrasal Semantics
6.4: Semantic Features
6.5: Pragmatics Definitions
6.6: Grice and the Cooperative Principle
6.7: Reading Between the Lines
6.8: Deixis
6.9: References

6: Meaning- Semantics and Pragmatics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.1: Arbitrariness and Compositionality

To start off, listen to the comedian ISMO's routine about how complicated the word *ass* is--at least, if you're trying to learn English. (The video is captioned)



Arbitrariness and Compositionality, from Sarah Harmon



Video Script

With respect to arbitrariness and compositionality, there are a lot of the pieces that you're going to see below in what Catherine Anderson wrote that are exceptional, but I think they need a little bit of setup now. When we talk about semantics, one of the things that is kind of important to remember, is the fact that this is really like philosophy in some ways that we're going to talk about really abstract concepts. Meaning is harder to imagine sometimes than the structure of a phrase or how we put a lexicon together with different morphemes. To start off, let me start talking about a few things that are going to be closer to what you may hear in a philosophy class; there's a reason for that and I'll get to it.

With respect to semantics, we always have to start off in the same place, and that is talking about an **ontology**. In linguistic terms, an ontology is pretty close to what you hear in a philosophy class: it is a box with all the possible meanings. Each language assigns

a lexicon or a phrase to that meaning; that's what we call **lexical meaning** or **lexical semantics**. Let's unpack that a little bit, because that might be a little hard to understand. I like thinking of an ontology as a box, a box for each given language and, in many cases, you can say a specific dialect. You can say that that box come has an entire litany all of the possible phrases and lexicons and the meaning associated with it, so they are connected in some way. What happens is that as fluent speakers of a given language, whether it is our native language or one that we have learned, we pull from the items out of that box and we import it into our **entailment**, our mental lexicon. That mental lexicon that we talked about in morphology comes in here, because that is how we have built our language, we have pulled items from the ontology we have entailed certain other pieces to that and have built up our mental dictionary or mental lexicon. This really encapsulates a lot of what we do in semantics, because everything that we talked about with respect to semantics, and even pragmatics, which we'll get to later, has to do with understanding what are the items in that box, what are the lexicon and phrases, as well as the meaning in that ontology.

We have talked about this gentleman before in chapter one, as well as when we got to morphology: Ferdinand de Saussure. As you can see, he lived from the middle of the 19th century into the early part of the 20th century. We already talked about his Principle of the Arbitrary, but let's come back to that to refresh our minds a bit. It really is the cornerstone of semantics; you have a hard time understanding anything in semantics without understanding the arbitrariness of language. I talked about him briefly earlier; he was Swiss mathematician, logician, and philosopher that used these tools to look at language. He is a fascinating individual; at a different point in time, I might get into it.

For now, his **Principle of the Arbitrary** is what we need to focus on. I talked about earlier that you have, for example, *sofa*, *couch*, *davenport*; three different terms, and for most of us we use at least two of those to describe the same piece of furniture. Or, we talk about a *groundhog* or *woodchuck* and it's the same animal. Even how *air*, *food* and *water*, the three things that human beings need to live, that none of those terms are the same in any two languages. We use different lexicon in every single language to describe the same three things that we all need. That is proof of arbitrariness, that any given language arbitrarily assigned a lexicon or a phrase to a specific meaning. That's exactly what the Principal the Arbitrary is: connecting the **reference** to the **meaning** or the **sense**. Those were Saussure's terms and when he did this; he had a different aspect for us to think about. He described accurately that when we talk about the *morning star* and when we talk about the *evening star*. It's the same heavenly body. What does that mean? You may have heard those terms before, the *morning star* and the *evening star* and; if not, the *morning star* is the last heavenly body, the last thing we see in the sky before the sun comes up and makes it so bright that we can't see any stars. The *evening star* is the first thing we see once the sun goes down is the first heavenly body that we see. It isn't actually a star—it's the planet Venus—but even still, we use this these two terms to refer to the exact same thing. As Saussure described, that is arbitrary; we arbitrarily name it a different thing, depending on the time of day. That this happens in English or maybe used to happen (there aren't too many people who say *morning star* and *evening star*), but arbitrariness does happen in every language.

Really what we're saying is that a reference or a lexicon or phrase that we use is really a set of **truth values**, meaning that, if it is correctly attaching or entailing to a certain meaning that it's a **true reference**: it's true for me, it's true for you. If it's not, then it's a **false reference**. Let me give you an example. If I hold up this device and call it my *phone* or my *telephone* or my *mobile phone* or my *cell phone*, all of those are true references, meaning that in any dialect of English this thing would be considered a 'phone'. It could be any or all of those lexicons. But if I hold this thing up and I call it a *pencil*, that's going to be a false reference, because in no dialect of English does this apparatus get connected to or entailed to a 'pencil'; it just doesn't happen. Let me give you another example. If I say, that is a true reference, because this thing that in English we call 'sky', it's the thing that is above us when we go outside right; get rid of the clouds and what you see is the sky. That thing up there that thing has a color to it, or at least that we perceive and that color is associated that hue is associated to the term *blue*, so that's a true reference. If I say *the sky is green*, for most all of you, that would be a false reference, because in no way does that color get correlated to the term 'green'. I will say, though, if you are from an area or have spent time in an area that is prone to tornadoes, then you know what that phrase means. If you have the unfortunate instance of being in an area when a tornado is about to drop, the clouds look green; the light bounces off the clouds in such a way that it looks green, and so you say *the clouds are g*reen or *the sky is green*. In that specific reference, it would be true, but for most everybody else, it would be false.

If we talk about Saussure here and the Principle of the Arbitrary, then we have to include Gottlob Frege, who also came from Switzerland. He's also a mathematician, logician, and philosopher of language and he worked alongside Ferdinand de Saussure. He came he came up with the **Principle of Compositionality** and it is equally important when we talk about semantics; you can't talk about semantics without either the Principle of the Arbitrary or Principle of Compositionality. This principle is actually pretty straightforward and I probably don't need to explain it very much. It is that the meaning of the whole is determined by the meaning of the parts and the way they combine. The components help to form and shape the meaning of the overall. That pretty much makes sense, but I'll give you a couple of examples. If we think about this morphologically, then certainly if we have a term *unbelievable*.

All we need to do is look at the individual morphemes that are associated with that lexicon, note the meaning of the different morphemes, and we can put together the meaning of the full lexicon. In the term *unbelievable*, we have a root *belief*, we have a prefix *un-*, which means ‘not’ and it's a negator, and we have a suffix, a derivational suffix, that means ‘able to’, *-able*. ‘Not’ ‘believe’ ‘able’; ‘to not able to believe’—that's what *unbelievable* means.

It even helps us to understand the history of how certain terms or certain phrases get changed over time. For example, if I say *holiday*, you have a certain image in mind for most English speakers, that means just a day off. It may include a vacation; certain dialects of English equate a *holiday* to *vacation*. However, if I told you that, historically, that it was a sacred day or a holy day, that may not be the meaning of it now, but you can see the history of that, and you can understand how it could come to mean that. These are examples of the Principle of Compositionality: the parts and the way they combined inform the meaning of the whole.

We can also talk about when there is a **lack of compositionality**. We already have in the previous chapter, when I talked about grammaticality and I talked about the statement that Noam Chomsky came up with: *colorless green ideas sleep furiously*. We talked about how it is a grammatical sentence structurally; it follows the phrase structure rules of English. Why it doesn't make sense is because it lacks compositionality; the meaning of the individual pieces, that they don't work together. In semantics we call that **anomaly**; an anomaly is when there is a lack of compositionality. Most of the time, anomaly produces a lack of understanding; you don't understand what somebody is trying to communicate.

However, there are times that anomaly can be massaged a little bit; it can be understood, given a context. One example is, **metaphors**. If I say *there is a fork in the road*, what I’m saying has nothing to do with an eating utensil in the middle of the roadway. In that sense, there is anomaly, because there's a lack of compositionality. As an English speaker, you understand what I’m trying to say; I’m describing a situation in which a single road forks or splits into two or more subsequent roads. Because you are an English speaker, you understand what that metaphor is trying to say. That is an example of a bit of **elasticity** with this Principle of Compositionality, that we can play with this a little bit. This also includes **idioms**, and it includes something called **collocations**. A collocation is a combo: X and Y. They are understood as a unit, and they have a little bit of a metaphor quality to it. For example, if I was describing somebody in their late 40s through early 60s, and they are a dark haired individual but there's quite a bit of gray sprinkled throughout their hair. A term we might use in English is *salt and pepper*, that that person has *salt and pepper hair*. If I just said, *salt hair* or *pepper hair*, you wouldn't understand what I was trying to say. If I say, *salt and pepper*, together, that is understandable, as a chunk of meaning. It stretches that Principle of Compositionality just a little bit, but it's understood because of the context, because that phrase is part of the ontology of English. I'll give you another example. I used to have a neighbor who did not tell stories very well, and one of the things they used to do was not give very much in the way of detail. If you talked to them after a trip, they would say they went *here and there*, and did *this and that*. *Here and there*, *this and that*, those are two collocations. They are general phrases to explain that there was probably nothing important going on: they didn't go to anywhere that they felt was interesting, or they didn't want to tell us where they went. Either way, they used two different collocations to describe his vacation.

All of these expansions of the Principle of Compositionality and the Principle of the Arbitrary inform everything with respect to semantics. From this point going forward, both of those principles will have a role in every aspect of semantics and pragmatics.

Elements of Word Meaning: Intensions and Extensions, in Anderson's *Essentials of Linguistics*



Video Script

We're now starting to consider how our minds represent the meanings of words. If someone asked you, "What's the meaning of the word *pencil*?" you'd probably be able to describe it — it's something you write with, it has graphite in it, it makes a mark on paper that can be erased, it's long and thin and doesn't weigh much. Or you might just hold up a pencil and say, "This is a pencil". Pointing to an example of something or describing the properties of something, are two pretty different ways of representing a word meaning, but both of them are useful.

One part of how our minds represent word meanings is by using words to refer to things in the world. The **denotation** of a word or a phrase is the set of things in the world that the word refers to. So one denotation for the word *pencil* is this pencil right here. All of these things are denotations for the word *pencil*. Another word for denotation is **extension**.

If we look at the phrase, *the Prime Minister of Canada*, the denotation or extension of that phrase right now in 2017 is Justin Trudeau. So does it make sense to say that Trudeau is the meaning of that phrase *the Prime Minister of Canada*? Well, only partly: in a couple of years, that phrase might refer to someone else, but that doesn't mean that its entire meaning would have changed. And in fact, several other phrases, like, *the eldest son of former Prime Minister Pierre Trudeau*, and *the husband of Sophie Grégoire Trudeau*, and *the curly-haired leader of the Liberal Party* all have Justin Trudeau as their current extension, but that doesn't mean that all those phrases mean the same thing, does it? Along the same lines, the phrase *the President of Canada* doesn't refer to anything at all in the world, because Canada doesn't have a president, so the phrase has no denotation, but it still has meaning. Clearly, denotation or extension is an important element of word meaning, but it's not the entire meaning.

We could say that each of these images is one extension for the word *bird*, but in addition to these particular examples from the bird category, we also have in our minds some list of attributes that a thing needs to have for us to label it as a bird. That mental definition is called our **intension**. So think for a moment: what is your intension for the word *bird*? Probably something like a creature with feathers, wings, claws, a beak, it lays eggs, it can fly. If you see something in the world that you want to label, your mental grammar uses the intension to decide whether that thing in the word is an extension of the label, to decide if it's a member of the category. The next unit will look more closely at how our intensions might be organized in our minds.

One other important element to the meaning of a word is its **connotation**: the mental associations we have with the word, some of which arise from the kinds of other words it tends to co-occur with. A word's connotations will vary from person to person and across cultures, but when we share a mental grammar, we often share many connotations for words. Look at these example sentences:

> Dennis is cheap and stingy.
>
> Dennis is frugal and thrifty.

Both sentences are talking about someone who doesn't like to spend much money, but they have quite different connotations. Calling Dennis cheap and stingy suggests that you think it's kind of rude or unfriendly that he doesn't spend much money. But

calling him frugal and thrifty suggests that it's honourable or virtuous not to spend very much. Try to think of some other pairs of words that have similar meanings but different connotations.

To sum up, our mental definition of a word is an intension, and the particular things in the world that a word can refer to are the extension or denotation of a word. Most words also have connotations as part of their meaning; these are the feelings or associations that arise from how and where we use the word.

Check Yourself

? Exercise 6.1.1

Think about the compound word *hockey player*. What kind of meaning do *Wayne Gretzky*, *Sidney Crosby*, and *Alex Ovechkin* have in relation to *hockey player*?

- Extensions.
- Intensions.
- Denotations
- Connotations.

Answer

"Extensions" and "Connocations"

The reason: They denote *hockey player* because they are all names of famous hockey players. Another word for 'denotation' is 'extension'.

? Exercise 6.1.2

Think about the word *champagne*. Which of the following words represents a **connotation** of the word *champagne*?

- sparkling wine
- Veuve Clicquot
- celebration

Answer

"Celebration"

The reason: We associate *champagne* with any celebration.

? Exercise 6.1.3

Think about the compound word *chocolate cake*. Which of the following words represents an **intension** of the word *chocolate cake*?

Answer

- too many carbs
- made with flour, eggs, cocoa, sugar
- birthday party

The reason: When we think of *chocolate cake*, and we think about what it is made of, its ingredients, that is the intension.

Ambiguity, in Anderson's *Essentials of Linguistics*

In Chapter 8, as we learned to draw tree diagrams to illustrate how sentences are represented in the human mind, we thought about Deep Structure as the place where meaning is assigned and calculated. For example, in a question sentence like, "*What are the kids eating for lunch?*", we claim that the word *what* is related to the verb eating in the same way that *eggs* and *eating* are related in the declarative sentence, "*The kids are eating eggs for lunch*." The relationship between *eating/eggs* and between *eating/what* arises at

Deep Structure, where *eggs* and *what* are both in the complement of the verb. In our theory, a sentence's meaning is correlated directly with the sentence's syntax.

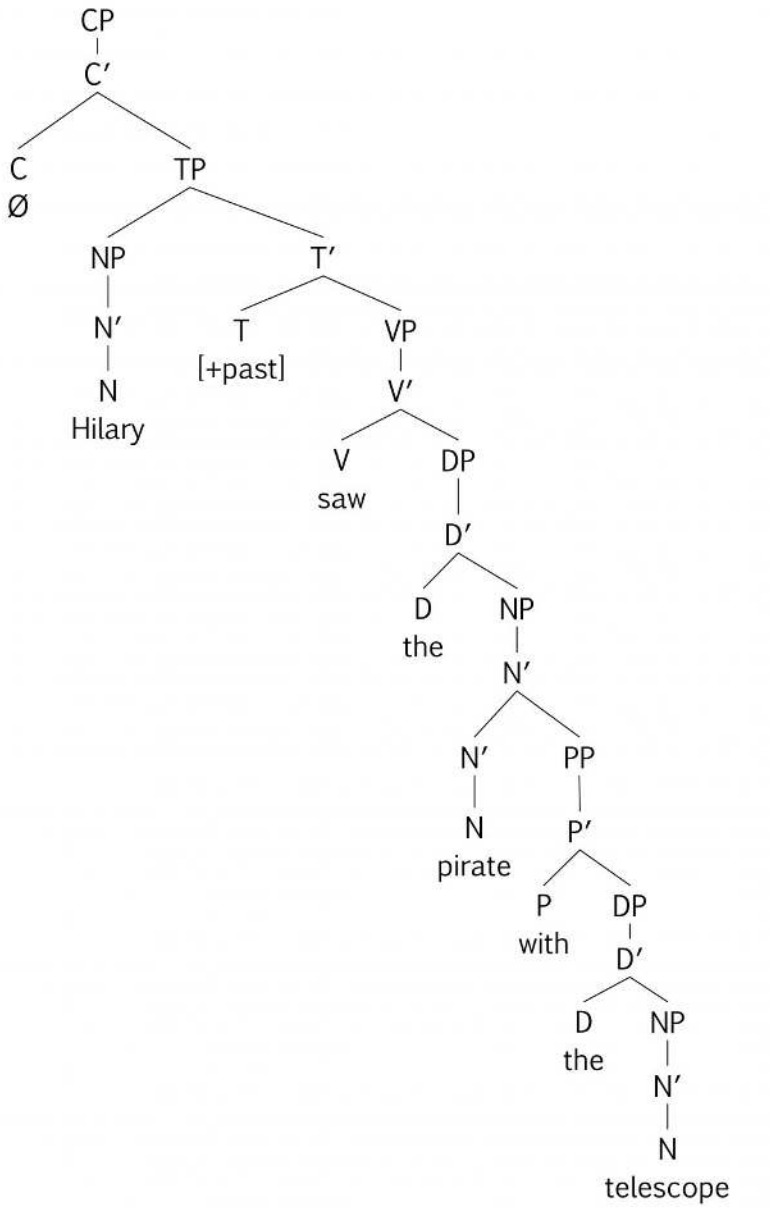
This idea is a core one in linguistics: the meaning of some combination or words (that is, of a compound, a phrase or a sentence) arises not just from the meanings of the words themselves, but also from the way those words are combined. This idea is known as **compositionality**: meaning is composed from word meanings plus morphosyntactic structures.

If structure gives rise to meaning, then it follows that different ways of combining words will lead to different meanings. When a word, phrase, or sentence has more than one meaning, it is **ambiguous**. The word *ambiguous* is another of those words that has a specific meaning in linguistics: it doesn't just mean that a sentence's meaning is vague or unclear. *Ambiguous* means that there are two or more distinct meanings available.

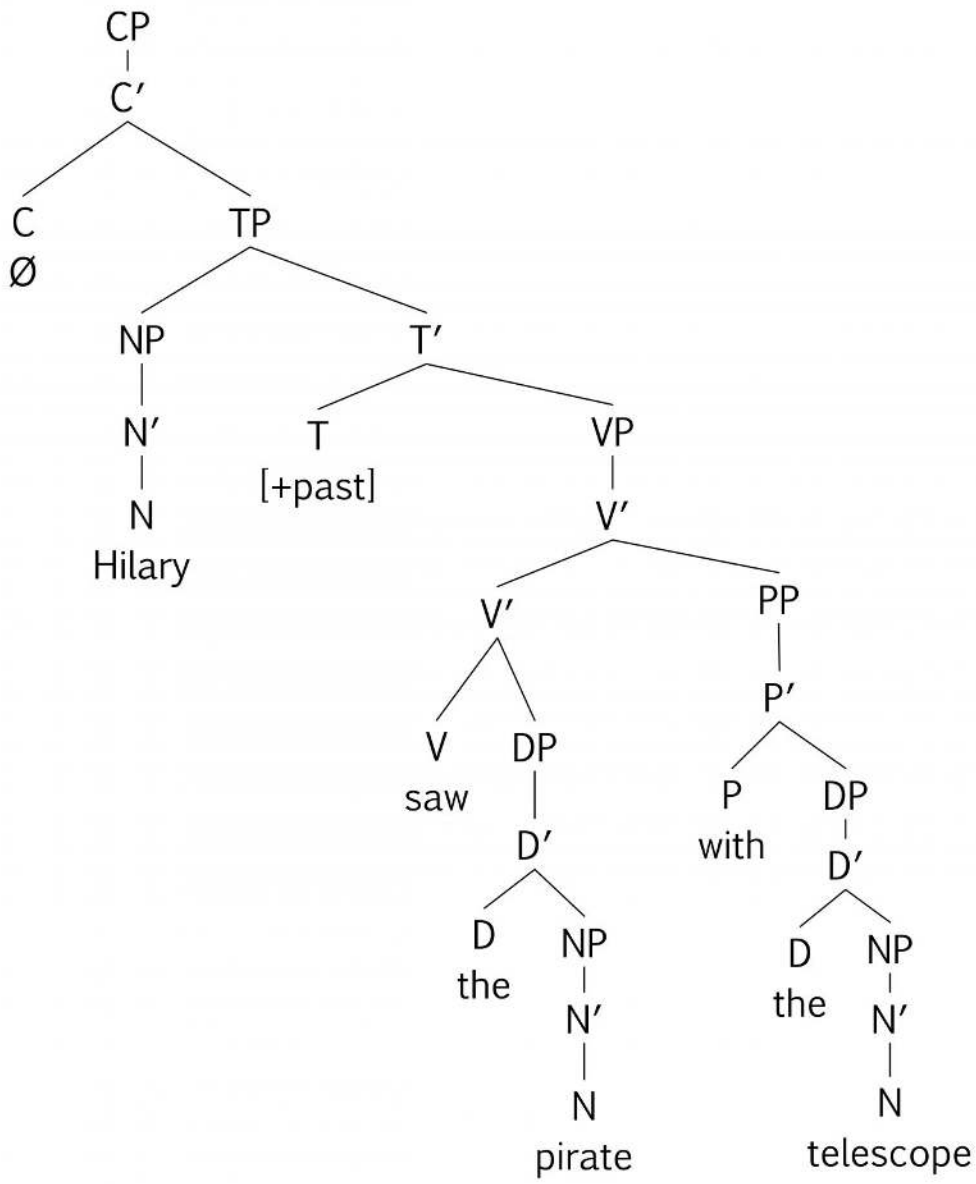
In some sentences, ambiguity arises from the possibility of more than one grammatical syntactic representation for the sentence. Think about this example:

> *Hilary saw the pirate with the telescope.*

There are at least two potential locations that the PP *with the telescope* could be adjoined. If the PP is adjoined to the N-bar headed by *pirate*, then it's part of the DP. (Notice that the whole DP *the pirate with the telescope* could be replaced by the pronoun *her* or *him*.) In this scenario, the pirate is holding a telescope, and Hilary sees that pirate.



But if the PP is adjoined to the V-bar headed by *saw*, then the DP *the pirate* is its own constituent, and *with the telescope* gives information about how the pirate-seeing event happened. In this scenario, Hilary is using the telescope to see the pirate.



This single string of words has two distinct meanings, which arise from two different grammatical ways of combining the words in the sentence. This is known as **structural ambiguity** or **syntactic ambiguity**.

Structural ambiguity can sometimes lead to some funny interpretations. This often happens in news headlines, where function words get omitted. For example, in December 2017, several news outlets reported, “Lindsay Lohan bitten by snake on holiday in Thailand”, which led a few commentators to express surprise that snakes take holidays.

Another source of ambiguity in English comes not from the syntactic possibilities for combining words, but from the words themselves. If a word has more than one distinct meaning, then using that word in a sentence can lead to **lexical ambiguity**. In this sentence:

> *Heike recognized it by its unusual bark.*

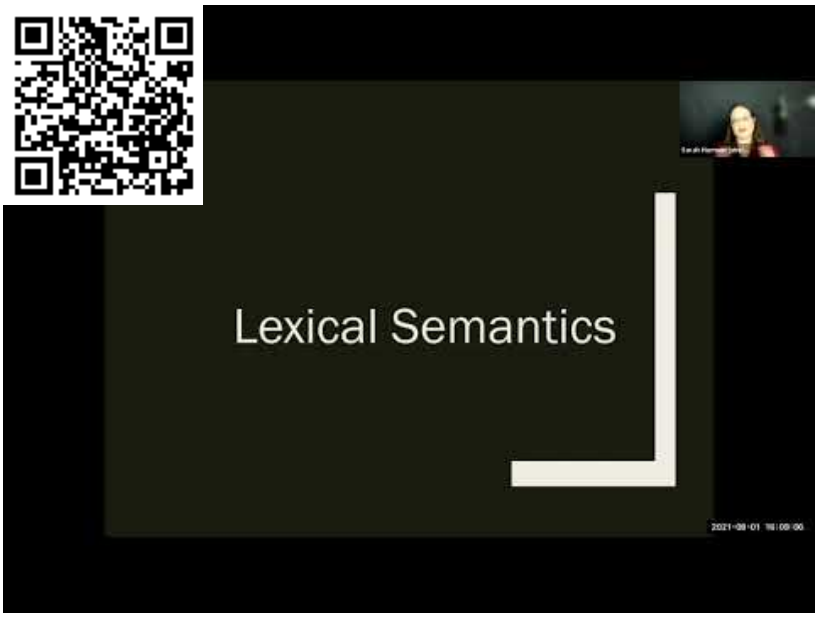
It’s not clear whether Heike recognizes a tree by the look of the bark on its trunk, or if she recognizes a dog by the sound of its barking. In many cases, the word *bark* would be disambiguated by the surrounding context, but in the absence of contextual information, the sentence is ambiguous.

6.1: Arbitrariness and Compositionality is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **10.1: Elements of Word Meaning- Intensions and Extensions** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

6.2: Lexical Semantics

Lexical Semantics, from Sarah Harmon



Video Script

Catherine Anderson does a really great job of setting up lexical semantics, but there's a few more pieces that I want to cover. One has to do with these connections between reference, so the lexicon or phrase, and the meaning they are associated with.

Almost assuredly you have heard of **synonyms** before. They are the quintessential proof that language is arbitrary, because if you have similar or same terms for the similar terms for the same meaning, that's arbitrariness.

Antonyms, you almost assuredly have heard of them before. There are three different kinds of antonymy: complimentary, gradable, and relational. **Complimentary antonyms** are this or that: *on/off, alive/dead* (sorry, zombie fans, there's nothing in between!). **Gradable antonyms** are the opposite; they are along a gradation or scale: *hot/cold* but that there is a whole bunch of other temperatures in between. **Relational antonymy** does have to do with some kind of relationship: *employer/employee*, *parent/child*, *teacher/student*.

Homophones; you may have heard them as homonyms before. Homophones is what we say in linguistics. They have the same sound, but two completely unrelated backgrounds or origins. In morphology I talked about this before, when I talked about the word *bank*; that lexicon has two different entries and they're unrelated. One is the financial institution, and the other is the side of a river. *Seal*, is another one, whether you're talking about the animal or the thing you use to close a document or paper. That is a homophone because they have two completely different meanings and they're completely different origins.

There's three more than I want to go through that you may have heard of before, but maybe not. The first is **polysemy**. Polysemy is different than homophony; remember homophony has to do with different origins. Polysemy, as the name implies, means they have the same origin just multiple levels of meaning. *poly-* = multiple, *-semy* = meaning. If I say the English term *sheet*, you think of something flat of something thin and flexible or pliable. *A sheet of paper; a sheet of linen.*

A **hypernym** is a greater subset. Instead of talking about a specific aspect of something like a *red, blue, pink*, or *purple* thing, you talk about a *color*: 'oh it's got a lot of *color'*. That's a hyponym. Instead of talking about a *tiger, lion, cheetah, leopard, lynx,* or *bobcat*, you talk about *felines*. That's a hyponym.

A **metonym** is a word substitute. Instead of talking about the *King of Spain* or the *Queen of England*, you talk about the *Spanish crown*, *the English crown*; that term crown is a metonym. Instead of talking about the action during the baseball game, you say the action on the *diamond*. That is a metonym because the diamond is the shape of the infield of a baseball field.

All six of these areas include some aspect of compositionality, but especially arbitrariness. Whether it's the actual origin or the usage in any given moment, when you're using these six lexical semantic categories, you're talking about arbitrariness, the arbitrary

combination of the term, and the meaning.

6.2: Lexical Semantics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.3: Phrasal Semantics

6.3.1 From 9.2 Events, Participants and Thematic Roles, in Anderson's *Essentials in Linguistics*



Video Script

We've spent a lot of time thinking about the structure of sentences. We're now turning our attention to what sentences mean. Sentences usually describe events or states in the world. And events usually have participants: the people or things that play a role in the event. Usually, noun phrases are used to refer to the participants in an event. It turns out that, even across events that are quite different from each other, some participants share some elements of meaning.

Take a look at the underlined phrases in each of these sentences.

> Mina tore the wrapping paper.
>
> Sam ran a marathon.
>
> The students studied for their exam.
>
> Neeraja waited for the bus.
>
> Carlos ate the rice.

We can see that the grammatical role of each of these is a subject: They're all in the specifier of TP. Semantically, the events that each sentence describes are quite different: tearing is different from running which is different from studying or waiting or eating. But even across these different events, the participants described by the underlined noun phrases all share some semantic similarities: all of them choose to take part in the event, all of them are causing the event to happen. Let's look at another few sentences.

> Mina tore the wrapping paper.
>
> A nail tore her skirt.
>
> The fabric tore.

All of these sentences have the same verb and they all describe a tearing event. And all the underlined phrases have the grammatical role of subject, but they don't share the same semantic properties. In the first sentence, Mina is the one who causes the tearing event to happen: you can imagine her gleefully tearing the paper open to see what's inside. In the second sentence, the nail is sort of responsible for the tearing, but it certainly doesn't choose to make it happen. And in the third sentence, the fabric is the thing that the tearing happens to, not the participant that makes the tearing happen. So even though all three of these NPs are subjects, they don't all share semantic properties.

Remember that we use grammatical roles to label the syntactic position of a noun phrase in a sentence. We're now going to introduce a new kind of label, called **thematic roles**. We can use thematic roles to identify common semantic properties of the participants in events. An important thing to notice about thematic roles is that they are independent of grammatical roles. In this pair of sentences,

> Kavitha cooked this lovely meal.
>
> This lovely meal was cooked by Kavitha.

the grammatical role for *Kavitha* is different: *Kavitha* is the subject of one sentence but an oblique in the other. But semantically, Kavitha's role in the cooking event is the same in both sentences. We say that Kavitha's thematic role is the **agent**.

The kinds of participants that we label as agents tend to have three properties: usually, they are **volitional**, meaning they choose to participate in the event. They're **sentient**, that is, they're aware of the event, and often they're the ones that bring the event about or **cause** it to take place. Let's look back at that tearing event.

> Mina tore the paper.
>
> The paper tore.

Again in these sentences, *the paper* has two different grammatical roles: it's the direct object in the first sentence but the subject of the second, but semantically its role as a participant in the tearing event is the same in both: it's the thing that the tearing happens to. Its thematic role is called a **theme**, or in some books, you'll see it called a **patient**. Theme participants typically **undergo** events, that is, events happen to them. They're **affected** by events, and often they **change state** or **position** as a result of an event.

Take a minute and try to think of some sentences that describe events that have agent and theme participants. They're probably the two most common thematic role labels, and in fact, one theory of semantics says that every participant is either an Agent or a Theme, just to a greater or lesser degree. But it can also be useful to have labels for some other kinds of participants, and the grammars of many languages encode other semantic properties besides those two.

Some languages make a morphological distinction between an animate agent and an inanimate **cause**. In a sentence like, *The hurricane destroyed the houses*, the hurricane is clearly responsible for the destroying event, but it's not sentient or volitional — the hurricane isn't choosing to bring about the destroying. Likewise, in *The movie frightened the children*, the movie isn't really a typical agent. We label these inanimate participants with the label **cause.** A cause participant shares the agentive property of causing an event to happen, but it's not aware of the event and doesn't choose to cause it, because the cause is inanimate.

In this sentence, *The knife cut the bread*, would you say that the knife is a cause participant? Certainly, the knife is inanimate, and it's not aware of the cutting event, but it's also not really causing the cutting to happen, is it? There's some unnamed agent who must be using the knife to cut the bread. We could label *the knife* as an **instrument**. An instrument is the participant that an agent uses to make an event happen.

Many languages have special morphology to indicate the **location** of an event, like in these sentences:

> The Habs won the game at the Forum.
>
> The kids ran through the sprinkler on the lawn.
>
> The parade travelled around the neighbourhood.

The noun phrases the *forum*, the *lawn* and the *neighbourhood* all have the thematic role of **location**.

So we've got labels like **cause**, **instrument**, and **location** to describe some of the roles that inanimate participants typically have in events. I want to return to animate participants to look at one more important role. Let's look at the human participants in these sentences:

> Phoebe tripped on the curb.
>
> Sun-Jin won the lottery.
>
> The movie frightened Farah.

If Phoebe tripped on the curb, it doesn't seem quite right to label *Phoebe* as an agent — presumably, she didn't choose to trip on the curb, even if she is aware of it, and she isn't really the cause of the tripping event; the curb is. And no matter how badly you might want to win the lottery, you can't really cause it to happen, so Sun-Jin isn't a great example of an agent either. Likewise, if we say

that the movie frightened Farah, Farah isn't exactly a theme; yes, the frightening is happening to her, but she's not necessarily changed by it, and she is aware of the event.

Let's label these participants with the thematic role of the **experiencer**. Experiencers are like the middle ground between agents and themes. They are animate and sentient, so they're aware of events happening, but they don't necessarily choose or cause events to happen; events happen to them. Because experiencers have this in-between status, they can show up either as subjects or as objects, like in these examples:

> The children were scared of the clowns.
>
> The clowns frightened the children.

And we could say that Phoebe and Sun-Jin are experiencers of their tripping and winning events: they don't cause the events to happen, but they are aware of the events happening.

To sum up, thematic role labels capture the semantic properties of participants in events, independent of the syntactic position of the noun phrase. Just because something has the grammatical role of a subject doesn't mean it will necessarily have the thematic role of agent and vice versa. There's a fair amount of argument in the literature about exactly how many thematic role labels are necessary to capture the relevant patterns of behaviour in the languages of the world, with proposals ranging from two to about fifteen thematic roles. We'll settle on the middle ground and use six thematic role labels:

- Agent
- Theme
- Cause
- Instrument
- Location
- Experiencer

Check Yourself

? Exercise 6.3.1

What label best describes the thematic role of the bolded NP?

The guard chased the intruder.

- Agent.
- Theme.
- Cause.
- Instrument.
- Location.
- Experiencer.

Answer

"Agent"

Hint: Who is doing the action? It's *the guard*.

? Exercise 6.3.2

What label best describes the thematic role of the underlined NP?

The wind slammed the door shut.

- Agent.
- Theme.
- Cause.
- Instrument.
- Location.
- Experiencer.

Answer

"Cause"

Hint: Is the wind causing a change of state of the door? Yes, it is.

Exercise 6.3.3

What label best describes the thematic role of the underlined NP?

The guard followed **the intruder**.

- Agent.
- Theme.
- Cause.
- Instrument.
- Location.
- Experiencer.

Answer

"Theme"

Hint: Who is being followed? It's *the intruder*

6.3.2: From 9.3 Thematic Roles and Passive Sentences, in Anderson's *Essentials of Linguistics*



Video Script

Many sentences describe events that involve two participants: an agent and a theme. And it often happens that the agent role shows up in subject position and the theme role in object position. These sentences illustrate that common pattern: the subjects are all agents and the objects are all themes.

> Ilona broke an icicle.
>
> Zainab introduced the guest speaker.
>
> The manager fired the receptionist.

It's a common tendency across languages for the agent to occupy the subject position, but of course not all agents are subjects, and not all subjects are agents. These next sentences describe pretty much the same events as the last three, but the noun phrases in

subject position are not agents.

The icicle broke.

The guest speaker was introduced by Zainab.

The receptionist got fired.

So while the usual pattern is for agents to be mapped onto subject position and themes onto object position, most languages also have a way of reversing that usual mapping. In English, the strategy we have involves both morphology and syntax and is called a **passive** structure. A passive sentence reverses the usual mapping between thematic roles and grammatical roles.

In this first sentence, *The police arrested the burglar*, the police are the agent and they're in subject position, and the burglar is the theme in direct object position.

In the second sentence, *The burglar was arrested by the police*, the semantic relationship of the police and the burglar to the arresting event is the same: the police are still the agent and the burglar is still the theme. But their grammatical roles are different. We can use this passive structure to reverse the usual pattern and focus our attention more on the theme than on the agent.

The reversal that happens in a passive sentence works the same even if the thematic roles aren't the classic agent and theme. Take a look at this pair of sentences,

The exhibit impressed the audience.

The audience was impressed by the exhibit.

In the first sentence, which is an active sentence, the usual mapping plays out not with an agent and theme, but with a cause participant in subject position and an experiencer in the object position. When we use a passive structure in the second sentence, the thematic roles of the participants don't change, but their grammatical roles do.

So how can you tell if a sentence is in the passive voice? It's easy: a passive sentence will always have some form of the verb *be*, followed by a past participle. All of these examples are passives.

The burglar was arrested.

The children were invited to the party.

This flight is expected to arrive on time.

The candidate is being prepared for the debate.

I am appalled by your behaviour.

But if you have the verb *be* plus a present participle, or if you have the verb *have* plus a past participle, then those aren't passives. All of these sentences are in the active voice:

The report is calling for changes.

The burglar was planning a heist.

The children were behaving poorly.

The hosts have invited several guests.

The dog had eaten all the Halloween candy.

A passive structure is a morphosyntactic strategy that English uses to reverse the usual mapping of thematic roles onto grammatical roles. Some languages accomplish this reversal with morphology on the verb or with morphology on the noun, but it's pretty common for a language to have a strategy in the morphology or syntax that has this effect in the semantics of a sentence.

Check Yourself

? Exercise 6.3.4

The patient was diagnosed with alopecia.

- Active.
- Passive.

Answer

"Passive"

Hint: Look at the verb, and note the use of the verb *to be* and the past participle. Also, think about the role of *the patient*--are they performing the action, or undergoing the action?

? Exercise 6.3.5

Is the following sentence in the active or passive voice?

Eileen was convinced that her appointment had been cancelled.

- Active.
- Passive.

Answer

"Passive"--in both cases.

Hint: Look at the verbs, and note the uses of the verb *to be* and the past participle. Also, think about the role of *Eileen* and *her appointment*--are they performing the action, or undergoing the action?

? Exercise 6.3.6

Is the following sentence in the active or passive voice?

The children had been invited to a tea party.

- Active.
- Passive.

Answer

"Passive"

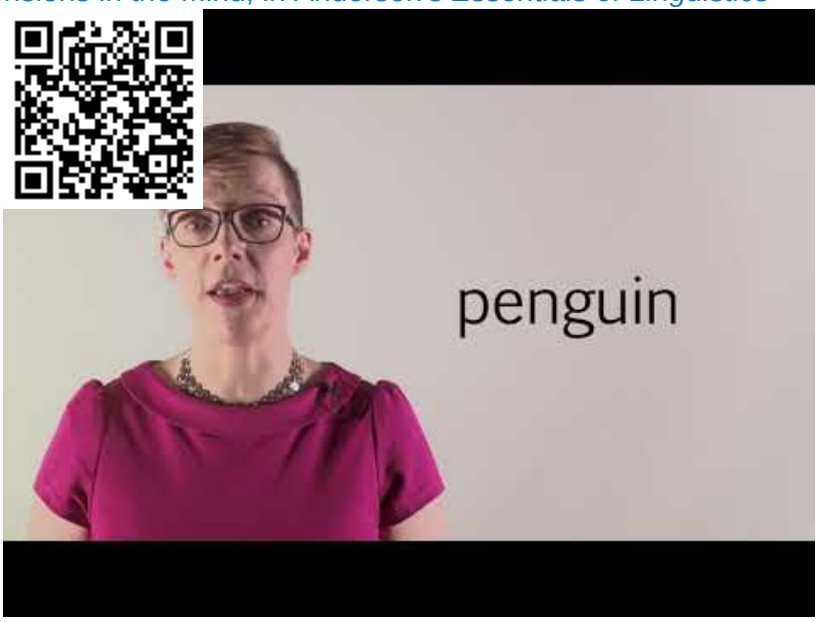
Hint: Look at the verb, and note the use of the verb *to be* and the past participle. Also, think about the role of *the children*--are they performing the action, or undergoing the action?

6.3: Phrasal Semantics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.2: Events, Participants, and Thematic Roles** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **9.3: Thematic Roles and Passive Sentences** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

6.4: Semantic Features

6.4.1 From 10.2 Intensions in the Mind, in Anderson's *Essentials of Linguistics*



Video Script

In the last unit, we saw that one important piece of a word's meaning is the **intension**: the attributes or properties in your mind that you use to decide whether a thing in the world can be labelled with that word. In this unit, we'll think about how those intensions might be organized in the mind.

One theory suggests that intensions might be organized in our minds as sets of **binary features**. So the intension for the word *bird* might be made up of features like [+living], [-mammal], [+wings], [+eggs], [+flying]. The intension for the word *fish* would have some features that are the same as the intension for *bird*, like [+living], [-mammal], [+eggs]. But the intension for *fish* would have [-wings] and [-flying]; instead, it would have [+swimming]. Some of these features could be shared across intensions for words that refer to quite different things in the world, so the intension for the word *airplane*, for example, probably includes [+wings] and [+flying], but [-alive].

The nice thing about using feature composition (also known as componential analysis) to represent intensions is that it can capture some of these similarities and differences across categories of things in the world using the simple, efficient mechanism of binary features. It may well be that our intensions for words describing the natural world are made up of some binary features. But can you think of any problems with this way of organizing meanings? Think about a penguin. A penguin is a member of the category of things that can be labelled with the word *bird*, and it shares some of the features of the intension for the word *bird*: it's a living thing, it has wings, it lays eggs. But a penguin can't fly. In fact, a penguin has the feature that's associated with our intension for *fish*: it can swim. So it's definitely a bird, but it definitely doesn't have all the features associated with the intension for bird. If our intensions were organized in our minds just as binary features, then we wouldn't be able to represent the meaning of the word *penguin* in our mind, but clearly, we do have an intension for the word *penguin*. So how might penguins be represented in our minds?

Another theory of intensions suggests that we have **fuzzy categories** in our minds. These categories contain exemplars, which are basically our memories of every time we've encountered an extension of the word. Some members of the category are prototypical **exemplars**: they have all the typical attributes of members of that category, so they're near the center of the category. For most North Americans, a robin is about as **prototypical** as it gets as an exemplar of the category *bird*. Some exemplars are more **peripheral**: they have fewer of the defining attributes and they might have some attributes that aren't typical. So a penguin, for example, is more peripheral because it doesn't fly, and an ostrich is peripheral because it's so darn big. Because the category has fuzzy boundaries, we might even have some exemplars in our mind that aren't really members of the category at all, but share some attributes with category members, like bats: they're small and they fly, but they're not actually birds. In the next unit, we'll talk about some of the evidence we have that our intensions might be organized in fuzzy categories with prototypes.

Check Yourself

? Exercise 6.4.1

Thinking about the category of *fruit* for most speakers of Mainstream American English, *pomelo* is probably:

- more peripheral than *apple*
- more prototypical than *apple*

Answer

"more peripheral than *apple*"

Hint: Is *pomelo* similar to an *apple*? Possibly not; while they are both fruit, they are different types.

? Exercise 6.4.2

Thinking about the category *pets* for most speakers of Mainstream American English, *tarantula* is probably:

- peripheral
- prototypical

Answer

"Peripheral"

Hint: Is a tarantula a typical pet in North America? Not so much.

? Exercise 6.4.3

Which of the following attributes are likely to be part of the average English-speaker's intension for the word *pencil*?

- ink
- writes
- lead
- erasable

Answer

"Writes" and "Lead" and "Erasable"

Hint: Think about the attributes of a pencil. The only one that doesn't link is *ink*.

6.4.2 Semantic Features, from Sarah Harmon



Video Script

Catherine did a really great job of explaining semantic features. This is an area that I've got quite a bit of experience. Therefore, I'm going to give you a little bit more richness, as it were, with respect to semantic features. Specifically, why it's so important when we are doing any kind of linguistic analysis on a language.

To start off, Catherine talked about **common/semantic features**. She talked about how we can combine them to describe all sorts of possible setups. One of the things that we use semantic features for is to analyze how a given language groups nouns together. In some languages that's based off of gender; biological gender is the origin and then grammatical gender follows. If you've ever taken a class in an Indo-European language that was not English, and you wondered why a table was feminine or a pencil was masculine or a scooter was neuter, that has to do with semantic features. It's usually based off of whatever patterns you see for biological gender, and then everything else follows suit.

That's not the only way that we can classify or group nouns together. For example, if you learn a Sino-Tibetan language or any of the languages of Southeast Asia, you will run into the setup where you have **classifier**. I'm going to use Mandarin as an example—I'm not going to pronounce it because I'm going to be terrible at it. For both of these examples, you have a count noun and you have a mass noun, and then in each case you have a different type of classifier that goes with it. If you want to talk about only having one book, you do not in Mandarin say 'one book'; you say 'one-the classifier that means that this is a countable thing-book'. However, if you're talking about something that needs more specification—meaning, you need to specify the quantity of some grouping or mass of that item—then you need to use a specific quantify. A classifier that describes what kind of situation you're talking about, so the word for 'one', followed by the classifier for box, and then 'lightbulb'. Note that because there is no inflection in Mandarin for plural, you don't say 'one box of light bulbs', instead you use different phrasing to say 'this is a specific grouping of lightbulbs'. Instead of tacking on additional information by phrases in Mandarin and other East and Southeast Asian languages, you use a classifier, in this case, the one that means box.

This is also done in a number of Niger-Congo languages, where you have what is frequently called multiple genders, but it really ties into semantic features. You have a morpheme that says that something is animate or not; you have another morpheme that says that it's a plant versus an animal versus some kind of mineral or rock. You might have another thing that says, whether it's female or male, and along it goes.

Semantic features are really helpful when we try to describe a situation that happens with respect to different types of nouns and how they combined with other aspects. You have here family terms at the bottom part of the slide, in the case you see Latin. In Latin, you actually specified what side of the family that the given aunt or uncle was from. That's still really common in a number of languages, although not the Romance languages, as it died out early in their formation. I'm showing you at the bottom of the screen the word for 'cousin'; there are three different setups. We have an English, which has no differentiation for gender or family side or anything.

In the case of Spanish, you do differentiate for gender, so you do say that this is a female or a male cousin; notice that the terms are exactly the same, just the different vowel the end which signifies gender. In German, you actually have different terms, depending on whether the cousin is female or male.

What is really cool is when this gets showcased in different ways, even within a given language. Many of you are Spanish speakers, either natively or not, and you're looking at these data and they are hard to parse out for you. This is Asturian Spanish. Asturias is the province in the north central part of Spain. They showcase something called a mass neuter, which means that a mass noun that we're talking about it has a different gender associated with it, and you see it in the direct object pronoun. Yes, Spanish speakers, in fact all Romance speakers, you're looking at this a little oddly because you are used to these pronouns going in front of the conjugated verb. That is not the case in Asturian Spanish; it actually goes afterwards, and we can talk more about this later. But if it's a ‘package’, which is masculine and count (you can count it individually, e.g., ‘one package’, ‘five packages’, ‘10 packages’), then the pronoun that gets associated with it is [lu]. I know, Spanish speakers, that's different; just bear with me. If it's a feminine, count noun, like a ‘mare’ (female horse), then the pronoun is [la]—Spanish speakers, you would expect this. However, if it's a mass noun, like ‘grass’—think about it, you can't just say ‘one grass’, you have to say ‘one field of grass’, ‘one blade of grass’—you have to quantify it somehow. The pronoun that goes with it is [lo]. For those who don't know Spanish, [lo] is typically the masculine direct object pronoun. It is not usually used with feminine anything'. In the history of Spanish, and also in a couple of dialects of South-Central Italian and in a couple of dialects of Rhaeto-Romance, which is a Romance language is spoken in the Alps, this phenomenon happens in these remote locations. It's not very common and we don't know where exactly it comes from, because Latin didn't do this and very few dialects of any Romance language do this. There is a history of it in the history in Spanish, going back to the 13th century, and in fact in South central Italian we also see it going back to the 12th and 13th century. We don't really see it much before that, and so we don't really know where it comes from.

What makes this fascinating is to be able to look at the semantic features and how they play in a given language, and in some cases, how one dialect will differ from another.

6.4: Semantic Features is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.5: Pragmatics Definitions

First, watch this video, which is an animation of part of a talk that the psycho-linguist Steven Pinker gave on what language tells us. (This video is captioned.)



Pragmatics Definitions, from Sarah Harmon



Video Script

Now we're going to switch gears we're going to go from talking about semantics to talking about pragmatics. We're still talking about meaning; meaning is still crucially important with respect to this. However, some of our definitions are going to change.

That video that you just saw with respect to Steven Pinker, he was really discussing about these nuances, these little hints that we don't actually say but we imply. They're contextual clues that inform how we might phrase something a little differently. This is all having to do with **pragmatics**. Pragmatics is really talking about meaning but within a specific context. Let's start with some basic definition changes little tweaks.

To start off, we're no longer going to talk about sentences and paragraphs; we're going to talk about **utterances** and **discourse**. This is important because when we talk about pragmatics, we're talking about dialogues between at least two people, maybe more and how people interact. You'll notice, if you ever stop and really listen observe two people having a conversation, that there are plenty of times when the speakers use full and complete sentences—perfectly beautiful full complete sentences—in what a given person says. However, just as often, if not more so, you will have interruptions where one person starts and the other person finishes a given sentence or thought, or stops the person in their tracks to change the subject. You also hear many non-lexical elements: ‘mm’ or ‘um’ or ‘hmm’. Those are all non-lexical. they're not actual morphemes, but I am conveying information. These are all types of utterances, and they are crucially important to look at the context, the entire dialogue, of what is being said. Instead of talking about truth values, we talk about **felicity conditions**, when we are talking about whether something is true for that given context.

The last thing, and at first it may not seem relevant, but it deals with **pronouns**, or in linguistic terms we call it **anaphor**. If you think about when you use pronouns, it's always within a given context. If I start saying, 'she' and 'her' without setting up who it is I'm talking about, you don't know who I'm referring to. In this way, the context clues come in. Felicity conditions also become involved; it depends on which pronoun gets used for a given structure. For example, if I say *Lana saw herself*, in that case 'herself' is the pronoun we're going to use if we want to say that is a reflexive action. That really just means that the anaphor, the referent, is in the earlier part of the same TP (tense phrase) or sentence. If we want to say that the person that was seen in the mirror was somebody other than the subject—not Lana, but somebody else—then we use a different pronoun.

The more we stack their sentences, the more complex it gets. For now, just understand that pronouns/anaphor (or anaphora) are always tied to a context; they're also tied to deixis, which we'll talk about soon enough.

6.5: Pragmatics Definitions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.6: Grice and the Cooperative Principle

6.6.1 From 10.5 Pragmatics and the Cooperative Principle, written by Bronwyn Bjorkman, in Anderson's *Essentials of Linguistics*

Probably all of us have had the experience of having our words misinterpreted, or taken "out of context". This often happens even if someone definitely understood our literal words — they may simply have misread our intentions or goals. This type of misunderstanding is precisely the kind of thing we might investigate in pragmatics — both what we *intend* to communicate, and how someone else might *interpret* our communications.

In this unit we focus on one particular type of pragmatic reasoning, the the calculation of **conversational implicatures** on the basis of what are known as **Gricean Maxims**—these maxims were proposed by the philosopher H.P. Grice in a (1975) paper that proposed that in conversation we adopt a **Cooperative Principle** when interpreting what people say.

Terminology: Entailment, Presupposition, and Implicature

Before discussing the Cooperative Principle and the individual maxims, let's introduce some terminology that will help us talk about the logical relationships between sentences—the different ways in which we can draw conclusions.

Consider the following pair of sentences:

- Jennice and Alice have both read *War and Peace*.
- Jennice has read *War and Peace*.

If the first of these sentences is true, the second one also has to be true! Put another way, if the second sentence is false, then the first sentence is also definitely false. This is the relationship of **entailment**.

Entailment

For two propositions (i.e. things that can be true or false) P and Q, P entails Q if whenever P is true, Q must also be true.

Entailment is technically a semantic relationship rather than a pragmatic one, but it's useful to have it in mind to contrast two other relationships with: **presupposition** and **implicature**.

Consider next the following two sentences:

- Nadim's brother is visiting.
- Nadim has a brother.

This might seem to be another case of entailment, but it works a little bit differently. If the first sentence is true, the second one also has to be true. But if the second one is false—if Nadim doesn't have a brother—then it's not just that the first sentence is false, it seems like we can't even really interpret the first sentence. Here we say that the first sentence **presupposes** the second one.

Here's another pair of sentences where the first sentence **presupposes** the second one:

- Lou stopped smoking.
- Lou used to smoke.

If someone asks you *Have you stopped smoking in the last year?* and you never smoked, you couldn't answer "yes" or "no"—instead you might say "Hey, wait a minute! I never smoked!" This temptation to say something like *Hey wait a minute!* is a sign of a presupposition that isn't satisfied.

Presupposition

For two propositions P and Q, P presupposes Q if Q has to be true for P's truth or falsity to be evaluated.

Finally this brings us to **implicature**, which is the relationship most relevant for our discussion of Gricean maxims.

Consider a final pair of sentences:

- Marie has two cats.
- Marie has exactly two cats.

If someone said to you "*I have two pet cats.*", in most contexts you would assume that they didn't have 10 cats—if it turned out that they did have 10 cats, you'd feel that they'd misled you somehow. But there's nothing about the first sentence in the pair above that logically *entails* that Marie doesn't have more than two cats.

Grice (1975) introduced the term **implicature** for the relationship between the first and second sentences in this pair.

Implicature

> For two propositions P and Q, P **implicates** Q if a listener would infer Q on the basis of someone saying P, despite P not entailing or presupposing Q.

Implicatures, unlike presuppositions or entailments, are **cancellable**—that is, you can negate them without contradicting yourself or saying something infelicitous.

- Marie has two cats, but (in fact) she has ten cats. (implicature → cancellable)
- #Lou stopped smoking, but they didn't used to smoke. (presupposition → not cancellable)
- #Nadim's brother is visiting, but Nadim doesn't have a brother. (presupposition → not cancellable)
- #Jennice and Alice have both read *War and Peace*, but Jennice hasn't read *War and Peace*. (entailment → not cancellable)

Grice distinguished two types of implicatures:

- Conventional implicatures: triggered by specific words
- Conversational implicatures: calculated based on the Cooperative Principle / specific maxims

We will be mostly concerned with **conversational implicatures**. Just for illustration, though, an example of a word that triggers a conventional implicature in English is the coordinator *but*. Consider the contrast between the following two sentences:

- Ruowen likes chocolate ice cream and Helen likes vanilla ice cream.
- Ruowen likes chocolate ice cream but Helen likes vanilla ice cream.

Logically speaking, both *and* and *but* mean the same thing—both of these sentences are true only if it's true that Ruowen likes chocolate ice cream AND true that Helen likes vanilla ice cream.

Grice observed that the coordinator *but* **implies** that there's a contrast between the two clauses, though, or that it's somehow surprising to assert the second one. This is the conventional implicature of using *but* instead of *and*.

Gricean Maxims and the Cooperative Principle

What are we trying to do when we have a conversation with someone? There probably isn't any one thing that we're *always* trying to do, but often part of what we're trying to do is exchange information.

Some of the things we communicate are not part of the logical or literal meaning of our words. For example, consider the following exchange (adapted from Grice 1975):

- A: How does your friend like working at the bank?
- B: Oh, pretty well. They like their colleagues, and they haven't been sent to prison yet.

What does B mean by saying their friend hasn't been sent to prison yet? They could mean a number of things: maybe B is given to telling jokes, or maybe they mean that their friend isn't usually trustworthy, or maybe they mean that if you didn't like working at a bank you'd steal money.

The actual meaning conveyed will depend on the context: what A and B both know, their relationship to one another and to B's friend, and other factors. Whatever B means, though, it's clear that by saying "they haven't been sent to prison yet", B is conveying something more than just the literal meaning of their words! Many many people haven't been to prison! Why is B bringing it up?

The meaning conveyed by a utterance based on these kinds of considerations is its **conversational implicature**: implicatures that arise from the structure of discourse or conversation, based on our understanding of how communication works.

Conversational implicatures arise in the context of a general **Cooperative Principle** for communication. The idea of this principle is that we assume, when we speak to people, that we are working towards a common goal (or common goals), and we can interpret what people say in light of that. So in the mini dialogue above, A can assume that there was some reason that B mentioned that their friend hadn't been sent to prison yet—they didn't simply add a true but irrelevant statement to the conversation for no reason!

Under the general heading of the Cooperative Principle, there are four more specific **maxims** that Grice proposed:

1. Maxim of Quantity
2. Maxim of Quality
3. Maxim of Relation
4. Maxim of Manner

In the following sections we will review how each of these maxims works when it is successful, before turning to two ways in which someone can fail to follow the maxims in conversation: either by **violating** a maxim or by **flouting** one.

If we **violate** a maxim, then we simply fail to follow it. At best, violating a maxim results in being a confusing or uncooperative conversationalist. At worst, violating a maxim involves lying or being intentionally misleading.

If we **flout** a maxim, by contrast, we **blatantly** fail to follow it—we aim to communicate something precisely by making it very obvious that we have chosen not to follow the cooperative principle, and trusting that our audience will draw the intended conclusions.

Maxim of Quantity

The maxim of **quantity** states:

- Make your contribution as informative as is required.
- Do not make your contribution more informative than is required.

To follow this maxim, we make the strongest claim that's both compatible with the facts and relevant in context.

For example, consider the following conversation:

- A: Does Elspeth have any siblings?
- B: Yes, she has a sister.

When hearing B's response, A assumes that B is *fully* answering the question—that is, that B is being as informative as possible. So A would naturally assume that Elspeth has *exactly* one sister, and doesn't have any brothers.

If it turned out that Elspeth has two sisters and a brother, A would feel that B had misled them—this would be an example of **violating** the maxim of quantity.

Changing the context can change how we calculate this implicature, though. Suppose that A needs to borrow a car in order to run an errand, and the following conversation ensues:

- A: Does Elspeth have a car I could borrow?
- B: Yes, she has a car.

In this context A will conclude that Elspeth has *at least* one car. Even if it turns out that Elspeth has two cars, A won't feel like B misled them—because the second sub-maxim above says that you shouldn't be more informative than a conversation requires, and in the relevant context all A needs to know is whether there's a car they can borrow.

Flouting the maxim of quantity can be done in a few different ways! Grice gives the example of a reference letter for a job as a Philosophy professor that says, in its entirety:

- "Dear Sir,[1] Mr. X's command of English is excellent, and his attendance at tutorials has been regular. Yours, etc." (p. 52)

This letter is ostentatiously much shorter than a reference letter would usually be, and so gives rise to the implicature that there is nothing else that the writer can say about Mr. X that would be positive.

A possibly more subtle example of flouting the maxim of quantity might be something like the following:

- Student: When is Assignment 2 due?
- Professor: You can find that information in the syllabus, which is posted on the course website.

In this exchange, the professor hasn't actually provided an answer to the student's question—in that sense it is an uncooperative response. The professor intends to communicate that the student should be able to answer their own question on the basis of information available to them. (The professor's response probably also involves flouting the maxim of relevance, since they have not directly answered the question asked.)

Maxim of Quality

The maxim of **quality** states:

- Do not say what you believe to be false.
- Do not say that for which you lack adequate evidence.

In some ways the first of these points is the most basic maxim for the Cooperative Principle: communicating in good faith seems to require that we are—or at least try to be—truthful.

The second point—don't say that for which you lack adequate evidence—is a bit harder to judge, and what counts as "adequate evidence" varies a great deal from context to context.

Violating the maxim of quality involves lying—intentionally saying things that are untrue—or else saying things that you don't have enough evidence for.

If your housemate asks you what day garbage is being collected this week, and you can't really remember but you think it might be Tuesday or Wednesday, you would be violating the maxim of quality if you confidently replied: "*Garbage pickup is definitely Wednesday this week.*"

Flouting the maxim of quality usually involves irony or sarcasm. For example, consider the following mini-dialogue between a child on a road-trip and their parent:

- **Child, asking for the 20th time:** Are we there yet?
- **Parent, fed up with answering:** Nope, we're just going to keep driving in this car for the rest of our lives.

In this case the parent doesn't intend their child to take their words literally; they're **flouting** the maxim of quality to convey an implicature that the question was unwelcome.

Metaphors or idioms are also cases of flouting the maxim of quality! If I say a scarf is *as light as a feather*, this is not literally true—but I don't intend for it to be taken as true!

Maxim of Relevance

The maxim of **relevance** states:

- Be relevant.

The idea behind this maxim is that when we converse, we shouldn't introduce irrelevant topics—we try to stick to the topic of conversation, and we assume that our contributions will be interpreted in that light.

Consider the following exchange:

- A: Are you visiting family this weekend?
- B: I have a term paper due on Monday.

A natural interpretation of this exchange is that B is saying that they do *not* plan to visit family this weekend, and that the reason is that they have to work instead.

But this interpretation is an **implicature**, because if we think only about the literal meaning of B's words, this interpretation is a bit mysterious—B doesn't actually directly answer A's question, but introduces new topic that doesn't have anything to do with travel or families.

If we assume that B does intend to be relevant, though, we can explain the implicature: for the term paper to be relevant to the question about travel, it must be that working on the paper controls whether B is able to travel to visit family.

Indeed, suppose we know that B finds it easier to write term papers at home for some reason. In that case we might interpret their statement above as meaning that they **do** plan to visit family. This illustrates the type of context dependency that's typical of conversational implicatures!

Violating the maxim of relevance means making irrelevant contributions. You might do this because you're absent-minded, or because you aren't actually paying attention to what the other person is talking about, but you can also violate the maxim of relevance more subtly. Consider a slightly different dialogue:

- C: Are you free to hang out this weekend?
- D: I have a term paper due on Monday.

Imagine this dialogue uttered in a context where D does actually have time to hang out, but for whatever reason doesn't want to spend time with C. Assuming D does have a term paper due on Monday (and therefore is not violating the maxim of Quality), their response would violate relevance: they're saying something true but irrelevant, in the hopes that C will draw the (incorrect) implicature that D doesn't have time to hang out because of the time needed to work on the term paper.

Flouting the maxim of relevance involves saying something obviously irrelevant, often to communicate that you want to change the topic of conversation.

For example, if a conversation starts getting awkward and you interject by saying: "*How about that hockey game last night?*" (when nobody had been talking about sports, never mind about hockey), then you would be **flouting** the maxim of relevance in the hopes that your audience would understand that you were trying to convey: "*Can we please talk about something, anything, else?*"

Maxim of Manner

The maxim of **manner** states:

- Avoid obscurity of expression. (That is, don't use words or phrases that are hard to understand.)
- Avoid ambiguity.
- Be brief.
- Be orderly.

This relates not to the content of what you say, but the way you express yourself.

It is easiest to discuss each of these sub-maxims in turn, because they have slightly different effect in conversation.

Avoid obscurity of expression

To follow this maxim, we try to use words and expressions that our audience understands. For example, the terminology you would use when explaining a topic in a university class is different from the terminology you'd use if talking about the same thing to a Grade 1 class!

Violating this sub-maxim involves using words and phrases your audience doesn't know. We often do this accidentally—for example, your linguistics professor might accidentally use a grammatical term that they haven't defined in class—but if you use a word that you suspect your audience will misinterpret, with the goal of misleading your audience, that would also be a violation of this sub-maxim.

Flouting this sub-maxim would involve intentionally using words that you don't expect your audience to understand—in most contexts, this would be with the goal of conveying something like: "I'm smarter / know more than you do", which is pretentious and rude.

Avoid ambiguity

To follow this sub-maxim, we try to avoid saying things that can reasonably be interpreted in more than one way.

It's very easy to **violate** this sub-maxim accidentally, because often you don't see the ambiguity in something you say until it's pointed out to you! But again, you can be intentionally ambiguous in the hopes of misleading people—this is an uncooperative way of talking.

Flouting this sub-maxim often happens in certain kinds of jokes, as in the following:

- A man walks into a bar. Ouch!

This joke turns on two things: 1. being familiar with the common joke set up: "*Someone walks into a bar.*" and 2. intentionally using the other meaning of the ambiguous word *bar*.

Be brief

To follow this sub-maxim, we avoid going on at great length when a shorter statement would do.

Violating this sub-maxim involves saying or writing something much longer than is needed.

Flouting this maxim is more subtle. One example is avoiding a single word and instead using a long paraphrase, as in:

- What did you have for dinner last night?
- Well, we combined all the ingredients listed in a recipe for *risotto milanese*, in the indicated order, and the result was edible.

By using this long paraphrase, instead of saying "*We made risotto milanese.*", the second speaker's response gives rise to the implicature that the recipe didn't turn out as intended, or wasn't very good.

Be orderly

To follow this sub-maxim, we list or relate things in an order that makes sense. For example, when telling a story, we usually start at the beginning and then relate events in the order they happened in.

Violating this sub-maxim can be very confusing, as you'll know if you've ever had to interrupt someone for clarification about the order of events in a story they're telling!

Flouting this sub-maxim is not something we would do very often. A possible example might be intentionally relating events out of order when writing a fictional story, to convey something about the mood or the narrator's state of mind.

References

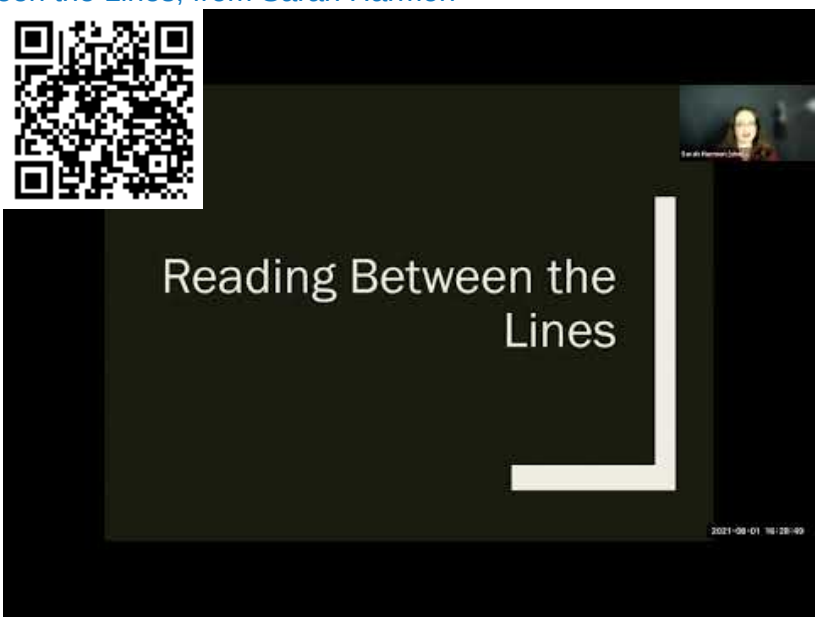
Grice, H. P. (1975). Logic and conversation. In *Syntax and Semantics 3: Speech acts*, ed. Cole et al. (pp. 41–58). Brill.

1. This is a now–outdated greeting for a formal business letter when you don't know the name of the person who will receive the letter. It would now be rude to assume that the recipient would be a man. ↩

6.6: Grice and the Cooperative Principle is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.7: Reading Between the Lines

6.7.1 Reading Between the Lines, from Sarah Harmon



Video Script

Catherine did a really great job and explaining implicature and presupposition and so I'm not going to go too much more into that. There is one more piece that I want to bring to this discussion about reading between the lines, and it has to do with performative verbs.

What is a **performative verb**? It is pretty much what you think it is: when you say the action and you are performing the action at the same time. It's a type of **speech act**. If that's a little hard to understand, close your eyes and what do you think of when I say: *declare, baptize, pronounce, sentence, promise*. All of those are verbs that in the present tense are speech acts. They are performing an action. The second I say *I declare you married*, at this point in time, the second I say *I declare*, that is when that couple is now married. That is what that means that is a speech act, or that is a performative verb. Every time that somebody says to you, *I promise I will do this*, or *I promise I will get your work graded by Tuesday*. The second that the person says, *I promise*, that is the promise or the contract, as it were, between the two people involved, that is a speech act. that is a performative verb.

We're going to see pretty soon, when we talk about deixis, that in some languages, there needs to be a specific inflection to say that this is a performative verb. What is true in all languages is that the speech act must be in present tense in order for it to be performative. The second you change to future tense or past tense; it is just describing what will happen or what has happened and nothing more than that. It is not performative unless it is in the present tense.

Speaking of deixis, let's move on to that.

6.7.2 Hesitation and Dialogues

The following video is from a TedEx contributor, Lorenzo García Amaya, in which he discusses the role of hesitations when we speak. Think about how you and your peers/friends speak, and how often you use hesitations to convey information. (The video is captioned.)



6.7: Reading Between the Lines is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.8: Deixis

Deixis: Meaning that Depends on Context, in Anderson's *Essentials of Linguistics*



Video Script

We saw in a previous unit that many words have extensions that can change over time while their intensions stay fairly constant. For example, the extension of the phrase *the Prime Minister* changes every few years, after elections. But there are some words whose extensions change all the time, depending on who says the words and what context they're in.

Think about two kids who are fighting over a ball. One kid says, "It's mine!" and the other says, "It's mine!" Both of them are uttering the same words, but they each have a different extension for the meaning of the word mine. When the tall kid says *mine*, they mean that the ball belongs to the tall kid. And when the kid with pigtails says *mine*, they mean that the ball belongs to the kid with pigtails.

This phenomenon, where a word's referent changes depending on who says the word, is called **deixis**, and words or phrases that allow deixis are called **deictic expressions**.

In every language, first-person and second-person pronouns are deictic. Whoever says the word *I* or *me* or *myself*, they're using the word to refer to themself. And when we utter the word *you*, we mean the person or people we're talking to, whoever those people may be. And the first- and second-person possessives are deictic too.

What about third-person pronouns? Is the pronoun *she* deictic? Let's look at an example. Suppose Sam says, "The prof said she would give us all A's." The pronoun *she* is ambiguous — it could refer to any feminine person, so it's possible that Sam means that the prof said that the TA or some other prof would give all A's, but the likeliest interpretation is that *she* refers to the prof. Now what happens if Tai says, "The prof said she would give us all A's"? The word *she* is still ambiguous, but in exactly the same ways — it could still refer to the prof, or it could refer to some other feminine person. The potential referent for the word she does not depend on who is uttering the sentence, so it's not a deictic expression.

So first- and second-person pronouns and possessives are deictic in every language. But that's not the only place that deixis happens in language. Lots of languages also have **spatial deixis**, whose referent depends on the location of the person who utters them.

Imagine this conversation between Sam and Tai, who live in different cities: Sam lives in Hamilton and Tai lives in Toronto. They've been talking about getting together on the weekend. Sam says, "Are you coming here this weekend?" and Tai replies, "No, I thought you were coming here!" Both of them utter the word *here*, but each one is referring to a different place — for Sam, the word *here* refers to Hamilton, but for Tai, *here* means Toronto. The referent for the word *here* depends on the location of the person who says it.

English has some pairs of deictic expressions that depend on location. *Here* indicates some relative proximity to the speaker, while *there* means something that is farther away from the speaker. The linguistics labels for this near/far distinction are **proximal** and **distal**. The English demonstrative determiners also make a distinction between proximal and distal: *this* and *these* refer to things that are closer to the speaker, and *that* and *those* refer to things that are farther away. English even has verbs that express this distinction: *come* and *bring* refer to moving towards the speaker, while *go* and *take* mean moving away from the speaker.

Many languages make a three-way distinction in spatial deixis. In Spanish, for example, *este* corresponds roughly to English *this*, while *ese* and *aquel* both get translated as *that*. But *aquel* is definitely far away, while *ese* is farther away than *este* but not as far as *aquel*. This intermediate spatial distinction is labelled **medial**. Plenty of other languages, like Arabic and Korean, also have a three-way distinction. In fact, English used to have a proximal-medial-distal distinction as well, with the word *yon* expressing the distal, but *yon* has pretty much vanished from modern English.

Languages also have ways of expressing **temporal deixis**. Suppose you go to your prof's office to ask some questions and you find a note on the door that says, "Working from home today. I'll be in the office tomorrow." You have no way of knowing what day they're working from home and what day they'll be in the office unless you know what day the note was written, because *today* means whatever day they posted the note and *tomorrow* means whatever day comes after that day. *Yesterday* obviously works the same way: its referent is relative to when it gets uttered, and the same is true for *now* and *then*, *soon* and *later*. English also has expressions like *three weeks ago* and *next year* that are deictic too.

In fact, even the tense morphology on verbs is deictic. Suppose you get a letter from your aunt in the mail and it hasn't got a date on it. It's a little beat up and it looks like maybe it got lost in the system for a while. The letter has some news about the family and includes the sentence, "Alex will spend the summer planting trees." Now, because this sentence has a future tense verb in it, you know that the tree-planting was set to happen some time after the letter was written, but without knowing when the letter was written, you can't know whether Alex has already planted trees or is still planning to do it in the future or is planting trees right this minute. The time that the future tense refers to depends on when the verb was spoken, or in this case, written.

To sum up, every language has deictic words, phrases or expressions that refer to something different depending on who speaks or writes them, and in what context. The most common kinds of deictic expressions are personal, depending on the identity of the speaker, spatial, which depend on where the speaker is when they say the phrase, and temporal, which depend on the time the speaker says the phrase.

Check Yourself

? Exercise 6.8.1

What kind of deixis is illustrated by the bolded part of this sentence?

*"I heard **this course** was supposed to be easy."*

- personal deixis
- spatial deixis
- temporal deixis
- not a deictic expression

Answer

"Spatial deixis" and "temporal deixis"

The reason: *This* usually implies a difference in space related to whoever is speaking, and/or differences in time related to whoever is speaking.

? Exercise 6.8.2

What kind of deixis is illustrated by the bolded part of this sentence?

*"Stan told Bill **his brother** had phoned."*

- personal deixis
- spatial deixis
- temporal deixis

- not a deictic expression

Answer

"Personal deixis"

The reason: *His* implies a difference in personal relationship related to whoever is speaking.

? Exercise 6.8.3

Which part of this sentence is deictic?

"Sam and Tai met five years ago."

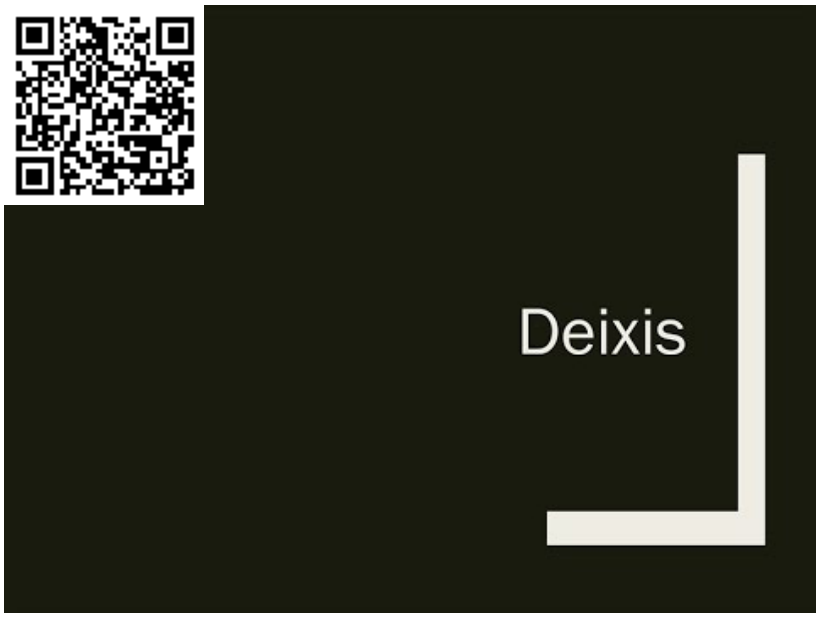
- Sam and Tai
- met
- five years
- five years ago

Answer

"Five years ago"

The reason: Both the amount of time and the term *ago* create a difference in temporal deixis.

Deixis, from Sarah Harmon



Video Script

What Catherine explains above is what you need to know about deixis. There's a little bit more I want to say, because this is an area that has fascinated me for a very long time. I've done a little research on it, published a couple papers on it, but mostly, this is just an area of complete fascination. Maybe it's because English doesn't do as much with deixis as other languages, with respect to verbal inflection. That's really where we're going to talk about this.

As far as what deixis is, Catherine has already explained this above. she does a really great job of it. Remember that **personal deixis** is having to do with first and second person pronouns and the demonstrative. **Place deixis** is directions, this concept of *here* versus *there* versus *yonder*. **Time deixis** usually is focused on two things: one is the adverbs that we use, like *now*, *then*, or *later*. It also involves more importantly, for the rest of this video, verbal inflection, or specifically what we call in linguistics TAM (tense, aspect, mood).

If you've ever taken a foreign language class—especially if it was an Indo-European language or Japanese or Korean or anything that has a lot of verbal inflection—then you've heard these terms before at least a couple of them, if not all three of them. When I teach Spanish, I do subject my students to this, because I think it's an important aspect to remember, that those little bits that we put on to a verb, all those inflectional affixes, they're doing something. Let's get into it a little bit, let's unpack what this mean.

Tense is the easy one, because tense only has three options at most: past, present, future. That's it, there is no other tense beyond past, present or future. This is one of my pet peeves when I read a grammar book that talks about 'the progressive tense' or 'the preterite tense'. Those are not tenses at all. We only have three possibilities. That does not mean that all languages have three tenses; in fact, that's not the case. Most Indo-European languages have all three, but many languages around the world, only have two: past and non-past is the typical definition. Some languages do not have any actual tense markers—think the Sino-Tibetan languages, where there are no markers of any kind for anything. They use adverbs of some kind, to tell you when something happened.

A lot of languages that have inflection—not the East and Southeast Asian languages that we've talked about but other languages—may not have tense markers they do have aspect markers. I'm thinking especially the languages throughout Central America and in parts of the islands in and around Australia. We're talking about languages that may not have a past tense, present tense, or future tense, as far as their inflections. Yet, they do have stuff like 'ongoing' or 'at a specific moment' or 'kind of in the past, but not really'. Those are all aspects**. Aspect** is really just a manipulation of time; you're describing how the time transgressed.

Some of these terms that I'm going to show you, you may have seen, especially if you learned a grammar of a foreign language. Many of them might be new to you, so I'll explain.

- **Iterative** or **Durative**: this is a really common one, especially in the Indo-European languages, but even outside. This is when we're talking about time that goes over a stretch, a period. During…duration. Iterative and durative, technically speaking, are slightly different but for your purposes they're the same.
- **Punctual** or **Aorist**: technically speaking, those are slightly different, but for your purposes we'll keep them together. Punctual, at a specific moment in time. For example, if you learned a Romance language or if you learned a Germanic language, you probably heard of the perfect or preterite, maybe the compositional past tense. You've heard of the imperfect, also. Well, the imperfect is durative and the perfect/preterite is punctual. If you think about that difference about how you would say in any of these languages, 'I went to the store at three o'clock yesterday' versus 'I was going to the market when and then there's something else happened.' Those two versions of 'go' in any Germanic or Romance language (and most Indo-European languages) are going to be two totally different forms and that has to do with aspect.
- **Habitual** or **Imperfective**: frequently they can get mixed into this iterative/durative. But it often also gets split out, especially in many languages in Africa and Niger Congo languages. Habitual is just what you think it is: something that used to happen or happens repetitively recursively. Imperfective is along those same lines.

All three of these aspects, English doesn't really do, at least not anymore. We used to, but that's for another time. What we do have are the next two.

- **Perfective**: At that time and before. 'I have eaten today' or 'I will have read that book for next week'.
- **Progressive** or **Continuous**: 'I am talking to my mother, right now', 'my brother is working right now'.

Those are both aspects that we do have an English; you can see them not just an English, but in other languages, as they're very common. The last one is a little less common.

- **Inchoative**: usually it talks about the start of something. Sometimes it gets folded into the punctual. In some languages, many indigenous languages in North America, have inchoative.

These are all different aspects. Some are pervasive; you frequently will see a progressive and a perfective, you frequently see either habitual or durative or both. You may not have all of them, and certainly English is a great example of this, because we don't have these as inflections anymore. We tend to have them as compound verb phrases; they use auxiliaries and that kind of thing. However, if you go back 800 years, we used to have quite a bit of this.

That's tense and that's aspect. The fun one, at least for me, is mood. Don't be scared off; you're going to see a lot of this information and think it's all on the semantics quiz or the meaning quiz. It's not I promise. I show you this to get you to understand all of these things. When you are in an environment where you having to learn a new language, and it has all these different things that you do to the verb, and you weren't sure what was going on, some of it is aspect and some of it is mood.

The way I like to think about **mood** is condition or context; it is the context under with something happens. We have three main moods and then of course they gets split out from there.

- The first one is **interrogative**, as in a question and, yes, in many languages it is an inflection on the verb. When you want to make a question, instead of a question word or changing intonation pattern, you put an inflection on the verb.

You also have realis and irrealis. Let's break this down.

- **Realis** = something real, so something that is observable in many cases. The one you are familiar with, is **indicative** or sometimes called **declarative**. I know you're aware of this, because that's pretty much what we use in English; we really don't have much else in the term of mood anymore. If you are studying another Indo-European language, certainly indicative or declarative is what we focus on.
- But there's one other realis mood and it's really common in the Semitic languages, and that is the **energetic**. Energetic is really cool; it is when you want to say that something is strongly believed by the speaker, or the speaker wishes to really emphasize it put energy to it. You see here that there is this energetic [la-], this affix that gets attached to a verb. Let's take the verb that means like 'a strong obligation', 'must' or 'should', something like that, and if you stick [la-] in front of it, it means that 'she must do something' so 'she must' right okay. There's that [ktb], by the way, that we saw earlier in morphology when we talked about infixation; I'm bringing up here again.

This is cool but where the fun really begins. This is a lot of information don't worry, you do not have to remember all of it, this is just to kind of give you some context to explain a little bit more.

- **Irrealis**: what funkiness happens in so many different languages, with respect to verbs, frequently involves irrealis
  - The one that probably most of you have heard of before is called the **subjunctive**. You've probably heard that before, especially if you learned an Indo-European language. Subjunctive is all over the place; everybody has a subjunctive. Technically, English still has a subjunctive, although it's dying very quickly. It went from about two generations ag, people regularly using a subjunctive, to almost no one does anymore unless you're speaking legal terms (that's a different story). Subjunctive is a hypothetical or unlikely events, along those lines.
  - You also can have in many languages an **optative**, if you ever go learn Ancient Greek or Sanskrit, you will learn about the optative. This is when you're talking about a hope or a wish.
  - A **desiderative** is also a wish, although there's not too many examples of those
  - A **dubitative** implies doubt.

In the Indo-European languages, pretty much all of those have been combined into what we now call this subjunctive; there are a couple of Indo-European languages that do have a dubitative, but not as often. To give you a couple of examples:

- Optative: I usually see this involved with a deity. You are praying to a deity that something happens so 'may you have a good trip', 'May God make it that you have a good trip' or whichever deity you want. For those that know Spanish and you ever wondered what ojalá [oxalá] is, ojalá is a borrowed piece from Arabic from when the Moors were in Spain and Portugal, and it is borrowed off of their optative. It just got co-opted into old Spanish and Portuguese, although it's died out now in Portuguese, but it did use to exist. Ojalá, Allah, as in God.
- Desiderative: This really only exists currently in Japanese; it used to exist in Sanskrit and then Proto-Indo-European and maybe a few other ancient languages. The desiderative is this concept of 'I really wish' or 'I really want something to happen'. To say, 'I desire to go there', 'I want to go there'. Watashi-wa asoko-ni ikitai; that -tai is the desiderative.
- Dubitative: this does exist in a few more languages; the only Indo-European language that has it is Bulgarian. It is thought that it got it from Turkish because in the Altaic languages, this does exist. What is really cool is that there are a number of Native American languages, mostly in northern North America, like Ojibwe, is one of them. We see the dubitative as implying doubt; in Ojibwe here's an example: 'he is in Baawitigon today' (I probably butchered that I apologize), and then instead of ayaa, which is 'to be' it's ayaadog, meaning, 'I guess he is in Baawitigong today' or 'he might be in Baawitigong today'.

That's the subjunctive and its counterparts, but there are other irrealis; imperative, hypothetical and inferential are the three main categories, and then they have subcategories off of that.

- **Imperative**: it's a direct command. Even in English, there's something that's a little different about a direct command.
  - The **jussive**, think justice, and frequently, this is a judge or some kind of ranking official imparting a sentence or something along those lines. When we talked about performative verbs, the jussive can come into play in certain languages. This

concept of you're imploring, like 'let him live', 'let her go', something along those lines. It does exist in quite a few languages.

- Within the **hypotheticals**, we're talking about things that are contrary to fact: 'if I had a million dollars, I would buy a house', that kind of hypothetical.
  - There's also the **potential**, which is the **conditional** in most Indo-European languages. Those are heavily tied together, and in fact most languages, if you have a conditional, you also have something hypothetical as well. The potential is also related to this.
- **Inferential**: this is a really interesting one, you see this in the Balkans and you see this in Turkish, but you don't see it too much beyond that. It's this concept of you did not witness something directly—that would be declarative—but you heard about it, you inferred information, you got it third party. That's a really interesting way to say that something happened from what you hear, not what you observed directly.

The big takeaway with respect to deixis is that it is highly contextual; that's why it's part of pragmatics. It is highly inflectional, that we use this to inflect all sorts of meaning in our verbs. Equally important, deixis is one of those things that as we learn a new language, it is incredibly difficult to learn. It is a nuance that is not something that is easy to pick up, especially if it is a type of deixis that we are not used to. Native English speakers, when you go to learn a language that has all sorts of inflection, that inflection is deixis. It explains why sometimes it's really difficult to learn what those little pieces mean. But more on that later.

How Many Verb Tense Are There in English?

Watch this TedEd video in which Anna Ananichuk explains a bit more about time deixis with respect to English verb forms. (The video is captioned.)



6.8: Deixis is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.9: References

As always, both *Introduction to Language* (Rodman, Fromkin, and Hyams) and *Language Files* (Ohio State University Press) have a long list of resources. But for this chapter, I consulted a wide variety.

To start off, the Wikipedia pages on grammatical mood and grammatical aspect have very good information, and the references they list are outstanding. They are regularly updated with the latest research in the field, and are great starting points.

The other texts that were consulted are considered 'handbooks' of semantics and pragmatics--if you take an upper division or graduate course in semantics or pragmatics, you will often have these are course materials and/or reference materials:

- Allwood, Jess, Lars-Gunnar Andersson, and Osten Dahl. 2012. *Logic in Lingusitics*. Cambridge University Press;
- Comrie, Bernard. 1995. *Aspect*. Cambridge University Press;
- Comrie, Bernard. 2012. *Tense*. Cambridge University Press;
- Levinson, Stephen C. 2014. *Pragmatics*. Cambridge University Press; and,
- Palmer, F.P. 2012. *Mood and Modality*. Cambridge University Press.

6.9: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

7: Language and Society- Sociolinguistics

1: Sociolinguistics definitions
2: Geographic dialects
3: Appalachian American English (AE)
4: African American English (AAE)
5: Other Social Dialects
6: Gender and Class Dialects
7: Prestige and Politics
8: Social Norms
9: Languages in Contact
10: Indigenous North American Languages (Optional)
11: References

7: Language and Society- Sociolinguistics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1: Sociolinguistics definitions

To start things off, watch this video, a clip from Trever Noah's "African American" comedy special. He is talking about his experiences in learning about life in America.



Sociolinguistics Definitions, from Sarah Harmon



Video Script

You have learned about phonetics, phonology, morphology, syntax, semantics, and pragmatics; those are the core areas with respect to linguistics. From here on in the textbook and in the course, we will take those tools and apply them to a variety of situations. For the remainder of the course and the remainder of the book, we're going to look at how language varies across social dimensions, across historical dimensions, and across psychological dimensions. First up is sociolinguistics, and then we'll talk historical linguistics, and then we'll talk, psychological or cognitive linguistics or neuro-linguistics. Those will have three separate sections; actually, the psychological aspect is split into two because we have to talk about how we acquire language and how we process language, and those are separate issues. To start off, let's talk sociolinguistics and let's set the stage a bit.

Up to now you've heard me use a couple of terms. Let's dig into that a little bit more. One has to do with a **dialect**, and the other way we think it as a **speech community**. When we talk about a speech community, it's exactly what you think it is: a community of people who speak the same way. They have the same dialect, if you will. It may be a mainstream dialect, or it could be something totally different; it depends. Later on, we're also going to talk about diglossia, which is what happens when you have two or more dialects that you use on a regular basis. That usually refers to somebody who is a part of more than one speech community, and that's what we're referring to here.

If you recall from the first chapter, we said that the difference between a language and a dialect was that intelligibility was the key. Not intelligence, but intelligibility, meaning if two people from different speech communities are able to generally understand one another, they're considered to be mutually intelligent dialects. However, if you have two people from two different speech communities, and there is a real difficulty for one to understand the other linguistically—not culturally, not politically, nothing else, just linguistically—then that means that they speak two languages.

In thinking about this, let's talk about some generalized and incorrect views about dialects. When I say incorrect or wrong, I mean from a linguistic standpoint.

- Many people think that an accent is a dialect and that's actually incorrect. Even though the next subsection for this page is from Catherine Anderson and she talks about an accent; watch what she says, I'll let you read it later.
- A lot of people think that some people speak a language and others speak a dialect of that language. That is quite incorrect.
- The third generalization that we wish to debunk is that some people think that dialects are impoverished variants of a language, that they're just nonstandard and perhaps not correct.
- The fourth big generalization that is incorrect is that people might think that languages are defined by political boundaries, as in geopolitical boundaries.

Every single one of those statements and generalizations is 1,000% incorrect, at least from a linguistic standpoint. They are all prescriptive in nature and remember in linguistics we do not want to be prescriptive. We want to describe; we want to describe what people actually do when they communicate with one another, both within their group and with others.

Let's take a modern linguistic twist on all of this:

- An accent is not the same thing as a dialect because ‘accent’ is really not a term that we use in linguistics. Frequently, when people talk about accents, they're really saying something like the sounds maybe a couple of words are different, which is absolutely not the case. When we talk about different dialects, of course there may be some differences in pronunciation and phonemes or allophones, and there frequently are different terms used for the same thing. But there's also frequently morphological and syntactic differences, pragmatic differences, and so many other things. Later on in this chapter, we will see examples of different dialects so you can understand what we're talking about.
- The fact that one person speaks a language but a different person might be a dialect is incorrect. Everyone speaks at least one dialect of a language and, in many cases, you might actually speak more than one dialect.
- We cannot determine whether certain dialects are superior to others because that's not what a linguist does. A linguist describes what they see; we observe and we document, but we never give a value of higher or lower to anyone style of speech. Sometimes we know that a certain dialect might be more prestigious than others, but that is an objective point of view.
- Languages are defined by cultural or linguistic boundaries, but not geopolitical boundaries. Sometimes they happen to coincide, but what we care about is what speech communities do and so those frequently line up with linguistic and frequently cultural boundaries. To give you an example, Norwegian and Swedish, and technically Danish, are three dialects of the same language. Even though those are three different nationalities, linguistically they're actually the same—although Swedish and Norwegian have a lot more in common and Danish is a little bit different, it's a definite distinct dialect. Another example is English; think about the number of countries where English is the primary and official language—it is not just England, and it's not even just the United Kingdom. It is a slew of countries. Spanish would be another example, and Mandarin would be another example.

When we talk about languages, we're talking about speech communities who use the language; that is what we're talking about.

One more really important thing to bring up about dialects versus languages. **Prestige** is an issue; it is going to be a factor and we'll talk more about this in its own chapter in this in this larger topic. There are always going to be certain dialects that will carry more prestige. Think about if you're in England or the United Kingdom, it's Queen's English that is going to be going to be the most prestigious; Received Pronunciation is often what it's called or RP. Frequently, and certainly since the 20th century, we can talk about the media, and especially national media and international media when we talk about the style of speech. Then it would be considered a **mainstream dialect**. In the past that was often called the **standard dialect**; we have switched topics nomenclature,

and we now call those mainstream dialects to reflect the fact that they are mainstream and used by a wide variety of speech groups, at least as a more prestigious dialect. It is also important to understand that prestige on a dialect is usually something that has carried over a number of generations, it can be modified.

It should be noted that just because you speak that mainstream or prestige dialect, it does not mean you necessarily hold more power than anyone else; it just means that you speak that dialect. It also means that if you speak a mainstream dialect, you are often perceived as having ‘no accent’ and that your manner of speech is the one that most everybody understands.

To repeat a point made earlier, dialects are mutually intelligent variance of a language. They can and do differentiate a person who's born and raised in California from somebody who's born and raised in Toronto, Canada; New York; London; Edinburgh, Scotland; Dublin, Ireland; Johannesburg, South Africa; Sydney Australia; Auckland New Zealand; you get the point. For each of these are all different dialects, there are systematic ways that they are similar, and yet systematic ways that they are also different. To repeat what we said in the first chapter, it is really crucial to understand that when we talk about these dialects, we take the geopolitical out of it. That means that there is no one ‘Chinese language’; there are about 210-ish Chinese languages, and not all of them are part of the Sino-Tibetan family; most of them, but not all of them. The fact that Serbian and Croatian are technically to dialects of the same language, Serbo-Croatian, but that their closeness or distinctness can vary between generations and that entirely depends on the relations between those two cultures. As we said, Norwegian and Swedish and, technically, Danish are all dialects of the same language; it is now considered Northern Germanic or Scandinavian. There are different dialects of it.

When you can't understand when somebody is talking to you—not because you choose not to, but because you actually cannot understand the words coming out of their mouths—that's when you have a different language scenario. If it's close, it's probably a dialect.

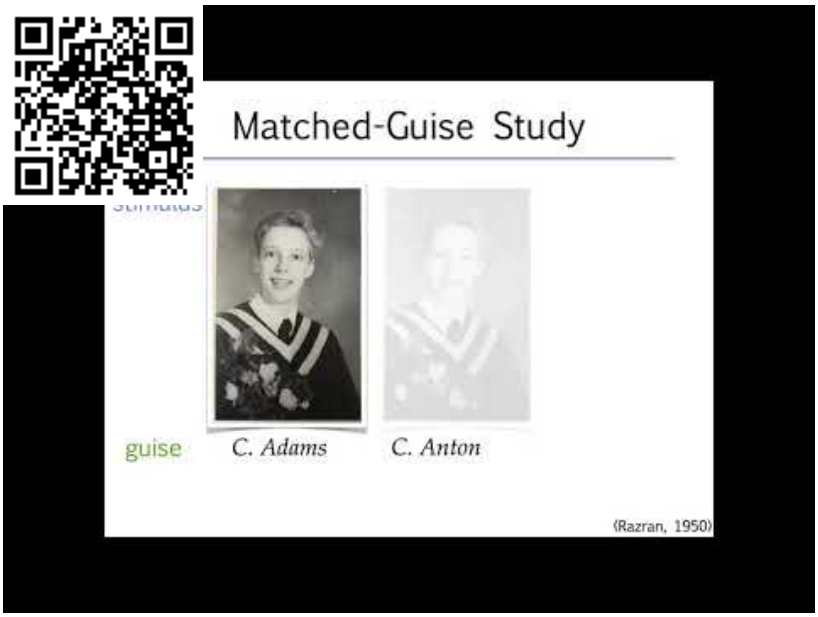
How exactly do dialects arise? The simple answer is there's some kind of **isolation** to speech communities. There was originally one, and there somehow became a break or situation that created isolation; that is how you get to dialects. It doesn’t happen instantaneously, rather it usually takes several generations. We can have a variety of ways to separate or isolate as a given speech community.

Certainly, geographical dialects are the first thing that most people think about with respect to dialectology. We're going to take a look at Appalachian English (AE) in a different section of this chapter, because I want to give proper respect to it. It's not just ‘hillbilly English’; it's not just some 'poor English', ‘mountain English’, whatever you want to call it. In linguistic terms, it's a very rich dialect and actually gives us a lot of insight as to how English came to be the way it is now.

We also have social dialects, gender dialects, and class-based dialects. We'll look at a few will look at African American English Vernacular (AAE). Certainly, the isolation there has to do with the institution of slavery and the segregation of anybody who was in part or in whole of African descent. We also will briefly look at a couple of different social dialects that have to do with gender and sexuality. We'll look a little bit into class-based dialects, and then finally we'll talk about scientific or jargon isolation. If you think about it, that makes sense. If you work in a particular industry, you tend to have a certain manner of speaking with your colleagues and with others in your industry—that's also a type of isolation and that can create a dialect.

As we go through this concept of sociolinguistics, remember what is a society is—what we're at the heart of and how they use language to differentiate themselves.

Attitudes about Accents, in Anderson's *Essentials of Linguistics*



Video Script

Let's talk about accents. Why do non-native speakers have an accent? Well, actually, everybody has an accent. It's just that if someone's accent is pretty similar to your own accent, you don't really notice it. You only notice accents that are different from your own accent.

So a better way of asking this question would be why do L2 speakers have accents that are different from L1 speakers? We saw in the last unit that the mental grammar of an L2 speaker is influenced by their experience of their native language, their L1. So the accent of an L1 Mandarin speaker in English is going to be different from the accent of an L1 Dutch speaker.

Now the thing about having an accent that is noticeably different is that people will notice it because it's different. When I moved to Chicago in 1998 after having lived in Ontario for 25 years, people said to me, "You sound weird. Are you Canadian?" My vowels were different from Chicago vowels, and people in Chicago noticed that difference. I didn't really experience any negative consequences of sounding like a Canadian while living in the US, but if you have an accent that's different from the people you spend time with, you might have experienced **stigma**. If an accent is stigmatized that doesn't mean it's bad or inferior in some way — remember that linguistics doesn't rate or rank languages or accents. But if it's stigmatized, that means people have negative attitudes and expectations about that accent. In places where the majority of people speak English, there's often a stigma towards people who aren't native speakers, who learned English as adults. But there are also some varieties of L1 English whose speakers experience stigma, such as African-American English, the varieties spoken in the southeastern United States, and in Canada, Newfoundland English.

For people whose accent is different from the mainstream, there can be many negative consequences. You're less likely to get a job interview, and your boss might not recognize your skills. It's harder to find a landlord who's willing to rent you an apartment. If you have to go to court, what you say won't be taken as seriously, and the court reporter is likelier to make mistakes in transcribing your testimony. Kids whose accents aren't mainstream are disproportionately labelled with learning disabilities and streamed out of academic classrooms into special ed. And probably Alexa, Siri, and Google won't understand your requests!

Why do these things happen? Well, in the case of Alexa, it's because the training data doesn't include enough variation in dialects and accents. But the rest of these situations arise from people's expectations, and their expectations come from their experiences and their attitudes. Now, for issues of stigma, it's hard to observe people's attitudes directly, because by and large it's not socially acceptable to express negative attitudes towards minority groups. So instead, researchers use a technique called a **matched-guise** study to try to draw conclusions about attitudes.

A matched-guise study works like this. The researchers present participants with some kind of stimulus. In the original 1950 experiment using this technique, the stimulus was yearbook pictures from a local university. They hold the stimulus constant, and change the guise that it appears in. So in 1950, the guise was the name that labelled the yearbook picture. One group saw the pictures with so-called American names, and another group saw the pictures with Italian or Irish names. Then the researchers asked their participants to rate the people in the pictures as to their Beauty, Intelligence, Ambition, and Entertainingness.

The core idea in a matched-guise study is that if you find a difference in your participants' ratings, that difference can't be because of the stimulus, because you've held the stimulus constant. Any difference in ratings must be because of the guise — the way you labelled your stimuli. I'll leave you to guess how the ratings in that 1950 study differed with the different guises.

Molly Babel and Jamie Russell, two linguists at the University of British Columbia, conducted a matched-guise study with UBC students as listeners. They recorded the voices of several people who were native speakers of English, who had grown up in Canada. These recordings were the stimulus. Then when they played these recordings to the listeners, they presented them either as audio-only, with a picture of the face of a White Canadian person, or with a picture of a Chinese Canadian person. For any given voice, the listeners rated the talker as having a stronger accent when they saw a Chinese Canadian face than when they saw a White Canadian face, and they were also less accurate at writing down the sentences the talker said. Apparently the faces influenced how well the listeners understood the talkers.

Dr. Babel interprets their results as a mismatch of expectations. In Richmond, BC, where they conducted their study, more than 40% of the population speaks either Cantonese or Mandarin. If you live in Richmond, you have a greater chance of encountering L1 Chinese speakers in your daily life than L1 English speakers. So when you see a face that appears Chinese, you have an expectation, based on your daily experience, that that person's accent is going to be Chinese. If the person's accent turns out to be that of a native speaker of English, the mismatch with your expectations makes it harder to understand what they say.

So we've seen that people's expectations, their experiences and their attitudes can lead to stigma for speakers with accents that are different from the mainstream. And that stigma can have serious, real-life consequences on people's employment and housing and education. But there can be consequences for listeners too!

If you're having a hard time understanding someone whose accent is different from yours, that could have serious consequences, for example if you're getting medical advice or trying to learn something new. It's pretty common for L1 English speakers to argue that L2 speakers should try to "reduce" their accents, but as linguists we know that that's hard to do after childhood, because your L2 grammar is shaped by your L1 experience. Fortunately, linguistics research also tells us that even though it's hard to change how you speak an L2, it's relatively easy to change how you hear someone speaking an L2.

Just as our experience and our expectations can lead to stigma, our experience also influences our perception. The more experience we have listening to someone, the better we understand what they say: this is called **perceptual adaptation**. Perceptual adaptation was first shown for a single talker: the longer listeners had to listen to an unfamiliar talker, the more they understood of what the talker said. Extensions of that research have also shown that experience listening to several speakers with a particular accent makes it easier to understand a new speaker with that same accent. And it turns out that listening to a whole variety of different unfamiliar accents then makes it easier to understand a new talker with a completely different accent. In short, the more experience we have listening to someone, the more familiarity we have with their voice or their accent, and the more familiarity we have, the better we'll understand what they're saying.

So if you are listening to someone whose accent is different from yours the best way to understand them is to listen more. And if you're talking to someone and they're finding your accent unfamiliar, you can say to them, just listen more!

1: Sociolinguistics definitions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2: Geographic dialects

Various Accents of English, in Anderson's *Essentials of Linguistics*

Adapted from

www.oercommons.org/courses/ho...nguistics/view

© 2006. Indiana University and Michael Gasser.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation.

URL: www.indiana.edu/~hlw/PhonProcess/accents.html

Edition 3.0; 2006-12-31

In this section, we look at various English accents and how they differ from one another. Remember that an **accent** is the set of pronunciation conventions of some speech community. Where we draw the boundaries between accents is pretty arbitrary; if we call General American a single accent, for example, we'll have to deal with the range of variation that exists among speakers within that large community. And any boundaries we draw will be wrong in another sense because the group of people who have one pronunciation convention may not coincide neatly with the group of people who have the other set of conventions that belong to the accent we're considering. For example, the group of speakers who pronounce the words *pin* and *pen* the same includes speakers of Southern US accent but also some speakers of General American, which is a very different accent from Southern US English in many other ways. The point is that conventions of pronunciation tend to cluster together; this is what allows us to talk about "accents" at all.

Another point to keep in mind is that in most countries there is a standard, prestige accent alongside a number of accents associated with particular regions, social classes, or ethnic groups. Each of these non-standard accents can be described in its "broad" form, the form that is most different from the standard in the country where it is spoken, but what many people are speaking much of the time is something in between a particular non-standard accent and the relevant standard. In this section, we concentrate mostly on broad variants of non-standard accents because they illustrate the range of possible differences best.

When comparing two dialects or accents, one possibility is to see one of them as deviating from the other. A biased view of non-standard dialects often starts this way: the speakers of these dialects are seen as just making mistakes with the standard when what they say is non-standard. But of course, this is not what is actually happening. Speakers of non-standard dialects learned the conventions of these dialects by hearing other speakers speak them, just as the speakers of standard dialects learned the conventions of their dialects. They are no more speaking the standard wrong than the speakers of the standard dialect are speaking their dialect wrong.

Overview of English Accents

Before looking at examples of differences between accents, it might help to have a sense of what the major accents are and where they're spoken.

The British Isles

There is no "British" accent. England, Scotland, Ireland, and possibly Wales all have their own unofficial standard accents, and the standards of Scotland and Ireland, in particular, are as different from that of England as American accents are. The standard, or prestige, accent of England is usually referred to as Received Pronunciation (RP). This is what the royal family, all recent Prime Ministers, and most BBC announcers speak. It is probably what most Americans think of as an "English" accent, though it is spoken as a native accent by no more than about 10% of the English population. It differs most noticeably from General American in the pronunciation of a few vowels and in the way [ɹ] is treated following vowels. For example, in RP there would be no [ɹ] sounds at all in the phrase *the northern fourth of the park*.

Within England, there are many identifiable regional accents, probably more than in the United States in fact. Among these, London accent (sometimes called "Cockney") stands out because it is familiar to many Americans through film and drama characters such as Eliza Dolittle in *Pygmalion/My Fair Lady* and because it has a number of very characteristic features. Many of the vowels in this accent differ considerably from RP and General American. Other very striking features are the loss of initial [h] ("*'e 'as an 'ard 'eart" = "he has a hard heart*") and the frequent glottal stops in place of other stops in other accents ("*iʔ'll taʔe a loʔ o' time to seʔle" = "it'll take a lot of time to settle"*). Perhaps the other major accent boundary in England separates the

accents of the north from those of the south. Americans may be familiar with the English of Northern England through the speech of the Beatles or the characters in films such *The Full Monty*. These accents can be identified fairly easily because they make no distinction between the vowels [ʌ] and [ʊ]; both are pronounced like [ʊ], so that the words *look* and *luck* are homophones.

Scottish and Irish English share one feature with northern England English; the tense vowels [i], [u], [e] and [o] are not pronounced as diphthongs, as they are in RP and General American. In addition, these accents are like General American, and unlike most accents of England, in how they treat [ɹ] after vowels.

The Western Hemisphere

The unofficial standard accent of the United States is usually called General American (GA) or Mainstream US English (MUSE). This is the accent of much of the Midwest and the West and the most frequent accent for US newscasters, though, interestingly, many of the more recent US Presidents have spoken regional varieties rather than GA. As the prestige accent, it has been encroaching on some regional accents, for example, in the northeast, but at the same time, changes within GA are creating what amount to new accents. One striking example of this is Northern Cities accent, spoken in cities such as Chicago, Detroit, Cleveland, and Rochester, and distinct from GA in the pronunciation of lax vowels. So for example, the word *socks* in the name of the Chicago White Sox is pronounced [saks] in the Northern Cities accent, as compared to [sɑks] in Canadian English.

The Southern US accent is spoken by people mainly in the southeastern part of the country. Like the London accent, this accent has strikingly different vowels from other English accents. African-American Vernacular English (AAVE) is a dialect associated with an ethnic group rather than a region, though of course, you don't have to be African-American to have learned it. The accent associated with this dialect is similar in many ways to Southern US accent, while the phonology, morphology and syntax of this variety have their own characteristic properties.

People from the northeastern US are often easy to identify by their accents; the accent of New York City stands out within this region, again mostly for its vowels. Some other US cities, especially Pittsburgh, are known for particular pronunciation conventions. In Pittsburgh, for example, [a] may be used where GA has [aʊ], so *downtown* may be [dantan].

Standard Canadian English (except in the province of Newfoundland) is very similar to General American, and it doesn't vary much from place to place. One characteristic of Canadian English is the pronunciation of [aɪ] and [aʊ] in certain contexts, which we'll learn about in Section 4.6.

English is the native language of much of the Caribbean, with some features common to the region and others specific to particular islands. As with other accents, there are characteristic vowels in these accents, and in addition, a tendency in the Caribbean, as there is in some US accents, to make no distinction between [t] and [θ] or between [d] and [ð]. Jamaican English in particular also has quite striking intonation patterns.

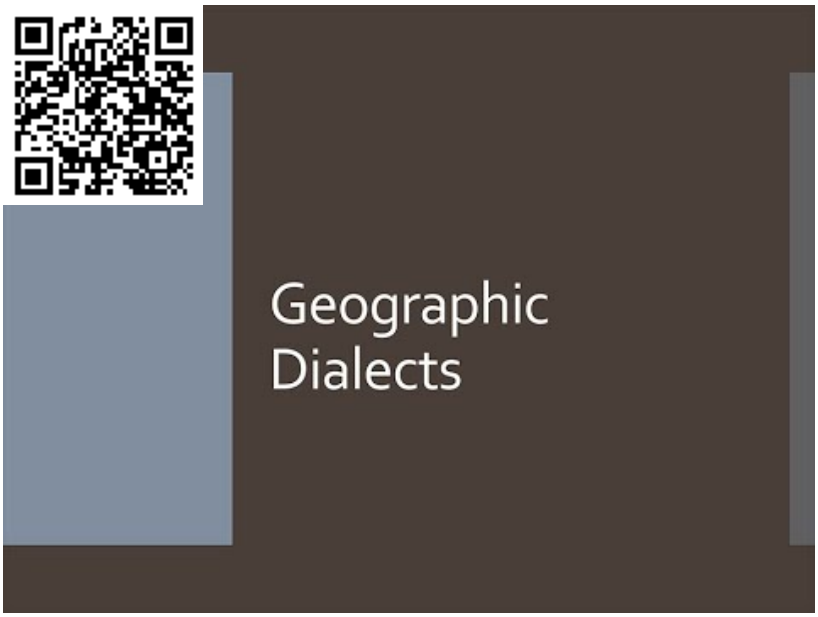
The Southern Hemisphere

English is the native language of most Australians and New Zealanders and a sizable minority of South Africans. While the standard English accents of these countries tend to approach RP, the broad accents of most English speakers in all three countries have tense vowels similar to those in the London accent. The lax front vowels of Australian and New Zealand English differ from those in other accents.

Non-native accents

English is spoken as a second language by millions of people, especially in regions that were once colonized by Britain in South Asia and Africa. In some of these regions, there are particular English pronunciation conventions that derive from the phonology of the local languages. For example, in the English of South Asians (Indians, Pakistanis, Bangladeshis, Sri Lankans, Nepalese, Bhutanese, and Maldivians), the alveolar consonants [t], [d], [n], and [l] tend to be replaced by retroflex consonants, which are common in the languages of this region. These non-native conventions are one of the ways that English is becoming even more of an international language.

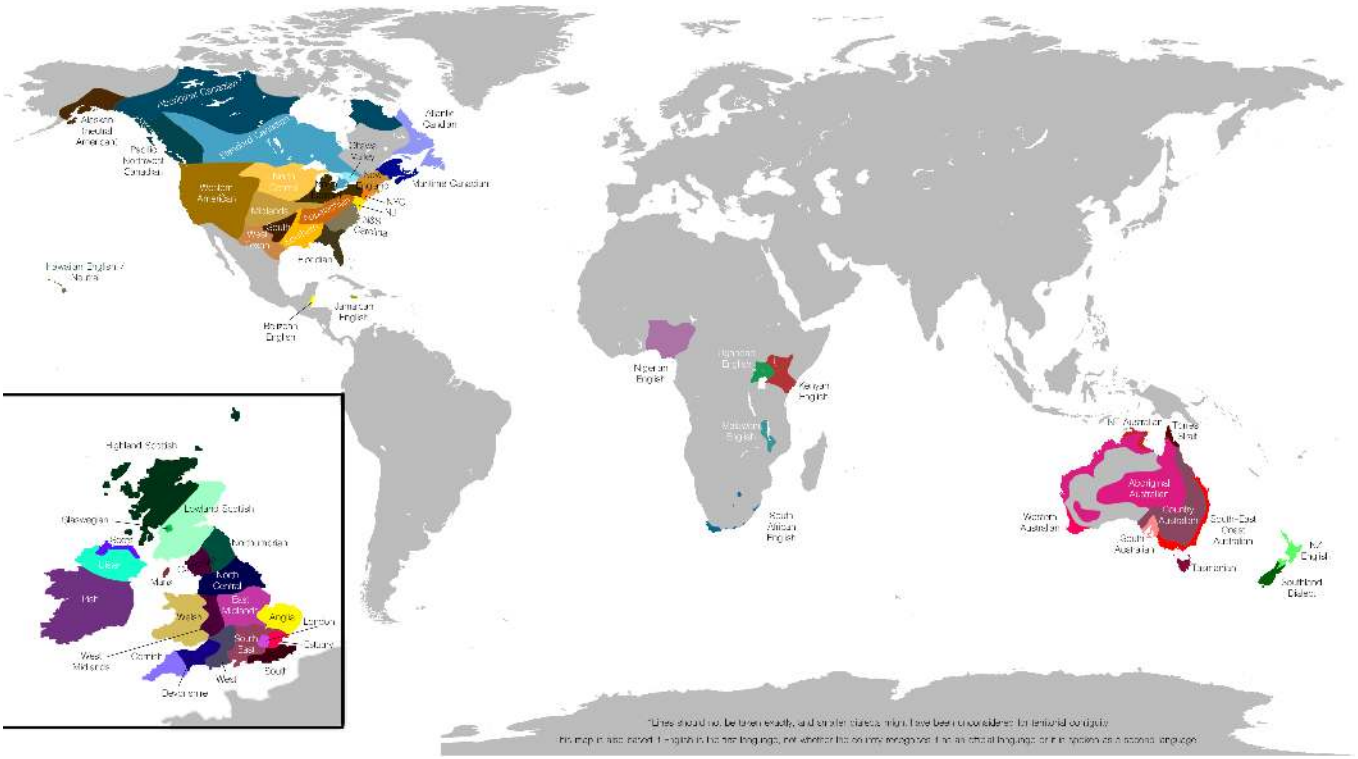
Geographic Dialects, from Sarah Harmon



Video Script

Catherine's explanation of a geographic dialect, I think, needs a little more fleshing out. Certainly, this is the case when we talk about English as a global language. I'm going to expand a little bit more than what she has in her section on geographic dialects, and instead focus a little bit more on our next few topics.

When we talk about geographic dialects, we're talking about the dialects that arise because of geographic isolation. Take a look at this map of global English.



Notice that you have two full continents where English was brought and dominated the entirety of the continent. When we talk about North America, and when we talk about Australia (including New Zealand), English has dominated those continents. We have to add Africa because, in certain parts, like over in eastern Africa where Kenya is, for the most part we have English is an

official language. In the area around Nigeria, of course, English is an official language, and then, of course, South Africa, especially the cities on the coast, English is pretty much a mainstay.

Those are just the colonies, to say nothing of the United Kingdom and Ireland, where, of course, we have had English for close to 2,000 years. Notice that just this one little island—and realistically, you can talk about southeast England where London is—for the most part, it is from there, and the policies that come out of that city, that you get English spreading, not just through the British Isles, but throughout the world, as part of the original British Empire.

Now, lest you think that that is the only thing to say about that, just think about North America alone, the United States and Canada. Think about how many different dialects that we have, not to mention Mainstream American English (MAE) and Mainstream Canadian English (MCE). Just looking at that section of North America is pretty stunning. It tells us that we have a lot of different geographic dialects within North America. We're going to take a look at some of that in a little bit more detail.

As you have no doubt guessed, this is not just an English-centric course; it's important to think about, for example, a language like Spanish, with the same kind of colonization of history. As a result of very wide spread of different geographic dialects from that one country of Spain, we can count two main areas, the provinces of Extremadura and Andalucía, as where the majority of Spaniards left to colonize the western hemisphere. Everything from Mexico down to the very tip of South America, all of these areas colonized mostly by people from the southern half and the eastern half of Spain. We have the Spanish-speaking islands of the Caribbean as well. Notice the different dialects that we have.

When we talk about a global language, English is not the only one to keep in mind. Spanish is definitely in there and, certainly, you can count Mandarin. It is part of that equation when you factor in the number of places throughout Asia and beyond where Mandarin is a major player linguistically. When we think about the isolation that happens as a result of colonization, frequently it's not just because you have left the homeland, the old country. Rather, it’s because you're moving to a place that may be cut off geographically or, certainly for several generations, you might be cut off from the others in the region. Eventually there's more spreading of the dialect.

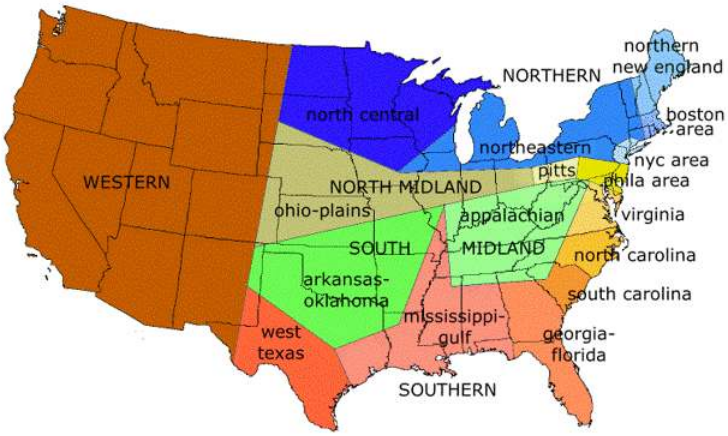
It is also important to understand that when we talk about a global language, there are still considered mainstream tendencies, and Spanish is a really good example of this.



There are differences within each one of these dialects, each one of these colors that you see. Latin American Spanish has certain characteristics that Spain (Castilian) Spanish, including the Canary Islands, do not share. The reverse is true, too. We aren't just talking about different sounds; we're talking about parts of the verb conjugation, and the structuring of syllables is entirely different depending on what dialect you speak.

One important note, if you didn't know this about the history of Spanish is that for about the last 150 years, the prestige Latin American dialect is not Mexican; it's Colombian. By Latin Americans themselves, they frequently state that Colombian Spanish—particularly up in the Andean region near Bogotá, Medellín, and that area—has the clearest Spanish that anyone has spoken. Linguistically, we don't hold that prestige; we can mark it, we can observe it, but we don't hold it. I find it interesting that even most Spanish speakers in the United States don't realize that Mexican Spanish is not carry quite the amount of prestige just certain others. However, I will say in the last 25 to 30 years it has changed, due to the power of the Mexican government and the Mexican economy, that they have held a little bit more prestige in the last 25ish years. Where we'll go, who knows.

Let's focus on the United States because I find this to be also an interesting tale. This is a geographic map of the dialects of North American English with specific reference to the lower 48 states of the United States of America.



I phrase it that way because of a few things. If you think about Alaska, which is way well probably off screen, and if you think about Hawaii, which is way probably off screen as well, those are two very isolated places, so we don't count them when we talk about Mainstream American English. We're not usually going to count Alaskan or Hawaiian English; we're going to talk about the lower 48. If you remember your American history, and if you think about the original 13 colonies, you see that is where the most fractured area with respect to linguistic data we have. It makes some sense; the established colonies had some isolation and then, as we expanded west, we grew more general and more general.

Appalachian English is something that we will focus on in this class to bring just some understanding of it. It is a full, completely rich and historic dialect of American English; in fact, it really connects us to our English ancestors.

I will also tell you that here in the Bay Area, we have our own dialect. It's actually a subdialect. As somebody who's not only born and raised in San Mateo California—notice I said [sæm:ətejə]. I didn't say [sæn məteʲoʷ] I said [sæm:ətejə]. I live, not in [sæn hoʷzeʲ] I live in [sænəoʷzeʲ]. My dad, he was not born in Sacramento [sækɹəmɛntoʷ] he was born in [sækəmɛnə]. This is San Francisco Bay Area English; there's just so much that I can say about this. I'll leave it for another time; perhaps I should have said there's hella things that I can say about this dialect. It obviously has a bit of a piece of my heart. It is true that those of us were born and raised here, and especially if our parents were also born and raised here—in my case, it was my mother—we're going to have a distinct dialect of our own.

Every place where there is some kind of isolation, there will be a dialect that springs up. Whether it stays or whether it disappears, that's always the question. For now, let's leave it at this: we're going to start looking at different dialects about what they do. We look at these dialects, but we do not pass judgment; we do not say they are inferior. In fact, they are on a level playing field; everybody deserves the same amount of respect so first up is Appalachian English.

2: Geographic dialects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.14: Various Accents of English** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

3: Appalachian American English (AE)

Appalachian English (AE), from Sarah Harmon



Video Script

Let's start talking about one of these geographic dialects that, I have to say, holds a little bit piece of my heart. Appalachian English is the English that is spoken in and around the Appalachian Mountains; think Tennessee, Kentucky, bits of the Carolinas, the southern and eastern part of Pennsylvania, and that area. It holds a place of my heart from my family but indirectly; I have three aunts and all three of the married folks from that part of the world. I remember hearing them and realizing that the way they spoke was really beautiful and, even though they downplayed how they spoke—in fact they often said, “Don't sound like me; sound like you. You're a Californian and you sound better than I do”—that basically started me down this road of trying to understand a little bit more about Appalachian English.

Despite what you may think, it is a distinct dialect from what you speak if you go into the Delta or the Gulf area, think of Alabama, Mississippi and that area. It is different than in the Ozarks; think Southwest Missouri, Arkansas, northeast Texas, and that area. It is different than what you speak in the Plains region; think Kansas, Nebraska, Iowa, and that area. It is different also from Midwest English; think Indiana, Ohio, and even parts of Illinois and Michigan.

Appalachian English is truly unique, and it is also really well studied; in fact, it is one of the best studied dialects in the United States. I don't just mean recently; I mean historically. We have about 100 years of data and analysis on this dialect alone. That's why I want to focus on it, and hopefully either you are interested in it, or it makes you interested in a different dialect for the same reasons that I love it.

I love starting off with this little graphic.



There have been many versions of this in Pittsburgh; they're very proud of their cultural background, and of course it's very tied to the Appalachians. In many ways, Pittsburgh is the northern tip of where the Appalachian dialect starts, and it pretty much goes south and east from there. They are so proud of it; they have all sorts of souvenirs with Pittsburghese graphics. What it reminds me of mostly is how unique Appalachian English is, so let's take a closer look at it.

Let's start with the phonology and build up. When we talk about Appalachian English, the big difference with the phonetics is the vowels. This is a theme you're going to hear with respect to English: English vowels are very different across the different dialects. The consonants don't change too much, although we will see some changes, but most especially it's the vowels that suffer or change from one dialect to another. In the case of Appalachian English, we have a modification of all vowels—not just front vowels or back vowels, not just high vowels are low vowels, but all vowels. A lot of times it has to do with going from tense to lax—if you think about the [i] in *pinch* and *pinch* it's not an [i], but an [ɪ]. In Appalachian English, it's going to go high, it's going to tense up: [pʰiɲtʃ]. Instead of [tʰɛn], like the number *ten*, it's going to be [tʰin]. Think [θiŋk] becomes [θæŋk]; rather [ɹæðɻ] becomes [ɹʊðɻ], and push [pʊʃ] becomes [puʃ]. You modify the vows, you making them more tense, most of the time, but you're also moving them higher or lower more front or more back, and it just kind of gets jumbled up.

Before you think that this is just random, know this is actually historical. Appalachian English, in many ways, has **fossilized aspects** of the language that was brought over by the colonizers of that area of that time. Most of those colonizers came from a couple different spots within England, and so they neutralize those two dialects, and then it gets fossilized because they're isolated within the Appalachian region. In many ways, what we see here with the vowels represents a lot of the sounds that we heard; if we have that magic time machine and we could travel back to when English settlers first came to the Appalachian Mountains.

We also have **metathesis**. If you recall when we talked about that, I said that the verb *ask* actually started off life as [æks] or [æksian]. We see it here in Appalachian English; it's kept, but it is not the only area that we see metastasis we actually see it in a lot of cases when we have r's and l's. For linguists, that's not uncommon; r's and l's, or approximates if you will do funny things in all languages and even across dialects. In fact, whatever your native language is, whether it's English or something else, I guarantee you the approximates have changed over time and they change from one dialect to another. Here, instead of [æskt] it's [ækst]. Instead of [pɹivel], it's [pʌɹvel]. Instead of [ælbəm], it's [æbləm].

We also have **syllable initial stress** in Appalachian English. If you think about Standard or Mainstream American English, you realize that sometimes it's the first syllable that is stressed, and other times it's not. But with Appalachian English, there's a real pull to make sure that that the primary stress syllable. You say these terms as: *Détroit, cigar, directly, Nóvember*.

For some of you, that just might sound like 'Hillbilly talk', or like I'm making fun of them. However, this is truly Appalachian English as its core. All of these three phonological processes are not unique to Appalachian English, and we certainly hear them in other dialects in the region, African American English Vernacular as one example. But these are characteristics that have a long history of documentation, at least with respect to American dialectology.

Let's move from phonology to morphology. When we talk about morphology, it's not just the pieces coming together; it's actually the morphology and syntax coming together. We have an ***a-* prefix**; this is something that happens in a few dialects of English and it's actually a preservation of something we saw in old English. When I was living in Texas, I certainly heard it down there. *He came a-running. I knew he was a-telling the truth. I was a-washing one day*. These are way to mark that something is in progress and frequently has an emphatic element to it.

A number of the irregular past tenses that are frequently connected to **umlaut**—that change of vowel that signifies that there's an inflection—remain in Appalachian English. Many of these irregular past tense forms have been changed over in Mainstream American English and many other dialects of English, to have the common *-ed* inflection. However, in Appalachian English there are fewer that have; quite a few that have remained: *climbed* for them is still *clumb*, *heated* is still *hit*, *raked* is still *ruck*, and *dragged* is still *drug*.

We also have a series of **double modal verbs**, meaning you use two different models—a type of auxiliary verb like *must*, *do* or *did*, *can* or *could*, *will* or *would*—and they can have multiple roles. Frequently this is connected to something that we call aspect, what we talked about when we talked about deixis. We're connecting two different types of aspect and again there's some history. In this case, it is connected to some things that we used to do in Old English that in a number of the dialects in more remote areas of England, they were still doing into the early 20th century.

Multiple negation is also something that we frequently do in Appalachian English. Realistically, when we talk about multiple negation, we need to remind ourselves that Mainstream American English and higher, more prestigious dialects of English have poo-pooed the idea of multiple negation. The reality is that it exists in many dialects, not just of English, but throughout the world. Certainly, if you have ever learned a Romance language or a Celtic language or many other Indo-European languages, you know that multiple negation is a common thing; it's actually a very common thing in the world languages.

Overarchingly, when we talk about Appalachian English, there are these are the seven key areas that we need to focus on. Combined along with the lexical differences, they helped to create a different dialect. In an upcoming journal assignment, you will be listening to a little bit of Appalachian English. At first it will be difficult to understand, but realize that it is English, and when they slow down, you can understand them. There might be differences and certainly it does not sound like what we say here in California, or even in mainstream English. But it is English, it is an American dialect of English, and we can understand it pretty well.

3: Appalachian American English (AE) is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4: African American English (AAE)

African American English Vernacular (AAE), from Sarah Harmon



Video Script

The next dialect that we're going to work on is African American English Vernacular, or AAE. Again, I have a bit of a story to tell them this one.

I grew up in San Mateo, California, and I was in high school in the early 90s when the topic of Ebonics came to the main stage. It was a really interesting time to be fascinated with language. At the time, I was learning Spanish, and I was learning Italian. Specifically, I was trying to understand the Italian that my family spoke compared to what I was learning in the classroom. I was listening to all of my friends speak all of these different languages: Spanish, French, Samoan, Tongan, Tagalog, Ilokano, Mandarin, Cantonese, Shanghainese, Japanese, and the list goes on and on.

When the Ebonics debate started, it fascinated me. It came out of Berkeley, and it was a revolutionary way to address the kids who were living specifically in Oakland, Berkeley and Richmond, also the greater Bay Area and the country at large, with respect to inner city African American youth. The struggles, specifically, that they were having with writing and reading in English. This perspective, to say, “Look, they don't just speak poor crappy English. They speak a different dialect of English, and it is a radically different dialect.” You have to treat it as such; you have to address it, and say, “Okay, that is really great for certain times, but if you want to be a ‘proficient English speaker reader and writer’ (massive air quotes, right?), then you need to be using Mainstream American English, so we need to switch your dialect. Not put your home dialect down, rather to raise it up. But understand that there's this other dialect that you need to use within society in different situations.” To take a formerly prescriptive idea and make it more descriptive, that was the power of the Ebonics movement. It absolutely fascinated me.

Growing up in San Mateo—even when I grew up in San Mateo, which is very different than what it is now—there were not too many folks there that were African American and spoke Ebonics, at least not regularly and out in the open. However, there were plenty of other languages that I was hearing on a daily basis, both within my family, in my community, and in my schools. What was really interesting was when I would go to East Palo Alto for something having to do with work, or I went to a Giants game in Candlestick Park—which is in Hunters Point—or if I visited my family out in Fairfield and other parts of Solano County—where there were (and still are) strong pockets of African American culture and African American English vernacular. It was really interesting to see to see this debate in the San Francisco Chronicle, in public forums, and to hear people using that dialect that was very different than mine.

With that let's talk about African American English Vernacular, about its historic ties to another dialect that we've talked about, and especially about what makes it rich and unique. It’s another dialect of English that has been studied well for over 100 years.

We're going to start with the phonology; that probably shouldn't be surprising by now. One of the things that we can see is a distinct reduction in diphthongs. Remember that diphthongs are when we have two vowels or a vowel and a semi consonant or a glide combined into the same nucleus of a syllable. We regularly see **monothongization**—so only one vowel instead of two. What is also really interesting as you see a lot of the same patterns in the dialects of Appalachian region and of Gulf American English—think of the dialects that are spoken in Alabama, Mississippi, the panhandle of Florida, and that area. You have first how these terms are written in Mainstream American English, and then you have African American English pronunciation, and then you have Standard or Mainstream American English pronunciation. There's a consistent monothongization with respect to these terms; it is not just these five terms, rather it is throughout the dialect.

We also see **neutralization** and **deletion** of certain sounds. We have **mergers**. Remember, if you think back to phonology and phonetics, we talked about how rare English is seven to 14 different vowel sounds; that is really rare. What we see with African American English Vernacular is actually a reduction to something that's closer to what most languages have. Depending on the subvariant, AAE or AAVE has between five and seven vowels; that's a little more typical.

We also see regular deletion of liquids, so r- and l-sounds, and we see a **labialization** of interdental sounds. Sometimes they can be just dentalized, but frequently their labialized. If you think about how *Ruth* and *brother* are said, in African American Vernacular English they're really frequently as 'roof' or 'root' sometimes, and not 'brother', but 'brudder' or 'brover'. We also see a **shift of the stressed syllable** to the first initial syllable of the lexicon.

These are really common and it's been interesting to see the historical connections. As I said, there's some real strong connections with what we see here and with Appalachian English. It does seem like the Africans who were brought over here as enslaved folks and forced to learn English, that they mostly learned it from folks who had an Appalachian dialect. Therefore, it is not shocking to see that, phonologically, these two dialects have many similarities.

Let's talk about the morphological and morphosyntactic aspects of AAVE. When we have **simplification** of consonant clusters at the end of a lexicon in African American Vernacular. This is also not uncommon; this seems to be the case in many dialects outside of the United States. It is also interesting to note that most of you who speak Mainstream American English, if you really get going and talk at a fast pace, you do the same thing, so this is a very common change that we see.

We also see some really important changes with certain verbs, and the main verb to talk about here is the verb *to be*. The fact that we have not just *to be* as a concept, but that it may exist or may not be used. This is not random; it is directly related to its function. If we're talking about habitual *be*, as in 'John is generally happy', then in African American English, for the most part, you would say, *John be happy*. However, if you're talking about a non-habitual or even stative, as in you're talking about his condition right now, that is not going to be the case. Instead, *be* is completely left out: John happy. Why would that happen, because this is something that does not happen in most English dialects. There are a couple of reasons: one is connected to the various languages that were spoken by the slaves that were brought over from Africa, especially Western and Central Africa. Most of the languages there do not have a stative *be*; the verb *to be* is only used when you're talking about something that is generally happening. However, there's a secondary explanation, and this has to do with something we'll talk about towards the end of this chapter with respect to creolization. Frequently when we see creoles, if there's a version or a use of the verb *to be* that does not get continued on, it's going to be the stative, non-habitual version. That's pretty interesting.

More morpho-syntactic and morpho-semantic aspects of AAVE:

- Multiple or double negation: again, we saw this with Appalachian English, and this happens in a number of dialects.
- Frequently, the determiner there, as part of the phrase there is or there are, is deleted in African American English and often replaced with either it or just left out entirely.
- We also have a lack of subject verb agreement, specifically with the third person singular. That is really interesting because if you think about the present tense conjugation of English, the only time we see inflection is third person singular. To have that completely be deleted is an interesting concept. It makes it more isolating in some ways.
- There's a completive or prospective done: it's not just you have messed up, rather, it's you done messed up. This is something that also connects that to Appalachian English.
- We have two different types of been, both stressed or unstressed; depending on what it is that you're doing.
- Add to this the concept of fixing to, meaning I'm about to do something or I'm going to do something. This is pervasive throughout the entirety of the South East of the United States—doesn't matter what dialect we're talking about-from about Oklahoma or Kansas, south through Texas, and all the way across to the Atlantic coast, fixing to is a phrase of is used and has been used for 200 years at least. In African American Vernacular English, that fixing to has changed to finna, and it is very common.

- You also have a deletion of the –‘s, that possessive or genitive marker.
- Again, there is a higher reliance upon word order, although English was pretty strict with respect to word order.

Just as with Appalachian English, we're not going to go into the lexical differences; that's too long of a discussion, but this is an interesting little tidbit with respect to African American Vernacular English: It is not a poor version of English. Quite the opposite, it is highly rich highly complex. In one of your journal assignments, you have the option of watching a little bit about Gullah English. what makes Gullah English so important is that it is a documentation of how the slaves used to speak their version of English, which has been crystallized. In many ways, Gullah English is the precursor to African American Vernacular English; it still has aspects of various Central and Western African languages that got mixed into the English as the slaves were forced to learn their new language. It's interesting to see how over generations and generations of use, and more contact with Mainstream American English, how this has been a dynamic dialect. It continues to change; it used to be said that there was really only one, maybe two sub-dialects of African American Vernacular English. Current investigation says otherwise, that there are at least 20 different sub-dialects. So, the next time you listen to an African American musician, rapper, artist of some kind, and they're trying to encapsulate so much of their culture in history, listen to how they talk listen, to that dialect. You will hear a lot of what we have been talking about in that. And that's just the tip of the iceberg.

4: African American English (AAE) is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5: Other Social Dialects

Other Social Dialects, from Sarah Harmon



Video Script

When we talk about social dialects, African American Vernacular English is definitely the big talking point. I honestly could go on for some time about it, but it would be better to leave that for a different discussion. It's important to bring in two other very big areas with respect to social isolation. Here in California, it's impossible not to talk about Latino English. It is a real thing, and it's not just some crude mish-mash version of Spanish and English; there's actually something going on. I also want to bring up something that probably most Americans don't even know ever existed, a dialect for the LGBT+ society.

Latino English

First up, let's talk about Latino English. Sometimes it's called Hispanic English, and there is a clear bilingual element to this dialect. (We'll talk more about bilingualism in a different section of this chapter.) It is important to bring up Latino English as a dialect in and of itself. Clearly, is going to have a mix of influences from Mezo-America, South America, and all of the different dialects that exist in those areas. There is some code switching—of course you would expect that—but what is important to remember is that it is a systematic combination, and we'll get to that in a minute.

If you ever learn about the history of Latino people in the US, you may have come across a term before called *Chicano*. The difference between *Chicano* and *Latino* and *Hispanic* is a topic for a different discussion. Frequently, Chicano is referring to a Mexican American from somewhere west of the Rockies: typically California, but it can include Arizona, New Mexico, Colorado, and Nevada. When we talk about these movements remember the root of that term is *move*, and this is an area that is moving. In the last 10 years, the term *Latinx* is gaining more and more attention. The use of that term, what it references and who accepts it (and who doesn't) is an ongoing discussion; I won't get into it here, although, if you wish to talk about it in Colloquy, I'm happy to bring it up. Do know that this is an ever-evolving dialect; it's a live dialect, and so things are changing all the time. Another reason to revisit this topic frequently.

When we're talking about Latino English—which is the term I'm going to use— let's talk about the phonology. Of course, it's going to take the English phonology and bring it a little closer to Spanish. As an example, in Latino English there is a reduction of the number of vowels—we've seen and heard that before—and there's also going to be fewer fricatives. English is a Germanic language—we love our fricatives—while Spanish doesn't. Spanish is a very Romance language and doesn't tend to have quite the number of consonant sounds that English does.

Speaking of consonants and reductions, there is a **reduction in consonant clusters**, especially at the end of the lexicon. Again, this is a very common phenomenon in a number of dialects in a number of languages. If you recall in phonology, we talked about the **epenthesis** rule for Spanish—the fact that you cannot start a lexicon with the combination of [s] and another continent, meaning

that [sk] is not going to happen in Spanish. We see that in Latino English, there's frequent epenthesis before that combination of an another consonant: *school* becomes [eskuL], *stripe* becomes [estɾaʲp], *Starbucks* becomes [estaɾbaks]. If you have native Spanish speakers around you or those who speak Latino English, you know that you have heard these tendencies many times.

When we talk about the morph-syntactic and morph-semantic features again we're going to see some really common themes. **Multiple negation**, which is common in a number of dialects we've seen it in the previous two chapters. We also have to remember that Mainstream Spanish allows for multiple negation—in fact, doesn't just allow for it, it encourages it—so we see it here in Latino English as well. There's a number of borrowings from Spanish.

What is interesting to point out is that this is Latino English, and as we talked about earlier, Spanish is a global language. Even if we just focus on the western hemisphere, from Mexico all the way down to the tip of South America, realistically we're talking about anywhere from 20 to 35 different dialects. What has happened with Latino English, there is a substantial amount of **neutralization** of dialects, and this is especially true with the lexicon. Whichever dialect is the dominant dialect for the area that will determine what gets used for Latino English. Here in the San Francisco Bay Area, Mexican English is very, very dominant and, depending on where you are in the Bay Area, it could be from Michoacán, it could be from Sonora, it could be from Morelia, it could be DF or Mexico City, it could be from a number of places. The Latino English variety of the local area depends on which dialect of Mexican Spanish you're going to encounter. By no means are Mexicans the only Hispanic population; here we have a thriving Peruvian community, as well as numerous cultures from all over Central America represented here., and we have quite a few Cubans here. Yet Mexican Spanish tends to win out in the local Latino English dialect. However, if you were to go to Florida, Cuban Spanish is going to be the dominant Spanish dialect; as a result, that's going to also be the dominant input for Latino English there. If you go to Spanish Harlem in New York, Puerto Rican Spanish is the dominant form. As a result, you're going to get predominantly Puerto Rican Spanish as the mainstay of the terminology in Latino English there.

Polari

Let's switch gears to talk about a specific dialect for part of the LGBT+ community: Polari.

Polari is something you very probably have never heard of before; in fact, when I started looking at Polari a little bit, I asked a number of my friends in the LGBT+ community if they'd ever heard of it. The only person out of a very large group that ever heard of Polari happened to be from Manchester, England, so he knew of it because, well, it was something that was talked about up there.

What is pillory? Well, it has a really interesting history, and one that I think should be studied, especially if you are into gender studies or anything having to do with the LGBT+ community. The term Polari seems to have come about in the 19th century, although the roots of this dialect start in the 16th century. We're talking about England, and we're talking specifically about folks who are male, although there may have been some transgender folks in that time as well, and all homosexual. There could have been bisexual male users of this dialect, but the documentation seems to indicate that Polari speakers were more homosexual males than bisexual or trans. We don't seem to hear Polari much with respect to the lesbian community, although it may just not have been documented.

Polari borrows heavily from certain languages, in particular Italian. In fact, the term Polari is a corruption of *parlare*, which is ‘to speak’ in Italian. There might be some other Romance languages sprinkled in, along with the language of the Romani. This is interesting because it is not a Romance language; it is in fact more closely related to Hindi than anything else. There's quite a bit of London slang, because it started up in London, and specifically Cockney slang, because it was mostly spoken in the east end of London. (More on Cockney soon enough.) Later into the early- to mid-20th century, you get even a bit of Yiddish in there.

Polari was a coded dialect; this was not something that people spoke in their everyday lives, necessarily. What they were doing was coding, to see if there were others out there like them. Specifically, it was a way to talk without giving away your sexuality, as it's important to remember that, until the 1960s or even the early 1970s, it was actually illegal to be homosexual in the United Kingdom. Punishments ranged from jail to forced castration, to many other types of ostracization, and isolation in the most grotesque forms. Therefore, if you know it is actually illegal to practice your sexuality or your sexual identity, you're going to find ways to express yourself that most mainstream Londoners or Brits would not understand—that's what Polari is.

What is really interesting is that once Polari outdid itself or became more known in mainstream, it started dying off. This coincides with the laws that had to do with banning homosexuality; they were repealed in the late 1960s. The movement started in the early 60s, but didn't actually get done till about 1969, and this is when realistically we see the end of Polari. It publicly outted itself in a radio program called *Round the Horne*, a really popular radio program that was listened to throughout the United Kingdom. In some people's eyes, it led to the Sexual Offences Act of 1969 being repealed. Whether that was actually the case, we might just

have a discussion about that later, but what is interesting to note is from that point onwards, Polari started to become part of the mainstream.

Let me show you some examples. Many of the numbers come directly from Italian.

Polari	English	Italian
medza	half	mezza
una, oney	one	uno
dooey	two	due
tray	three	tre
quarter	four	quattro
chinker	five	cinque
say	six	sei
say oney, setter	seven	sette
say dooey, otter	eight	otto
say tray, nobber	nine	nove
daiture	ten	dieci
long dedger, lepta	eleven	
kenza	twelve	

The numbers 11 and 12 have an unclear history; I would hazard a guess that that might be based a little bit more on Polari having connections with Romani, but I'm not quite sure.

Here are some examples; this is the original outing, as it were, of Polari in *Round the Horne*, written in a sketch by Barry Took and Marty Feldman. (Marty Feldman might be a name you recall; he was a very legendary comedic actor and writer. If you ever watched *Young Frankenstein*, the Mel Brooks movie, he was Igor.)

Example 5.1

Omies and palones of the jury, vada well at the eek of the poor ome who stands before you, his lallies trembling.—taken from "Bona Law", a *Round The Horne* sketch written by Barry Took and Marty Feldman

Translation

"Men and women of the jury, look well at the face of the poor man who stands before you, his legs trembling."

That kind of sounds like English and kind of doesn't sound like English, so let me translate: 'Gentlemen and ladies of the jury'; *omies* are men, *palones* are women. 'Vada well'—look well—'at the eek'—at the face—'of the poor ome'—the poor man—'stands before you, his lallies trembling'—with legs trembling. *Vada* is directly from Italian; *vedere* is the Italian verb for 'to see', and if you're going to command somebody to see well, it's *veda*; *veda bene*.

I'm also going to give you a little bit of a Morrissey song that I absolutely love; it's called “Piccadilly Palare”—and the chorus of it is this one:



‘So good to see you, oh your lovely face and your lovely hair.’ *Riah* is hair, *eek* is face, and those are both more connected to Cockney.

There is also a sketch from the BBC comedy, “Are You Being Served?” I have to admit, this is probably one of my all-time favorite situational comedies or sitcoms; it was a very long running sitcom in the UK on the BBC, no less. It was all about this fictional store called Grace Brothers, which was similar to Saks Fifth Avenue or Nordstrom, very high end with the highest level of service. As you can see, this thumbnail is at a meeting: in the middle you have Mr. Rumbold the department manager for the men's and ladies wear; next to him is Captain Peacock, the floor walker who greets and directs folks as they come off the elevator; Mr. Granger is sitting on the other side of Mr. Rumbold, and he's the old head of the men's wear department. Mr. Grace, who is the owner of the store, is supposed to be in his 90s, but is very young minded and decides that there needs to be a change to attract more customers, specifically more younger customers. His change is to talk more like the youth, including some Polari. Take a listen.



That's Polari. It's also important to point out that the guy in the glitter suit, his name was Mr. Claiborne Humphries. He was considered camp in those days; the truth was they always made innuendos about his sexuality and eventually the character came out as gay, but not for 20 years. He had to play camp in order to be acceptable and get through the sensors at the BBC.

The final video that I'm going to play down here is a little bit from Rowan Ellis, a prolific YouTuber in the UK. She works predominantly to talk about LGBT+ issues, everything with respect to the culture. She didn't know about Polari much and so she talked about it.

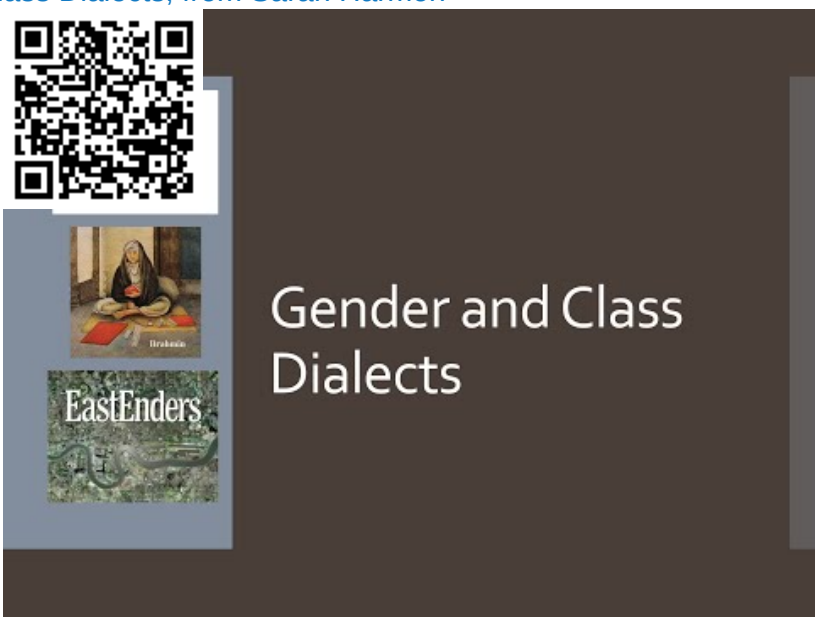


What is really interesting is that certain aspects of Polari have become a part of mainstream, just like Cockney has certain phrases and lexicon that have entered Mainstream British English. This is also similar to how various aspects of Appalachian English and African American English have become part of Mainstream American English. This has to do with **normalization**; at some point, a dialect that was either secretive or definitely considered substandard in prescriptive terms becomes accepted, or becomes normalized. Part of that is an evolution of the language and the culture along with it. Those of us that are old enough to remember when rap first started coming out, we remember that there was a huge shock at the language being used—this is before gangsta rap started, back when the Sugar Hill Gang, Run DMC, Big Daddy Kane, and all of these original rappers that were coming out and getting mainstream play. There was a big shock, mostly amongst white Americans, who were shocked at what they were hearing—that African American Vernacular English had been glorified. Within 10 years, to the early- to mid-90s, African American vernacular English became more mainstream. I remember having a fierce debate with my dad because he was aghast that I was listening to rap. He couldn't understand how I could connect with it; I just liked the rhythm, the beats; white suburban girl here liked it. I have to say my dad changed his opinion about rap pretty early on; he was pretty cool with it mostly because he had the same discussions with his mother when it came to rock and roll. When I brought up that comment, he backed off; he started listening. The same conversations happen with every generation or two: 'So-and-so's language, so-and-so's dialect is terrible; it's an impoverished blah-blah-blah… that's not what we do with respect to linguistics. We mark the trajectory; we observe and we document. What is true is that marginalized dialects, when they become less marginalized and even folded into the mainstream, pretty interesting things happen to the then-mainstream dialect.

5: Other Social Dialects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6: Gender and Class Dialects

7.6.1 Gender and Class Dialects, from Sarah Harmon



Video Script

In the video above, I have a little bit of something called Cockney, which we'll get into that in a minute. Cockney is part of a gender or class dialect that first started off with the men of Eastern London, and then quickly spread throughout the East End of London. When we talk about gender dialects or class dialects, they tend to be well connected to one another in a variety of ways.

Gender-based Dialects

Let's first talk about gender-based dialects. We're not just talking about women being a little more polite than others, or men being a little bit more straightforward than others or anything along those lines. When we're talking about a gender-based dialect, we are specifically talking about how there are phonological, morphological and sometimes even syntactic differences between how an identified male talks versus and how an identified female talks. Many of these differences can be seen, not just within the gender itself, so if a man talks to a woman, there could be differences there too.

I'm going to start with Muskogeean, a language in the Algonquin language family, and is spoken in the general northeast of the United States, a little bit closer to upstate New York. If we have these verb phrases

Male Speaker	Female Speaker	Gloss
kan	ka	'he is playing
molhís	molhìl	'we are peeling it
lakawhos	lakawbol	'lift it!' (command)

Notice that there is actually a difference between how a man says it versus how a woman says it, Notice, too, in these examples, the data that we have do not include anybody who is not such gendered. That is a hole, but sociolinguists are working to address these areas of need. It is really interesting to note that there is an actual difference in the inflection depending on whether the person identifies as male or female.

Traditional Japanese has some of the same characteristics, although not with a verb phrase, but with personal pronouns: the term for 'I'. If you have learned Japanese, probably you do not speak Traditional Japanese; you speak Modern Japanese or Mainstream Japanese, which is usually based out of Tokyo. Traditional Japanese is what the upper class used to speak up until about 40 years

ago and it's based mostly out of Kyoto. Traditional Japanese basically was isolated as a class difference, but then within that there were also gender differences in certain cases.

> **Note**
>
> Traditional Japanese (not general population)
>
> - First-person singular pronoun 'I'
> - Female dialect: *átashi*
> - Male dialect: *watákushi*
> - Sentence particle, 'it's a matter of'
> - Female dialect: *no-*
> - Male dialect: *n da*
> - Neutral dialect: *n desu.*

The first-person singular pronoun is an important pronoun, linguistically. Females tended to refer to themselves not as 'I', and if you know modern Japanese, then you know it's *wátashi*. Instead, they said *átashi*. Notice there's a difference between that and how an adult male almost always would say *watákushi*, not *wátashi*. In many ways the modern Japanese almost seems like a combination of the two: take the [w] sound from *watákushi* and put it to the front of *átashi* to get *wátashi*. This does not include the fact that in many dialects of Japanese, the first male of the family refers to himself as *boku*. This boku is used by many boys, although it can be continued on into adulthood, but *boku* is also different because others, especially the mother or grandmother of the family, will use that term with the first male child, instead of their name; there is more to this story, but it will be saved for another time. Much of the gender dialect distinction that existed in Traditional Japanese does not exist anymore in Mainstream Japanese, given that the regular use of Traditional Japanese mostly died out about 40 years ago and had been in decline for some time prior.

There is a sentence particle when you want to say 'it's a matter of', as in 'it's a matter of respect'; 'it's a matter of pride'; 'it's a matter of (something). There was a neutral version usually used in mixed gender groups, and there was a male version and a female version.

By the way, this is not the same as Portuguese or any other language that uses grammatical gender on nouns and adjectives; for example, if you speak Portuguese, you know that the concept of saying 'thank you' is actually tied to the adjective for 'obligated', such that when you say 'thank you' to somebody you're saying 'I'm obligated to do this for you', but all wrapped up into one lexicon: *obrigado* or *obrigada*, depending on how the speaker identifies. There is a third option that's starting to come out and there's a couple different flavors of it, much like what is happening with Spanish. But this is a different thing; this is only with isolated cases in the case of like *obrigado* or *obrigada.* In the case of a gender dialect, we're actually seeing a systemic difference between how a cis-gendered female speaks and a cis-gendered male speaks. It has to do with syntax and morphology, not just phonology.

Linguistic Behaviors of Men and Women

We cannot leave this topic without talking about how men and women speak, and not just in English, but in all languages. Cis-gendered males and cis-gendered females definitely speak in different ways, both culturally as well as linguistically. It is interesting that this study really doesn't start until the early 1970s, in the US and actually specifically Berkeley. Robin Lakoff and Deborah Tannen are the two originators here, but there have been many others that have followed sense and have done studies, not just on American English, but throughout the world. There is a tendency, certainly in western cultures, to talk about women having certain linguistic characteristics and men tending to have certain other linguistic characteristics. Notice that I'm hedging a lot, using qualifiers like *tending*. Generally, we are not saying that every single woman uses more proper educated language, nor hedges to start an utterance, nor uses tag questions, nor uses more politeness techniques. We are not saying that no male does any, let alone, all of these. When Lakoff and Tannen were doing this evaluation this observation. They were not including the LGBT+ community at that time; that wasn't even on the radar at that era, although it has been interesting to see the evolution of area of sociolinguistics and the inclusion of various gender and sexuality groups. This is true both in the observation and the study of gender-linguistic tendencies, and in how all these different groups interact with one another and within themselves. It has been an interesting just observation lab for anybody who's a linguist. There are tendencies—not just with American speakers and not just with English speakers, but worldwide.

There is a tendency for women, especially those who are mothers and those who live and work in male-dominated societies, to have the following linguistic characteristics:

- use more prestigious language examples;
- hedge a little bit more;
- add tag questions—"You understand?" “Right?”—at the end of their statement; and
- use more politeness techniques.

Please note: this is not to say women can't be rude—we are very capable of being rude, we can very capable of being frank, speaking directly, using all sorts of jargon and slang and everything else in between.

Men have a tendency, especially if their fathers to have the following linguistic characteristics:

- use more colloquial speech;
- take stronger positions in language, although that can change;
- use fewer tools to check in with the audience like a tag question; and,
- use more direct in their statements.

It has been interesting to watch how empathy training in particular has made this change, or hasn't in some cases, and the reactions to it. This area of gender-based linguistic differences is evolving, and the data coming in now are fascinating. But in 1971 this was a pretty big revelation; in fact, Robin Lakoff and Deborah Tannen have both written extensively, not just on the subject, but on how their seminal work took off and let everybody explore this concept of what it is to use language in different ways, based off of your gender identity.

Why would any of this happened in the first place? Why would women and men have different language patterns? Frequently, especially in male dominated societies, women frequently have to negotiate a little bit more, we have to be a little nicer, we have to be non-confrontational; we have to double-down on that if we're in male-dominated arenas. Some of the research lately has shown that these tendencies are true even if in an all-female environments, but if there's a really dominant female that shows a lot of male linguistic traits, many women will take the other side of that coin, as it were.

Have these studies changed? Yes; we're starting to see quite a bit of change, especially in the last 10-15 years. Is this going to stay, or is that change going to entrench itself, or is it going to change further? We don't know is the answer, but that's what we're going to be doing as linguists: we observe language as it is used.

One of the things that we have noticed is that this concept of face is important to women, and not just here in the United States or in English-speaking countries or communities. It is important, almost universally in almost every culture, in particular if a woman is a caretaker or mother of children. **Face** is the concept of wanting to put your best face forward. Women tend to put a lot of emphasis on face, although it's not to say that men don't; certainly, some men do and some men don't it, and emphasizing positive face doesn't make them more feminine. I will be going back to this topic and later point in time, but I made a point when I first started working on this topic, maybe about 15 years ago, to write observations and current research at the time in a little journal and I told myself in 20 years I wanted to come back to that. We're almost up on those 20 years and I’m starting to cultivate a little bit of that research now, both in journal articles and books, but also in my own observations. In my case, I’m not looking at English; I’m looking at Spanish and I’m looking at Italian, which are the two languages I most often work in. I’m looking at how this is changing, especially given that Italian culture is heavily macho with male dominance being very strong; certainly, you can say for most if not all Spanish speaking cultures machismo is an element as far as how male driven or male dominant varies from one culture to the next. Spain is fairly macho, for example, and that's where I’ve been doing a lot of this initial investigation, looking at the dialects of Spain. I’m also looking at what's happening in Latin America, because there is so much information there and, in many ways, Latin Americans are a little bit ahead of the curve with respect to the Spaniards on this topic; there seems to be a bit more equity in language there. It's an interesting discussion; come back in about five years and we'll talk more.

Class-based Dialects

While there is some connection between gender-based dialects and class-based dialects—a connection that we’ll talk about later—it’s time to focus now on class-based dialects. In this section I’ll showcase two very different examples.

If you know anything about the cultures of South Asia, you know that, at least at one point and perhaps still, there was a caste system, which is a social stratification with different social levels. The highest caste was frequently called the Brahmin caste or the Brahmin group; they were usually priests, the royal families, the highest of the highest. That group frequently did not speak with lower caste members, which meant that there was isolation. In that way, the Brahmin dialect of many of the languages of South Asia, whether they are Indo-European or Dravidian, there are distinct differences. We see some examples from the Brahmin dialect of Tamil, which is spoken in Sri Lanka. It's a Dravidian language so it's not connected to Hindi, Gujarati, Urdu, or any of those languages, as they are all Indo-European. Dravidian languages actually dominated what is now considered India, and then got pushed south to the southern part of India, as well as Sri Lanka.

We see differences of pronunciation, as well as morphological differences; while I don't have it here, but we do have some syntactic differences. Why is this important? These castes for the most part officially have gone away, although you could argue that unofficially they still exist. What is clear is the Brahmins that were of the upper echelon over the last several generations, they have mixed in more with many other of the general populace. As a result, many of these class-based dialects have started to neutralize and started to disappear, but others still exist. This is not just true of Tamil, but of most of the South Asian areas where there was a very distinct and very separate Brahmin caste or speech community.

The class-based dialect that I think for most English speakers is really the perfect example of a caste-based dialect is Cockney. You may have heard of Cockney before this class, certainly if you are into British soap operas and British media in general, you know of Cockney. It is probably best explained as a rhyming word play, but in truth it's almost a dialect. It's a coding, much like Polari is was a coding for gay men in particular of the 19th and 20th centuries, but Cockney in this case is more for those who are Eastenders. The term ‘Eastenders’ refers to those who live in a very specific region of the eastern part of London and, specifically, it has to do with if you are within sounding distance of the chime of a very specific church in that part of London, I believe it's Westgate. If you are within the sound of that bell, you are an Eastender and you probably know Cockney. There are two long running dramas in the UK—one is called *EastEnders* and the other one is called *Coronation Street*, often called ‘Corie’. They were examples of Cockney in action, as almost every character talked using Cockney.

The exact start of cockney is unknown, although we suspect that is probably somewhere in the late 18th-early 19th century; definitely by mid-19th century it is strongly entrenched into the East end of London, where we probably see its most with respect to the Irish laborers who lived in that part of the city. However, the roots of Cockney may go as far back to the Middle Ages. What we do know is that this concept of rhyming was used to code or not talk about certain things, as a way to get one on get one over on their landlords or their managers. It definitely has some distinct phonological and lexical aspects.

The phonological aspect is just something that's very common to eastern London. Examples include:

enition, or weakening, of that r-sound at the end of a word; sometimes it goes to a [w] and sometimes it completely disappears;

eduction of a [t] to [ʔ], especially when intervocalic; think of that neutralization rule that we saw where [t] and [d] would combine to use that little tap, as in that we do not say [lætɻ] and [lædɻ], instead saying [lætɻ]. For Cockney, that [ɾ] is going to change to a glottal stop more often than not;

eletion of an [h]; so this isn't a [haʷs], but [aʷs];

ıe velarized [L] changes to a vowel; and,

onting of interdental [θ, ð] to labiodental sounds (something we saw in African American vernacular English)

Again, this concept of changing sounds reducing them down, it's a really interesting comparison with what we hear in so many other dialects throughout the English-speaking world.

Of course, if it's a rhyming slang, we clearly have to talk about lexical changes. Its seemingly random and yet not; it seemingly has some connection to mainstream media, but not always. It's a type of mental gymnastics you have to play. Instead of saying ‘a bottle’ of something, like a bottle of Scotch or a bottle of wine, you say ‘an Aristotle’, because ‘Aristotle’ and ‘bottle’ rhyme. Instead of saying something is in your ‘pocket’, you say something is in your ‘rocket’; again, the rhyme. You can also take a rhyme and then shrink it: if you want to say you're going ‘up the stairs’, you can say you're going ‘up the apples and pears’, or you can just say you're going ‘up the apples’ and just knock off the ‘pears’ part. If you are in trouble, you can say you're ‘in Barney Rubble’, or you can just say you're ‘in Barney’. Everybody knows what you mean. ‘Face’ as a ‘Chevy Chase’; ‘pub’ is a ‘nuclear sub’; ‘chest’ is a ‘bird's nest’; you get the point. There's a slew of these rhymes, and it's all as a way to code, so that the person you're talking to knows what you're saying while everybody else may not.

There's also some metaphoric slang, especially with respect to certain bodily functions or fluids. For example, like blood is claret, which is a type of red wine blend that is really popular and has been for hundreds of years in in England.

Below you have an example taken from the movie *Lock*, *Stock* and *Two Smoking Barrels*. Guy Ritchie is a filmmaker from the UK. You may have seen some of his work before, and he has a love affair with Eastenders; he loves including at least one, if not most of this cast, having Cockney as part of the repertoire. In the clip below, you have an example where somebody is describing a fight that happened in the bar the other night. It is captioned and by captions I mean it's actually translated subtitles as to what the person is trying to say in more mainstream British English; the closed captioning should also take that into account.

Have fun with it; Cockney is kind of fun.



7.6.2 Do women and men use language the same way?, from Anthony Pym (optional)

Anthony Pym is a professor of sociolinguistics and translation at Rovira i Virgili University in Tarragona, Spain. He has a series of lectures on his YouTube channel on language. This is one of his talks on how women and men use language. (There is no video script associated with the video, but it is captioned via YouTube's autogenerated captioning.)



6: Gender and Class Dialects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7: Prestige and Politics

7.7.1 Prestige and Politics, from Sarah Harmon



Video Script

I made it clear at the beginning that there are certain dialects that are given more prestige than others. When I say they're given more prestige, it is not by linguists; we don't give prestige to anyone. However, in a society, a certain dialect or manner of speaking will get more prestige than others and frequently, it has to do with power and therefore politics so let's briefly talk about that.

In every single culture, there is considered a prestige dialect. What does that mean? The prestige dialect is the one that everybody emulates frequently, although not always; it is the dialect of those who are generally in power—not just one person, but several people. That being said, I have to reiterate the fact that we as linguists do not mark something as supposed to be prestigious. We observe. The other thing we observe is that frequently non-prestige dialects are considered substandard—not just nonstandard but substandard—and then many times, even being banned. That can be true not just of dialects, but of minority languages that are used in that society.

Let me give you some examples of languages and dialects that have been banned. This has to do with prestige and power.

If you know anything about the history of Spain, especially in the modern times, in the in the 30s there was the Spanish Civil War, and Francisco Franco came into power. He banned all other dialects of Spanish outside of Castilian Spanish, because the capital Madrid is in Cast, therefore, that is, the mainstream and of story, and so, for the 42 years that he was empower if you spoke anything other than Castilian Spanish, you had problems. This attitude was doubled down on the other languages that were spoken in Spain; Basque, Catalán, Gallego, Valenciano, these were all banned. When I say banned, I mean you could not speak them or read them or write them in public. Music, radio, TV, schools, print media, none of it.

That is a pretty extreme example; by no means is it unique. When we're talking about modern times, so think the Renaissance forward, there have been multiple cases of a group put into power that banned all other dialects and languages. In France, part of the French Revolution in the 1780s and 90s had to do with the marketing of Parisian French as the standard, as the prestige dialect, as the language of the people and of the language of the revolution. As a result, all other dialects, as well as all other languages, spoken in France were not allowed in schools, exactly because they were just not prestigious enough. That continues to this day. You have Basque, Breton, Occitan, Provençal, Gascon, all of them—if they have not already died out, they are severely threatened. The Basque community is probably the strongest of that group, mostly because of the Basques who are in Spain helping the ones in France. Breton is still around, although very much endangered. Provençal on its very last legs; there probably are only literally a handful of native speakers left. Gascon and Occitan have already disappeared.

The Celtic languages in general, but especially in the United Kingdom and the British Isles, including when Britain also included Ireland as part of its realm. The Celtic languages were all suppressed by the government and by the churches, first the Catholic

Church and then the Anglican Church. It continued all the way through the early-mid part of the 20th century; it's really only since about the 1970s that there has been a concerted effort to revitalize the Celtic languages of the British Isles. The revitalization movement started in Ireland, once it gained its independence in the 1910s. It was able to start establishing Irish Gaelic as an official language to start building up the speakership, and it has gotten to the point that most folks, depending on where you grow up in Ireland, are exposed to Irish Gaelic. All are taught Irish Gaelic throughout their schooling years.

Certainly, if we talked about this topic in modern times, we do need to bring up the USSR or Soviet Russia. Although the seeds of this really start before the revolution, at the start of the Russian Empire; being that Moscow was the capital, Muscovites were given prestige for their dialect. Any other dialect of Russian was not held in such esteem, and, of course, the same was true for any of the non-Russian languages spoken in the Empire. This is particularly true of any of the non-Indo-European languages. We still see this battle in the Ukraine, with respect to Ukrainian versus Russian, and that Ukrainian is definitely coming back strong. We also see the other side in Belarus, as Belorussian is very endangered, all thanks to the close ties that the current government regimes have had with Russia. As you can see, it also with Lithuanian, Georgian, Armenian, Kazakh, Uzbek, and so many other languages in what first was part of the Russian Empire and then later Soviet Russia.

Why it's important to bring this up is the fact that when we talk about language policies, one of two things happens: either the language that is being affected—the minority languages, at were—is being tamped down and goes underground. This mostly happens with languages, but it can happen with dialects; African American English is actually a great example of this. When these folks were enslaved, they were pushed to not speak these dialects but then, as there was more and more freedom to do so, more and more folks spoke African American Vernacular English in their daily lives. It can also lead to revitalization efforts down the road. For example, once Franco passed away in 1978, one of the first things that happened was a revitalization, especially of Basque and Catalán. This was not surprising, because the Basque Country and Cataluña are the two richest provinces in all of Spain, the two biggest ports and the two biggest longstanding cultural areas. Gallego and Valenciano have had more of a struggle. Gallego is starting to get some traction back and certainly connections with Portuguese help, but it hasn't come along the same way that Catalán and Basque. Valenciano is even more endangered, especially given that, if you think about where Valencia is in Spain, it is halfway between Castile on the one hand and Cataluña on the other. In many ways, Valenciano is almost a in-between language that encapsulates aspects of both Castilian Spanish and Catalán. It's not having as much success as either of those languages in it's going forward.

In the case of the Celtic languages, we see a very strong showing with respect to Welsh, we see a pretty strong showing with respect to Irish Gaelic, although there are questions about sustainability. However, we see less sustainability with Scotts Gaelic, although there may be a reinvigoration again. Manx completely disappeared and then completely came back, all because of the will of the people. Breton is very much endangered, and Cornish is completely gone, although you still see vestiges of it now and then, if you go to Cornwall. Those are just the languages; when we talk about dialects, it's the same route as well.

All languages evolve, of course, but at some point, something's going to happen and all languages at some point will have a death. It's just inevitable, but how that happens depends. For example, most of the time it just gradually falls out of use, with generation after generation after generation opting for something else; this is **gradual language death**. Sometimes that something else is forced upon them because of a change in government or society, but often it's also just due to historical evolution; think of Latin, which had a fairly gradual death. If you think of Classical Latin, that actually was not being spoken by most folks when Augustus Caesar came into power. The upper elite might have been using Classical Latin, but certainly not the common folk; they were using Vulgar Latin, and that evolved into the modern Romance languages. Sometimes you do have a **sudden language death**, where you have a total extinction of the population all at once; think of a genocide, in particular. **Radical language death** happens due to political influence, usually within one to two generations. It's maybe not quite as sudden as the first one, but definitely is within a couple generations. We also have **bottom to top language deaths**, when the language just falls out of use, but it's still being used on some levels; this is actually the case of Latin, as it's a combination of three and four. It just started gradually not being used, and then it comes back a little bit with respect to Medieval Latin and Clergy Latin, which is the Latin that was used in the monasteries all throughout England and mainland Europe right through to the Renaissance.

It is important to also realize that when we're talking about how languages can become extinct, a lot of it just has to do with circumstances. A great example of this right now is Covid-19, and the fact that we have seen devastating losses in so many speech communities where the language may have already been on the way out, if not highly endangered. We're seeing it throughout parts of Asia and the Pacific Islands; we're seeing it throughout Australia; we're seeing it through the Americas as well with indigenous populations; and, to a lesser extent, we're also seeing it in Africa. It's important to understand that when we have language

extinction, yes, there is a sadness to it. However, as a linguist, it's also a natural part of evolution; this is what happens. As long as we try to document what happens and what how that language has been spoken for some time, we at least do what we can.

It is a reminder, by the way, that just because a language becomes extinct, that does not mean that we can't revitalize it. The Celtic languages are really great example of this; another one is Hebrew. The Modern Hebrew that is spoken today has really only been in existence for about 150 years. Classical Hebrew died out many centuries ago. It was revitalized because a group of people wanted to do it, and they did so in a systematic way: taking Classical Hebrew and adding modern aspects, including aspects of Modern Arabic and Amharic, two closely related or sibling Semitic languages. Manx is another example; essentially, it was extinct for at least three generations, but then the people brought it back with a concerted effort.

No matter what you do language changes. It evolves, period, end of story. Every time a child is born and they acquire at least one, if not more than one language, there are changes that happened. There are changes genetically, sociologically and linguistically that just happen. When you have a change in population and/or speech community—mass migration, massive illness that takes out a very large group of people—we also have change with respect to analogy. I'll explain a little more of that in historical linguistics; suffice it to say that when we talk about analogy, we're talking about weeding out irregularities. We also just see other changes in society; take Latin for an example. In Classical Latin, you have very distinct, very crisp inflections—almost always suffixes—that are full of consonants. By the time we get to Vulgar Latin and Late Latin—Vulgar Latin is the Latin of the people, Late Latin starts with the century of the Roman Empire, roughly the 5th century CE—we start seeing constant deletion, especially at the word final position. However, if that consonant is part of how you distinguish one inflection from another, that's going to have a lot of repercussions, and one of them is the loss of case system—no more fancy inflection for the subject versus the object. You have more fixed word order, and this all happens at the same time, so it becomes a chicken in the egg paradox. You can't really tell when one change starts and then the other change is affected next; they all seem to happen all at the same time. It reminds me of watching metal, especially like gold or silver, in a crucible. It goes from a solid state to a liquid state lickety-split and all at once. A lot of these changes seem to happen that same way, or at least they start that way, and then they continue to evolve as speakers continue to use the language.

7: Prestige and Politics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8: Social Norms

7.8.1 Social Norms, from Sarah Harmon



Video Script

When we're talking about social norms, we're talking about how these social norms affect language in this case. Believe it or not, there's actually quite a bit that we can say about this.

Let's start off with talking about taboo. **Taboo**, of course, is that concept in which we talk about something that is unacceptable in society. In most every culture, there are certain cuss words or swearing or profanity words that you're not supposed to use…yet, everybody uses them. We use them as some kind of exclamation, or some kind of intensifying agent, and very frequently those profanities have something to do with blasphemy (having to do with the god or gods of whatever religion is being practiced in the area). They could also have to do with body parts or body fluids or bodily functions. Think about the cuss words or the profanities that you know—not just in English, but in any other language you speak. Chances are they have to do something, either with blasphemy, or something with the body. It is really interesting, by the way, to note that in so many cases in English, the profanity comes from an old Germanic word, but the acceptable term is frequently a Romance word. My favorite example is *shit* vs *manure*; *shit* is an old English word, very Germanic word, while *manure* is definitively French and got borrowed. In many cultures, there's an avoidance of death to the point of not saying any word having to do with death: not saying the names of dead people or dead ancestors, there's a number of ways to go about this. The most famous example, perhaps, has to do with the Chinese languages. Because of how the writing system reflects usually a number of things, the character and the term for the word *four* is very close to the character and the pronunciation of the word for *death*; it's either the exact same term just different tone or very similar. As a result, the number four is a very unlucky number. This is spread, not just from China and the various Chinese languages and cultures, but throughout most of East and Southeast Asia, simply due to the influence of the Chinese and their geopolitical status. By contrast, in most languages of the Sino-Tibetan realm, the word for *eight* and the word for *fortune* or *luck* are either the same or very similar, and so the number eight is good luck.

Euphemism and dysphemism. You probably have heard the term **euphemism** before, when you are making something better. The prefix *eu-* means ‘good’ so a euphemism. Instead of talking about how shitty somebody looks, you talk about how different they may look, how unconventional they may look; for some people, that would be a euphemism. **Dysphemism** is the opposite; it is purposefully including taboo, profane language or anything negative into something My favorite example of this is the word *fuck* in English, not just because it's a word I tend to use, but because its usage is amazing. This is true not just in American English, but in all world Englishes, that term has a life of its own. We can even make it an infix to make something even more intense. For example, take the adverb *absolutely*, and infix -fucking- into the middle of it, and you get *abso-fucking-lutely*. We even have a couple of famous acronyms, FUBAR and SNAFU—I'll let you look them up—that are examples of dysphemism.

Humor is always a part of social norms. I told you I’m a fan of *Archer.*



I like *Archer*, and the fact that not only the joke is pretty straightforward with respect to humor, but the name of the character: Fuchs. It's meant to be a play on a German last name, but you get the point. Humor is always a part of social norms in language, as it always taps into what is taboo and makes fun of it somehow. It's also very culturally specific; what works in one dialect or one speech community may not work in a different one.

Other elements to talk include **register**, which describes the changes that arise when we speak to different speech communities. If there is a significant difference—a class difference, role difference, along those lines—we measure the difference in register. When we think about going formal or informal in our speech, that's an issue of register. In many languages, you actually change morphology, syntax, semantics, and/or the lexicon when you use formal language; think of different types of inflection, different forms of address, along these lines. **Politeness** is also factored into register; how you are polite to your mother may not be how you were polite to your sibling or how you reply to your boss, or how you are polite to somebody you really don't like. In many ways we modify our language. Even the use of **jargon**—when it's appropriate to use jargon versus not when you have to explain it, and when you can just leave it as-is.

These are all elements of what we like to call social norms. Below is a is a video by Anthony Pym on diglossia and that really is going to hit with respect to register. **Diglossia** is when we maneuver between different speech communities and we modify our language accordingly. Sometimes it's being in a different dialect, but often is not; it also includes register and how we maneuver between these different groups.

7.8.2 Diglossia, from Anthony Pym (optional)

Anthony Pym is a professor of sociolingusitics and translation at the Rovira i Virgili University in Tarragona, Spain. He has a series of lectures on his YouTube channel. In this talk, he discusses diglossia. (The video is captioned via YouTube's autogenerated captioning system. There is no video script.)



8: Social Norms is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9: Languages in Contact

Languages in Contact, from Sarah Harmon



Video Script

The last section, at least the last required section for this chapter, is languages in contact, and this is one of my favorite sections. In part, there's a little bit of history in here, which that's why I like it. However, it's also because it really encapsulates what it means to be a human being and use language. Very rarely does a single person, certainly in the modern era, only surround themselves with people in their own speech community. Whether is mass media or trade or travel or anything else, we are constantly in contact with different dialects and different languages. Let's talk about what happens when two languages come together.

I'm going to start off by setting the scene. Picture it: You have this group called the Lasinki. The Lasinki leave their home island of Lasinkia and they go off to find a new home. About a month later, they finally land in a different place; they call it Walonkia. There are indigenous folks in Walonkia, but they speak a language that has nothing to do with the language of the Lasinki; they're completely different languages. Initially, things are good, the Lasinki and the Walonkians are trading and they're conversing. But then, things turned sour and the Lasinki end up conquering the Walonkians, and in doing so, the Lasinki become the group that are in power. They impose a number of elements of their way of life onto the Walonkians, and this includes the language. On the side, the Walonkians will speak their language to each other when the Lasinki are out of earshot, like at home or perhaps in a small gathering of fellow Walonkians. However, when they're in public they're forced to speak Lasinki.

We know the situation; what happens linguistically?

It's a three step process. When the two groups are doing okay, they're the see each other as equals and their trading, they're using what we like to call a **lingua franca**. That term is something you may have heard of before, but let me put a linguistic spin on it. Specifically in linguistic terms, a lingua franca is a trade language. There's no shift in power; the two groups see each other as equals, therefore it is just a little bit of this and a little bit of that, with a lot of gesturing and not an exact language. I think of a lingua franca when I think about my dad whenever he goes into a place that he doesn't speak the language. He loves traveling; he's gone throughout parts of Latin America and parts of Europe. He's a salesman and he likes to go into markets to buy things. He never brings an interpreter nor a translator; he always goes by himself. He will phrase it that he uses 'the language of trade', which means he's using a lot of gestures. If he knows a little bit of the language—like I have taught him a little bit of Spanish and a little bit of talent of Italian—so that when he's gone to some of these places, he can say, "What is that?" It may not come out perfect, but he knows a little bit. That would be like a lingua franca.

The second that the Lasinki come into power they conquer the Walonkians, and they come into power and they impose their language on them, then you have a shift in the paradigm. At that point, they are no longer equals politically; you have the Lasinki higher in society and you have the Walonkians lower in society. Because there is that difference, that strata difference, the Lasinki

become the dominant or superstratum language, the Walonkians become the substratum language. When they start mixing, you get what is called a **pidgin**. It must be said, a pidgin is not a native language; it only comes about for a single generation, that initial change in situation. However, it is systematic; when you force two groups to combine their languages, there ends up being a systematic reordering and combining of those languages. Always when we see a pidgin, we see a very strong reduction in everything: we see fewer consonants, fewer vowels, we see fewer puzzle pieces or morphemes. Those morphemes now take on different roles, well more than what they used to. There are patterns, but there's slipshod, meaning there isn't a strict word order or strict set of inflections that are being used. There's a lot of play going on, a lot of multiple roles doing multiple things. Lexical entries undergo the same systematic shock; a lexicon can mean a wide variety of things. For example, a *bottle* can be anything that holds liquid, whether it's something like a water bottle or a baby bottle or a giant jar or jug.

The second that you have a new generation learning that pidgin then that becomes a creole. A **creole** is basically a standardization of that pidgin. As each child is born into that society, they learn this pigeon and subconsciously start putting order to the chaos. This happens regularly, and what is really interesting is that within two generations of a pidgin forming, you have a fairly stable creole. While there is more stabilization in an early creole, there can be still be quite a bit of linguistic flexibility. You can have a lot of possible meanings tied to one specific morpheme for up to 10 generations, it can happen or it can just change and be done.

There's a very interesting debate that's been going on for some time, and not just among linguists. Anthropologists have gotten in on this debate regarding at what point does an early creole standardize, as it were, becoming more regular and, dare we say, become a language. There's a lot on this topic, and I don't have nearly enough time to get into it, but suffice it to say that in its earliest stages know a creole is not a language. There are still too many possibilities that are in play, too much manipulation and not as much standardization. However, what happens if you have a creole that has been spoken for five generations or more, that has become much more stable and has become much more regular? At what point does it stop becoming a creole—a ‘lesser than’, if you will—and starts becoming a language? There are different camps in this arena: there's one camp—and I will openly admit that I’m part of this camp—that says that once you get a creole that has been spoken for about four or five generations, it's pretty much a language. The reason is because you have regular rules, a growing lexicon, a system of derivation, and, perhaps most crucially, you have native speakers for multiple generations. A child hears not just their parents talking in this creole, but their grandparents and maybe great grandparents at that point. You probably have a language. Certainly, if we're talking about a creole that has been around for 150, 200 years—and we do have those—then we're probably talking about a language.

There is a sizable number of linguists and anthropologist who say, even at that point, it's still a creole and not a language. Their rationale is that because there's too much variation, the grammar can get played with a little bit more, meaning that there's a lack of standardization at times. This argument is changing in favor for those of us who believe that after a certain number of generations a creole is really a language, and it gets back to that point of variation. The fact that when we talk about dialects, we're talking about variance—and I’ll go back to the African American Vernacular English. There aren't one or two dialects of that; there could be upwards of 20 or 30, maybe even more dialects and sub dialects of African American Vernacular English. Who are we say that a creole is no different? In many ways, if we really want to look at the objective landscape of a creole, most would have to agree that there's quite a bit of regular derivation and regular evolution of these creoles.

I'll give you some examples of what I’m talking about when I say long-term creoles, in the table below.

You see Hawaiian Creole, Haitian Creole, and Sranan. Hawaiian Creole and Sranan are both English-based creoles; that means English was the predominant **super-stratum language**, and then the indigenous and other languages were the **substratum languages**. Hawaiian Creole, of course, has the Hawaiian language as the substratum. Sranan is spoken in Suriname, South America, so the Arawak languages that were spoken in that area and form the substratum base. Haitian Creole as a French-based creole that has some Taino and other Arawak languages, but mostly, it is a number of central and western African languages that make up the substratum. Exactly because of the Slave Trade, because Hispaniola—the full island where Haiti is—was a processing center for a number of slaves. As a result, a lot of the substrate languages that we see for Haitian creole come out of central and western Africa.

In this table, I have the verb ‘to walk’ in third person singular.

Form	Gloss	Hawaiian Creole (English-based)	Haitian Creole (French-based)	Sranan (English-based)
Base	'he walks'	He walk	Li maché	A waka
Anterior	'he has walked'	He bin walk	Li té maché	A ben waka
Irreal	'he would/will walk'	He go walk	L'av(a) maché	A sa waka
Non-Punctual	'he is/was walking'	He stay walk	L'ap maché	A e waka
Anterior + Irreal	'he would have walked	He bin go walk	Li t'av(a) maché	A ben sa waka
Anterior + Non-Punctual	'he was/had been walking'	He bin stay walk	Li t'ap maché	A ben e waka
Irreal + Non-Punctual	'he will/would be walking'	He go stay walk	L'av ap maché	A sa e waka
Anterior + Irreal + Non-Punctual	'he would have been walking'	He bin go stay walk	Li t'av ap maché	A ben sa e waka

Yes, we're using the masculine pronoun here. Realistically, the data still suggest that this is going to be the case for masculine or feminine, especially with Sranan and Haitian, but there is change here too. I've given you the grammatical explanation of the verb conjugation, and I've also given you the gloss, which describes exactly what verb denotes. Notice 'he walks': *He walk, Li maché, A waka.* We see 'walk' in that conjugation—Hawaiian English is using it with no change, same with Sranan, and Haitian Creole has a simplified pronunciation of *il marche*.

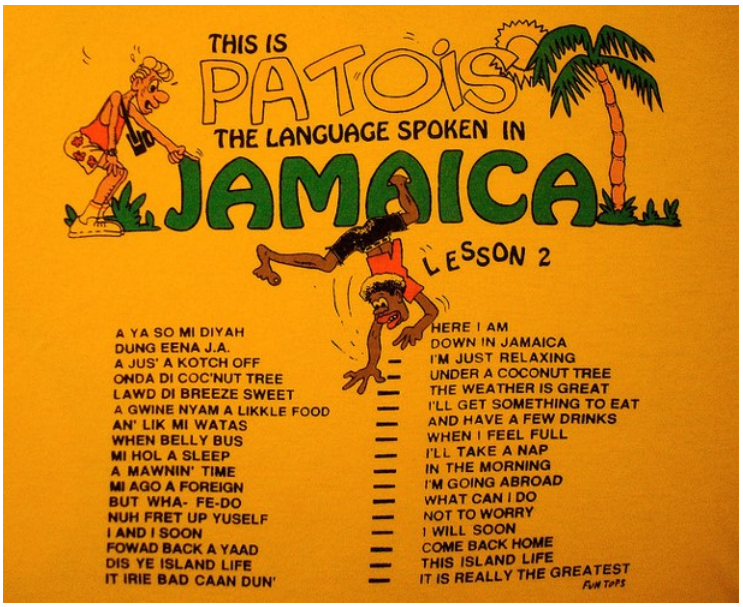
Notice that any tense, aspect, or mood that we would use an auxiliary for in English, you have it in Hawaiian creole, but you have modified or simplified versions, in some cases, and the same in Sranan. If you know, French, you will recognize some of the same thing here in Haitian Creole.

All three of these creoles have been in existence for over 200 years. At what point do they stop being creoles or lesser-than and start becoming languages? The best example of that is Haitian Creole. While it does have a couple of different variants, depending on how much French versus how much other substratum languages are being used, by large Haitian Creole has a standard mainstream form. It's the language of the Haitian government and the Haitian media, and it has been this way for over 100 years. Haitian Creole has existed for 250 years. Additionally, you have a full written version; it is taught in schools and has different dialects. So again, at what point is it no longer 'not a language'? At what point have we graduated it to become a language?

For those of us who believe that Haitian Creole is probably a language at this point, you could also argue the same for Hawaiian creole. Although it is dying out in favor of English, as you would imagine, but also Hawaiian. There is a concerted effort that has been going on over the last 40 years to reinvigorate the Hawaiian language. It is still considered an endangered language. The University of Hawaii in Manoa on Oahu is leading the charge and has been for 40 years on preserving Hawaiian. You still have Hawaiian creole, and they are still showcasing it so that it may continue to grow, but Hawaiian creole is being less favored, compared to Hawaiian English (the American English spoken in Hawaii) and Hawaiian, which is pretty interesting.

There are creoles around the globe; these are some of the areas that we see a number of creoles. Not shockingly we're talking about places that have been heavily colonized. It is interesting that creoles are not just based in English and French; we have Dutch-based creoles, especially in some of the Pacific Islands and as well as in the Caribbean. We have creoles in Latin America, which is not a place that most people think of. When we talk about that region, Belize would probably be the biggest example. Belize has an English-based creole, with the superstratum being English and the substratum languages are not the indigenous Maya languages of the region and some of these African languages from Central and Western Africa—because Belize was another processing point, as well as an escape point, for slaves. We also have Spanish-based creoles that are mixed with a number of indigenous languages, especially throughout the Andean region of South America. Additionally, the one that nobody thinks of is Portuñol, which is a creole of Spanish and Portuguese and depending on which side of the border, whether you're in Brazil or Paraguay, depends on which is the dominant language.

If I say 'creole', many people think of Jamaican Creole or Patois, as called in Jamaica. Here's another great example a creole that has been around for over 200 years—close to 300 years even—and is part of the mainstream. It is part of the government, the schools, and the media. Of course, when we're talking about music and culture from Jamaica, you can't not talk about Patois. I love Patois because it's a great example of trying to document from an early stage how a Creole came to be. This is a graphic taken from a lesson that you can find on on the Internet for how to speak Patois.



You have the standard English on the right side and then you have Patois on the left side and for most speakers of English that looks very foreign but if I say it, maybe it'll be a little easier.

Example 9.1

A ya so mi diyah, dung eena J.A. A jus'a kotch off, onda di coc'nut tree. Lawd di breeze sweet. A gwine nyam a likkle food, an'lik mi watas. When belly bus, mi hol a sleep. A mawnin' time mi ago a foreign. But wha- fe-do, nuh fret up yuself. I and I soon foward back a yaad. Dis ye island life it irie bad caan dun'.

A bit of translation and correlation:

- Here I am down in Jamaica: dong eena J.A
- Just relaxing: a jus' a kotch off.
- Under the coconut tree: onda di coc'nut tree.
- The weather is great; lawd de breeze sweet. 'Lord the breeze is sweet'.
- I'll get something to eat: a gwine nyam a likkle food. 'likkle', 'little'. Food. You got going, gwine.
- and have a few drinks: an' lik mi watas. Lick as in to lick it up.
- When I feel full; when belly bus; 'when belly bust',
- I'll take a nap; mi hol a sleep. 'Going to my hole to sleep.'
- In the morning; a mawnin' time.
- I'm going abroad; mi ago a foreign, so 'going foreign'.

And you can go on from there.

What I love about creolization and this process that languages go through as they connect and combine—especially as we start to document this and how it changes over time—is that you get a little window into the past. You can start seeing how maybe some of these standard or mainstream languages that we talked about—English, Spanish, Russian, Japanese, Mandarin, Swahili, Arabic, Hebrew—how maybe they came to be when they got confronted with other languages. It's important to realize that, as we stop focusing so much on hierarchical evaluations and when we just step back and observe language, creoles and especially well-established creoles, really are languages.

If we talk about creolization and we talked about languages in contact, the other piece of this puzzle is bilingualism. Many of you are bilingual—in my case, I'm a learned bilingual. I learned Spanish and I learned Italian, so I'm multilingual in my case. Many of you may be natively bilingual or natively multilingual, meaning you learned all of your languages while you were a child. It's important to understand that being bilingual is not a bad thing; it's not going to make you stupid or unintelligent, and it certainly does not mean that you're not capable of anything. I say that because that mentality still exists to this day, even here in sunny California, in the United States—in this progressive area, it still exists. It also existed and has existed well from time *in memorium*.

The reality is quite different, and we have started to change that narrative, beginning in the 1980s. Now, it is getting even stronger, because we can prove that bilingualism is not a bad thing. Rather, it's probably the best thing you could do.

Codeswitching, which frequently gets lumped in with so-called 'poor language skills', in fact is nothing of the sort. We're increasingly showing that those who code switch actually have a really strong grasp of more than one language in multiple areas. Code switching is something we do when we go back and forth between two languages, and we do it for a variety of reasons. Sometimes it's just what happens to fly out at that moment; other times there's a reason for it. For example, I like to talk about this phrase that we have in Spanish: the *ganas*, which is part of a longer phrase, *tener ganas*, 'to have *ganas*'. If you are a native Spanish speaker or a learned Spanish speaker, like myself, you've heard that term and you know that term. There isn't a good translation for it into English. It gets translated as 'motivation', sometimes it's 'the will to do something', if you have the *ganas* to do something, you want to do it. However, it's not really that. It's every fiber of your being wants to do it; or, if you do not have the *ganas* for something, every fiber of your being does not want to do it. How do I translate that? well I don't a lot of times if I'm talking to someone who speaks Spanish, even if somebody just learned a little bit of Spanish and that's what I'm feeling, then I'll just say, "I just don't have the *ganas* to do it today." Or if it's something I really want to do, like, I have a lot of *ganas* for this, I'll just switch it.

Education is both a hindrance and a motivator with respect to bilingualism. It is formal education that has been the main arena that most children have been forced to not be bilingual anymore, and some of you may have experienced some of that in your own lives or in your family's lives. Certainly, that was the case in my family with respect to Italian, and it was the case of my husband with respect to Japanese, and there are many similar stories. Yet, what we have shown over the last 40 years of research is that if you have a strong bilingual education system, the children that come out of that system not only are they very strong in two different languages, but they have critical thinking skills that far surpass anything that a monolingual student will have. What is more, if you are talking about one or both of those languages being related to the heritage of the child, then that is a huge positive boost for the child; they start to see pride and then feel that pride in their respective backgrounds. This is where that discussion on Ebonics also comes into play. It was so important to have educators recognize that the children who spoke Ebonics or African American Vernacular, that the worst thing you could do was to tell them they were dumb because they couldn't speak proper English. Rather, if you showcase them and said, "This is one dialect; let's get good at that. Here's the other dialect you need to know; let's get good at that," and then showcase when we use the different dialects with which groups of people with which context—when you do that, there's a lot of pride that comes in to the child. They grow up being a stronger person culturally. We've seen this now with African American Vernacular English and we have seen it in other areas as well. There have been studies recently about diglossia education with respect to Appalachian English. While for a number of generations people eschewed their Appalachian past and were ashamed of it, now people are starting to be proud of their culture and their history, at least as far as being proud to be a speaker of that dialect.

I know for me, I'll sometimes start sentence in Spanish, y termino en español.



This is pretty much what we all do when we codeswitch; those of us who are multilingual, we do this all the time. There's nothing to be ashamed about; it’s just what happens when you have two cultures, two languages, and a lot of positive vibes.

9: Languages in Contact is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

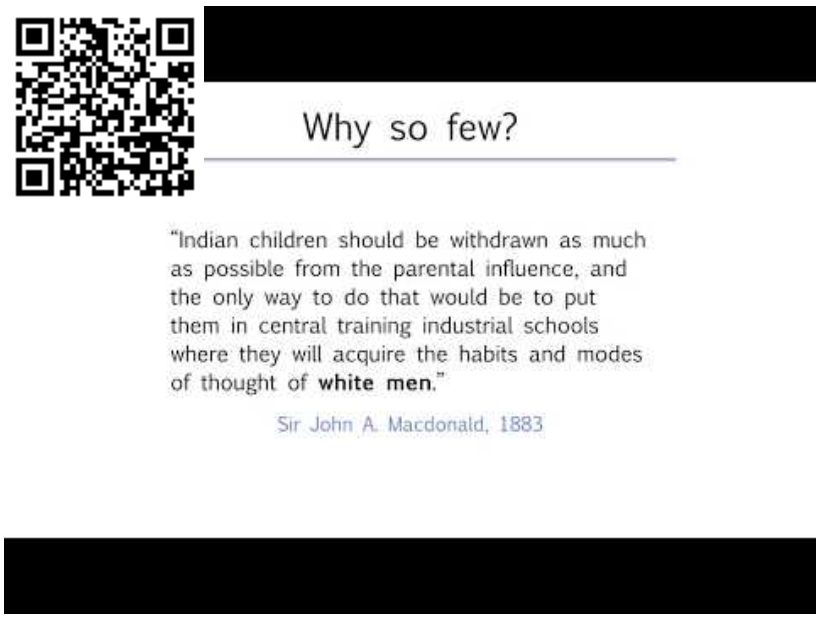
10: Indigenous North American Languages (Optional)

This section is optional, in that it won't be on any module quiz nor does it have to be incorporated into your term project. But it is an interesting look at how various indigenous North American languages work, as well as the efforts into recovering and maintaining them. Catherine Anderson did an exceptional job in this--it's an entire chapter, and I'm bringing it in here as a single 'page'.

These pages are excerpts from a conversation that Catherine Anderson had with David Kanatawakhon-Maracle, a Mohawk instructor at Western University.

11.1 Indigenous Languages and the Legacy of Residential Schools

Many of the Indigenous languages spoken in what is currently Canada are quite endangered, because of deliberate strategies by the settler government. For over 100 years, children from Indigenous communities were forced to attend residential schools where they were severely punished for speaking their home languages. The consequence was that fewer and fewer Indigenous people were able to maintain fluency in their languages.



Video Script

A lot of our attention so far has focused on English, which is convenient because it’s a language that we all know, but we can learn a lot about mental grammar by looking at other languages. Canada has an incredibly rich and diverse history of languages that were spoken by Aboriginal peoples long before European settlers arrived. Linguists estimate that there were more than two hundred different Indigenous languages spoken in this region, and these languages were quite different from each other — they formed about 15 different language families.

At the time of the 2016 census in Canada, there were still about two hundred and thirteen thousand people speaking about 64 different Indigenous languages from 12 different language families. Some of these languages, like Cree, Inuktitut and Ojibwe, are quite healthy, with thousands of speakers. But many more Indigenous languages are critically endangered — they have only a few hundred or a few dozen speakers who are quite elderly. When those speakers die, the language could die with them.

Why have so many of the Aboriginal languages been lost? It’s tempting to attribute it to economic and cultural pressures — TV shows and books and music are all in English, and everyone wants to speak English to get a job — but it’s not as simple as that. From the time that European settlers first arrived in this region, they engaged in deliberate strategies to try to eliminate First Nations people and their culture and language. The settlers engaged in war with the Indigenous people and brought new germs that caused devastating epidemics. The Europeans took over fertile land to grow their own crops and forced Indigenous people to live in small, confined reserves that could not sustain the crops to feed their people.

And these strategies aren't just from hundreds of years ago: between the 1960s and the late 1980s, the Canadian government seized thousands of Aboriginal children from their homes and placed them forcibly in foster homes and adoptive homes largely with white families, which meant that the children did not learn their parents' language. This forced adoption is sometimes called the "sixties scoop", and it continued the tradition of the residential schools.

The residential school system existed in Canada for more than 100 years, and the last residential school closed in 1996, not very long ago. Aboriginal children were taken from their families and forced to live in quite appalling conditions in schools that were run by the government and by the churches. The person who initiated the system was Sir John A Macdonald, Canada's first prime minister. He was quite clear that the whole purpose of taking children from their families was to make sure that they grew up without knowledge of their history, language, and culture. Here's his attitude about children who grow up in their families and communities:

"When the school is on the reserve the child lives with its parents, who are savages; he is surrounded by savages, and though he may learn to read and write, his habits, and training and mode of thought are Indian. He is simply a savage who can read and write."

And here's what his plan was:

"Indian children should be withdrawn as much as possible from the parental influence, and the only way to do that would be to put them in central training industrial schools where they will acquire the habits and modes of thought of white men."

He was completely open about his goals: he wanted Aboriginal children to stop thinking and speaking in the ways they learned in their families and communities, and to start thinking and speaking like white men.

In 2015, Canada's Truth and Reconciliation Commission issued its report, after years of consulting with survivors of residential schools. The executive summary begins this way:

"These residential schools were created for the purpose of separating Aboriginal children from their families, in order to minimize and weaken family ties and cultural linkages, and to indoctrinate children into a new culture—the culture of the legally dominant Euro-Christian Canadian society."

The TRC's Calls to Action acknowledge the crucial role that Indigenous languages will play in achieving reconciliation between Aboriginal people and the larger Canadian population. Here are just some of the Calls to Action.

- We call on the federal government to draft new Aboriginal education legislation including protecting the right to Aboriginal languages [and] the teaching of Aboriginal languages as credit courses.
- We call upon the federal government to enact an Aboriginal Languages Act that incorporates the following principles:
- Aboriginal languages are a fundamental and valued element of Canadian culture and society, and there is an urgency to preserve them.
- The federal government has a responsibility to provide sufficient funds for Aboriginal-language revitalization and preservation.
- The preservation, revitalization, and strengthening of Aboriginal languages and cultures are best managed by Aboriginal people and communities.
- Funding for Aboriginal language initiatives must reflect the diversity of Aboriginal languages.
- We call upon post-secondary institutions to create university and college degree and diploma programs in Aboriginal languages.

So there is a need for language preservation, revitalization and teaching. What can linguists do to help with these efforts? Of course the most important thing is to work with Aboriginal communities, to listen to the community members and find out from them what they think would be most valuable.

Some linguists have helped to document Indigenous languages, recording and transcribing speech and stories from native speakers. This is especially crucial when the speakers are elderly and the language is critically endangered. Documenting a language also involves doing phonological, morphological and syntactic analysis to be able to write grammar books and dictionaries for the language. Some linguists have also helped to develop writing systems for languages that didn't have any written form. If a language has been documented, then linguists can help to create educational resources and curriculum material that language teachers can use, and can help to train language teachers.

11.2 Preserving Mohawk

In this unit, Dr. Kanatawakhon-Maracle talks about how, for a language to survive, it must be widely spoken.



Video Script

I've been a language teacher for years, you know, and trying to teach and there's always been this segment of people out there who "really support what you're doing" and stuff like that and I've gotten into the habit of just ignoring them because their support is verbal; they're not in my classes; they're not learning the language. Real support for an Aboriginal language is getting out there and learning that language and learning to speak it, you know, so, to help bring that language back into its own. I don't expect any community to work towards, you know, sole monolingualism — that's that's just not doesn't make sense — however, bilingualism is a fairly normal way to be with a large percentage of the world's population. And, and, for, you know, my grandparents, my grandfather was bilingual, you know, and my great-grandparents were bilingual and they could use English when they needed to and they used Mohawk when they needed to and, or by choice or whatever.

I speak Mohawk and English so the thing is I can also, you know, use both languages. I … the difference I guess is that I also read and write Mohawk as well. So I'm a speaker and I'm literate which is the sort of thing that we would want to teach students especially at the university level because there's a lot of stuff written in Aboriginal languages that are presently not available in English or, you know or probably doesn't necessarily have to be available in English if they're speakers of the language.

People have always said you know, "Oh yeah, we know that the language should be in the home." No! The language should be in the street! If the language is surviving — if the language is truly an important part of being — it's in the street; it's in the stores; it's outside of the home. When you keep the language in the home it dies, because the speakers of the language eventually leave that home and then they go into the street where they're speaking English all the time and they meet somebody else who is also speaking English and eventually … the next generation is being raised by two English-speaking people and of course then the language is, is gone.

> [CA: Would you say that that attitude that says "oh the language is for at home," is that, is that another legacy of colonialism where it was shameful to speak an Aboriginal language?]

Yeah well let's keep it.

> [CA: Yeah, it's private, but not outside the house.]

Yeah and the real problem is that when when the language is only spoken in the home especially in contemporary society where people are, spending more and more time at home in front of some sort of technical device — in times past they were out going from home to home and all people were speaking the language and visiting and the language is very much alive — but once it becomes ensconced within the home and people get to the point where, sure they can talk to their parents, but they can't really understand their neighbours.

11.3 Learning Mohawk

Some Mohawk people have learned to speak Mohawk by growing up among fluent speakers. Many others are trying to learn to speak Mohawk in school or in university. Having some knowledge of linguistics can make learning a second language easier in some ways.



Video Script

One of the things about language is — I understand Mohawk and I speak Mohawk because I've heard hundreds of people speaking it.

[CA: Right. So that's how you learned it, was growing up in the community where it was spoken?]

Yeah I basically grew up with Mohawk and English and as I got to be a teenager I spent a lot more time with the older folks because they were more inclined to be speakers. Also, I find old people a lot more entertaining than younger adults, you know, they no longer have the sort of worries that, and the stress, that younger people have. And the stories, I mean, the really funny thing is that I find with older people, with the old folks, they're not very trusting, you know you basically have to visit them a lot before you kind of crack that shell and you get access to, to what they what they know. And I suppose to a certain degree that's self-serving but at the same time these people have got, they've got tradition to pass on; their responsibility is to be passing on this stuff and if they've decided that well nobody wants to hear that you know and so they stop telling it then it, it dies with them. So I found that spending a lot of time with the older folks you hear a lot of stories.

And the funny thing too about, one of the things that I found with the language is they were a lot more fun in Mohawk than they were in English. They get cranky and grumpy and unhappy when they speak English and I don't know why because, it just, I had an uncle used to visit and, well, boy I mean this he had a complaint about everything he had a gripe about everything and stuff and it always seemed to be so grating in English but when he was speaking Mohawk we spent a lot more time laughing. Because that was the language of his childhood and that was how he grew up and he's, you know, he would tell stories from that time where it's often you know were, were a lot more amusing or a lot more interesting than what he was having to deal with presently in English.

I've had students that have come in that have taken a number of years in the immersion program at Six Nations. They seem to have a sense of the language but they're not speakers. The course that I teach right now is a fairly heavy grammar-based course because I find that if I mean you can you can learn all the vocabulary you want, but if you have no sense of the grammar of the language, how are you going to utilize that vocabulary? What are you gonna do with it? And the thing is that the students that have come in that have some language, they know a lot of vocabulary… can't do a thing with it! They know how to say expressions; they know dialogues; they know a whole lot of things that I find is, okay, great, so you at least know the pronunciation, which is a good place to start. Anglophones can't get past what they see written. "Well, but that's a 't' it's written as a 't'." I says yeah, it's pronounced [d].

[CA: And here's where, actually, thinking of my students the having a bit of Introduction to Linguistics would help to say, well, look, this is a, this is an allophone and it's voiced in these circumstances and voiceless here, someone who has Intro Linguistics might get that.]

Yup. I had students who will comment on the fact that they've taken a linguistics course and it has helped their pronunciation it does make them more aware of it. […]

11.4 Mohawk Culture and Language



Video Script

What I also do too is that I include culture with the language. I'll be teaching them a particular word or phrase or expression but then I'll tell them where it comes from — why it is this way — why we say it that way — why we don't say this word. I mean, the word *nyaweh* in Mohawk gets interpreted as 'thank you' in English but that's kind of the beginning and the end of it. We don't, the reality is, if you follow older tradition, which is the way I was raised, you don't say *nyaweh* for every little thing; you don't use it the way it's used in English. In English it's just thank you thank you thank you — it becomes meaningless; it becomes a grunt, quite literally, in the English language, because people just use it so freely that it starts to lose its meaning. In Mohawk, *nyaweh* is used, or should only be used, between yourself and the Creator even when you say *nyaweh*, you know you see something beautiful, you see a sunset, a beautiful flower, nice majestic scenery or whatever, stuff like that, then you say *nyaweh* because now that *nyaweh* is directed towards the Creator and it's showing appreciation for what you're dealing with. When we sit at the table and we eat, the first one that gets up says *nyaweh*, not for the food, you know, but for, for the opportunity to sit with other people and share food. We're getting into the habit of using it much the same way it's used in English — you'll hear young, young speakers, more contemporary ones that are using or learning the language, they'll use *nyaweh* the same way they do and I said, No! (laughter).

[CA: Well I wonder, is there, is there a tension there that, so, on the one hand, you want to honour the traditions and the things you've learned from the Elders and from the older people and, on the other hand for a language to stay alive it has to change, right? Is that, so, if people are changing the language some, it's because it's still a living language…]

I don't know — there are certain things that we don't, we don't want to change, that we don't particularly want to update because then it starts to erode our uniqueness. If you're going to speak Mohawk the way you speak English, why don't you just speak English? You can update certain things but other things you can't. Like negating a future situation — in English, you can say, "Oh, it will not snow today!" How presumptive you are! (laughter) Because, just because the sky is blue, but there's a cloud and you know the clouds — if there's one cloud there's another cloud and another cloud and another cloud and by the end of the day we could see snow, which is all within the realm of the "will" because the "will" is in the in future. We cannot negate the future.

[CA: So that's, so that's a cultural attitude that shows up in the grammar of Mohawk? That you don't use negation with the future?]

You can construct a negative future. Nobody does. Fluent speakers don't. I mean, why would you? You know, because it interferes… We have other ways of kind of getting around it right but … most people just wouldn't. You just would not say, "it will not snow." We can create what amounts to it, a negative sort of thing and we use it with the non-definite, which is like saying "it would not" or "there's a possibility that it won't" It would not… but the thing is that you cannot directly say, "it *will* not" so we go to a very fuzzy sort of a non-definite situation and we negate that. Cheating in a way but at the same time, it is an important cultural part of the language. I mean, the fact that you have a culture that that doesn't negate the future — they deal with the future in a different sort of way — so those sorts of things, I think they have to be kept in language because they are the sort of things that add to the uniqueness of a particular language.

Word order in Mohawk. English has a set word order: subject-verb-object. Mohawk… Mohawk's word order is, is quite literally whatever comes out of your mouth. What joins it all together are pronominal prefixes and that works. But because the pronunciation of Mohawk, the pronunciation of a word in Mohawk is set; however, due to the situation in which that word may occur within, within a statement or sentence, the accent on that word may shift. So okay fine, so if I say *kahiatónhsera* for "book" then the accent is on *tón*. *Kahiatónhsera*, okay fine, but if I say *kahiatonhseráke*, "on the book" that accent shifted to the penultimate syllable. So accent shifts on a word depending upon where that word occurs. English is a language blessed with one or two syllable words which actually puts English speakers in an odd situation since most of them seem to have a hard time pronouncing a word that has more than one, more than two syllables, (laughter) which makes my name really hard for them, "Oh, Kanatawakhon, oh, I can't say that!" It's worse if they see it written.

But the thing is, in Mohawk, word organization, word position is dependent on emphasis, so if I want to say, "The boy is walking on the road," what am I saying?

"The BOY is walking on the road?" "*raksá:'a ire ohaháke*".

or am I saying,

"The boy is WALKING on the road"? "*ire raksá:'a ohaháke*".

Or am I saying,

"The boy is walking on the ROAD"? "*ohaháke ire raksá:'a*".

So I shift my words around, there's actually six arrangements of that, the three words and it's all depending on, on emphasis. Also dependent upon if you're answering a question. Because, "What did you buy?" "A COAT I bought." Because the question *what* is asking, is asking for information which is then placed first which puts it in an emphasized position. That's a very important part of the uniqueness of a language. So there are three very unique things with the language that we don't want to, we can't lose by modernizing it or contemporizing it. The language is set.

The culture that goes with the language … if you stop using a stone axe then eventually the word for stone axe is going to disappear unless for some reason … And then some vocabulary we've created in the past that we've carried through into the future like *oháhsera*, "a light" now is used primarily in reference to artificial lighting but originally it referred to something that looked very much like this bone *oháhsa* with the *–ra* suffix so then *oháhsera* just kind of gives the impression or gives the appearance of this particular bone and if you look at that bone and you look like a candle — yeah — so the thing is so we called the candles *oháhsera*, then lamps showed up. Well, more or less the same shape, *oháhsera*. Then lights, lamps, you know, living room lamps and stuff showed up, okay, *oháhsera*. Nowadays *oháhsera* refers to anything that throws artificial light, you know, ceiling lights, wall lights, the whole bit. So that is a word that has followed through time because even though we had your basic application but the shape kept changing.

[CA: That's a natural semantic drift that happens in most languages…]

Yeah.

11.5 Creating Materials for Teaching Mohawk

Dr. Kanatawakhon-Maracle has developed his Mohawk textbooks and exercises over his many years of teaching Mohawk, but when he started teaching, there were very few books or materials available. In this unit, we talk about the kind of linguistics work that can be helpful in creating textbooks and other teaching materials.



Video Script

Over the years I've been teaching Mohawk since '91 … '90-91 and I wanted to, I initially taught the course here, got the opportunity to teach the course because I wanted to see if the text material that I had developed for a language course would actually work. So I had developed a textbook and we used it the first year in that language course — mmmyeah, it did — as I modified it and I did things to it. That one sitting there that's the most recent within the last five years. There is an audio that goes with that on a USB stick. That's just the teaching text; there's another one that's called Supplements and it's divided into ten, ten supplement areas, where you've got everything about numbers, everything about locations, everything about… And it's equally as thick as that…

And then the band council back home, they finally got the opportunity to offer the Mohawk language in the public school there, so they wanted me to come home and teach it because I was quite literally the youngest speaker there and we didn't have a lot of old folks in the community to draw on. So I said okay fine, I'll, well I went back home and of course they said, "Okay there's, here's the Eastern school, here's the Central School, here's the Western school and then here's the main school one, Grades 1, Grade 2, Grade 3 and then 4 to 8 and I drove to each one every day for a half an hour, about a half an hour of language in each one, sometimes 40 minutes. So my first year of teaching was Kindergarten to Grade One — Kindergarten to Grade 8, you know of course having to deal with an attitude all the way through, worst at the 7/8 level because they started to be a lot more like their parents, and a lot more annoying. And then your little guys you know, just soaked it all up and were a lot of fun.

But then, the band council says well here's, here's your job, here's where you're teaching, and here's a hundred dollars a week. (laughter) No materials whatsoever, so I bought, I had to buy any materials that I needed for flashcards, for doing things within the class to help the kids learn and that sort of thing but one thing that there wasn't was a textbook. There was, there was no available materials that they were using and that was pretty much everywhere, so you pretty much had to develop your own material so I started doing more and I got thinking well the language is a lot more than just words. And there are a lot of words that seem to be very much the same.

I taught for five years and about the fourth year came across this book published by Günther Michelson called *A Thousand Words of Mohawk* and it was all about the roots. And I bought a copy of this thing and I started looking at it, "Oh, wow, this makes so much sense." So then I started, well maybe that is the better way to teach the language course.

[CA: So he had done the linguistics research to assemble the roots?]

Yeah, he was a linguist himself and of course, a lot of the linguistic work done on Iroquoian languages at the time was for the most part unreadable. (laughter) I didn't have the education to deal with all that sort of weird and wonderful vocabulary. If I got a book, I had some linguistic stuff on the language, but as long as it provided enough examples that I could, then I could figure out what they were talking about … but … I really needed to have the examples. So I started doing his stuff and then gradually working this into, to doing the class, classroom stuff.

[CA: So it sounds like there could be a valuable role here for people who may not have, who may not know languages, Indigenous languages but know some linguistics to work with speakers of the language to create materials?]

Oh yeah, I think, if they, and if they're going to be working with speakers they really do have to be somebody who has a sense of the grammar of the language. One of the things I learned how to do by trial and error was, learned how to ask the right question. Because speakers will tell you what comes to their mind. So we say, well what's the word for "tree"? "Oh, *kerhitáke*". Okay, eventually I learned that means "on the tree". *Kerhitákon*, eventually I learned that was "in the tree", and then, you know, they would give me all of these, *tkerhitoke*, "There's a tree standing there," and eventually I figured out that, *kérhite* was the word for "tree". Oh, yeah, yeah, you're telling us *kérhite*, yeah, that's "tree". (laughter)

But the thing is that, and when you're asking them something, you know, a question like, you know, "I trust him." You trust him for money? You trust him for what he says? You trust him for what he's doing? You trust them to get the job done? What? Because those are all different, you know?

And, and the business of using pronominals — we have a subjective, objective and transitive, and they would mix them. Now this is a problem that was happening in the, in the language programs in the schools, is because the fluent speakers were suddenly, oh Aunt Maisie there, she's a fluent speaker. Yeah, she's 85 but she can teach these kids — what a horrible thing to do to an old woman (laughter) — but anyway, she could use the money, so. But the thing is that she had, as a fluent speaker she had no sense of the grammar of the language.

[CA: Right, didn't have the metalinguistic awareness.]

Constantly mixing categories, constantly mixing, you know, mixing things up. "How come you said, *wahahní:no* yesterday and today you said *rohahní:no* for 'he bought'?" They both kind of mean that, *wahahní:no*, "he just bought it", *rohahní:no*, he bought it, but quite a while ago".

[CA: Yeah, that's something that my students in, in first-year linguistics struggle with making this, this unconscious implicit knowledge about how their language works and making that explicit. It's a real challenge. Whatever, I mean, we mostly do it in English but whatever your native language is, it's hard to become conscious…]

Yeah, so learning – learning to ask the right question, you know and, even when, when doing a sentence, you really have to pay attention to, to how they're organizing the sentence, how they're putting it together, and even though when they would say things, I would, "yeah, yeah, I know". I have a sense of what they were saying, but (laughter), it just, sometimes it was so confusing, sometimes very frustrating, sometimes you would ask three or four different speakers the same thing and they would all tell you something different. Is it because there are four ways to say the same thing or is it not really the same thing but simply refers to similar situations?

Now, I mean, that, that textbook there is all about the grammar of the language — what you use where, how you organize it, what you say when, and stuff.

[CA: And it's over your years of experience that you've assembled…]

Yeah. Over the years I've written five different language learning textbooks complete with, with exercises, drills and all the sort of stuff. That also has a book of exercises and drills to go with it. The difference is that is on, on audio, so it's … you can, you'll find the exact same textbook on screen which you can highlight the audio and get a pronunciation.

11.6 Speaking Mohawk and Reconciliation

In this unit, Dr. Kanatawakhon-Maracle shares his view that language will play a vital role in reconciliation efforts, and talks about some of the challenges of maintaining a language when English is so dominant in Canadian society.



Video Script

> [CA: I mean I was thinking about your students, like, are, are they going to speak Mohawk to their kids when they have kids, do you think?]

I think it's really up to, up to the student, I mean, they may have thought you know learning the language is difficult. Finding a compatible partner who also speaks the language is going to be the real test. And some of them have, have found partners who, and they have raised kids, they're raising kids together.

> [CA: So are there kids who are growing up who are learning to speak it as they're growing up?]

Yeah.

> [CA: That's starting to happen more?]

Yeah.

People don't realize that learning language is a lifelong… I, every now and then I'll run across vocabulary — oh yeah, wow this word — and then just … learning new vocabulary, words I hadn't heard before or words that I'd heard but I didn't have time to figure out the context and so, always in a state of language learning you know even after you're a speaker.

We need language. I don't know how we function without language. And nowadays with Native people you know, so much of this, this reconciliation thing going on and that sort of thing. Years ago in my home community, a lot of farmers that lived around the territory also spoke Mohawk. Very minimal in a way and stuff like that because they hired a lot of people from the territory to work on their farms. So they learned Mohawk. They learned it to a degree — you could go into stores in Deseronto and shopkeepers would, you know, would deal with, with the people in Mohawk. Now, it was Mohawk that would be related to the whole buying and selling and this sort of thing that but they did that. Now, to me, that's an aspect of reconciliation. That's where two groups have reconciled with each other — okay you're there, you speak your language but I will learn to speak with you mostly because I want your money — and … we, you know, we'll speak your language because we need your goods. I mean there's always a give and take on any, any two groups that have reconciled with each other but I think at the same time, too, when people take time to learn your language, they do have a certain respect.

> [CA: That's what I was thinking — it certainly shows respect, that I value interacting with you enough to do it in your language.]

Yeah, and the ones that want to interact more, learn more of the language.

11.7 The Future of Indigenous Languages in Canada

In this unit, Dr. Kanatawakhon-Maracle speaks eloquently about how vital each Indigenous language is to the identity of each First Nation, and about the long-term effects of colonization by the English and the French.



Video Script

Canada is a bilingual country … and I think if Canada learns to extend itself, we'll start including more and more Aboriginal languages. I think Canada is, would be better to tout itself as a multilingual country. Because I think when doing that, even if they, if their definition of multilingualism is, is the two founding languages and then the Indigenous languages, you're still looking at you know like fifty-five languages. And then because the difference, I think too, Native people have to be included in all that because we were here when English and French showed up. Okay, English and French showed up and they created the present-day institutions and stuff like that, okay fine. Everybody else who has come to this country basically has read the brochure and understood that English or French are the languages, therefore come to this country understanding that, okay I'm going to have to learn English or French because that's how the company is organized, country is organized. Or if maybe I'll go learn an Indigenous language as well. But the Indigenous languages have to be at the table.

> [CA: The English and French arrived and didn't say oh well we're gonna have to learn the language that people speak — they said we're here, now you're gonna speak our language.]

Yeah, but you know initially they did learn our languages to deal with us because they didn't have much choice. We had what they didn't have and when we got to a point where we no longer had what they wanted, and of course, they wanted the land which meant pushing us off anyway, so the respect for the languages and stuff, kind of went out the door. Then it became well everybody here speaks English only. … I find that we need we need to have the languages. I think that gives us a greater sense of who we are — the language!

And, and the thing is that for, for Native people in this country we have spent so many years under the colonial thumb and so many years being convinced that our own languages and our traditions and everything that's about us is inferior or not as good as … And the thing that I've found is that if you're a Native person you can work your butt off to become as much like you know the non-native Canadian; at the end of the day your skin is still brown and that's not going to change. If I focus on speaking Mohawk then in the process of learning my language I'm also learning Mohawk culture and what it is to be a Mohawk person and that's, I think is something that is very important. We've gotten into the habit of being Indians or Natives or Aboriginals or Indigenous. Nowadays the word's 'Indigenous'. I think I keep telling my students, I said you know, I said, when I was born I was born an Indian but then I became a Native and then I became Aboriginal and then I became First Nations and now I'm Indigenous, yay! You know, Indigenous is the word of the 21st century and I'm sure they'll find another one, but the word I would really like them to find and stick to is Mohawk, Oneida, Ojibwe, Chippewa. Know us by our nationalities, know us by our distinct cultures, know us by what makes us unique in the world.

10: Indigenous North American Languages (Optional) is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **11.1: Indigenous Languages and the Legacy of Residential Schools** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **11.3: Preserving Mohawk** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **11.4: Learning Mohawk** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **11.5: Mohawk Culture and Language** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **11.6: Creating Materials for Teaching Mohawk** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **11.7: Speaking Mohawk and Reconciliation** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **11.8: The Future of Indigenous Languages in Canada** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

11: References

11: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

8: Language Change- Historical Linguistics

1: Language Change Facts and Definitions
2: Reconstructions and Analysis
3: Indo-European and Romance Phonological Reconstruction
4: Stages of English
5: Morphological Stages
6: Syntactic and Lexical Changes
7: Typology
8: Language Families
9: Writing Systems
10: Arabic and Romance- A Comparison (Optional)
11: References

8: Language Change- Historical Linguistics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1: Language Change Facts and Definitions

Language Change Facts and Definitions, from Sarah Harmon



Video Script

Okay, folks: we're getting into the fun bit. Alright, for me it's the fun bit. Historical linguistics is my specific area of expertise, so this entire chapter of our text is all me, as Catherine Anderson doesn't really go into historical linguistics. I also should mention that this is not all my research; in fact, this is quite a number of pieces of people's research that are used to put together this chapter.

To set the scene a little bit, let's talk about some basic definitions and some basic facts. Up to this point, we've talked about those core areas of linguistics that we did earlier: phonetics, phonology, morphology, syntax, semantics, and pragmatics. We've already talked about sociolinguistics, where we focus on synchronic language change. **Synchronic** means 'at a given moment', usually tied to the current moment. When we talked about dialects and we talked about register and all those areas, those are all synchronic.

Now, we're going to move from the synchronic to the diachronic. **Diachronic** means 'over a long period of time', not one or two or three generations; now we're talking hundreds of years. This is an area that has always fascinated me because I love history and I love how language changes over time. One of my early influences to this area was within my own family, when my mother's family came over from Italy at the turn of the 20th century. At that time, Italy was technically a unified country, but to be very clear it wasn't. If you know anything about modern Italian history, you know that the reunification movements that happened, starting in the 1860s and through the 1880s and 90s, the country was not actually unified very much, not like when we think of like the Roman Empire. In truth, everybody spoke multiple dialects and, in fact, there were actual multiple languages of Italian spoken all the way through really the 1930s. That's when things changed due to a dictatorship (and that's a story for another time). However, when my great grandparents came to the United States, they were still speaking *lombardese* and *Genovese*—the dialects specific to Lombardy and the area around Genoa, which are both in the north west portion of Italy. They spoke very different dialects; in fact, at the time, you probably could still say that *lombardese*, as it were, might have been a different language, or certainly one of the Galo-Italic dialects. In addition, the Galo-Italic language was its own thing—the languages of the Northwest corner of Italy, and you can include Provençal and maybe even a little bit of Occitan, which were both spoken in the southwest portion of France. When I started learning Italian, I started learning standard Italian., which is based off of the Florentine regional dialect, and what I was learning was very different than what my two family dialects were—that has everything to do with history and with the change of language over time.

To give us a little bit of a background before we go into analyzing language change over time, let's talk about three important facts with respect to language:

Language change is not random. I know it may feel like it, and certainly as we go on to study how to reconstruct and how to analyze language, over time, some of these changes may seem random. Then again, every time you have learned a language, there

might have been aspects that may have seen to you to be random language change and language as a whole is not random, it is not we have proven that pretty well over the last 75-80 years of diachronic analysis. Rather, it is rule governed at all stages and across all stages, that means that there's always a reason for it, and at times, when we cannot find that rule or pattern, we say, “Okay, we'll come back to this,” and rethink this, re-analyze this later on when we have some more tools and, especially, if we come across more data. As linguists, again, we want to be objective, we want to describe what we see and part of that is saying that randomness doesn't really exist; there's always a reason.

Historical change always affects the components of a grammar, meaning every aspect of it; it affects the phonetics, the phonology, the morphology, the syntax, the semantics, and the pragmatics. You can say that it affects not just one specific dialect, but frequently affects others as well.

Historical changes also relentless. In fact, one of the banes of my existence is when I hear people say, “These young kids today, they don't respect the language like we used to do, like we do now. Our older generation really knows this language much better than this new one.” That just does not fly with respect to linguistics; that is not an objective, descriptive mentality. It's highly prescriptive, and it also flies completely in the face of what we know to be true, which is that all languages change. They all change because when they stop changing, they've died out.

Another term that you're going to hear is the term **proto-language**. A proto-language is a reconstructed, undocumented earlier version of a language; frequently, we’ll call it a **parent language** to another set. Focus on those two terms: **undocumented** and **reconstructed**. When we create something in linguistics, we are very careful to mark it as either a tested or hypothetical and, just like in the sciences, a hypothesis is an idea that needs to be tested and until it gets enough testing and data behind it. It remains a hypothesis, then it could become a rule or law or a phenomenon that is described when there are data behind it. When we're able to test it out, the same is true for proto-language or proto-languages; it’s our best guess based on the data that we have available and how we know languages in general, as well as within this family, with respect to change. You will hear me use the term **Proto-Indo-European (PIE)** quite frequently in this chapter, and that is because we have no written documentation of it. We have no direct, written data of PIE to suggest that this is definitely what the parent language of Latin, Ancient Greek, Sanskrit, Aryuvedic, Hittite, Old Germanic, Old Celtic, you get the point—we don't know what that actual parent language did. We have really good guesses; we have hypotheses that are backed up with data that we have found. But that is it; we do not have actual written documentation of Proto-Indo-European, and that's why we have that *proto-* prefix in front of it.

As we go through the rest of this chapter, these three facts must be kept in mind as well as the information about a proto-language, and the reason is this: we need to continuously be objective and descriptive and understand that this is a social science.

1: Language Change Facts and Definitions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2: Reconstructions and Analysis

Reconstruction and Analysis, from Sarah Harmon



Video Script

In the previous section, we started talking about a proto-language. In this section, we'll talk a little bit more about what reconstruction is and how we start to analyze historical linguistic data. This is just a primer; this is just the beginning. If you ever have a chance to take an historical linguistics course, usually it's an upper division course, and you're going to learn quite a bit more. For now, suffice it to say that this is your first step into being able to see diachronic change in action.

To start off let's talk about language reconstruction. What is this when we say that we reconstruct a proto-language? What does that mean?

It starts off with the **comparative method,** which is pretty much what you think it is: it is when you compare data that you can trust, and you come up with your best reconstruction of what the earlier language might have sounded like. Notice, I said, "sounded;" we tend to use the comparative method most with respect to sounds, phonetics and phonology. We can use it for morphology and syntax, although it's less used. We do compare and contrast, but not quite in the way that I'm going to show you. We don't really do this with respect to lexical reconstruction, or talking about how lexicon change over time, because that's a different animal altogether. When we deal with lexical reconstruction, or specifically historical linguistics focusing on semantics and pragmatics, that's a different type of analysis. Mostly we use the comparative method for sound, and the sound is going to help us understand the other parts as well.

When we look at a data set, we have a specific series of steps that we follow. The comparative method always focuses on an individual sound at each corresponding position. I'm going to walk you through a data set soon enough, but just to kind of walk through the process, let's say we are analyzing a language family. We have data from four different languages and the data include 25 lexicons. These are **cognates**, meaning that they are pretty close with respect to sound and meaning. For example, if you know, a Romance language, the word for 'water' is pretty close, if not exactly the same, across the Romance languages. In some cases, they sound exactly the same, and others well there's been a few changes to the sounds. Regardless, the word for 'water' is the same, and so, if we were to take *agua* from Portuguese, *agua* from Spanish, *acqua* from Catalan, *eau* from French, *aqua* from Italian, and we were to spread those across, we would be able to reconstruct the original lexicon. We start with that [a] first, that first vowel, and go across, and it's the same vowel for pretty much everybody, except for French—probably the earlier version also had an [a] at that same place. That consonant sound right afterwards is close but not exactly the same; sometimes it's a [g], sometimes it's a [k], and sometimes it has disappeared, like in French.

As we continue to reconstruct, we go across, analyzing each individual sound; we see where there's similarities and where there are differences. When there are differences, and specifically if there isn't a 'majority rules' situation—meaning all the data but one

show the same sound in the same place—if there's a majority, then we know the majority is probably going to win out. If you can't tell, then you use the **consonant hierarchy**. This is something I showed you when we talked about phonetics; I talked about how certain sounds are stronger than other sounds, and that we go from top to bottom on this list. In the case of *agua, acqua, eau*, and the other words for 'water' in the Romance languages, the big difference consonant-wise is that you either have a [g] or [k] or you have nothing, it's been deleted. We see that deletion at the bottom, so clearly, there used to be some kind of consonant and French just deleted it. Going between a [k] and a [g], you're going to go with the sound that is higher in that list, meaning that is higher in the hierarchy. In that case, voiceless sounds are higher than voiced sounds; [k] is a harder, higher sound then [g]. That means in the earlier version, probably, it was a voiceless [k] that existed and, in fact, that is what we see both in Classical Latin and then Vulgar Latin; that that consonant is a [k], it's voiceless, so we know that it's right.

With respect to vowels, it's hard to exactly reconstruct those, but sometimes we can. For example, in the Polynesian languages, most of them have between three and five vowel sounds, sometimes up to seven. Chances are, in Proto-Polynesian there were probably a few more vowel sounds that merged together; across human languages, we tend to see more merger. That isn't always the case, and I will come back to a really famous example in the next section. Vowels are very hard to do with respect to comparative method; it's much more suited to consonants, but sometimes we catch a break.

The comparative method is one way that we start reconstructing languages proto-languages, we start giving our best hypothesis for what the earlier version looked like or sounded like. However, it's not the only thing that we do in historical linguistics. One of the other areas that we focus on is **classification**. We'll talk more about classification when we get to topology, which is a further section in this chapter. When we try to see which languages might be related to other languages, there are certain core facts core data that we are looking for. The first is the numerals, and most specifically it's numerals one through 10 for mildly obvious reasons: most of us, typically, are born with 10 digits on our hands, so the numbers one through 10 are going to tend to stay pretty close; you're not typically going to borrow any number one through 10. We also look at nuclear family terms; as we talked about in semantics, family terms can range and sometimes those terms can be borrowed. Close family members—such as parents, children, and siblings, frequently grandparents (not always) and grandchildren—those terms do not tend to be borrowed from one language to another. There are exceptions to this rule; in fact, one really big one is many of the Filipino languages borrowed a number of family terms from Spanish. While I don't actually know the reason why, I would suspect it could be that the Spanish terms gave more specificity than what the then Filipino languages offered up so. For example, in Spanish, you have specific terms for 'aunt' and 'uncle', for 'cousin', for some of these other kinds of family terms, and it could be that in at the time of contact that many of the Filipino languages just didn't have as many specific terms. To be clear, that's hard to understand as to why that is the case. The third area that is really important with respect to genetic classification are the core elements of syntax and morphology. Notice it is specifically those two areas, and if we want to get very specific, it's how a given language embeds a clause or derives nouns and verbs. Why would that be? For example, English is a Germanic language. If you've ever learned a little bit about the etymology of many words in English, you'll know that we have borrowed a very large number, starting from the eighth and ninth century, when the Vikings started raiding the shores of the British Isles. The Angles and Saxons and Jutes, who were there and had established themselves, they were speaking a Germanic language, but it was a different Germanic language than what the Viking spoke, which was Old Norse. And so there were terms borrowed from the Vikings. Then, fast forward to 1066, when William the Conqueror came over from Normandy, which is in France, and he spoke Old (Norman) French—and Old French became the court language, which meant that a slew of terms for anything having to do with court life, government, aspects of a castle—the fact the term *castle*—all borrowed from Old French at that time. Fast forward a few more years to the Renaissance, you have a massive influx of Latinate terms that everybody shares. English has accumulated a very large number of lexicons from other languages, most of them Indo-European and well beyond with the colonization of various parts of the world. Yet the core makeup of English—specifically how we derive our nouns and verbs, how we rely on certain processes, how we create our phrases—that is very Germanic, and that has not changed in the history of our language with respect to external influences. Internal influences, yes, there have been changes and we'll talk about those coming up. For these reasons, we don't focus on the lexicon with respect to core elements of syntax and morphology. We look at the lexicon as a whole, and by that same token, we don't look at the phonology as much. Certainly, there are characteristics that are typical—the sounds that we have an English are very Germanic. We have a number of fricatives and a lot of vowels; that's very Germanic. However, we can't look specifically at the sounds themselves, because we have borrowed or modified so much over history.

The fourth piece, as I said, the external influences; we do have to look at the migration patterns, we have to look at how the given language has gone through the world. What I have here is the key migrant routes from Africa to Europe, because this explains not just how a number of these African languages have changed over time with their influence or colonization or being conquered by European groups, but also just in the major influences. When we look at the migration patterns and we see a consistent migration

pattern, we know that there's going to be some kind of influence of one language over the other. There's so much more to this part of the story, and we'll get to some pieces, as we go through these next several sections.

I could go on forever talking about classification, and I will come back to more when we get to typology. For now, just know that when we talk about a language family, they have quite a few pieces in common. This is how we start figuring that out.

2: Reconstructions and Analysis is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3: Indo-European and Romance Phonological Reconstruction

Indo-European and Romance Phonological Reconstruction, from Sarah Harmon



Video Script

Now that we have studied a bit about reconstruction, let's start talking about reconstruction in action with some data. This is going to focus primarily on the Indo-European languages and the Romance family specifically. There are a few reasons, not just because this is my area and I'm going to be able to explain this best. With the Indo-European language family, we have a very special set of data—it is the language family with the longest continuous deciphered writing data. We have a about 3,500 years of continuous deciphered data from a variety of sub-families. Technically speaking, the oldest writing systems are not in the Indo-European families. Certainly, the Chinese writing system has been around for about 4000 years, give or take, and that isn't even touching Egyptian Hieroglyphics, or especially what we consider to be the first writing system, the Sumerian writing system.

However, there are some problems: the Sumerian writing system is excellent and it got borrowed from the Sumerians to later generations, specifically the Akkadians, who spoke an Afro-Asiatic language, and the Hittites, who spoken an Indo-European language. We're still trying to decipher a lot of pieces with respect to just Sumerian, let alone any of the others. There's the unfortunate aspect of the Sumerian language being an isolate; to this point, we have yet to connect it to another language family or even subsequent languages. It's very difficult, so we have to put that one aside. With respect to the Egyptian Hieroglyphics, we definitely know how most of them were pronounced throughout their history; we've gotten pretty good data on that. As far as the oldest hieroglyphs, that's a little harder to decipher. We're not entirely sure how they were pronounced and so I'm going to put that aside as well. Finally, with respect to Old Chinese: no question that the ethnic groups that have spawned and continued to thrive in that region have had a writing system for a very long time. But we aren't entirely sure how everything was pronounced. While a lot of work has been happening in China with respect to that, and there is a good amount of research in that area, there's still a lot of controversy with respect to that data. For now, I'm going to also put it to the side; who knows, maybe in a few years, I will be able to include it. Therefore, I will stick with the Indo-European languages, and specifically Romance, which is my actual specific area of expertise.

We have not just the modern Indo-European languages, and not just the classical languages of Latin and ancient Greek. We have to include Sanskrit, which has a very long history, and Vedic or Aryuvedic, which has an even longer history of being written down, and then the granddaddy of them all, which is Hittite. Hittite is the earliest attested Indo-European writing; they borrowed the Sumerian writing system buccaneer form, but it is the earliest documentation and it has helped us to understand some of the potential sounds of proto-Indo-European.

Let's take a walk into the past, shall we?

This first slide shows a couple of different reconstructions that we can do. The top is Romance, and this is modern Romance. This should be in into international phonetic alphabet (IPA), but it's not, although most of you probably either have some knowledge of how the Spanish or French or maybe both are pronounced. There may be some of you who are Italian speakers, Portuguese speakers, and so you can pronounce them. I will go through them just the case.

When we reconstruct, we say, “Okay, let's look at each sound individually across the board, and let's see what the sound could be in the earlier parent of that language.” Admittedly, we are cheating a little bit with Romance because we have the parent is written down—I don't just mean Classical Latin, but I mean Vulgar Latin. There's really only a window of about 150 years that we have no writing, or precious little writing, in any Romance language. It really goes all the way back to Old Latin.

In reconstructing these sounds, we go across. The word for ‘dear’, like ‘a dear person’ or ‘a dear friend’: *cher* in French, *caro* in Italian, *caro* in Spanish, *caro* in Portuguese. Therefore, we know that that first sound in each modern word is a [k, k, k, ʃ]. One of these things is not like the other. (Come on, you know that song, right?) Because of those data, we know:

- That first sound is probably going to be a [k], because three out of the four languages, the data here suggests that that is the sound.
- The next is the vowel: [a, a, a, ɛ]. Again, one of these things is not like the other, so probably it's going to be [a] because that's pretty close.
- As far as the R is concerned, I don't have them written out here and IPA but technically speaking. Spanish and Italian both use that alveolar tap, that real quick flick that hits the alveolar ridge. If you speak Spanish, you know what this is; Italian has the same sound. It's not like the trill; that's not the term that we use. What is interesting is that both French and Portuguese, depending on the dialect, that ‘r’ rhotic, has intensified and velarized; it has gone to the back of the mouth, and almost always it's the velar region. It is either a trill or a kind of fricative, depending on the dialect. For right now, I’m going to actually keep it as an alveolar trill, because there's so much play and the alveolar trill is higher up on the hierarchy than the others.
- This final sound—the [o, o, o], deletion—clearly, French is not going to represent the old pronunciation. It is going to be a lot closer to what we hear in Spanish or Italian.

We can do the same thing for the term for ‘field’. If we go across:

- [k, k, k, ʃ], so we know that that first sound is a [k] because three of the four have it.
- [a, a, a, a], okay all four languages show that it's the same sound, so it's probably going to be an [a].
- we have an [m] sound, and then we have a [p] sound at all for the languages
- again, three out of four and in a vowel is probably an [o].

The word for ‘candle’, you can do the same thing, right?

- Three of the four have a [k];
- All four have an [a];
- All four have an [n];
- The next sound is almost certainly some kind of a [d]-like sound;
- There's a lateral liquid;
- Three out of the four have an [a], the fourth has an [ɛ], so the chances are that it's going to be [a].

The great news is, as I said, these are Romance languages, and we actually know that this was how both Classical Latin and Vulgar Latin sounded, so this is a pretty good clue. In fact, many of the Indo-European languages are frequently test runs for any historical linguistic theory; if you can prove it on any of the Hellenic family (that's Greek), anything with respect to Romance to Latin, and if you can do it with Sanskrit (Hindi or Gujarati and the languages in northern and central India, Pakistan and parts of Afghanistan). Those languages have such a long and rich attestation, and we have so much data for those three language families in Indo-European alone. If your theory holds water there, it might hold water elsewhere. This is one of the reasons I got into Romance linguistics; I wanted to be able to test out theories. If a theory can work to explain something in the Romance languages, then it might be a pretty good theory for how languages change. If not, well, we might have to find something else that might work.

With respect to Proto-Indo-European, as I said, there is a reconstruction based off of the oldest attestations of the Indo-European languages—not just Classical Latin, but Old Latin, for which we do have good records, as well as the various ancient Greek language, including Attic, especially as we get to know more about Mycenean, which is written in Linear B. The more we start understanding Hittite, the best data seem to come from there, because it is the oldest attestation of an Indo-European language. As we get those languages together, and as we start comparing and contrasting and seeing how these correspondences work, we can start building how Proto-Indo-European might have sounded like.

Here are some correspondences. if you think of the term for ‘father’. Remember that English is a Germanic language, so ‘father’ is frequently with that labiodental fricative. *Pater* in Classical Latin—actually, most attestations of Latin—and notice that the word for ‘father’ in most of the Romance languages, if not all of them, begin with a [p]. In Sanskrit, it's very similar, and if you speak Hindi, Urdu, Gujarati, or any of the Indic languages, you know that your term for ‘father’ is very similar to *pitar*. What we're going to say is that the Proto-Indo-European sound is almost assuredly that same [p], that there's a sound change that happens in the Germanic languages. As you go around, the word for ‘three’ and most of the other languages do not begin with that interdental fricative; it begins with a [t]. Chances are in Proto-Indo-European it probably also started with a [t].

When you see consonants that have an ‘h’ afterwards, if you speak an Indic language or even an Indo-Aryan language as a whole—Farsi still has some of the sounds, too—these are what we like to call breathy sounds. It's not just a [b], but there's a <glottal vibration/exhale> with it: [b̤]. They tend to be voiced obstruents, almost always stops, but you can have some fricatives and affricates in there. We see them in Sanskrit, Vedic (Farsi’s much older ancestor) and we see something similar in Hittite. Therefore, the chances are that Proto-Indo-European might have had it, too.

One more thing to point out: notice that the Proto-Indo-European sounds all have an asterisk in front of them. If you recall when we were in syntax, we said that the asterisks meant that it's ungrammatical, that it doesn't follow the phrase structure rules for the language. There is a different meaning here—arbitrariness, right? In historical linguistics, we use the asterisk to mark that something is reconstructed. This means that we do not have evidence or data of it being that way, but based off of the data of the child languages, the chances are pretty good that this is the sound that was in the proto-language.

When we talk about phonological languages or phonological changes as a whole, we're going to talk about some of these things that we saw way back when we talked about phonology. The same rules apply; in fact, most of what we saw with respect to phonological change all the way back in chapter three is due to an historical process, meaning there's some change over time that has happened. Whether we're talking about loss or deletion, apocope is part of that, insertion or epenthesis, assimilation and dissimilation rules and metathesis—all of these are going to come back in a big way. We know that these are processes that happen, not just within one generation, or maybe spilling into a second, but we're talking over 3, 4, 5, up to10 generations where we see the change happen. Most language change is gradual; it is not instant by and large. There are some exceptions to that rule, but we'll talk about that at another time. I’ll show you a couple of phonological changes that we see in the Indo-European languages, and I’m going to start with Grimm’s Law—by the way, Grimm is Jakob Grimm, as in one of the Grimm Brothers. We know them for collecting fairy tales throughout Central Europe. Jacob Grimm was also comparing languages, specifically the languages that were spoken throughout Central and Eastern Europe, as they were going out into the countryside and collecting tales. What he ended up showing was a little bit about why the Germanic languages sound different than the other languages of Central and Eastern Europe—Slavic languages, Romance languages, and the like. He came up with some interesting facts. We have subsequently found more data, and we still call this Grimm’s Law, even though he didn't fully flesh it out; he gave us the idea and we attribute it to him. Specifically with respect to Grimm’s Law, focus on this top right corner of the graphic.

Grimm's Law

$b^h \rightarrow b \rightarrow p \rightarrow \phi$
$d^h \rightarrow d \rightarrow t \rightarrow \theta$
$g^h \rightarrow g \rightarrow k \rightarrow x$
$g^{wh} \rightarrow g^w \rightarrow k^w \rightarrow x^w$

Proto-Indo-European	Meaning	Non-Germanic (unshifted) cognates	Change	Proto-Germanic	Germanic (shifted) examples
*pṓds	"foot"	Ancient Greek: πούς, ποδός (poús, podós), Latin: pēs, pedis, Sanskrit: pāda, Russian: под (pod) "under; floor", Lithuanian: pėda, Latvian pēda	*p > f [ɸ]	*fōt-	English: foot, West Frisian: foet, German: Fuß, Gothic: fōtus, Icelandic, Faroese: fótur, Danish: fod, Norwegian, Swedish: fot
*trit(y)ós	"third"	Ancient Greek: τρίτος (tritos), Latin: tertius, Welsh: trydydd, Sanskrit: treta, Russian: третий (tretij), Lithuanian: trečias, Albanian: tretë	*t > þ [θ]	*þridjô	English: third, Old Frisian: thredda, Old Saxon: thriddio, Gothic: þridja, Icelandic: þriðji, Danish, Swedish: tredje
*ḱwón- ~ *ḱun-	"dog"	Ancient Greek: κύων (kýōn), Latin: canis, Welsh: ci (pl. cwn)	*k > h [x]	*hundaz	English: hound, Dutch: hond, German: Hund, Gothic: hunds, Icelandic, Faroese: hundur, Danish, Norwegian, Swedish: hund

It's the concept of Proto-Indo-European [b, d, g] and [p, t, k], all changing. It's a specific change within the Germanic languages, because we see a number of shifted sounds—and by shifting, I mean they go from voiced and breathy to voiceless—and then they get fricated—meaning they turn into fricatives. The term for a 'foot' and you see here: *foot*, *foet*, *fuß*, *fōtus*, *fótur*, *fod*, *fot*. All the term for the thing at the end of your leg that you stand on. Notice that in every other Indo-European language, almost without exception, that term starts with a with a [p] sound: *poús* or *podós*, *pēs* or *pedis*, *pāda*, *pod*, *péda*, *pēda*. Pretty important work. It should be noted that Latvian is thrown; Latin and Lithuanian are both Indo-European languages, although they have been heavily influenced by their Finno-Ugric neighbors, especially Estonian but even Finnish. Yet, their core vocabulary tends to be Indo-European.

Why do we have these languages up here? That is what Jakob Grimm looked at; he compared all of the Germanic languages. Most of those you probably have heard before. West Friesian is part of some of the islands off of Scotland; the Faroe Islands are between Scotland and Iceland. The others you probably recognize. He went and compared them to Ancient Greek and Classical Latin, because he knew those languages, and he had access to Sanskrit. Given his status, he would have had access to it. Grimm knew plenty of Russians, and he came across those who spoke Russian, Lithuanian and Latvian, even included some Albanian, and for kicks and grins included some Welsh. All of these cases—the [f] sound, the [θ] sound, and the [h] sound that we have in the Germanic languages—many times have more obstruent counterparts. They were more occlusive; they are stops in almost all cases in the more modern Indo-European languages.

To take it a step further, what Grimm hypothesized was that there would be in a parent language something a little bit harder, probably a little bit more involved, and so what we see is that in the Proto-Indo-European, we see that it is occlusive, it is a stop. It's also a very hard stop; it has softened over time, and softened even further when we get to the Germanic languages. The Germanic languages showcase a really great example of what a language family describes: this tendency to exist because there was an initial group or speech community that pulled off, got established elsewhere, and then spread out on their own. Think of

migration patterns, as they are the same thing. In the case of the Germanic languages, what we have learned subsequent to Jakob Grimm and his work, is that the Germanic languages, the Germanic peoples seemed to have broken off later than other Indo-European groups. In doing so, they started interacting with the other peoples in and around Eurasia. Because of this late separation and intermingling with other linguistic communities, their patterns are is going to be very different than what everybody else ended up doing with respect to the sounds.

That's a little bit about the Germanic languages, let me show you a little bit about the romance languages. Again, this is a little bit of my specialty and I have to give a shout out to one of my professors at UC Davis, Robert Blake, who was mentored by one of my professors at the University of Texas who also became my mentor, Carlos Solé.

When we talk about Romance languages and we talk about palatalization—remember that term? That the sounds are going more towards the palatal region of the mouth—if you think about the different Romance languages, there's a lot of palatalization. It's not exactly the same palatalization rule; we have different forms of it. What we see in the history of this language family is different waves of palatalization. I've put up here Spanish and Italian, and then I've given you what we have in Classical Latin, but slightly modified, and that's what we're going to see with Vulgar Latin. Vulgar Latin was the language of the people. It wasn't vulgar, like full cuss words; Vulgar Latin just means it was the language of the people, what people tended to speak, instead of Classical Latin, which was studied and written, but not really spoken much once you get to the end of the reign of the Latin kings—before the Republic. By the point that we get to the Republic, even Classical Latin is not really spoken that much anymore; certainly, by the time you get to Julius and Augustus Caesar, Classical Latin is not spoken hardly at all—maybe only by the Emperor and that class, maybe a couple generals, but that's about it.

The 'yod' palatalization from Latin into the Romance languages

Classes of yod		**Effects on various vowels**				
1st yod: oldest change; doesn't affect vowel	ty/ky → [t:s, z, ɵ, tʃ]	fortia → Sp. fuerza; It. forza	pettia → Sp. pieza; It. pezza	lutea → Sp. loza It. ---	malitia → Sp. maleza; It. ---	minacia → Sp. amenaza; It. minaccia
2nd yod: influences open vowels only	ly, k'l → [l: à x/ʎ]	folia → Sp. hoja; It. foglia	reg(u)la → Sp. reja; It. ----	cusculiu → Sp. coscojo; It. ----	cilia → Sp. ceja; It. ciglia	palea → Sp. paja; It. paglia
	(mnà) ny, gn → [ɲ, n:, n]	somniu → Sp. sueño; It. sonno	ingeniu → Sp. i/engeño; It. ingegno	cunea → Sp. cuña; It. cuneo	ligna → Sp. leña; It. legno	aranea → Sp. araña; It. ragno
3rd yod: regularly affects open vowels, sometimes closed vowels	gy/dy → [j, dʒ]	podiu → Sp. poyo It. poggio	pulegiu → Sp. poleo It. puleggio	fugio → Sp. huyo It. fuggio	fastidiu → Sp. hastío (It. fastidio)	exagiu → Sp. ensayo It. saggio
	by/my → [j, bj/b:, mj/m:]	fovea → Sp. hoya (It. fovea)	nerviu → Sp. nervio; It. nervo	rubeu → Sp. ruyo/royo It. ----	vindemia → Sp. vendimia; It. vendemmia	labio → Sp. labio; It. labbro

Classes of yod		Effects on various vowels				
4th yod: regularly affects vowels	kt, ks → [tʃ/t:, ʃ à x]	nocte → Sp. noche; It. notte	lecto → Sp. lecho It. letto	tructa → Sp. trucha It. trota	strictu → Sp. estrecho It. stretto	facto → Sp. hecho; It. fatto

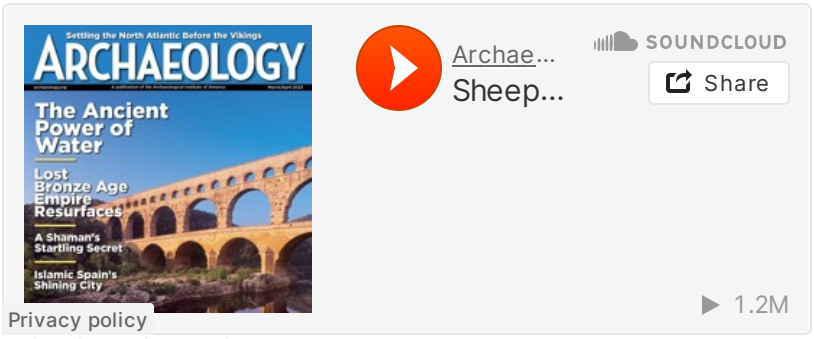
We see four different waves, and you see this term **yod**. This is actually a letter of the Hebrew alphabet, but, more specifically, it has this [j] quality to it in Classical Hebrew. Whenever you had that letter written, you knew that the sound that came before it was palatalized somehow. That's why that term was borrowed, to explain what happens in the Romance languages; this does not mean that Hebrew has any effect on the Romance languages, at least not in this way. It also should be noted that these four waves all happened before the fall of the Empire, so this first yod actually starts during Imperial times. This is the oldest change; it doesn't produce any effect on the vowel. There were seven vowels in Classical Latin, and they were both long and short. Realistically, the only modern Romance language that has anything similar to that is Sardinian; most all of the other Romance languages have five vowels, and they sound very different. To explain this table, we're talking about when you have a [t] or [k] next to this yod, and this is almost always a high vowel or a front vowel. In this case, *fortia*, which means 'strength'; notice that is next to a [i] sound, so there's your yod that's the palatalization effect. We have *pettia*, *lutea*, *militia*, *minacia*. If you speak Spanish, you'll recognize most, if not all, of these terms; some of them are older terms and may not be used very often. I'm giving you the Italian because there's a slight difference in how Italian palatalized over Spanish; if I had more room, I would include Spanish and Portuguese, probably Romanian as well, just to give you an idea. We'll start with *fortia* because the term *force* that we have in English is borrowed from French, and it comes from the same root. In Spanish, this becomes *fuerza*, so this [t] softens to a fricative, and may, depending on your dialect even front a little bit to [θ]. If you're in Spain, this is not [fueɾsa] in Spain, this would be [fueɾθa]. In Italian, it's [foɾtsa] so becomes a affricate [ts]. You see that here: *amenaza* for *minacia*, which is a 'threat' in Classical Latin, and [amenasa] or [amenaθa], depending on your dialect in Spanish, same thing; [minatʃa] in Italian, same thing. Notice it's a [tʃ] sound, not a [t] sound. A little bit later, maybe four to five generations later, you get the second yod, where you start opening up the vowel a little bit, and you have more consonants that are affected by it. You have the [l, ɲ], and the 'ng' together. Think, for example, the term for a 'leaf' or 'sheet' is a *folia* in Classical Latin; notice that I'm palatalizing that a little bit. In Italian it's [foʎja]; it's a fully palatalized sound. In Spanish, it's an [oxa], so that [x] is an even further palatalization. Yes, I know there's an 'h' there and that's another discussion. Yet, what you can see is the various progressions of this yod and this palatalization rule.

With respect to Grimm's Law and the Romance palatalization rule, you do not have to remember the individual steps at this time. In an upper division historical linguistics course, that's a totally different thing. What is important to understand is that these changes happen, that they are relentless. In fact, even if we go back to this first y'all and we think about the Spanish pronunciation that pronunciation has changed, and is still changing and continues to go on. For example, the pronunciation of the written 'j' in Spanish depends greatly on what dialect you're speaking.

When we talk about Proto-Indo-European, what we're really thinking at this moment—and I say 'at this moment', this is a map from 2007 and reflects the latest that we know with respect to historical linguistics and specifically the Indo-European group—this family started most likely in this Ukrainian Steppe region, near the Balkans and Greece. It seems to start with the Yamna culture or Kurgan culture. 'Kurgan' refers to the fact that they had burial mounds; it's a really specific type of burial that you learn about in anthropology. We know that this culture existed 4500 to 2500 Before the Common Era (BCE)—we don't say 'BC'/'AD' anymore, we say 'Common Era' (CE), 'Before Common Era' (BCE). This Kurgan or Yamna culture, we're talking for 6000 years ago, Proto-Indo-European is what we believe, to our best estimation, was spoken between 3000 and 5000 years ago. It could be older, it could be a little younger, but that's our best attempt at dating it right now. We do know that there are early splits so the earliest split would be to the Afanaseco culture, which is in Central Europe, or excuse me, Central Asia, and then the Maykopf, which is down to the south. We know that there's also the Anatolian branch, which splits off very early; this is where the Hittites are. We know that that branch was existent at least 1000 to 2000 years before the Hittite's kingdom comes to be, so we believe it to be about 4000 BCE. It's in the second millennia BCE, about 1700 BCE, that we start seeing the Hittite kingdom. Notice that we have multiple expansions into Asia; it's not just Europe. Remember, this is the Indo-European language family. This family also includes the Tocharians, which I still find to be the most fascinating group because we still only know pieces about them. This language family was in Central Asia and Western China; the Wusun are kind of in that realm, as well. These two families had a kind of writing system, but we are still trying to decipher it. The Ancient Chinese writings very rarely talk about them, so there's a lot more that we still have to learn; I'm hoping that there's some archaeological dig yet to be found that has more data.

When we're talking about the Indo-Aryan languages, we're talking about Persian or Farsi and its ancestors that's the Aryan branch—Vedic or Aryuvedic is part of that, as is the Iranians who are a different branch. The Vedic branch is where you get Sanskrit and then the subsequent Indic languages. Armenian is an Indo-European language—people forget about that, but it is—and then you have the various branches in the early first and second millennia BCE, you notice that there are various splits.

Below, I have included two SoundCloud links, examples of what we believe Proto-Indo-European sounded. They are a couple of stories. This is a really fascinating just piece to hear. Again, it's hypothetical, but our best guess so far.



Archaeology · Sheep And Horses



Archaeology · King And God

3: Indo-European and Romance Phonological Reconstruction is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4: Stages of English

Stages of English, from Sarah Harmon



Video Script

Let's start talking a little bit about why English is so weird—okay, not exactly weird, but you get the point. When we talk about why English has a very different setup, one of the big areas has to do with the vowels. We tend to say that vowels change but not radically over the course of the history of the language. English does not follow that rule; in fact, it's one of the perfect examples of not following that rule, and that is the Great Vowel Shift. Yet, a lot of that has to do with how English came to be in the first place, so let's talk about the various stages of English, what that means, and specifically get a little bit into the Great Vowel Shift.

When we're talking about English, we're talking about Old English, Middle English, and Modern English. Despite what you might think when you try to read Shakespeare, that's actually Modern English. Before that, if you want to know what English used to sound like before the Norman invasion, read *Beowulf*. *Beowulf* is **Old English**; in fact, *Beowulf*, when you look at it, it is very similar to Old High German, which is, if you will, the parents or ancestor to Modern German. It is not easy to read at all; in fact, if you know German you might pick up some words, maybe, but, honestly, not much it's very, very difficult to read. Old English has a full case system, it is highly complex with how it combines its terminology, it is very synthetic and realistically, very fusional. Everything changes in 1066, and if you know the history of the British Isles, you know what happens: William the Conqueror leaves Normandy, which is the northern part of France; the Normans spoke Old French—although interestingly, they were genetically more Viking than French—so they brought Norman French to the Court. This meant that every part of the upper half of English society or British society, if you were in middle class or higher you were learning and you were speaking French. You would be bilingual, especially in the middle classes, in both Norman French and the English that was being spoken. However, there's a big change that happens; if you were to read, for example, *The Canterbury Tales*, that is **Middle English**—in fact, in the middle of Middle English. It really encapsulates what happens when you have two mildly, but not really, related languages pushed together and combining. If you think back to sociolinguistics, when we talked about languages in contact, we talked about of lingua franca that goes to a pidgin that goes to a creole. There are a number of researchers and linguists who say that Middle English shows a number of examples of creolization:

- The case system reduces drastically;
- The vowels change;
- The word order becomes stricter; and,
- The language becomes less synthetical and little more analytical in how we combine our morphology and syntax.

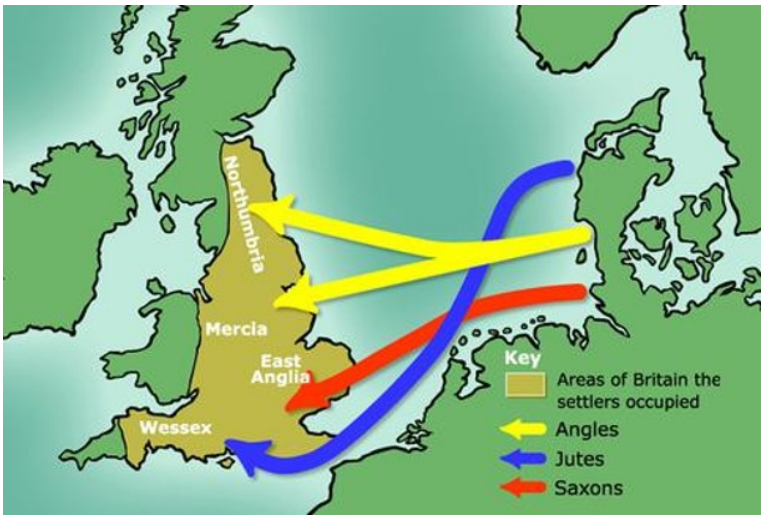
If you were to compare a page of *The Canterbury Tales* to a page of *Beowulf*, they would be different, miles apart. That is why there is a thought that maybe this stage of English is a kind of creolization. Personally, I'm not going to entirely subscribe to that theory; it is an interesting theory, and I like pieces of it, but I think there's plenty of evidence to the contrary of that. Suffice it to say

that there are certain major changes that happened with Middle English. Notice that's only a 400-year period, yet there is approximately 600 years of Old English. OE was a pretty stable language—not is not 100%, but pretty stable, you can see it through the history. That 400 years of Middle English is raucous; there are a number of changes happening. By the time you get to the Elizabethan era, now you're talking Shakespeare; this is Modern English. **Modern English** is usually set at 1500 CE, to say that that is when you start seeing a language that resembles the modern English language: there's no case system anymore. Word order fully replaced everything in the syntax. Shakespeare's difficulty is legendary, and I will not say that it's not a difficult thing to process and read—in part, it's because he moved stuff around to make sure he had the right rhythm, the iambic pentameter. However, if you were to read a document written in Elizabethan English prose, you would get a large chunk of it; you may not get everything, and in the last almost 500 years there's been some changes, but realistically it's pretty close. It's also about at that point that you start getting much more dialectical spread—clearly, if we think about the history of England and, right around that time is when you start getting this concept of Great Britain, in addition to colonization of the world (first of the Americas, and then throughout the world). One of the things that we see with respect to human language is that as we colonize, there's more stabilization in certain realms with respect to language, and that probably is tied into the stabilization of **Early Modern English**. When we talk about Early Modern English, we're talking about think Shakespeare. When we're talking about **Late Modern English**, which is 1700 CE and beyond, now we're talking more modern stuff; think Charles Dickens, which is difficult at times to read, in part that's because of what Victorian prose was like. Yet, the syntax, semantics, and morphology hasn't changed much from the time of Shakespeare. There have been changes—specifically, with respect to semantics, of course, but pronunciation of the vowels has changed, as well as the consonants. It's not that different.

What you're going to hear below are examples of all four really of these stages and you'll hear the difference between them. It's also really important to remember with respect to English, we have heavily borrowed from the Romance family. I tend to say Italic family because Latin is in there; Latin is the parent of the Romance languages, but it is not itself a Romance language. It is an Italic language, so one step up, therefore I tend to say either from Romance or Italic origin. It started really with 1066 CE, the Normans coming in and taking over the English court, but it continues through all the way into the Renaissance and beyond. For example, the language of science is frequently Latin; by that I mean when we create a scientific term, we're using either Latin or Greek as a root. There's been a lot that has been borrowed in, and not just to English, but to all of the Indo-European languages. This goes beyond and into technology as it goes forward. Note that, if you take the 1000 most common words in English, 30% of it are Latinate or Italic or Romance. This is only the lexicon; remember what I said a couple of sections ago that when we reconstruct the language and we try to classify it within the realms of what family it's in. We're going to look at syntax and morphology more than the sounds, and especially more than the lexicon.

When we look at the syntax and morphology of English, it is Germanic; the only thing we don't do anymore really is case—we're the only ones that don't, although Dutch has a reduced case system—and we don't switch the word order of subordinate clauses. If you go to learn another Germanic language, you'll know that if you have an embedded clause, the word order switch becomes SOV, that the verb is at the end. English stop doing that by Modern English, and it was already starting to phase out in Middle English.

When we're talking about the waves of migration, this is just a quick map to show you this.



With respect to the origins of Old English, we really have three main waves: we have the Saxons and the Angles, and they are predominantly the first groups to come over. The Jutes came over and went into southern England, to what is now considered southern and southeast England. If you know anything about the various dialects of England, you'll notice that the dialects of Wessex have certain examples that we do not see in the rest of England, and that predominantly is the influence of the Jutes. What's not represented here are the Vikings and, as I said, the Vikings came from what we now call Scandinavia: Denmark, Norway, and Sweden. They raided all of the British Isles, such that it's not only Old English that got affected. The Celtic languages—Scottish Gaelic, Manx, Irish Gaelic, Welsh, Cornish—all were severely affected by those raids. We have now seen in the last 150 years more archaeological evidence that suggest that while many of the Vikings raided and then went back, but many others stayed. The more they stayed, the more Old Norse mixed with the local language, whether it was Anglo-Saxon, which is frequently another term for Old English, or Scottish Gaelic, Irish Gaelic or Old Manx, all of them have various aspects that were brought in by those Old Norse speaking Vikings.

Just to prove to you that English really is a Germanic language—and I know you may not always believe it but it's the truth–think back when we classify languages, and we look at the core lexicon or vocabulary: the numbers 1-10 and other basic, core vocabulary. I have examples: *man, hand, foot, bring, summer*. These are core vocabulary about people, body parts, basic verbs (*to go, to come, to bring, to be* if it's existing, *to eat, to drink*), and phenomenon that are shared, such as flora and fauna that are shared (*summer,* in this case seasons. having to do with whether, basic terms for like *flower* or anything like that). If we look at these examples, the term for *man* in Dutch, German, Danish, Swedish or Norwegian, you have: /mæn, man, manʔ, man/. It's pretty much the same as 'man' or the vowel has changed, but nothing else. The word for *hand*: /hant, hɔʔ, hand/, .so pretty much the same across the board. Danish is just going to have a lot more glottal stops, even though Danish is technically a dialect of the Swedish/Norwegian. There's some big changes for phonetically for Danish, as compared to Swedish/Norwegian. *Foot*: /vu:t, fu:s, foðʔ, fo:t/. *Bring*: /breŋe, brɪŋə, breŋə, briŋa/. It's the same term, right? *Summer*: /zo:mer, zomer, sɔmər, sɔmar/. In fact, near where I live—I'm at the southwest corner of San Jose—and up in the Santa Cruz Mountains, you have Solvang, which is a Swedish community and Midsomar is a big festival in June. For comparisons, these are all Germanic languages, and here's Turkish down here; very clearly not a Germanic language, very much not an Indo-European language either. The word for 'hand' is /adam/--yes, Adam as in Adam and Eve, more on that another time—/el/ for 'hand', /ayak/ for 'foot', /getir/ for 'bring', /yaz/ for 'summer'. It's very different.

As I said, English has changed over time. This is a copy of The Lord's Prayer.

The Lord's Prayer Through Time

Late Modern English

Our Father who are in heaven, blessed be your name. May your kingdom come. May your will be done, on earth as it is in heaven. Give us this day our daily bread. And forgive us our sins as we forgive those who sin against us. Do not lead us into temptation, but deliver us from evil.

Early Modern English (Authorized King James Version, 1611)

Our father, which art in heaven, hallowed by thy Name. Thy kyngdome come. Thy will be done, in earth, as it is in heauen. Giue vs this day our daily bread. And forgiue vs our debts, as we forgiue our debters. And leade vs not into temptation, but deliuer vs from euill.

Middle English (Wycliffe's Version, 14th century)

Oure fadir þat art in heuenes, halwid be þi name; þi reume or kyngdom come to be. Be þi wille don in herþe es in heuene, yeue to us this day oure bread ouir other substance, & foryeue to us oure dettis, as we forgeuen to oure dettouris, & lede us not in to temtacion: but delyuer us from yuel.

Old English (West Saxon, ca. 1000)

Fæder ure þu þe eart on heofonum, si þin nama gehalgod; tobecume þin rice; gewurþe þin willa on eorðan swa swa on heofonum; urne gedæghwamlican hlaf syle us to dæg; and forgyf us ure gyltas, swa swa we forgyfað urum gyltendum; and ne gelæd þu us on costnunge, ac alys us of yfele.

Anybody who's a Christian knows this prayer, although you might have a different version of it, but you get the same idea. Even most non-Christians are aware of this prayer. The Late Modern English is from a few generations ago:

Example 4.1

Our father who art in heaven, blessed to be your name. May your kingdom come; may your will be done on earth as it is in heaven. Give us this day our daily bread and forgive us our sins, as we forgive those who sin against us. Do not lead us to temptation, but deliver us from evil

If you are Christian in any way, or if you were brought up as Christian, you know that prayer, and it's one of the first prayers that Christians learn. This is the King James version, so this is 1611.

Example 4.1

Our father, which art in heaven, hallowed by thy Name. Thy kyndgdome come. Thy will be done, in earth, as it is in heauen. Gius vs this day our daily bread. And forgiue vs our debts, as we forgiue our debters. And leade vs not into temptation, but deliuer vs from euill.

Notice I modified my pronunciation a little bit. I'm okay at Elizabethan English, although not great; there's better examples and I'll give you some examples of that. OK, some spelling changes, but you see that that's pretty much the same poem, that you can read it. I am not going to read Middle English or Old English; I'm going to give you those links below. Just look and compare; even Early Modern English from 1611 to the Wycliffe's Version, which is 14th century, so 300 years before. Notice that the Late Modern English version was probably written around 1940. Three hundred years before that, it's pretty much the same. But 300 years before that? The language is very different. And then 400 years before that? It's even more different.





One real quick note: this letter that looks like a 'p' but with an extra extension up? That is called thorn, and that letter is the voiceless interdental fricative, the sound [θ]. You will see in the Old English a letter that looks familiar, at least with respect to IPA. This is edth, right? It's the 'd' that has that cross, that's kind of written on a slant, and has that cross through the upright. Edth is the voiced interdental fricative, just like an international phonetic alphabet. Play those recordings and listen to them, read along with them. Listen to the amazingly different sounds.

The biggest change…are the vowels. Let's talk the **Great Vowel Shift**! This is going to explain a whole lot of weirdness, as it were, with respect to English. As I said, most of the time, vowels can change, but not that much. And then you have English…. This is the vowel system that we have in Old English: /a:, e:, i:, o:, u:/ as long vowels; they could also be short vowels but those didn't change as much. We're going to focus on the long vowels because the Great Vowel Shift is predominantly in these long vowels.

What happens between here and over here in step one, which is the beginning of what we call Middle English, around 1100 CE, you basically get an up and out movement. Think of the trapezoid of vowels, and the movement goes up and out. The long [a:] moves up and out to become long [ɛ:] and long [ɔ:]. [e:] and [o:] more or less stay the same, but now we don't just have [i:] and [u:] as long, we also have diphthongs. Now we go from a five-vowel system to seven. Do you see where this is going? A few more generations of up and out and we get what you see in step two. You start having a little bit more change, a firming of the sequence; notice there's no long low vowel. There is [e:, o:, i:, u:], and then [aj]. This is what we're hearing right around 1500 CE, in the Elizabethan era. Then, one more change up and out, so that for long vowels in Late Modern English, by the time you get to 1700, they have firmed up to what we have now.

This is going to explain a whole lot with respect to the spelling system of Modern English. The spelling system is archaic, meaning it is trapped in time. It reflects a pronunciation that we have not had in many hundreds of years; realistically, it is a lot closer to what Old English really sounded like. For example, *bite*, *beet*, *beat*, *made*, *foal*, *fool*, *foul*, that is how we pronounce those lexicons. Yet, the vowels like *beet* and *beat* are said exactly the same, but they're written totally different. Why? Because we have the Great Vowel Shift.

- *Bite* really was said originally as *beat*, and then [bi:t] went to [bəjt] and then went to [bajt].
- *Beet*, as in the root vegetable, was really sad as [be:t]; that's why it's written with two e's is because it was a long [e:], and the [e:] when [i:] and then shortened to [i].
- *Beat* was really pronounced as [bɛ:t]. Clearly, there was a difference, but that difference merged over time; that's Great Vowel Shift.
- *Made*—ever wondered what the heck that silent 'e' is in English? It represents the fact that in Old English, that was a schwa sound and it got deleted pretty early.

You can see how this goes.

When we say, "English is weird," it is true, at least from a bystander's point of view. We can explain all of that so-called weirdness, and the reality is there's a reason for these changes. The Great Vowel Shift happens because there's a radical shift in the population. Could these changes have happened on their own, if the Normans never came over and if Angles, Saxons and Jutes with some Vikings around continued in the British Isles? Sure, it's possible. But what we know to be true is that when you have big clashes of language contact, such that there is now a superstratum and substratum set of languages, any changes that might have occurred normally get accelerated hugely. It's possible that that could have happened, but probably got accelerated during that Middle English period. It is one of the most fascinating aspects to the history of the English language.

4: Stages of English is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5: Morphological Stages

Morphological Stages, from Sarah Harmon



Video Script

Okay, folks, it's not just phonology that changes over time. In this section, we'll talk a little bit about morphological changes, and then in the next section we'll talk about semantic and lexical changes. It is important to note that at this level, talking about each of these types of changes gets pretty complicated, and I don't want to go too far into this. But I will give you some examples and showcase what happens. Again, I'm going to be using Indo-European and especially the Romance languages as examples. This is in part, again, because they are so well documented over the history of these changes. That's not to say we don't have similar documentation elsewhere; it's just to focus a little bit on something that everybody seems familiar with. You all speak English, so looking at the history of English for some of these changes make sense; you can see how things have changed over time in a language you all speak. So many of you speak a Romance language, whether it's Spanish, French, Italian, Portuguese, Catalan, Romanian or any of the Romance languages. So many of you either natively or learnedly speak a Romance language, so by showing you the history of how these languages have happened, you can start applying it to other languages.

As we start going into morphology and morphological changes, one of the big ones when we're talking about Indo-European languages and many other major language families, has to do with **case marking**. What is case? Remember, we talked about it with respect to morphosyntax when you are using some kind of inflection to show the role that the noun phrase or propositional phrase may have, whether it's a subject, a direct object, etc. We see here Old English versus Late Modern English, so what we speak today. In Old English, you had four main cases: nominative, genitive, dative, and accusative. **Nominative** case is always the subject; **genitive** is always possession. **Dative** is always what we like to term the indirect object, so the person or entity that benefits from the action. But dative also can imply other roles that involve propositional phrases, like *to* or *towards*, like if we have the sentence, *I throw the rock at the building*, the noun phrase *the building* might be marked with a dative. **Accusative** usually is what we call the direct object, so the person or thing that is undergoing the action, but it also can imply some kind of direction away.

In a lot of languages, and certainly in old English, we had different forms of the noun—singular versus plural—according to different case, so that whatever role it had grammatically in the sentence was marked with the inflection. There wasn't just a 'stān' or stone; *stān* was the nominative singular, as well as the accusative singular. But if you wanted to say, *the stone's color*, you couldn't say *stān*; you had to say *stāne* because it was the color of the stone. If it was plural, you had different forms.

This is old English, and I chose to show you this instead of say Latin because Latin was a lot more complicated. There was something here in this data set that you don't see as much in Latin. Compare everything in this data set, notice how many forms are similar, and I mean very similar. This is written in modernized English writing, which is based off of IPA. Notice that the nominative singular and accusative singular are the exact same form, so whether it's the subject or the direct object, you don't know. Notice that this is true just not in singular, but in plural. Also, notice that *stāne* versus *stāna* are very similar; just that one

little vowel is different. When we see something like this, historically one of two things typically happens: either there is a merger of the roles, so the morphosyntax, specifically the case system, collapses and goes away. Or you start dissimilating every individual piece. In the case of English, we clearly see that case is being deleted, so that now we really only have genitive and non-genitive. By the way, genitive is that apostrophe-s (-'s) that we have as the possessive. Everything else is a different form, and then we have singular and plural accordingly. Also, notice that this plural genitive is starting to go away; more and more folks do not use this form, so there is further evidence of the case system further eroding.

As I said, this is all in English, but we see this in almost all Indo-European languages, where the case system has modified over time. Some languages, like English like the Romance languages for the most part, and in certain other cases, you see either severe reduction or a deletion of the case system. Other examples, like the Slavic languages and the Baltic languages, or the Aryan languages like Farsi/Persian, we see a dissimilation. We start seeing strengthening of those case markers in some way.

Remember that when you have a reduction in case marking, that isn't the only thing that's affected. Everything in the entire language is affected, which means you're going to affect the phonology, the morphology, the syntax, and the semantics; it's a big circle.

There are other examples that I wish to show you, and analogy is kind of the big one. **Analogy** is what we start seeing when there are mergers or combinations, and roles gets shifted. Phonology has a great impact on what happens; if the sounds start blending and sounding the same in a given environment, then you're going to have a whole slew of other effects in the rest of the aspects of the language. For example, if you talk about a *cow*, it used to be that the plural of *cow* was *kine*. Now it's really only *cows*; we have we have used analogy to apply the standard pluralization. The same is true with the present tense and the past tense. We saw that there are dialects that still showcase older forms, older pronunciations and older relics. When we talked about umlaut or vowels changing to either pluralize or make something past tense. We see it in Appalachian English, we have seen it in other dialects of English that are more remote.

Reanalysis part of this as well; when we're talking about reanalysis, we're also talking about how the morphology is affecting the syntax. If you have a reduction in case system, meaning you don't separate out the parts of speech, you have to start relying on word order. That change is going to affect the syntax and everything else with respect to the language. We see this with Latin and the Romance languages; we see this with the old English to Late Modern English—actually see it in Early Modern English. There's some interesting with versions of adpositions—prepositions or postpositions, depending on where they are with respect to the root. They can get reanalyzed case markers, and case markers can get realized as adpositions. With respect to the case markers that we see in Russian, some of them started off as adpositions, I believe prepositions. We see this from Traditional Japanese to Modern Japanese, as well. We also see the reverse; some of those case markers in Latin turned into prepositional phrases, for the most part course, although if we talked about Romanian that language has postpositional phrase.

When we talk about these changes, we aren't just talking about a change in one area. Language is not in a vacuum; it is going to affect everything else. So, I started off talking about phonological change, but notice that it impacts the morphology, and then in the next section we'll see how it impacts the syntax and lexicon.

5: Morphological Stages is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6: Syntactic and Lexical Changes

Syntactic and Lexical Changes, from Sarah Harmon



Video Script

Syntactic change and lexical change. Again, everything impacting everything else. Now that we've already explored phonological change and we've explored morphological change, let's look at how the syntax and the lexicon can also change. I'm mostly going to focus on English and the Germanic languages, as well as Spanish and other Romance languages, because these are languages that most all of you either have constant contact with natively or you have learned at least one of them, if not multiple.

One of the big things that happens when you have morphological change is that it impacts word order in some way. Either it becomes more loose, or it tightens up and becomes more strict. Certainly, when we talk about the loss of the case system like we see in English, and we see in the Romance languages as a whole, you're going to rely more on word order. To be fair, it's not to say that we cannot mess with the word order at all. If you are a poet, especially one that really focuses on meter, then you know that you are going to play with that word order to get the meter right for your lines; you know that's going to happen and certainly when we're talking not just about poetry, but song lyrics, you're going to play with the word order a little bit. But even within that, English has a set of rules; you cannot really mess with the word order too much, otherwise you break grammaticality. The same is true with the Romance languages to a varying degree. French has more strict word order because of the phonology; if you speak French or have tried to learn French, you know that the phonology of French is such that a lot of lexicons sound like one another. That means you're going to need more word order to keep everything straight, versus a language like Spanish, Italian or even Portuguese. You have a little bit more play sometimes because of other syntactic and morphologic rules in place, so that you can sometimes put the subject after the verb—which is not something you would think you can do, but in truth, you can.

All of those changes with respect to English, as well as the Romance languages, are very early. When we talk about these early changes, we're really saying that they became established, not a few generations ago but several hundred years ago. Honestly, with respect to the Romance languages, this change happened about 1500 years ago, at the fall of the Roman Empire. By that point, Vulgar Latin was already changing a lot with respect to case and word order, so that by the time you get the earliest forms of the Romance languages—Old French, that's around 850 CE, Old Spanish is around 900 CE, Old Italian is around 900 CE and so on—you really only have a period of about 400 years where things are in flux. By the time we get to the early Romance period, the word order is set as Subject-Verb-Object almost always; yes, there were some tricks that you could do then. By and large, that word order has not changed in about 1000-1200 years.

We see also with respect to syntactic change—and this is where Creoles come into play—we can see borrowing of grammatical constructions, although not as often. This frequently is in very isolated cases; we see more of it with respect to Creoles, of course, but as we talked about with creolization in the sociolinguistics chapter, if you have a creole that has been spoken for 300-400 years, you have a lot more stabilization then. Typically, there is some borrowing early on, but it doesn't tend to happen after a certain

point. There is reanalysis of grammatical structures, which we saw with the previous section on morphosyntactic changes; the loss of case usually happens at the same time as other types of analysis. Reanalysis happens throughout the entire language: phonology, morphology, syntax, semantics, pragmatics, everywhere.

We also have **extension of grammatical structure**. I like giving this example of the passive in Spanish. Native Spanish speakers, you probably don't even think about that pronoun *se* very often as to how many different ways, you use it. Those who learned Spanish, you think about this a lot, because it turns out this thing has a litany of uses. Let me just show you a little bit. When we get to Old Spanish—that's up until about 1400 CE, roughly when the Renaissance is a big change for a number of European languages. Both Indo European and Finno-Ugric languages, we see a big change once we get to the Renaissance for a variety of reasons, and we can talk about that at another time. But Old Spanish used *se* as a reflexive—that is still true in Modern Spanish: 'Johnny dressed himself', *Juanito se vistió*. This is the same now as it was then. But that pronoun *se* can be used to do different things, and it's in the Middle Spanish period in that 1300-1400 CE through about 1700 or so CE where we start seeing it being used as a passive. Spanish speakers, you know that you can say, *Se captaron casi 2.000 personas*. This is the Middle Spanish version: *Cautiváron-se quasi 2.000 personas*. Notice I changed my pronunciation a little bit, and certain terms like *quasi* instead of *casi*, *cautivaron* instead of *captaron*. Things have changed a little bit over time, but you can see that *se* being used as not a reflexive, but as a passive. This is something we can still do in Spanish, in fact, most of the Romance languages use *se* or whatever version of that pronoun for a passive construction where you're not mentioning who did the action. Notice the translation is, 'almost 2000 people were captured'; we didn't say by whom. We just said they were captured. In English, that's just implied knowledge; we don't think anything more of it. We use context clues to get more insight. In the Romance languages, that's not the case; you have to somehow include the subject, and this is how they frequently do it, by using this reflexive pronoun. Not all Romance languages do it this way, but a vast majority of them do, and there are ties to Latin; again, I'll explain it in a different time.

When we talk about semantic change or lexical change, there's some pretty big, obvious pieces. Clearly, adding or losing lexicon to a language is going to be part of any historical process, and that makes sense, right? Notice that we have **loss of lexicon**; these are just three of a slew of different terms that have been lost in most English dialects. I will say that a couple of these still exist in other dialects of English, specifically Scottish and Irish English. *Hie* is a really great example; this still exists in Scottish English, especially out in the highlands. *Wight*, the Isle of Wight, still exists; it's an island off of the Scottish coast, which is a pretty important island for a number of historical reasons. *Leman* as a sweetheart doesn't really exist anymore, but it could. Certainly, the **addition of lexicon**, whether it's through **coinage** or **eponyms**, and these are all ways that we create new terms. A lot of these terms, we saw them when we talked about morphology. I do want to expand on a word coinage; it has to do with brand names, or inventions, that is a little bit more specific. With respect to **borrowing**, I wanted to tease this out a little bit more. It's actually a really interesting process, and languages tend to rely more heavily on one process versus the other, although both can exist. **Calquing** is when you translate the word; I remember very distinctly being in high school learning Spanish, and hearing that a 'hot dog' in Spanish, was a *perro caliente*—literally 'dog hot' so they've just translated 'hot dog'. I thought that was hilarious. You can just borrow a **loanword**, so you've just taken the term wholesale, pronunciation at all—at least, to the best of your ability. Spanish has a tendency to calque; Italian, by the way, has a tendency to do loanwords. Here are some examples that we have brought into English: *joie de vivre*, *coup d'état*, *pizza*. We may not say these terms exactly as they are pronounced in their original languages, but we get pretty close. As I said, Italian likes loanwords: a 'weekend' in Spanish is *el fin de semana*, 'the end of the week'. In Italian, you could say, *il fine di settimana*, which is the same thing, but most people haven't said that in about 75 years; they say *il weekend*. Literally, that's what they say. Spanish says *deportes* for 'sports', *los deportes*. In Italian, it's *lo sport*. Different languages are going to choose one or the other, either more calquing or more loanwords. It's not the same language does not do both; clearly, if you are a speak fluent speaker of Spanish, you know the various terms have just been straight borrowed, pronunciation and all. My absolute favorite example of this is there is a textbook or two out there for Spanish, and students who are learning Spanish the term for 'jeans' as *los blue jeans*. No, nobody says that; you either say *los vaqueros*—*vaquero*, for those who don't know, is a 'cowboy', so *los vaqueros* implying that these pants that cowboys wear—or you say *los jeans*.

There are also, of course, a number of meaning changes. We talked earlier in semantics about **broadening** and **narrowing**. You have pretty good examples here of that. Broadening, of course, meaning the term has expanded; narrowing, that the term has contracted. You also have **shifts**, which again makes sense, because we no longer use the term for *knight* as a 'youth'; it is a specific designation. *Fond* does not mean 'foolish'; it means that you care about somebody. Within the meaning changes, there's actually more nuance: pejoration and amelioration. If you know a Romance language, those might look a little familiar. **Pejoration** is worsening, meaning the term has become a term that is not used for negative reasons; **amelioration** is to improve, that the meaning has improved over time. A *mistress*, a *hussy*, a *slut*, a *moron*, a *madame* (sometimes), these all have some kind of negative connotation to them—some more than others. But that wasn't the case even 200 years ago, even 100 years ago in the case of

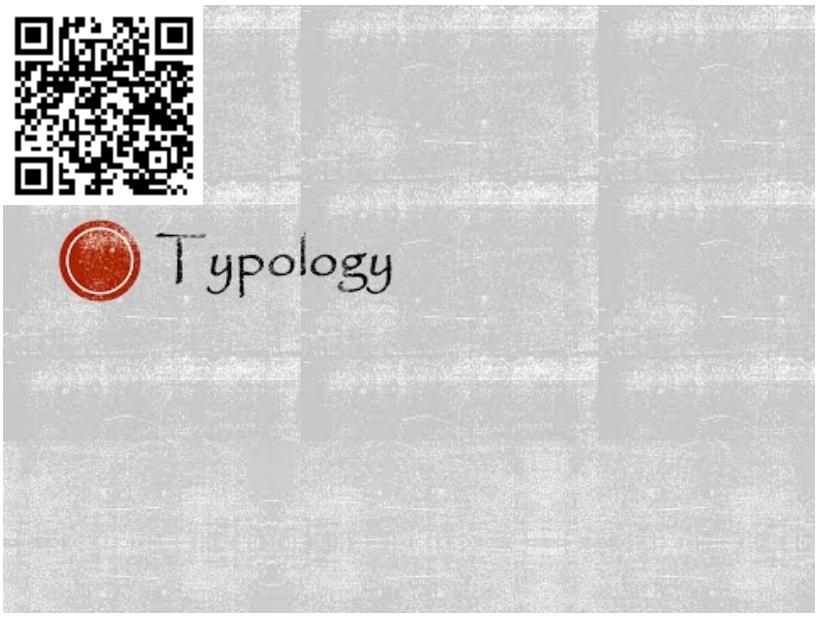
mistress and *madame*, they had no negative connotation whatsoever. In fact, you were the *master of the house* or the *mistress of the house*, meaning you were the lady of the house—nothing negative about that. *Fond*, *knight*, those are actually amelioration. Hyperbole and understatement: **hyperbole**, of course, meaning bigger or grander, **understatement** mean lesser. *Terribly, horrible*, *starve*, *quell*, these are all terms that have augmented their meaning—they have gotten to encompass much more than previously. *Terribly* was not something that was not an adverb that was used to mean anything more than *terrorizing* in some way. Now, that's taking on a much greater role; *terribly* can be either catastrophic, and in some dialects, it can mean 'very': "*Oh, it's terribly important*," 'it's very important'. Understatement is when we make the meaning less than it was before. *Kill*, of course, means literally to take the life of another animal, human or otherwise, but you could also say *I could kill for some chocolate right now*. Clearly, when we say that, we are not planning to commit a murder, which of course is also under become an understatement.

It really is the case that in language, we change meaning all the time, it affects everything that we do. The phonology effects the morphology and the syntax and semantics; the morphology effects the syntax and semantics; the syntax effects morphology, syntax, and semantics. They all interchange and they all work together as language changes over time.

6: Syntactic and Lexical Changes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7: Typology

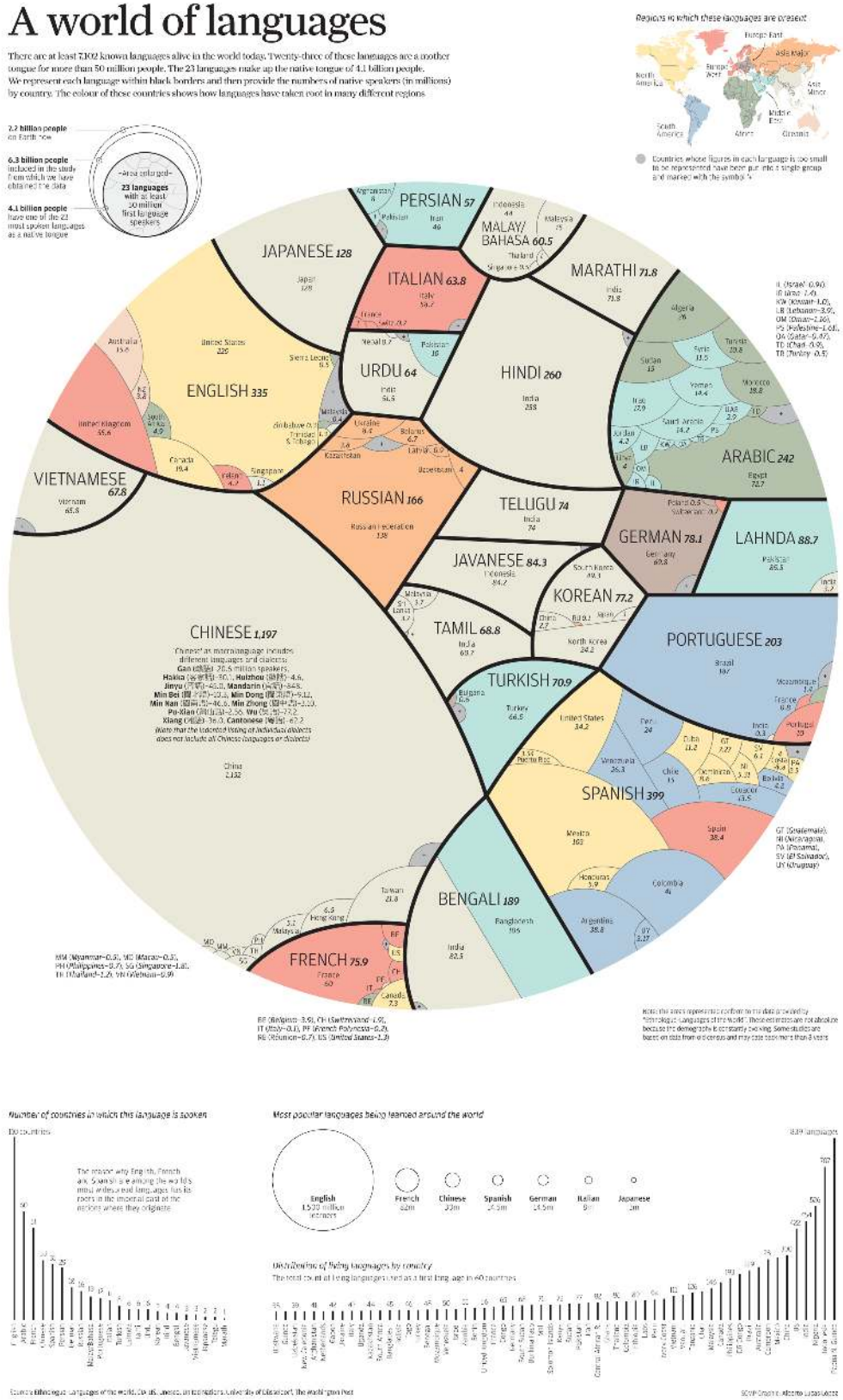
Typology, from Sarah Harmon



Video Script

Now that we've talked a little bit about how language change affects every aspect of language, let's go back to that concept of language families and how we know language families come together. More importantly, those aspects of language that seem to exist in more than one family. This is an area called **typology**; it is loosely connected in some ways, strongly connected in other ways, to historical linguistics. As an historical linguist, you have to know some patterns with respect to human language, especially within a given family, and how we know those patterns to exist is through the study of typology. Let's take a closer look.

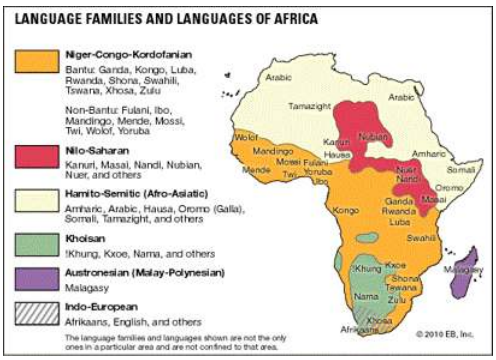
I like starting off with this infographic:



This is a way to view world languages by number of native speakers. These data were based off of data that was collected in 2015. You read the graphic by following the pieces: the bigger the piece of that circle, the more native speakers that language has. Notice

that the biggest number is Chinese, although remember that's Mandarin, at over 1 billion people; what you see here in the numbers is the number of millions, so one almost 1.2 billion as of 2010. It gives you some of the countries where the language is spoken. After Mandarin, of course, the other big players that you see are English, Hindi, Bengali, Portuguese, and Spanish. When we're talking about these languages, we're not just talking about the actual whole language; we're also talking about major dialects. Notice that, in English, you have a number of dialects that are represented here. The same with Spanish and Portuguese; you got a few Arabic, of course, too.

When we talk about world languages and we talk about tendencies, this is not just straight observation, for the sake of observation; this is also looking across the world and noticing different patterns. When we talk about these patterns within the language family or in most languages, this is typology. If we're going to talk typology, we have to talk about this person: Joseph Greenberg. He was a long-time professor at Stanford; I'm not just bringing him up because the college I teach at is overlooking Stanford, but because this is a very important individual. Joseph Greenberg started his research in the 1960s, with respect to patterns and really focusing on the number of patterns we can see within a family or within language is a whole. One of his biggest claims to fame, and I think one of the biggest reasons we need to remember him, is that until him, it was just thought that there were maybe two language families with respect to Africa: The North African languages or the Maghrebi group along the northern coast of Africa, and then everything else. He wasn't having it; he was not thrilled with that concept at all. It comes out of colonialism; it does not actually take into account anything with respect to the languages of the African communities. Greenberg basically used a type of comparative analysis, recording all of these terms for numbers 1-10, nuclear family, basic terms for flora and fauna. He threw out anything that was borrowed, especially from an Indo-European language or from Arabic because, clearly, there was some coloring there. He did take note of these borrowings, but not for the main classification. The main classification was on native terms; if a term was borrowed, it was borrowed locally. What he came up with is in this infographic:



This infographic shows there aren't two language families, by any stretch of the imagination. The top third, if you will, of Africa, often called Saharan Africa, you have Arabic as a main player, although there are a number of indigenous languages that are spoken in that in those areas. You also have Afro-Asiatic languages being the main player; we think of Arabic and Hebrew, but we have to include Somali, Amharic, Oromo, the languages of the Cape of Good Hope. We're talking about languages spoken in Somalia, Ethiopia, Eritrea, that area of Eastern Africa. There's also the Nilo-Saharan group, which is related to Afro-Asiatic but seems to be its own language family. These are the languages of the Upper Nile, as it were, and interior parts of Eastern Africa; we're talking Nubian, Massai, and many others. You have Niger-Congo, which is a very dominant group, but it is not the only group there are realistically, potentially two language families. These Khoisan languages, which are also called Bantu languages, are a separate language family. They may be part of Niger-Congo; when he was researching, but even to this day, that is still under contention. Then you have Malagasy, which is an Austronesian or Australio-Pacific language; it's very close to the languages of indigenous Australia, especially the western half of Australia. Down in South Africa, you have Afrikaans, English, and other Indo-European languages, but if we just talk about the native languages, you have, at the very least four sizable language families.

Even with that the sheer variety of languages, this was what Joseph Greenberg brought to the table: not just creating typology, as we know it, but in analyzing through objective, descriptive processes: describe what you see, not what you want to see. It was his work, specifically with the languages of Africa and the various language families of Africa, that launched a huge movement within linguistics. It helps us to really analyze linguistic isolates. He was one of the very first to pair linguistic genetics to actual biological genetics, as he started teaming up with leading geneticists from around the world. He started doing DNA tests, especially once we started understanding and decoding DNA, and combining the genetic DNA history with the linguistic history. It was an amazing experience to learn from him; I took a couple seminar courses at Stanford from him exactly so I could learn more about this, and it is his work that has continued on through the generations. There is an entire collection of his work at Stanford, and if you are

interested at all, I suggest going to check it out. It really encapsulates a big pivotal moment for linguistics, when we start saying that language is not in a vacuum. Rather, language is about the human experience, and unless we start understanding that human experience and how we different cultures and speech communities interact with one another, you never really understand what humanity is about. He's one of those people, and I do take his work very seriously.

That's Jospeh Greenberg, and what did he in creating typology. One of the main areas with respect to typology is this concept of **absolute versus non absolute universals**. When we say something is an absolute universal, we say this pattern we're talking about exists in all languages, period and end of story; we have yet to find a human language that does not follow this pattern. A non-absolute universal is that there's a tendency to follow that pattern. As you will probably guess, there's more non-absolute universals than there are universals, and that makes sense; languages have been spoken for who knows how long truthfully, maybe 100,000-200,000 years. They have been spoken not just by *homo sapiens sapiens*; we're pretty sure that our Neanderthal cousins also had a language and potentially other hominids as well. When we talk about a universal, it's really important to tease that distinction out: an absolute versus a non-absolute universal. Let me show you some examples.

- Absolutely universals all languages have syllables that have consonants and vowels. That does not mean that all syllables have a vowel; it just means that there's a phonological component in that nucleus that exists in all languages.
- Every single human language that has been observed or recorded has at least one stop; other options are optional, but at least one stop.
- When we're talking about lexical elements, there's always some kind of lexical content, and there's always distributional content: think open class and closed class systems.
- All languages have a concept of a word or a lexicon. It may be phrases and maybe clauses, but these things exist in all human languages.

Those of universals. When we're talking about tendencies, with respect to universals, these are the patterns that we tend to see:

- Languages tend to have syllables that have an onset and a nucleus. Codas are frequently optional. In most languages, you have a constraint as to what that coda can be; it is not a universal. It is a tendency, but it's a strong tendency.
- There's a tendency for languages to have nasals, and it's not just a regular tendency, it is a very strong tendency. There are a few languages that have been recorded over time, and including are still spoken today, that do not have a nasal sound, either as a nasal stop or as a nasal vowel. But it's exceedingly rare.
- Alveolar stops also seem to be there, whether it's [t] or [d] or some version of it.
- There seem to be high front vowels, the [i, e, y, ø]. That sound tends to be somewhere in every single language, generally speaking. It is possible to get through life and not have one of those vowels but it doesn't tend to be the case.

Now there are universals and there are **implications**, meaning if you have this, you will also have that. As you can figure, there are some absolutes and non-absolutes and there's many more non-absolutes or tendencies than absolutes. Absolute implications include:

- If a language has a mid vowel, it also will have a high vowel. We see this in every single language. It doesn't mean high front or high back.
- If you have a voiceless nasal, then you will have a voiced nasal. The voice nasal is the default form and some languages, Myanmar or Burmese comes to mind, you might have a voiceless nasal as well.
- If you have a dual, then you also have a plural. Indo-European speakers, we tend to think only of singular and plural, but many languages also have a dual, which is an inflection that says that there are two of something, not just one or more.

Those are the universal implications, the absolute implications; they if you have one, you will have the other. Then you have the non-absolute implications, which again are tendencies; there are languages that exist without them, but not many. Some examples include:

- With respect to front round vowels—think French and German—there's a number of languages that have [y, ø], then those languages also have both front spread vowels, like [i, e] and back round vowels, like [u, o]. There's a tendency; it’s not always the case, but it frequently is the case.
- If the language has a nominal affix, it tends to be that it will have a plural. Nominal in this case just means on a noun. If a language is going to have any kind of affixation on nouns, at the very least it will be a plural affix of some kind. Not always, but almost always.
- Again, if a language has bound morphemes, and it has case, then the case either goes at the beginning, or at the end, not in between the root and the number. That's really interesting; there seems to be an order for how human beings like to put their

derivational and inflectional affixation.

With respect to morphology, we talked about affixation and it's ordered; there are patterns that we see. With respect to syntax, we actually have talked a little bit about how word order seems coincide with certain other patterns: whether you have prepositions or postpositions, with the head first or head last, what kinds of phrases, does a language have agglutinating or fusional or polysynthetic phraseology. In phonology, there are also typological patterns that we can see with respect to the phonemic inventory. In fact, there's a really interesting one that involves tones. What you learned a little bit about in phonology and phonetics, and I want to go into a little bit more, is one of the great contributions that Joseph Greenberg gave to linguistics with respect to typology. He noticed that certain patterns only happen in certain language families. When we talked about tone and pitch and certain other aspects, you only see it in certain language families; contour tones—think the Sino-Tibetan languages or the various languages of Southeast Asia—they all have contour tones. This is the Vietnamese vowel system; notice that you have strong and weak syllables, and you have different tones for different things. This only exists in one part of the world: in East and Southeast Asia. Nowhere else have we observed contour tones in human language. With respect to register, high tone versus low tone, this is something you see in the Niger-Congo languages—most but not all—and that does include some of the Bantu languages. But only in those areas nowhere else in the world.

As for the reasons why we have these patterns, this is where topology divorces itself from historical linguistics. Typology is all about recognizing patterns; it doesn't normally go into the history of the language or language family, and certainly with the case of contour tones and register we don't really know. We're not really sure why these languages have them, but nowhere in the world who else do we see that. What we do know is that we see these two phenomena only in these regions, and nowhere else.

We do have the aspect of pitch. I'm giving you Japanese here; Korean and Japanese are related languages and Korean also has this. We really only see this in a couple other languages modernly; there is one language in Papua New Guinea called Una, and it's a Trans-Guinean language, and it seems to have this phenomenon. Ancient Greek seemed to have it, but then died out completely, not long after the Macedonian period—think of Alexander the Great. Here we have Japanese. All syllables are said with the same amount of stress, or you have the first or second syllable being stressed. That term that is spelled in Romanji, which is the Japanese writing system using Latin or Roman characters. The set of sticks that we use to eat are called *há-shi* so notice, I stress the first syllable *ha*: *há-shi*. If I say *ha-shí*, then that is a ‘bridge. If I say *ha-shi*, with the same amount of stress on both syllables, a flat pronunciation, that's an ‘edge’. The one I love bringing up is when you go out to your favorite Japanese restaurant, there is a difference between *sá-ke* and *sa-ke*. The second one is flat and it's the one you drink; the first one is one of my favorite sushi or sashimi. Even the term for the country itself—Japan and versions of that is an anglicization of Nihon and it's *ni-hón;* you stress the second syllable. If you stress the first syllable, you're saying ‘two sticks of…’, like ‘two sticks of dynamite’, ‘two sticks of gum’; it's *ni-hón*. This is a pitch language, and there are a few of these around; Japanese and Korean tend to get the most attention, there are a few others.

Let's start talking about broader picture with respect to typology. These are your dinner facts, as it were, about language. I love this one: 40% of the human population speaks at least one of these languages natively and they are in order of number of native speakers:

1. Mandarin
2. Spanish
3. English
4. Hindi
5. Arabic
6. Portuguese
7. Bengali
8. Russian
9. Japanese

Forty percent of the human population speaks one of those languages. It’s not a coincidence that a number of those languages are languages of powers that have colonized a lot of places; Mandarin, Spanish, Hindi, English, Arabic, and Portuguese are certainly among those. Why all of those languages in that list? Some of it has to do with where they are globally; the highest concentration of different languages is in the equatorial regions around the world. That's where there is the most diversity of languages, and the most languages spoken overall. You also have a story of human migration; as humans have migrated around the world, they bring their language with them, so when we talk about the number of people who speak one of these languages natively, we're talking

about migration. Colonization is part of that, but colonization is a subset of migration as a whole; as human beings move around the world, they bring their language with them.

In the next section, we'll talk about different language families, and this is to give you a bit of an idea of how these languages interact and how they are unique.

7: Typology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8: Language Families

Language Families, from Sarah Harmon



Video Script

Let's talk a little bit about those language families. When we talked about typology and historical linguistics, even throughout the entire course, I kept referring to these language families. So, what are they and what kinds of things do we know about what they do? It's really interesting to see how all of this plays out.

This is a map of the world, and you will notice how many different colors there are all of those colors referred to major language families.



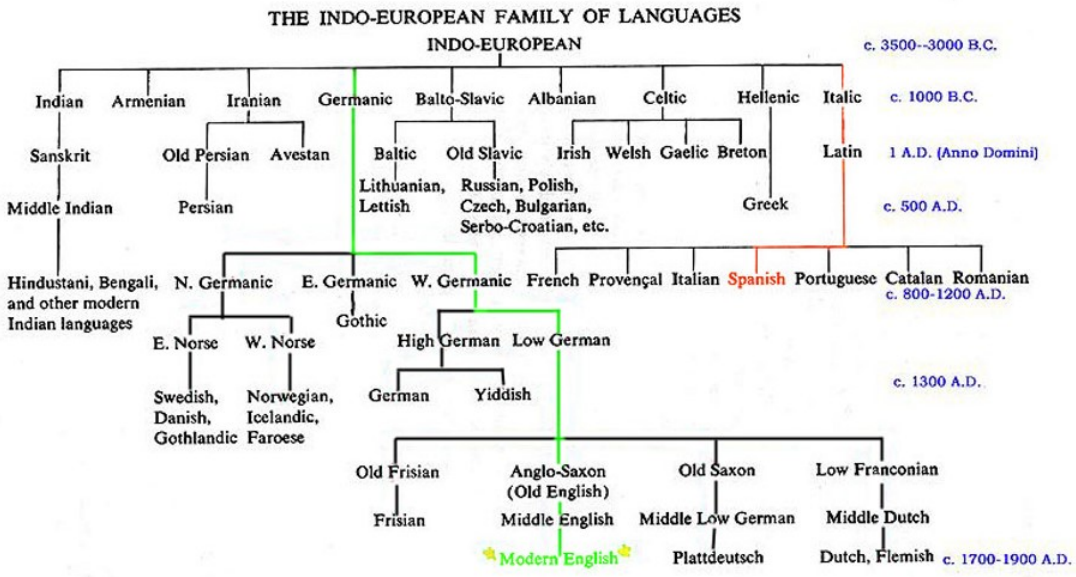
In a couple of cases, specifically the ones that you see listed out over here on the left—Indo-European, Sino-Tibetan, Afro-Asiatic and Altaic, we've broken them down a little bit more into specific languages. Even if you consider all of this list in the right side, look at the sheer number of colors that you have. It is unreal when we think about human migration; that stat that I ended in the previous section, with about 40% of the human population speaks Mandarin, Spanish, English, Bengali, Arabic, Portuguese, Russian, and Japanese. You see how widespread these languages are, this is an amazing graphic. It really speaks to the human existence, that we go all over the place, and as we go and create more isolation, those initial dialects that we speak eventually

evolved into different languages. Then those languages are part of a language family; those language families have certain characteristics.

Alright, let's get started.

Indo-European

The language family I have probably brought up the most, and the one you're probably most familiar with, is the Indo-European language family.

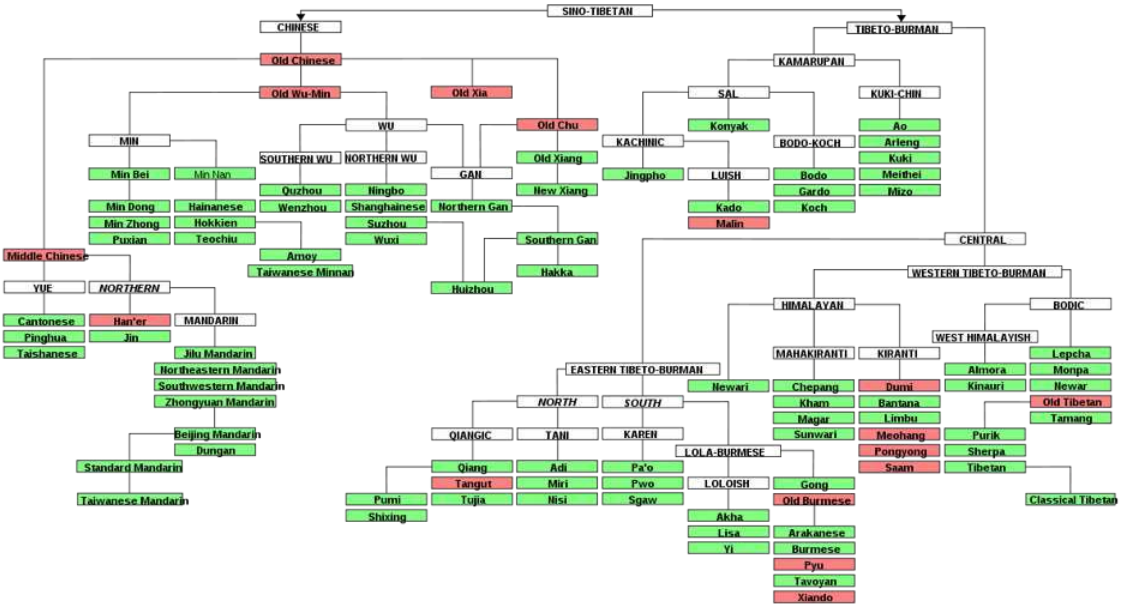


This map shows you where modern English is, and I'm just highlighting Spanish right here, but you get the point. This is pretty much most of the Indo-European languages, at least the major players. We believe that Proto-Indo-European was spoken around 3500 to 3000 BCE. It might be a little older than that; we honestly are starting to lean more towards 4000BCE. Even still, it is most probably the case that it took about 2000 years of migration to get to these major language families within Indo-European. We have Indian or Indic, Armenian, Iranian or Vedic, Germanic, Balto-Slavic, Albanian, Celtic, Hellenic, and Italic. Those are the main language families within Indo-European. It is interesting to note a few things

- Albanian and Armenian are linguistic isolates with in the Indo-European family. To our best understanding they seemed they seem to be independent migrations that just self-isolated, and we don't know why.
- Everything else shows a period of migration. For example, in the Indic group, we have Sanskrit as the main ancient language that you probably have heard of. When we talk about the modern languages of Hindi, Bengali, Gujarati, and Urdu, these are all Indic languages of some kind, and they have all come about, certainly in the Common Era.
- Iranian or Vedic is also part of this subfamiliy. This is also called Aryan. We have Avestan, which is a sibling language to Old Persian/Old Farsi.
- Balto-Slavic seem to have split up very early. I love it when Slavic speakers of some kind, say, “Lithuanian and Latvian looks totally different sounds totally different; it can't be the same family as Slavic.” Keep in mind that it split off very early; we do know that Latvian and Lithuanian in particular are very similar but have some very big differences.
- Celtic seems to have been split up between Insular Celtic and Continental Celtic. Breton is part of Continental Celtic and Welsh seems to be pretty closely related to that, although it tends to split the difference. Irish and Scots Gaelic seem to have more in common with each other.
- From the Hellenic branch, we only have Greek anymore, although there were other languages in this group.
- The Italic group, which is what I brought up earlier with respect to Latin. Latin had sibling languages; Oscan and Umbrian are two of them. If you go into the history of the early Roman society, before the Empire and before the Republic; we're talking about the Roman kings, and then even before that.
- Of course, the Germanic branch has a wide range as well.

Sino-Tibetan

We cannot talk only about Indo-European languages. One of the other big groups that I've talked about quite often is the Sino-Tibetan family. Again we're talking about a giant enormous group. The bulk of the attention goes to Mandarin and the other languages of China. But there are so many, as you can see.



One of my absolute favorite things to point out is that if you go from Old Chinese, you have four major branches, you have Middle Chinese, old Wu-Min, Old Xia, and Old Chu. This explains so much of what modern languages are in this part of the world. When you talk about Hakka, which is a language that is spoken throughout a lot of places, including Taiwan, parts of southern China and going into the diaspora in Indonesia, Singapore, Thailand, and Burma. That is a radically different language than even Mandarin or Cantonese. Cantonese, it should be pointed out, is very different than Mandarin, and Mandarin has had multiple versions; not just the dialect of Beijing, but you have Mandarin this book and in Taiwan, that is infected with indigenous Taiwanese you have the Mandarin spoken in the diaspora as well. Old Wu-Min, you have languages like Teochew with and Hokkien and Hainanese, these are very represented in again Indonesia, Singapore, and throughout Southeast Asia, as various peoples have left what we consider to be China and gone south or East and Taiwan. Hakka is over here, and Huizhou is over here as well. Don't forget the Tibeto-Burman side; in fact, it is hypothesized that this language family got its start in the area known as Tibet, not what we consider to be the rest of China. It is actually in the mountains of Tibet that we believe this language family to have started. Of course, Tibetan is part of this branch, and while you probably have never heard of many of these languages before, but most of these that you see here, of course, Tibetan and Classical Tibetan is here. Burmese is right there; Arakanese is right there as well. You have a number of languages from Southeast Asia, but you'll notice that there are no languages from Vietnam, Thailand, Laos and these areas; those are coming up.

Languages of the African Continent

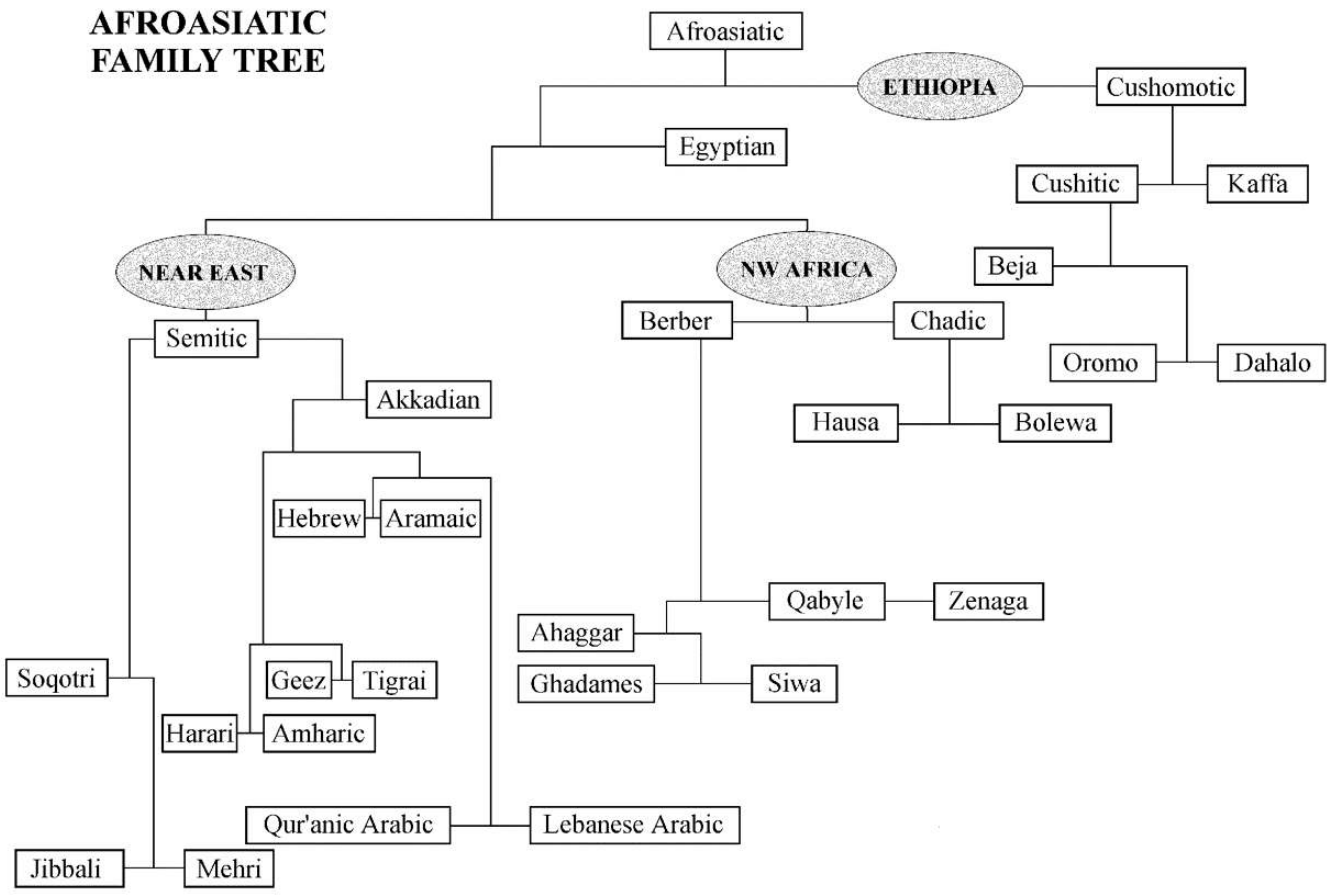
We talked a little bit earlier about the various African languages.



We see the Afro-Asiatic family in the northern Saharan area. In Sub-Saharan Africa we see the Niger-Congo family, for the most part, Nilo-Saharan kind of in the middle. Whether Khosian is a part of Niger Congo or not, that is something that's still being debated; I'm not going to go into the debate, but suffice it to say that it seems to act very differently. The Bantu languages are part of this area; they are connected to the Khosian, but they're different. They're a Niger-Congo language, but they're different and this area is still being worked on.

Afro-Asiatic

I brought up also many times the Afro Asiatic family.

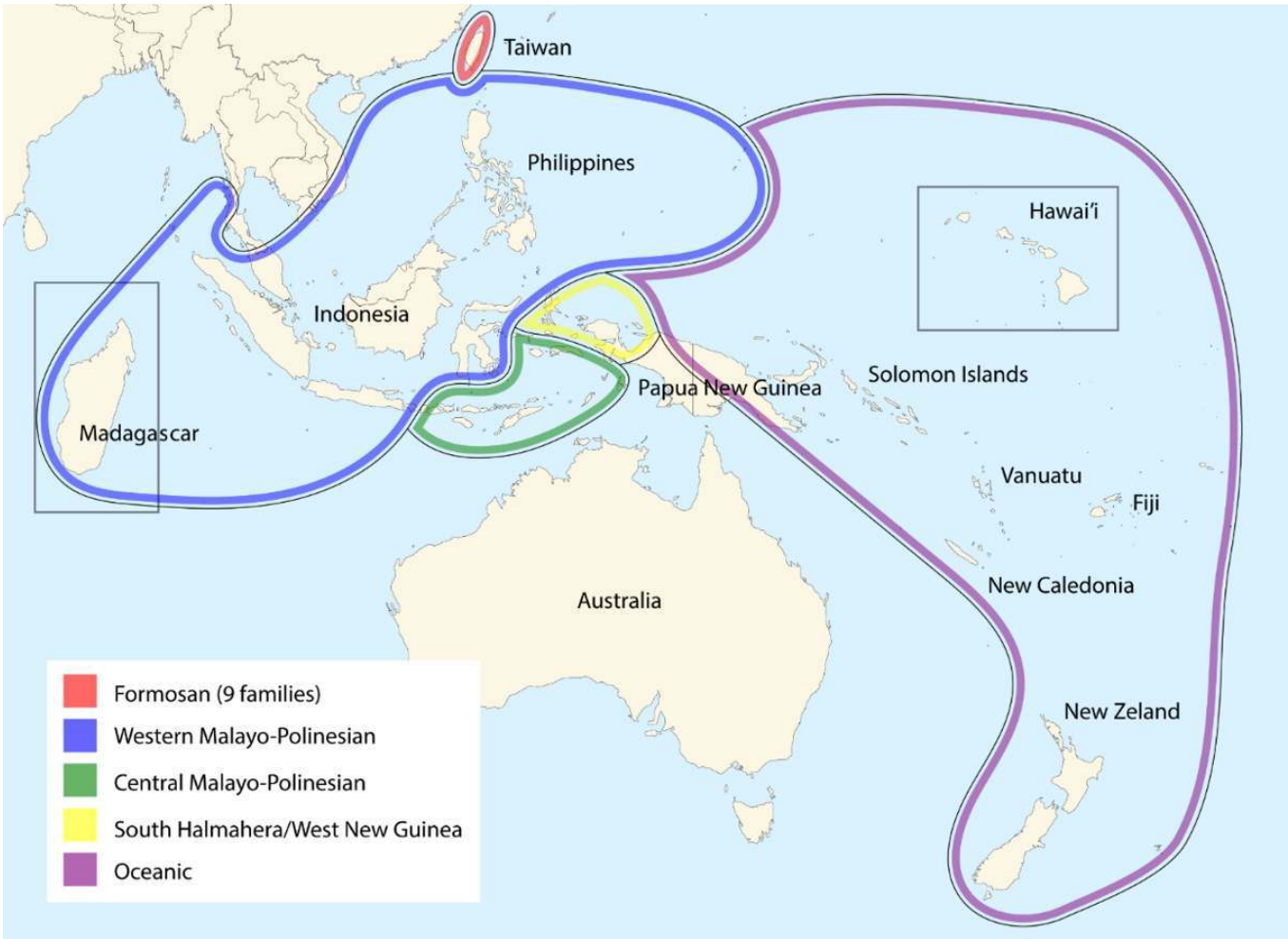


One of the very big hallmarks of this family are those continental roots. For example, [ktb] is the root for 'to write' in both Arabic and Hebrew, and I believe Amharic also, and that the vowels get stuck in as part of the morphology—the derivation and the inflection is through the vowels. This family also includes ancient Egyptian; it's important to understand that it was an Afro-Asiatic

language. Cushitic languages, now we have Oromo, but we had others as well, and they are historic languages of that North-East African quadrant. When we talk about the languages of a number of the southern part of the –Hausa, in particular, this is a Northwest African Saharan language. We have the Semitic languages of Akkadian, Hebrew, and Aramaic—if you are into a number of Old Jewish or Old Testament readings, you know that Aramaic is what was frequently spoken. Of course, Arabic is one of them, and it should be noted, there is an optional lesson just on Arabic languages—yes, I said Arabic languages, I didn't say Arabic, and there's a reason for that.

Austronesian/Australio-Pacific

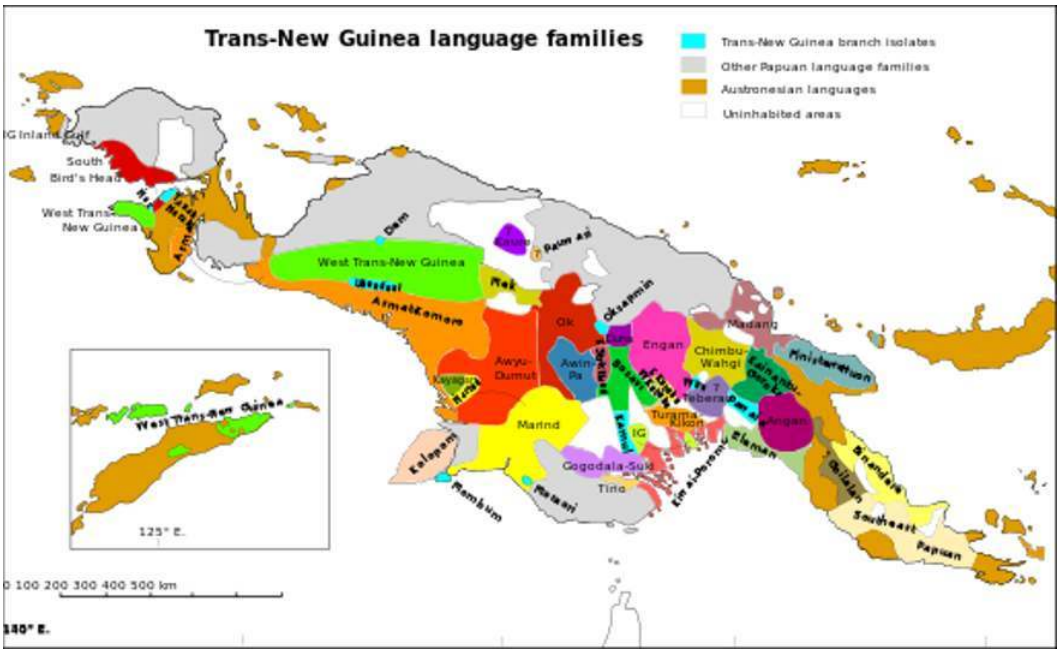
I brought up many times Austronesian or Australio-Pacific; those terms are used interchangeably.



It is a very fascinating language family; this is the language family that has infixation and tons of duplication. It has so many amazing aspects to it. The only one that you do not see here is the language of the Rapa Nui, which is the indigenous name for Easter Island and that area kind of off the coast of Chile, although, to be honest, it's well into the Pacific. What is absolutely fascinating is that this language family starts in Taiwan. When we think of Taiwan, we think of Mandarin for the most part. But the indigenous language is Austronesian or Australio-Pacific. It is from there from that little island that we get nearly all of the cultures and languages throughout the Pacific; there's very few places that that are not home to an Austronesian language. This fascinating group liked migrating, which they did spectacularly, and we know this from the similarity within each of these regions. For example, the Western Malaya Polynesian, we're talking the languages of the Philippines, Indonesia, and includes Malagasy in Madagascar. You have a lot of similarity here within the Oceanic languages, and then within the Indonesian and Papuan New Guinea regions; you have a couple other smaller groups, as well.

Trans New Guinea/Trans-Guinean

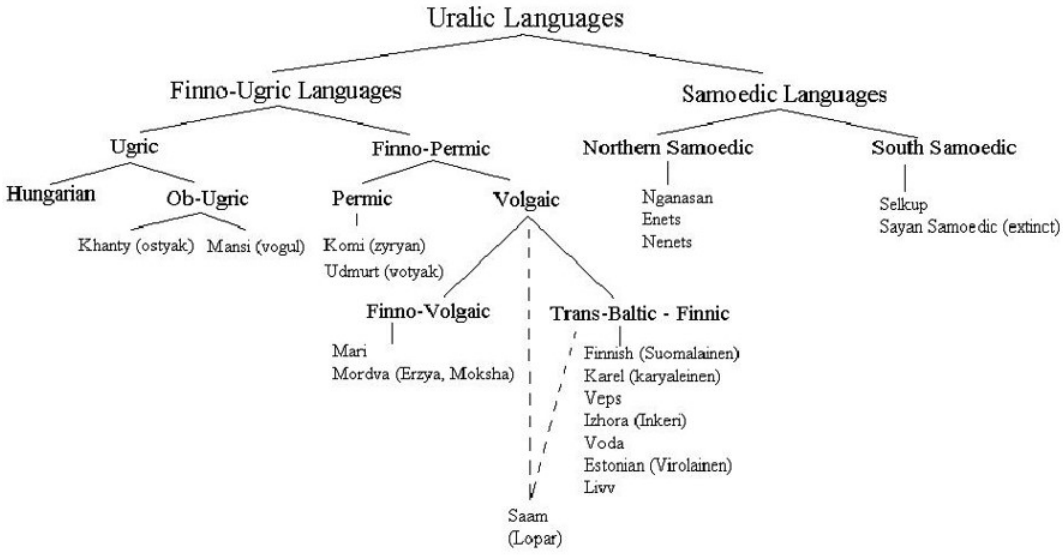
New Guinea, though, is not home just to Austronesian languages; it is also home to Trans New Guinea or the Trans-Guinean languages.



At this point, we think it is a separate language family, that it is not Austronesian nor Australian; it is something in between. We do know that the languages of the northern portion of Australia are very similar to this to these languages. It is absolutely mind boggling to think that an entire archipelago has this many different languages, but it is the case. If you know a little bit about the geography of this region, you know how incredibly remote these places are. Not surprisingly, a group of folks would go to a region and be completely cut off from everybody else; it's the story of human existence in many ways.

Uralic

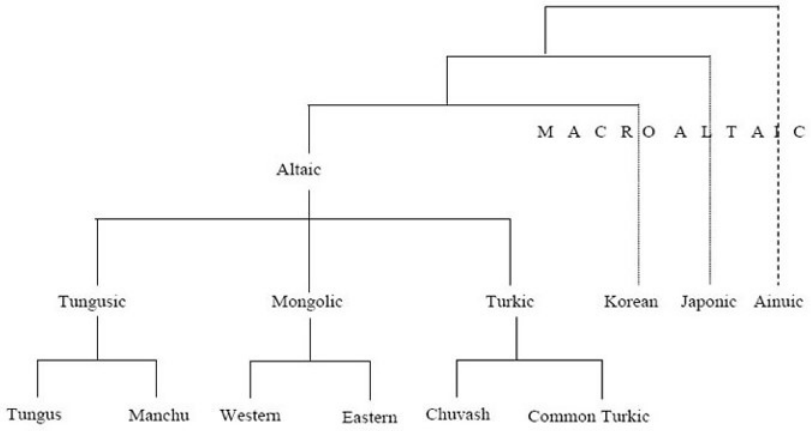
Within the Uralic family, I've talked frequently about the Finno-Ugric languages, but really it's part of the larger subset of our larger family called Uralic.



The bulk of the Uralic languages, though, are the Finno-Ugric group. When we're talking about Hungarian, Estonian, and Finnish, this is the Finno-Ugric family. We also have to include the Samoyed languages, which are spoken in the Arctic regions of both Scandinavia and Russia. This is a series of languages of indigenous languages that are very endangered. Estonian is technically considered endangered, although it is increasing in native speakers. Finnish and Hungarian are not endangered at this point, but pretty much everything else you see here, if it is still in existence, and that's a big if, it's highly endangered.

Macro-Altaic

As I said, there are some language families that are controversial, and Macro-Altaic is one of them.



Altaic take itself is not controversial; we pretty clearly can trace the history of the Turkic languages. Turkish, of course, is one of them, and a number of the languages that are spoken in the western part of Asia. Mongolic languages, Mongolian is one of them, of course, and then there are others, like some of the languages of Western China, these are all related. We have them pretty well documented for at least 2000 years, if not longer in the case of Mongolian. They seem to take up the old Chinese writing systems fairly early on, modified them to what they wanted, and then went from there. In some cases, when Islam came into those areas, they started writing the languages down in the Arabic writing system. As such, we have a lot of old documentation.

As I said, Altaic is not the problem; the problem is over here on the right, when we're talking about the native languages of Japan. Ainu is one of them; it's about the only one that is still in existence. All of the languages that existed of the peoples who lived in Japan before a group of Koreans decided to hop over the Korean Sea. Japanese and Korean are very closely related, and it all has to do with a group who got kicked out of the Korean court about 1200 years ago. They hopped over to what we now call Japan and became the group the Yamato people; they pretty much conquered the rest of the islands. Those three groups, the Korean groups—there are a couple other Korean languages, historically—the Japonica groups—that's Traditional Japanese, as well as Modern Japanese and Okinawan—and then languages that are Ainuic—Ainu and the other indigenous languages of Japan. We aren't sure of their origins. This was an area that Joseph Greenberg was interested in; he never really got to explore too much, not as much as the languages of Africa, the languages of the Americas and Basque. Many of us have been trying to figure out where these three groups come from. We know that the Japanese and Korean languages are pretty much the same group; we know that part. We suspect that the indigenous languages of Japan might have come out of Siberia, but we aren't sure. Was there a connection with the Uralic languages? Maybe. The closest seems to be with this Altaic group.

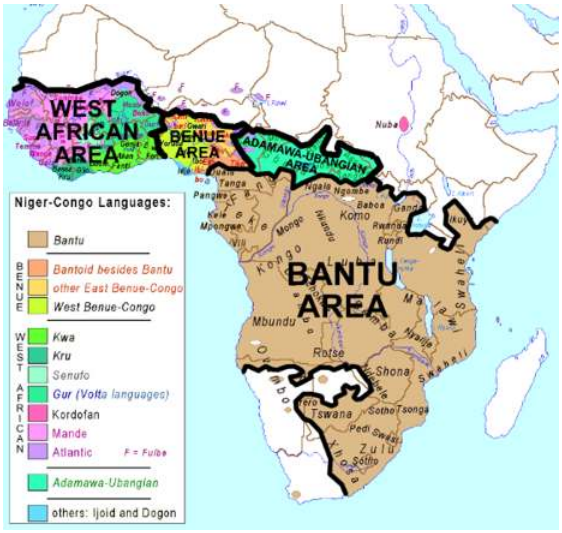
But there are things that the Altaic languages do that these other groups don't, and so we don't really know yet. More work is being done on genetic analysis, as in DNA analysis, and combining it with historic analysis of these languages. In the case of Japanese and Korean, we have a long written history of over 2000 years; in fact, the Old Korean writing system is based off of an earlier Chinese writing system, so we have a pretty good length of documentation. But it's really difficult to decipher this one. If you were to look up Japanese or Korean or Ainu in a reputable linguistic database to find out what language family they are in, we don't have a good answer, and no two databases will be necessarily alike. As an historical linguist, I tend to say they are Macro-Altaic, knowing the controversy that is involved with that and understanding that I have many times also called them isolates. I will use both terms, because we are right now really not sure.

Linguistic Supergroups

Before we leave this topic of language families, it's incredibly important to talk about super groups. This is where we have mass migration over several thousands of years, all mixed up such that we see certain trends that exist in those areas and nowhere else. In fact, we've already talked about two of these groups: Southeast Asia and the Bantu languages.



When we talked about Southeast Asia, we talked about specifically the fact that they have contour tones and that nowhere else in the world do we see anything similar. Contour tones, we believe, may have originated in the Sino-Tibetan languages—remember we're talking about Tibet, which is kind of over here—and the Sino-Tibetan languages seem to have come out of that area and spread over what we now call China and what we call Burma, or Myanmar. We noticed that the various language families that we see in Southeast Asia—there are at least two, maybe three, depending on how you qualify it. They all have these contour tones…so where did this come from? Did this come from Sino-Tibetan, or did it come from these other languages and they got pulled into Sino-Tibetan? We're pretty sure it comes from Sino-Tibetan initially, and then moved south and east. But we aren't entirely sure. All of these languages are highly isolating; there is little to no inflection. Certainly, Sino-Tibetan languages have no inflection, but even the languages of Southeast Asia—we're talking about the various languages spoken in Vietnam, Cambodia and Laos, as well as northern Myanmar/northern Burma. There are some languages that have a couple of inflections, but for the most part there's none. Where did that come from? What is more, they are so isolating compared to any other language family on the planet. They also use sentence particles and numerical classifiers; we saw an example of classifiers in semantics with Mandarin. The extent to which those elements are used is staggering in this region.



We talked about the Bantu languages, including some of the Khosian languages along with some of the Niger-Congo languages. We call them Bantoid for now because we're not entirely sure if they are a separate group. In some ways, they are similar; in some ways, they are different to Niger-Congo. This is where we have those register tones, those high-low tones. we have clicks in these areas. The Bantu area or Bantoid languages all have clicks, but the languages of Western Africa or the Banue area do not have clicks. Why is that the case? You have a high level duplication in these Bantu languages that you do not see in the other Niger-Congo languages or some of the other Khosian languages. You see nominal classes—we saw that with Swahili, in particular—that

we do not see another Niger-Congo languages, and the sheer number of them is staggering at times. This is very clearly the Bantu area, a very clear linguistic supergroup it covers multiple language families.



There's a third supergroup, one that was a little closer to my world as far as an historical Romance linguistics, and it's the Balkan area. Frequently, we use the German term, *Balkansprach*, so ‘Balkan speak’, but it includes the languages of Hungarian, Romanian, Bulgarian, Greek, Serbo-Croatian—yes, I'm going to combine them—Bosnian, Slovenian, Slovakian, Czech, Austrian-German (not all German, but Austrian-German), and even includes a bit of Turkish ,although Turkish is an Altaic language and certainly has its own thing. There are certain aspects to Turkish that you don't see in the other Altaic languages, and they are aspects that are very similar to this other Balkan area. In these cases, we have a reduced case system; they all have a case system, but less so compared to their other languages in their language families. For example, Austrian-German has fewer cases then Standard German or Swiss German. Romanian actually has case, which is different than the other Romance languages, but the Slavic languages in this area have way fewer cases than, say, Russian, Ukrainian or Polish. Even when we compare Hungarian to Finnish or Estonian, there are way fewer case markings, they have a post article; instead of saying ‘the cat’, that is an inflection and it is a suffix: ‘cat-the’. They are also highly analytical languages; they're not very synthetic. We're going to have more isolation, but there is still some synthesis, some inflection, but there's just less of it. There are loanwords that you find in this region, but you don't find elsewhere; Romanian, a great example of that, as there are a ton of Slavic and Turkic words that are brought into that language that you don't see in any other Romance language. As I said, most of this has to do with migration, conquering, wars, and the like. One of the classic statements we say in Romance linguistics is, we really need that time machine. So, one of you, please, set up that time machine, so that we can find out what happened to Romanian. We have an almost 800-year period where there's no documentation whatsoever, from the time that the Romans left, which is in the third century CE (around 400) until about 1400; we have about 1000 years where we have very little documentation. We don't really know how Romanian came to be versus the other Romance languages. The same is true for all of these languages; we have documentation, but it's later in its history. If you look into the history of that region, the sheer number of wars, conquering, language suppression, all of it that happened in that region, well, it makes sense as to why this is a supergroup. The same can be said for the Bantoid area and for Southeast Asia.

Therefore, when we talk about languages and language families, always there are patterns. It's just a matter of finding it.

8: Language Families is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9: Writing Systems

Writing Systems, from Sarah Harmon



Video Script

Let's talk writing systems! We can't talk about the history of language without talking about writing systems, and, yes, part of this is going to cover historic writing systems or writing systems that have been used across the history of human language. But we can't do this without also bringing in the modern writing systems, especially having to do with the smart phone.

Let's start off by talking about the hierarchy, if you will, of certain reading systems. When we say hierarchy, we're not saying one is better over another, or that one is preferred one over another. Frequently, what we're saying is something is more iconic versus more alphabetic; we're creating a spectrum. It is interesting to know that, with a couple of exceptions, you are talking about a historic evolution of writing over human language history.

When we talk about, for example, **ideograms** or specifically cave drawings or **petrographs**, we're talking about the earliest instance of humans trying to put something about what they want to say in a written format. From there, these **pictographs** evolve frequently into **iconographs**, where an icon doesn't just represent the thing that obviously it represents—that it maybe represents a sound or a syllable or versions thereof. The two best examples are Egyptian Hieroglyphics and Mayan Hieroglyphics. There are a slew of videos on YouTube, recent documentaries and examples of hieroglyphics, both in Egyptian and Mayan languages. These documentaries showcase how they have evolved over time, where you saw a picture of a bull that literally only meant 'bull'—a little bit more pictographic—to what it turned out to be—the term for 'bull' in Egyptian, whatever that sound was, that symbol started representing the sound, not just the meaning. It was really cool.

From there we get **cuneiform**, which is what we see with Sumerian, and then that gets borrowed by a few other generations and few other cultures, principally the Akkadian culture and the Hittite culture. Cuneiform is really the first instance of a written language with non-representational characters. If you think about a hieroglyph, they represent the thing; if you see a picture of a bowl, it's going to represent a bowl, along with other items. When we think about cave drawings or ideograms or even emoticons, these are pictures that represent a concept. When you get to cuneiform—and specifically Sumerian cuneiform—they are literal wedges. Those combinations of wedges represent a sound; at that point, you're getting to a writing system that represents sound and less meaning.

From there we go into a **syllabary** and then frequently an **alphabet**. There are a number of not just historic languages, but modern languages, that use either a syllabary or an alphabet. We have some ancient examples, like Phoenician, Ancient Greek and Latin—Old Latin along with later Classical and Vulgar Latin. When we talk about Japanese or Korean, or we start talking about Hindi or Cyrillic or Pinyin or Abjads, now we're talking about modern language and modern writing. It has changed over time. For example, abjads have a connection to the writing system for Sanskrit, which itself has been the basis of a number of writing systems for

various languages, not just the Indo-Aryan languages, but also a number of the Afro-Asiatic languages. Chinese is listed here is a syllabary, but not the way you think; the Chinese characters themselves are representations of syllables. The way we think about the Chinese writing system now it is technically a syllabary, and then Pinyin is the Latin-based representation that's more alphabetic. Korean has an interesting combination; it's technically an alphabet, but it acts like a syllabary. so there's ways to play with this.

There are variations to writing systems. We talked about **abjads** earlier; these are a writing system where the main symbol is a consonant, and then you have certain diacritics being vowels. We mostly think of the Semitic languages—Arabic and Hebrew as the two principal languages—using abjad syllabaries. It is based off of the writing systems that you see with Sanskrit and some of the other Indo-European languages of South and Central Asia. They have a fascinating history, and if you want to learn more just let me know, and I can send you some references.

Then you have Japanese, which can be seen as a bit confusing if you are not Japanese. There is not one nor two, but four different writing systems. In fact, I didn't put one of them up here; it's actually five. **Kanji** is the one that probably you know about; kanji are based off of Chinese characters that are very similar to traditional Chinese characters. They were in use from the earliest stages of Japanese; in fact, if you can read kanji, you have a pretty good idea of what the Chinese writing system says. It is still very formal, and there are, to my knowledge, very few if any new kanji being created. It is more a reference to certain characters that are word-based, and they frequently have to do with something very similar that you hear in Mandarin or one of the Chinese languages. Then you have the **kana**. The kana are the syllabary; there's two different versions of it: **hiragana** and **katakana**. Hiragana are always the native Japanese terms, and it's more formal; katakana are loan words. If you pick up anything that is written in Japanese, and you see a mix of characters, it's almost always going to be one of these two kana. It will be a mix, depending on whether the lexicon is a native Japanese word or if it's a borrowed word. The one I forgot to put up here is **romanji**, which is Latin alphabet spelling of everything.

Korean technically has an alphabet but it's written out as clusters or syllables, so it's almost like a syllabary.

When we talk about diacritics, this is also an interesting important piece, especially to the Latin writing system—many of the languages that you use and know of use a Latin writing system, like English. Most of the Indo-European languages of Europe, though not all, use the Latin rating system. A number of them use Cyrillic., and there is a separate writing system for Armenian and Greek.

Digraphs: This might be a term you've heard before. Digraphs are two symbols that combine; think about a couple of those sibilants that we have an English: [ʃ, ʧ]. While we write the sound using two different symbols in the writing system ‘sh’ or ‘ch’. The same thing with that velarized nasal, that [ŋ] sound frequently spelled with two letters; the combination of them we know to represent a specific sound. We do this with vowels, as well.

Examples of Writing Systems

Let's take a look at some of these writing systems. I'm going to show you combined Cyrillic and Ancient Greek; you might find that to be confusing, but there's a reason

Α α	Β β	Γ γ	Δ δ	Ε ε	Ζ ζ	Η η	Θ θ
ἄλ.φα	βῆτα	γάμμα	δέλ.τα	ἒψιλ.όν	ζῆτα	ἦτα	θῆτα
alpha	beta	gamma	delta	epsilon	zeta	eta	theta
a	b	g	d	e	z	ē	th
[a/aː]	[b]	[g]	[d]	[e]	[zd/dz]	[ɛː]	[tʰ]
Ι ι	Κ κ	Λ λ	Μ μ	Ν ν	Ξ ξ	Ο ο	Π π
ἰῶτα	κάππα	λ.άμβδα	μῦ	νῦ	ξεῖ	ὄμικρόν	πεῖ
iota	kappa	lambda	mu	nu	xi	omikron	pi
i	k	l	m	n	ks/x	o	p
[i/iː]	[k]	[l]	[m]	[n]	[ks]	[o]	[p]
Ρ ρ	Σ σ/ς	Τ τ	Υ υ	Φ φ	Χ χ	Ψ ψ	Ω ω
ῥῶ	σῖγμα	ταῦ	ὗψιλ.όν	φεῖ	χεῖ	ψεῖ	ὦμέ.γα
rho	sigma	tau	upsilon	phi	chi	psi	omega
r/rh	s	t	u/y	ph	kh/ch	ps	ō
[r]	[s/z]	[t]	[y/yː]	[pʰ]	[kʰ]	[ps]	[ɔː]

On the right side is Ancient Greek, and Modern Greek is a variety of this.

А	Б	В	Г	Ґ	Д	Ђ	Ѓ	Е	Ё
A	Be	Ve	Ge	Ge upturn	De	Dje	Gje	Ye	Yo
Є	Ж	З	Ѕ	S	И	І	Ї	Й	Ј
Yest	Zhe	Ze	Zje	Dze	I	Dotted I	Yi	Short I	Je
К	Л	Љ	М	Н	Њ	О	П	Р	С
Ka	El	Lje	Em	En	Nje	O	Pe	Er	Es
Ć	Т	Ћ	Ќ	У	Ў	Ф	Х	Ц	Ч
[ć]	Te	Tshe	Kje	U	Short U	Ef	Kha	Tse	Che
Џ	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	
Dzhe	Sha	Shcha	Hard sign (Yer)	Yery	Soft sign (Yeri)	E	Yu	Ya	

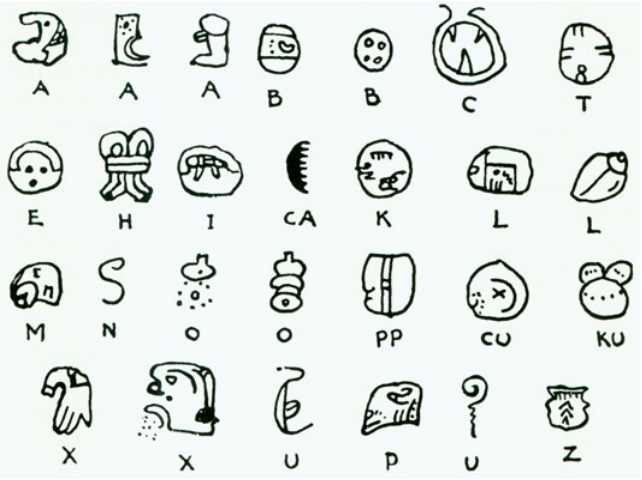
Cyrillic non-Slavic letters

Ӏ	Ә	Ғ	Ҙ	Ҫ	Ҡ	Қ	Ң	Ө
Palochka	Cyrillic Schwa	Ayn	Bashkir Dhe	Bashkir The	Bashkir Qa	Ka with descender	Ng	Barred O
Ү	Ұ	һ						
Straight U	Straight U with stroke	He						

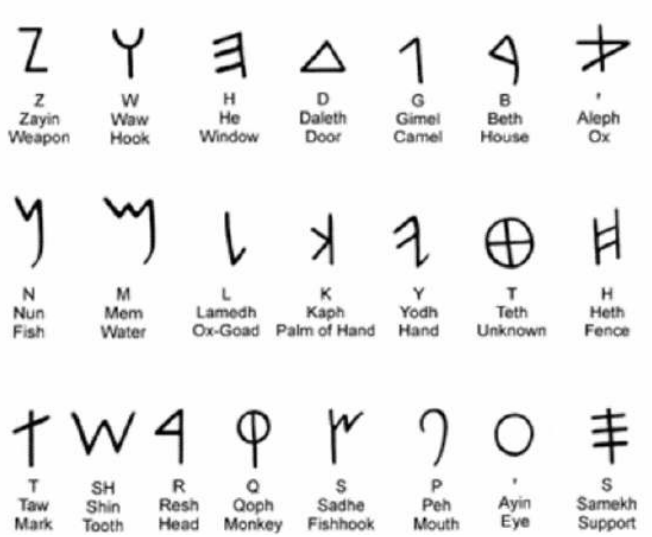
Cyrillic letters used in the past

Ѣ	Ѥ	Ѧ	Ѫ	Ѩ	Ѯ	Ѱ
A iotified	E iotified	Yus small	Yus big	Yus small iotified	Yus big iotified	Ksi	Psi
Ѳ	Ѵ	Ѷ	Ҫ	Ꙋy	Ꙗ	Ꙍ	Ѣ
Fita	Izhitsa	Izhitsa okovy	Koppa	Uk	Omega	Ot	Yat

Cyrillic is based off of a more modernized Middle Greek. If you notice the letter for that ‘a’, it's just like alpha. The B and V sounds are based off of beta. There are specific symbols that you see that came out of the Ukrainian or Russian branch of the Slavic language family. I phrase it that way because Ukrainians say they came up with Cyrillic first, and the Russian say they came up with it first; I am going to split the difference and say, Ukraine or Russian area of Slavic. We have a number of Slavic letters that get borrowed throughout the Slavic languages. It is interesting to note that some Slavic languages have moved away from Cyrillic, especially given the fall of the Soviet Union.



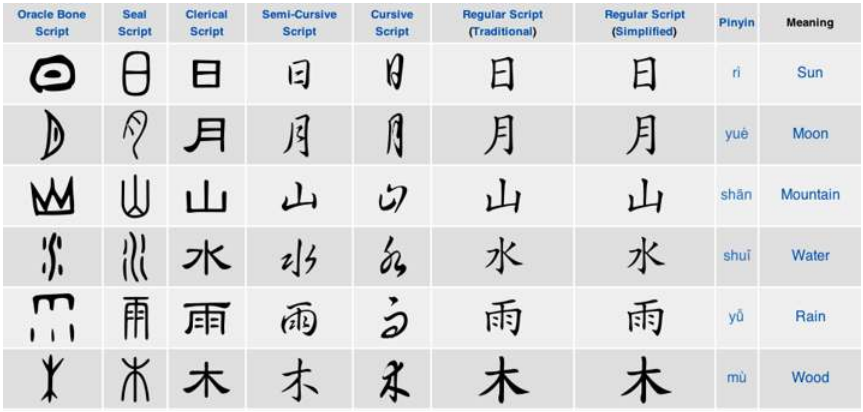
It is important to show hieroglyphics; you have seen Egyptian hieroglyphics many times; these are the Mayan ones. We're still trying to decipher these; what you see here are the ones that we have deciphered. There are so many more to come, and many, many more that we're still trying to work on.



I love showing folks Phoenician because in many ways it's an example of taking a later version of Egyptian hieroglyphs and changing them. There was a really fascinating documentary that was released in about how the ancient Egyptian hieroglyphs morphed, and the late hieroglyphs was almost identical to the Phoenician alphabet. Phoenician, to be fair, was a syllabary; it combined these symbols into syllable combinations. However, this is the basis for the Greek alphabet, so we frequently say that it is an alphabet, even though technically it's more of a syllabary. Notice that it is also a symbol that represents a concept. I brought up the concept of bull for a hieroglyph meaning not just an actual bull but the sound that that might also represent. If you turn your head to the side, as it were, or spin that icon for 'aleph' so that the point is down, it looks kind of like a bull; not surprisingly, the term 'aleph' in Phoenician meant 'ox'. There's a real connection to this; as I said, there's a full documentary on this.

خ	ح	ج	ث	ت	ب	ا
kha	haa	jiim	thaa	taa	baa	alif
ص	ش	س	ز	ر	ذ	د
saad	shiin	siin	zaay	raa	thaal	daal
ق	ف	غ	ع	ظ	ط	ض
qaaf	faa	ghayn	ayn	thaa	taa	daad
ي	و	ه	ن	م	ل	ك
yaa	waaw	ha	nuun	miim	laam	kaaf

I brought up abjads before; this is Arabic in the graphic. Arabic symbols are not just the actual main symbol; it is the diacritic that goes with it—the little dots or dashes. The abjad system has many more symbols. Abjads are syllabaries, because it is not just a consonant or a vowel; it's a combination, an entire syllable. Notice that most of these are onset and nucleus, not as many codas on here.



Oracle Bone Script	Seal Script	Clerical Script	Semi-Cursive Script	Cursive Script	Regular Script (Traditional)	Regular Script (Simplified)	Pinyin	Meaning
					日	日	rì	Sun
					月	月	yuè	Moon
					山	山	shān	Mountain
					水	水	shuǐ	Water
					雨	雨	yǔ	Rain
					木	木	mù	Wood

I love this graphic; I have no idea where I got it, but I love it. It is a history of the Chinese writing system. These symbols on the right are the earliest forms, and they go back about 4,000 years. From the Oracle Bone Script and Seal Script, and then when you start getting into Clerical Script, you're talking about old Chinese. Starting with Clerical Script, you see more systematic representations, going on to Regular Simplified Script that is used today in most of China and many, many parts of the diaspora. In Taiwan and a few other places, you tend to use Traditional Script, but you can start seeing the similarities from Clerical Script onwards, and to a lesser extent with the Seal Script. But they get less and less iconic; Oracle Bone Script is very highly iconic at this stage, and start getting more and more abstract, with more and more of a syllabary-type concept. Over here on the left, by the way, is the Pinyin, the Latin writing system that represents the sounds.



Thinking back to the Chinese Traditional and Cursive Scripts, compare those with kanji. Now you can see the connection between kanji and either Cursive or Regular Script; Depending on the country depends on which one is a little bit closer. You can see the very straight connections between kanji and the Chinese writing system. It is very straightforward, to the point that, like I said, if you can read kanji, you can get a pretty good idea of what is being said in the Chinese writing system just based on that connection.

Hiragana

あ a	い i	う u	え e	お o
か ka	き ki	く ku	け ke	こ ko
さ sa	し shi	す su	せ se	そ so
た ta	ち chi	つ tsu	て te	と to
な na	に ni	ぬ nu	ね ne	の no
は ha	ひ hi	ふ fu	へ he	ほ ho
ま ma	み mi	む mu	め me	も mo
や ya		ゆ yu		よ yo
ら ra	り ri	る ru	れ re	ろ ro
わ wa				を wo
ん n				

Katakana

ア a	イ i	ウ u	エ e	オ o
カ ka	キ ki	ク ku	ケ ke	コ ko
サ sa	シ shi	ス su	セ se	ソ so
タ ta	チ chi	ツ tsu	テ te	ト to
ナ na	ニ ni	ヌ nu	ネ ne	ノ no
ハ ha	ヒ hi	フ fu	ヘ he	ホ ho
マ ma	ミ mi	ム mu	メ me	モ mo
ヤ ya		ユ yu		ヨ yo
ラ ra	リ ri	ル ru	レ re	ロ ro
ワ wa				ヲ wo
ン n				

Compare kanji to kana. We have hiragana, which is for traditional and native Japanese terms, and then katakana, which is for borrowed terms.

A quick story: One of my very good friends from childhood is Japanese, and after college she moved back to Japan. She's very Japanese, and when I got married to my Japanese-American husband, who's half Japanese and half American, his last name is Smith. Jokingly my friend Mika sent me a gave me a new name: ‘sumisu-sama’. ‘Sama’ is, if you don't know in Japanese it is ‘ma'am’ or ‘madame’; it's very high-ranking term, and is frequently used for married women, especially if your family is from a higher echelon of Japanese society. ‘San’ is the one you're probably more familiar with, and that can be used as well, but ‘sama’ is a little bit higher still. So, she called me ‘sumisu-sama’, which just bowled me to see how she wrote it: in katakana, because Smith is not native Japanese. She also called me 'Maehara-sama'; 'Maehara' is hiragana, because it would be traditional, an old Japanese last name. It's one of the last names of my mother-in-law. (She actually should have used the kanji for Maehara, but she didn’t know it.)

As you can see, it’s not exactly kanji; it is very different. The kana are much more simplified; they are based off of kanji originally, but severely simplified.

The Korean Alphabet

Consonants \ Vowels	ㅏ [a]	ㅑ [ya]	ㅓ [ŏ]	ㅕ [yŏ]	ㅗ [o]	ㅛ [yo]	ㅜ [u]	ㅠ [yu]	ㅡ [ŭ]	ㅣ [i]
ㄱ [k,g]	가	갸	거	겨	고	교	구	규	그	기
ㄴ [n]	나	냐	너	녀	노	뇨	누	뉴	느	니
ㄷ [t,d]	다	댜	더	뎌	도	됴	두	듀	드	디
ㄹ [r,l]	라	랴	러	려	로	료	루	류	르	리
ㅁ [m]	마	먀	머	며	모	묘	무	뮤	므	미
ㅂ [p,b]	바	뱌	버	벼	보	뵤	부	뷰	브	비
ㅅ [s,sh]	사	샤	서	셔	소	쇼	수	슈	스	시
ㅇ (silent)	아	야	어	여	오	요	우	유	으	어
ㅈ [ch,j]	자	쟈	저	져	조	죠	주	쥬	즈	지
ㅊ [ch']	차	챠	처	쳐	초	쵸	추	츄	츠	치
ㅋ [k']	카	캬	커	켜	코	쿄	쿠	큐	크	키
ㅌ [t']	타	탸	터	텨	토	툐	투	튜	트	티
ㅍ [p']	파	퍄	퍼	펴	포	표	푸	퓨	프	피
ㅎ [h]	하	햐	허	혀	호	효	후	휴	흐	히

This is an example of the Korean alphabet. We say it is an alphabet but it's also a syllabary, because it's always a combination of a consonant and a vowel. For example, this is the symbol for [k, g], and you combine it with the vowel that goes with it, and you see those vowels here. The Korean alphabet, or *hangul* as it's called, is only about 100 years old; less than 150 years ago they switched to this writing system; before that they had a different writing system that was based off of the Chinese writing system.

अ a	आ ā	इ i	ई ī	उ u	ऊ ū
ऋ ṛ	ॠ ṝ	ऌ ḷ	ॡ ḹ		
ए e	ऐ ai	ओ o	औ au	अं aṃ	अः aḥ
क ka	ख kha	ग ga	घ gha	ङ ṅa	
च ca	छ cha	ज ja	झ jha	ञ ña	
ट ṭa	ठ ṭha	ड ḍa	ढ ḍha	ण ṇa	
त ta	थ tha	द da	ध dha	न na	
प pa	फ pha	ब ba	भ bha	म ma	
य ya	र ra	ल la	व va		
श śa	ष ṣa	स sa	ह ha		

When we talked about Arabic abjads, I said that there's a connection to Sanskrit. Hopefully, you can see a little bit of a here; this is the Hindi alphabet. A number of the Indic languages use a version that was originally based off the Sanskrit. (I couldn't find a really good graphic that I liked for Sanskrit so that's why I don't have it here.) With Hindi, you can see some similarities, where there is a main character along with certain diacritics. But it is evolved, so while this is not an object, you can see the connections, and there's a stronger still connection to ancient Sanskrit.

Punctuation

This is to write sounds and letters, but that's not the only part of a writing system. Clearly, we have to also include punctuation. In modern written languages, we have to include punctuation; you need to know when we stop and when we start. It's one of the frequent frustrations of the transcripts from these videos is there's bad or no punctuation.

Punctuation is a modern concept; it is not ancient. In fact, one of the challenges when you read a script or a document that is more than a couple hundred years old, is the punctuation is either not there or it's inconsistent; there isn't much in the way of standardization.

Writing in general is archaic, meaning that we always keep an older pronunciation with respect to how we write a term. English is not the only example of this, and you saw with the Great Vowel Shift a strong example of this, but it's not the only one. We also have **reform spelling**, which happens in a number of writing systems, English included. Reform spelling is when a linguistic society reverts back to a supposed previous pronunciation or spelling. In the case of *debt*, at no point in the English language was it [dɛbt]; there was no [bt] spoken in the history of English language with respect to that lexicon.

Text-speak

Speaking of punctuation and writing, let's do text speak. It is probably the bane of most English teachers' and professors' existence. Most feel that it's only been around for about 20 years; in fact, it's been around for longer than that. The use of acronyms, for example, is thousands of years old; as long as we have had an alphabet, we have had acronyms. In most languages, there's always acronyms of some kind. Once we started having instant communication, like on smartphones, computers and the like, this is changed rapidly to the point that I remember when FML l didn't just stand for 'fuck my life', it also was 'fuck my luck', getting that rhyming in there.

Sometimes we are using acronyms to create a euphemism—WTF being a great example of that. I'll give you an example with LOL: I remember when we started using these text-speak acronyms, I used to use LOL a lot, and I used to 'loller' a lot . My mom did not for the longest time have the same concept of LOL; it wasn't 'laugh out loud', because for her it stood for 'little old lady'. Acronyms will change from speech community to speech community and across generations. There were several artists in the late 70s and early 80s that frequently used numbers or truncation to express the same thing in written language. For example, instead of saying *ate*, they wrote the number *8*. Whether it was the *four* or *for*, it was always written with the numeral 4. Again, this is not new; humans have been doing this as long as we have had a written language, and certainly alphabets, in particular, seem to lend themselves to this truncation more easily.

Emoji! Let's talk about emoji and all these various things that we love about texting.

When we talk about text speak, we cannot be prescriptive, we have to remain objective and descriptive. There have been a number of complaints that text speak is ruining our lives, that we don't know how to write anymore, because we only think about life in however many characters we have in our text box. No, this isn't the case. Prescriptivists will do this; they will say, "Raaaaaar my God there totes is an effect!!! What the fuck-Oh my gawd-BBQ!!" (Sorry folks, I've been doing text speak for a long time.) The truth of the matter is that isn't the case; linguists, remember, describe what happens, and we can describe a number of things. When you work on your term paper, if you choose the language and tech option, one of the papers you're going to read is by Hawley Turner, who in 2009 had a very important study. She showed that if you teach writing and texting as different registers, much like we do with different dialects and different registers different modes of communication. If you teach students in that same way, they absolutely can split out that difference of when you have to use formal writing, and when you can use text speak, emojis, etc. It's an important reminder.

Language changes all the time. All aspects of language change all the time, and that includes the writing system. How generations from now will not only speak the language, but write the language that you're fluent in, that will change.

Are you willing to change?

8.9.2 More on Lost (?) Writing Systems (optional)

Watch this TedEd video in which Susan Lupack discusses the race to decode what (we think) is Linear B, the ancient written language that the Minoans may have used. (The video is captioned.)



Next, watch this TedEd video in which Jesse Byock explains what we know about Viking runestones. (The video is captioned.)



Next, watch this YouTube video in which Cecilia Pardo of the British Museum explains the khipu system that the Incan Empire (and their predecessors) used as both a record keeping device...and a writing system? Really interesting! (The video is captioned.)



9: Writing Systems is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10: Arabic and Romance- A Comparison (Optional)

Arabic and Romance: A Comparison (Optional)



Video Script

The last piece on historical linguistics—and this is purely optional; if you're not interested at all in this, you can skip watching or reading this, and it will not be on the historical linguistics quiz. I really want to go into this concept of Arabic and the fact that it is not a single monolithic language. Actually, if you look at the history of Arabic, and if you look at the modern usage of it, it's a lot more like the Romance family. So, we can talk about the Arabic family or the Arabic languages, and there's actually a good amount of comparison.

When you go to study Arabic, usually what you're learning is **Modern Standard Arabic (MSA)**; it is the Arabic that is used in writing, broadcasting, speechmaking, education, and many regions have adopted it as its official language. I phrased myself very purposefully; MSA is not necessarily what people speak on a day-to-day basis. Many of you who speak 'Arabic', the truth is you probably speak an Arabic that is specific to the region or speech community based on where you and your family come from.

Let's go back to that sociolinguistic definition of a language versus a dialect, and talk about mutual intelligibility. Is Arabic a single language with many dialects? Or, are they many related languages? Or are they something in between, something in progress? That's the main question.

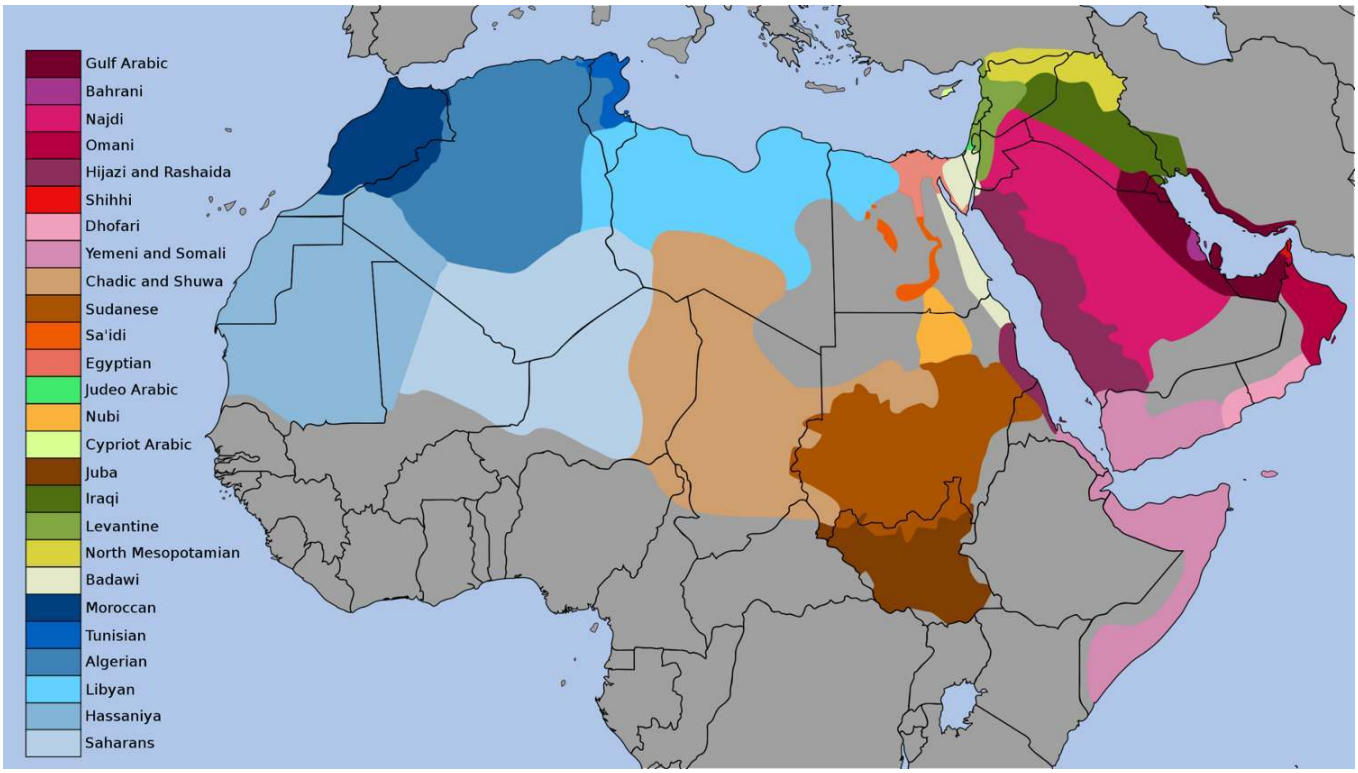
This wasn't my initial question; this wasn't an initial thought that I had. This question and topic came from a student from fall of 2020 who wanted to know more about Arabic and why, for example, his mother spoke one version and his father spoke a different version; they were very different.

Here's the reality when we're talking about Arabic: there are multiple versions of this. Classical Arabic is the language of the Quran, the holy book of Islam. It is also the language of Islam, as far as anything liturgical or anything having to do with religion; it is almost always in Classical Arabic. The old writings are in Classical Arabic. Modern Standard Arabic is what you learn in a classroom; it is what you hear in mass media, as I said. I lingua franca, if you will, of the Arabic-speaking world. It is the one that everybody learns. It is a modernization of Classical Arabic, but a specific component of it is the fact that it is not of one place or one region. It is kind of like the overarching language that people will use in an international setting, or in an official government setting. But then in their day-to-day lives, they have their colloquial Arabic. In many cases, it's written down, although not always; when you see written Arabic, it is Modern Standard Arabic almost always, especially in a textbook or something like that.

It is also important when we're talking about the speech communities, if you know anything about the cultures associated with them, you know that there is a big difference between a **sedentary speech community** and a **Bedouin speech community**. The Bedouin in some cases, especially anthropologically, refers to a specific culture. But I'm going to use it in its more macro concept, which is a group of people, a speech community, that do not stay in one place. They move around; it is part of their pattern, to be

migratory. In northern Africa, in parts of Arabia and the Levant, even Central Asia, you have Bedouin societies. These are all migratory societies; that's going to come into play with respect to any language family, but certainly when we're talking about Arabic. The sedentary speech communities are the ones that are established in a city or region; they don't move around, they don't migrate.

Let's talk about Arabic. When we say 'Arabic', let's focus on Colloquial Arabic.



The major regions of Colloquial Arabic are as follows:

- Saharan Africa;
- A little bit of Sub-Saharan Africa
- Parts of the Levant
- Parts of Arabia
- A little bit of the Persian Gulf

Keep in mind: we're not going to include any of the Arabic that spoken outside of this region, because that is more having to do with immigration patterns going out, the diaspora.

Notice some holes: we've got the Nilo-Saharan languages in this region, also Aramaic, and a number of the other languages of eastern Africa. In the Arabian Peninsula, we have a few more Afro-Asiatic languages. What is not marked in here is Hebrew; where modern Hebrew is spoken, we also have Arabic, specifically Levantine Arabic. You can see the different colors in the representation here

To give you a more concrete example, let me show you how to say a straightforward sentence

Examples of various Arabic versions of the same four sentences. Taken from https://en.Wikipedia.org/wiki/Variet...al_differences

Variety	I love reading a lot.	When I went to the library,	I only found this old book.	I wanted to read a book about the history of women in France.

Variety	I love reading a lot.	When I went to the library,	I only found this old book.	I wanted to read a book about the history of women in France.
Modern Standard Arabic	أَنَا أُحِبُّ القِرَاءَةَ كَثِيرًا ʾana ʾuḥibbu_l-qirāʾata kaṯīrā ʔana: ʔuħibːu_lqiraːʔata kaθiːraː	عِنْدَمَا ذَهَبْتُ إِلَى المَكْتَبَة ʿindamā ḏahabtu ʾila_l-maktabah ʕinda maː ðahabtu ʔila_l maktabah	لَمْ أَجِد سِوَى هٰذَا الكِتَابِ القَدِيم lam ʾaǧid siwā hāḏa_l-kitābi_l-qadīm lam ʔad͡ʒid siwaː haːða_lkitaːbi_lqadiːm	كُنْتُ أُرِيدُ أَنْ أَقْرَأَ كِتَابًا عَن تَارِيخِ المَرأَةِ فِي فَرَنسَا kuntu ʾurīdu an ʾaqraʾa kitāban ʿan tārīḫi_l-marʾati fī faransā kuntu ʔuriːdu ʔan ʔaqraʔa kitaːban ʕan taːriːχi_lmarʔati fiː faransaː
Maltese	jien inħobb naqra ħafna	meta mort il-librerija	sibt biss dan il-ktieb il-qadim	ridt naqra ktieb dwar il-ġrajja tan-nisa fi Franza.
Tunisian (Tunis)	nḥəbb năqṛa baṛʃa	wăqtəli mʃit l-əl-măktba	ma-lqīt kān ha-lə-ktēb lə-qdīm	kənt nḥəbb năqṛa ktēb ʕla tērīḵ lə-mṛa fi fṛānsa
Algerian (Algiers)	ʔāna nḥəbb nəqṛa b-ez-zaf	ki rŭħt l-əl-măktaba	ma-lqīt ḡīr hād lə-ktāb lə-qdīm	kŭnt ḥayəb nəqṛa ktāb ʕla t-tārīḵ təʕ lə-mṛa fi fṛānsa
Moroccan (Casablanca)	ʔāna kanebɣi naqra b-ez-zāf	melli mʃīt el-maktaba	ma-lqīt ḡīr hād le-ktāb le-qdīm	kunt bāḡi naqra ktāb ʕla tārīḵ le-mra fe-fransa

Variety	I love reading a lot.	When I went to the library,	I only found this old book.	I wanted to read a book about the history of women in France.
Egyptian (Cairo)	ʔana baḥebb el-ʔerāya awi	lamma roḥt el-maktab a	ma-lʔet-ʃ ʔella l-ketāb el-ʔadīm da	kont ʕāyez ʔaʔra ketāb ʕan tarīχ es-settāt fe faransa
Northern Jordanian (Irbid)	ʔana/ʔani kṯīr baḥebb il-qirāʔa	lamma ruḥt ʕal-mekteb e	ma lagēteʃ ʔilla ha-l-ktāb l-gadīm	kān baddi ʔagra ktāb ʕan tārīχ l-mara b-faransa
Jordanian (Amman)	ʔana ktīr baḥebb il-qirāʔa	lamma ruḥt ʕal-mekteb e	ma lagēt ʔilla hal-ktāb l-gadīm	kan beddi ʔaqraʔ ktāb ʕan tārīχ l-mara b-faransa
Lebanese (Beirut)	ʔana ktīr bḥebb l-ʔ(i)rēye	lamma reḥt ʕal-makt(a)be	ma l(a)ʔēt ʔilla ha-le-ktēb l-ʔ(a)dīm	kēn badde ʔeʔra ktēb ʕan tērīχ l-mara b-f(a)ransa
Syrian (Damascus)	ʔana ktīr bḥebb l-ʔirēye	lamma reḥt ʕal-maktbe	ma lʔēt ʔilla ha-l-ktēb l-ʔdīm	kān biddi ʔra ktāb ʕan tārīχ l-mara b-fransa
Gulf (Kuwait)	ʔāna wāyid ʔaḥibb il-qirāʾa	lamma n riḥt il-maktab a	ma ligēt ʔilla ha-l-kitāb il-qadīm	kint ʔabī ʔagra kitāb ʕan tarīχ il-ḥarīm b-faransa
Hejazi (Jeddah)	ʔana marra ʔaḥubb al-girāya	lamma ruħt al-maktab a	ma ligīt ǧēr hāda l-kitāb al-gadīm	kunt ʔabǧa ʔaǧra kitāb ʕan tārīχ al-ḥarīm fi faransa

Variety	I love reading a lot.	When I went to the library,	I only found this old book.	I wanted to read a book about the history of women in France.
Sanaani Arabic (Sanaa)	ʔana bajn ʔaḥibb el-gerāje gawi	ḥīn sert salā el-maktab e	ma legēt-ʃ ḏajje l-ketāb l-gadīm	kont aʃti ʔagra ketāb ʕan tarīḵ l-mare wasṭ farānsa
Mesopotamian (Baghdad)	ʔāni kulliš ʔaḥebb lu-qrāya	min reḥit lil-maktab a	ma ligēt ḡīr hāḏa l-ketab el-ʕatīg	redet ʔaqra ketāb ʕan tārīḵ l-imrayyāt eb-fransa

I have Modern Standard Arabic as the top, so that you can see how it is set up, this is the script, this is the, shall we say, Latin alphabet version of it, and then you have the IPA is the third. In the rest of the column is all IPA. You also see Maltese, Tunisian, Algerian, Casablanca-based Moroccan, Cairo-based Egyptian, northern Jordanian, Amman-based Jordanian, Beirut-based Lebanese, Damascene Syrian, Gulf Kuwaiti, Hejazi in Saudi Arabia, Sanaani Arabic from a different part of Arabian Peninsula, and then Baghdadi or Mesopotamian Arabic. There are lot of differences! Notice a few things:

- There's a case system that you see in Modern Standard Arabic, that may or may not exist in the other versions.
- You will also see word order changes; look at *I love reading a lot*, at the order of the second and third lexicon. Sometimes there's a fourth, by the way; if you look down this column, sometimes it's three and sometimes it's four, and there's a different order.
- Although you don't see it in here, there's a passive voice that is different; it's set up different in all these different regions.
- Some of these varieties of Colloquial Arabic have singular versus plural, and some have added a dual.
- There is a shift in the vowels.
- There are consonant cluster differences.

If you go back to the map, it kind of clusters; Realistically, you have six main regions.

- In northern Africa, this is called Maghrebi. There is more influence from the indigenous languages of that region, Berber in particular. You have also a little bit of Romance languages because notice it's right on the Mediterranean; it's going to have contact with Spanish, a little bit of Portuguese, it's going to have contact with French, especially given that in this region French is an official language. These were some of the last territories that France gave up, so the French connection is very strong. Because of that, you're going to have more borrowings, especially from French, but maybe even from Spanish or Italian. You also have Punic languages as an influence—that would be Phoenician, or more specifically Carthaginian—historically derived languages from those regions.
- Sudanese: you have more contact with Nubian languages from the Nilo-Nubian languages in particular. Sudanese and Chadian are the two main colloquial Arabics from that region, but there's some others.
- Egyptian Arabic: there's going to be more influence from Coptic languages and other Afro-Asiatic languages that are indigenous to that region.
- Mesopotamian: think Mesopotamia, the Tigris and Euphrates rivers. This area is going to have influence from the languages from eastern Turkey—some of them are Turkic and some of them are Anatolian, based off the Hittites—as well as influence from Iranian languages, especially Farsi. The Arabic spoken in Baghdad, for example, is going to be different than the Arabic spoken to the south.
- The Levant: This is frequently the area that we considered to be Israel, Lebanon, Syria, a bit of Jordan, and parts of the southern and central areas of Iraq. In those cases, you've got some Turkish and some Greek; that makes sense, historically. There is some

influence from the indigenous languages from that area. Aramaic comes in, so another Afro-Asiatic language that's historically been in that region for some time.

- Peninsular Arabic: The big obvious one would be Saudi Arabia as a country, but you've got a little bit of the southern edge of Iraq and of the eastern half of Jordan. The indigenous languages of the Arabian Peninsula filter in here.

What does this mean?

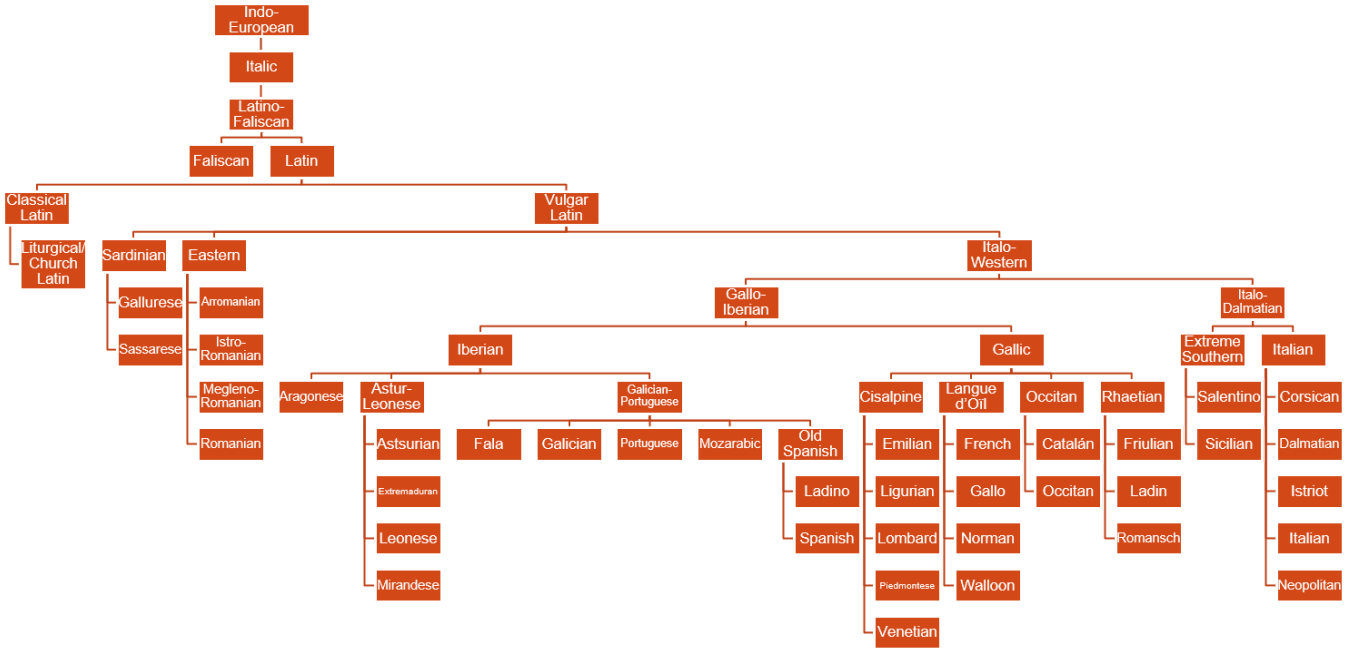
If you have somebody from Morocco trying to some talk to somebody from Syria, those are two different Arabics, and they can be radically different. In fact, Tunisian, Algerian, Moroccan, even Maltese, they're pretty close. Compare this to Lebanese or Syrian, which are both very different to the Maghrebi Arabics. There are entire combinations of sounds that you don't see in the other section.

Why bring this up?

Historically, folks like to think of Arabic as a monolithic concept that just has multiple dialects. Remember when we talked about a language versus a dialect, and we were talking about mutual intelligibility. I don't know about you, I certainly am not an expert in Arabic, I do not speak Arabic and anyway, I don't speak any Afro-Asiatic language…but I can look at these data, and I can compare and contrast. Even if I don't have the glosses, the morpheme, or lexicon-by-lexicon, morpheme-by-morpheme translation of what these concepts are, I can just see that Levantine is very different from Modern Standard Arabic, or from Maghrebi, or from Mesopotamian. I can use my eyes and ears, and see and hear that. This is not just in the minor intelligibility difficulties; this is not like saying, “It's kind of hard to hear and understand somebody from Australia or Scotland.” If they slow down and they speak clearly, we can understand them as Americans. But somebody from Casa Blanca in Morocco cannot necessarily understand somebody from Damascus, Syria. That's a question about mutual intelligibility.

What does this have to do with the Romance languages?

Here is the full Indo European tree.



Remember that when we're talking about the Romance family, we're talking about the languages that derived from Latin—Latin is not a Romance language, rather it's the parent of the Romance languages. If you take somebody from Madrid, Spain and you plop them into Rome, Italy, there are some things that are in common, but there's a lot that's different. Some of the sounds might be the same, but the some of the morphology isn't; the syntax might be generally the same, but there might be some differences. If you take that same person from Madrid and put them in Sardinia, now you got something totally different; Sardinian is a very distinct language. (Unfortunately, it is dying out; I wish it weren't, because it's a very beautiful language.) It encapsulates, it keeps so many of the archaic elements that we see in Vulgar Latin. Therefore, if you're going from Madrid, Spain, to Sardinia, that's really difficult, linguistically. If you take that same person from Madrid and put them into Paris, now you're talking big changes. Granted, most all of them are phonological, but there are morpho-syntactic changes there are semantic changes. You can't just go by the cognates, because sometimes they are radically different. Now, take that same person, and put them in to Bucharest, Romania. That is a radically different linguistic landscape, even more different than anything else they've heard before. Yes, these are all Romance languages and, believe it or not, if you can read one Romance language, you've got a pretty good idea of what's going to happen in the other Romance language. It may not be exact, but it's pretty good. I'm fluent in Spanish and Italian, and I speak a little bit of Portuguese. I can look at French and get 70% of what's there; I can look a Romanian and get about 65% of what's there. When French and Rumanian have been spoken to me, I get a chunk of it; some of it sounds familiar, but a lot of it doesn't.

Why is this a good comparison to Arabic?

Because it's a lot of the same concepts.

For the Romance languages, we have no problem at all saying that those are different languages, that the two dialects of Sardinian, Romanian, the various Spanish dialects, Italian and its various dialects, Galician, Portuguese, French, Romansch, on and on. We have no problem saying that those are different languages. Why can't we say that for Arabic? By the way, this is not just a European-minded person saying that; even within Arabic scholars, many of them refused to say that those are different languages.

Let me show you another way that we can say that those are different languages in Arabic. If we look at the history of the Romance languages, when we go from Latin to the Early Romance period—remember that Latin as a concept, even Vulgar Latin and Late Latin, more or less end with the fall of the Roman Empire. In total for the Romance languages, we have 2000 years of history. At the end of second century CE, or the end of the 100s CE, that's when we get Vulgar Latin turning into Common Latin. When we say Early Romance, we're talking around 850-900 CE; the fall of the Roman Empire is in the early sixth century, or the early 500s CE. That means that there are about 300 years where we have little bits and pieces, but not books, written in whatever the local languages are—not in Latin. This is the overall timeline for the Romance languages.

Notice the same timeline for the Arabic languages: Old Arabic is in the first century CE, and then Classical Arabic starts around third century CE, and then all the way through to the modernization, which is the 19th century. Pretty similar.

In the earliest of the Romance languages, we already started seeing a reduction in the case system, we see a fixed and changed word order. Although you could play with it, Latin was predominantly a subject-object-verb (SOV) word order; Romance languages early on fix that to SVO because they don't have case anymore. Guess what you see with Arabic? A reduction or loss of case system from Classical to Modern Arabic, and it changed word order. There is a more synthetic morpho-syntax for both groups. See where I'm going?

You have various changes semantically and morpho-semantically. This includes the loss of the middle voice, which is something I can explain in a different time. (If you have active voice and passive voice, Latin used to have a middle voice.) There is a loss of the neuter gender actually in vulgar Latin. There is a different passive construction from Classical Arabic to Modern Arabic, as well as a different number system. There are changes that reflect both the phonology, syntax, and the semantics. There are changes in deixis, different moods; we talked about the energetic that we saw in Modern Arabic, but in most Colloquial Arabics, so you don't have the energetic. The variations in lexicon, influence of the morpho-syntax, influence on the morpho-semantics, they all pattern the same way and across the same amount of time.

Again, I bring up this question: If we can say that the Early Romance languages started separating out and started individualized across 2000 years of history, why can't we say that for Arabic and the Arabic languages, if you will?

I'll bring up one more point: This use of Classical Arabic and then later on Modern Standard Arabic, these all have to do with formal education and formal governments. The same is true with Latin; first it was Classical Latin and then what was called Late Latin or Medieval Latin. Those were used at very high levels, the educated highest educated people only. Common folks did not use this, and what we're seeing—certainly with Classical Arabic, but even with Modern Standard Arabic—in many places, people are eschewing the use of Modern Standard Arabic. They understand it; if they turn on the radio or TV, or if they pick up a paper, they understand it. But they do not speak it. This really is suggesting that these Colloquial Arabics are different languages.

When I say that linguists have to be objective, we have to pull back, we have to be descriptive of what we see and observe and hear, and we have to document it accordingly, this also means revising our definitions. There isn't an Arabic language. There are many Arabic languages. And this is just Arabic; there are many other languages that we could also revisit with this kind of observation and comparison. We can say, "Okay, we don't have one language; we have several."

10: Arabic and Romance- A Comparison (Optional) is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

11: References

Just as before, the Fromkin, Rodman, Hyman text of *Introduction to Language* (2019; Cengage) has information and many materials. Below are the others that were consulted for this chapter, and can serve as good tools for further investigation.

Common textbooks

- Coulmas, Florian. 2012. *Writing Systems: An Introduction to their Linguistic Analysis*. Cambridge University Press.
- Croft, William. 2012. *Typology and Universals*. Cambridge University Press.
- Greenberg, Joseph. 2005. *Genetic Linguistics: Essays on Theory and Method*. (Edited by William Croft.) London: Oxford University Press
- Hock, Hans Henrich. 2021. *Principles of Historical Linguistics*, 2nd edition. Berlin: de Gruyter
- Payne, Thomas. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge University Press.
- Song, Jae Jung. 2010. *The Oxford Handbook on Linguistic Typology*. London: Oxford University Press

Peer reviewed articles and books

- Baines, Lawrence. 2012. "A Future of Fewer Words? Five Trends Shaping the Future of Language." *The Futurist*, March-April 2012: 42-46.
- Croft, William. 2006. "Evolutionary Models and Functional-Typological Theories of Language Change." In: van Kemenade, Ans and Bettelou Los, Eds. *The Handbook of the History of English*. Blackwell Publishing.
- Crystal, David. 2009. *Txting: The Gr8 Db8*. London: Oxford University Press.
- McCulloch, Gretchen. 2019. *Because Internet: Understanding the New Rules of Language*. Riverhead Books.
- Hawley Turner, Kristen. 2009. "Flipping the Switch: Code-Switching from Text Speak to Standard English." *The English Journal*, 98 (5): 60-65.
- Nevalainen, Terttu. 2006. "Historical Sociolinguistics and Language Change." In: van Kemenade, Ans and Bettelou Los, Eds. *The Handbook of the History of English*. Blackwell Publishing.
- Perley, Bernard C. 2012. "Zombie Linguistics: Experts, Endangered Languages, and the Curse of Undead Voices." *Anthropological Forum*, 22 (2): 133-149.

Websites

These websites are crucial in analyzing and compiling data on any language.

- Ethnologue
- World Atlas of Language Structures (WALS)

11: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

9: Learning Languages- Language Acquisition

9.1: Universals and Definitions
9.2: Child Language Acquisition Theories
9.3: Child Language Acquisitions Stages
9.4: Child Language Acquisition Linguistics
9.5: Child Language Acquisition Bilingualism
9.6: Second Language Acquisitions Theories and Components
9.7: Second Language Acquisition Teaching Methods
9.8: References

9: Learning Languages- Language Acquisition is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.1: Universals and Definitions

9.1.1 Universals and Definitions, from Sarah Harmon



Video Script

As we talk about language acquisition, both child language acquisition and adult language acquisition, it's really important to keep a few universals and a few definitions in mind. Before we actually start off with those universals and definitions, I'd like to tell a couple of stories, both of which were told to me and have been documented by someone else; they actually come from journal articles publications. They encapsulate so much about what it's like to learn a language from a child's point of view.

Let me start with the first one. This little boy, who was about three, has this plastic fish toy, but he doesn't say fish with that post-alveolar fricative; he says [fɪs], so he swapped [ʃ] for [s], right? That's what he produces. Whenever we do these kinds of data collections with children, we always have adults present, usually the researcher, perhaps the parents, etc. There's always some kind of interaction, and, in this case, a researcher worked with the little boy. So, the child is playing with this plastic fish, and the adult decides to use the same pronunciation as the child.

- The adult points to the toy and says, “Is this your [fɪs]?”
- The little boy says, “No, this is my [fɪs].”
- The adult tries again: “Yes, this is your [fɪs].”
- The child is still confused: “No, this is my [fɪʃ].”
- The adult pivots and says, “Oh, this is your [fɪʃ].”
- The little boy says, “Yeah, this is my [fɪs]!”

You probably have come across something very similar to that with either your own children, your siblings, or any other young child who's about three or so.

A different scenario, in this case you've got a bilingual situation. The child is about the same age, about 2.5-3 years old, and is a French-German bilingual. Usually in those cases, one parent speaks one language, while the other parent has the other language. You can think of eastern France, western Germany, and Switzerland as well, where French and German bilingualism would be very common. This child is with his father; this is another again very common setup where the investigators are just observing the interactions between the child and the parent. In this case, the little boy has a bunch of different things that he's playing with, and one of the things that he has in his box of toys is a button. Normally the child uses German with the father.

- The father points to the button and says [lnopf], which is the German word for ‘button’.
- The child says [nop].
- The father tries again [knopf].
- The child tries again [nop]. He's not quite getting the pronunciation down.

- Father tries a third time and says [knopf].
- The child thinks, “To heck with this, I am switching languages,” and he says [butõ].

This concept of the child maybe not able to produce, but as able to understand, and in the second case, with some more tools under his belt. I want you to keep in mind both of these scenarios throughout all of this chapter, whether we're talking about child language acquisition or adult language acquisition, because they are very much an exemplar of what is typical in language learning situations.

To start off with this section, and with this entire chapter, we really have three main questions that we're going to answer:

- We're going to talk about how humans learn their first language;
- We're also going to talk about humans learn an additional language, both as a child and as an adult.

Those three are different setups or different questions for a reason and we'll come back to it later; anyone that tells you that to learn a language as an adult is exactly the same as learning it as a child is lacking an extreme amount of information. We'll get to that soon enough.

Both of the stories that I brought up. and, in fact, everything having to do with language, comes down to this question of **competence** versus **performance**. I talked about that in the first chapter, but just to refresh your memory a bit, competence is also considered knowledge. If we think about it that way, knowledge or competence is what you know of a language. The performance is how will you use it. To put it another way, language is a skill. Using this skill, using this tool, is part understanding what you have to do, but also part performing the action. If I want to learn how to cook, I have to understand a couple of basic principles: how to turn on a heat source, how do extinguish said heat source, how to gain water and make it do different things. You don't need tons of knowledge—the old saying is, if you can boil water, you can cook—it's not entirely true but certainly it's a basic first step. The same is true from language; language is a tool and using it is the skill. We have to understand basic concepts before we can produce them. Just because we can't produce them does not mean we don't understand them; I’m a very advanced cook and can basically cook anything—I don't just mean an American sense; I mean literally I cook cuisine from around the world and have learned techniques from around the world. I can cook just about anything. Most of you probably do not have that skill set; you just haven't learned it yet. But you can taste everything for the most part. You can go to a restaurant, regardless of its cultural background, regardless of the techniques that they use, and you can taste the different textures and flavors, and smell the aromas. You can understand the food; you may not be able to reproduce that soufflé or that barbecue, or anything else that you're eating, but you can understand the food and the flavors. Language is the same thing; language is a tool, and using it is a skill. You have to understand it before you can perform it. You can understand a lot more than you can perform for any language that you're still trying to acquire.

It always makes me giggle in my Spanish classes, when a student tells me that they can understand Spanish better than they can speak it. I tell them that's because they're a normal human being; this is what we do. We comprehend before we can perform.

9.1: Universals and Definitions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Sarah Harmon.

9.2: Child Language Acquisition Theories

9.2.1 Child Language Acquisition Theories, from Sarah Harmon



Video Script

When it comes to theories of how children learn their first language or languages, there has been an evolution in this area. We're going to start off with some very common and old theories of how languages are acquired by children, but we're also going to point out what's wrong with them and how we have evolved in our thinking. It's important to understand that this is an area we are still trying to figure out; we don't have all the answers yet. We definitely are working with hypotheses and theories, but we don't know for sure. For reasons that are obvious, we don't fully understand how the brain works yet; we're still learning that too. Therefore, what we learn now could change—in fact, probably will change—over the course of your lifetime. It certainly has changed over the course of my lifetime.

To start off, let's talk about some theories that used to be very popular, and that used to be considered 'absolute truth'—as it turns out, it's not the case. These are all old and erroneous, in part, if not in whole.

The big discussion has always been on imitation versus reinforcement versus analogy. **Imitation** refers to the idea that children imitate language that they hear and just spit it back out. Well, yes and no; clearly, we have seen in data over the last 50 years that definitely there is something to the concept: children do hear things and they try to emulate it to the best of their ability. Think of the [fɪs, fɪʃ] scenario, think of [knopf, nop]; they were trying. But that is not the only way that they learn language, because otherwise they wouldn't make mistakes. **Reinforcement** is the theory that when you positively reinforce good behavior, the child will continue doing it, and that negative behavior is chastised or punished somehow. The problem with this theory is that children make all sorts of combinations with respect to their languages, and doubly so if they're in a multilingual environment. It doesn't matter how many times you tell a child that the plural of *foot* is not *foots* and that it's *feet*, because the child is not going to produce until some point in the future. Suffice it to say that it's not just the reinforcement; you can't just reward good language behavior and castigate bad language behavior; it doesn't have an effect. The same is true with **analogy**. Most people used to say that children acquired language and they built everything via analogy; they just made the same mistakes, and then learn to correct them at some point. Again, some of this is true; certainly, some of the errors that are frequently done by children with so called irregular forms in morphology and syntax could be examples of analogy. But there are plenty of examples where that doesn't happen. We come back to what seems to be the issue: analogy doesn't exactly describe most of what happens with respect to child language acquisition. **Conditioned response** is a little bit like reinforcement; it's the BF Skinner version of it. If you've taken psychology, you probably know what this concept is: a reaction to stimulus, reinforcing the positive while chastising or castigating the negative. Again, we know that doesn't work, let alone the ethical issues that come with conditioned responses examples.

We know that none of these explain what happens with respect to children as they learn their first language or languages. The question becomes: What do we know?

These are four theories and hypotheses that we have currently. While they may not cover everything with respect to child language acquisition, it does seem to be the case that they have a lot more in common than not.

The **Innateness Hypothesis** is the first one to talk about, both in general and chronologically. It's also the first one that tapped into Chomsky's Universal Grammar concept, the fact that we have this innate ability to speak a language, and that, so far, we have yet to observe the exact same thing in any other animal on the planet. I'm hedging that hugely for a reason; when we get to the next chapter and talk about animal communication, I'll explain a little bit more, but think about those hallmarks of language, because that's what the Innateness Hypothesis comes back to. It also refers to what we talked about with respect to topology, the fact that there are universals that we see, as well as strong tendencies and non-absolute universals. We see so many trends with respect to languages that are very common, if not ubiquitous, so clearly there's something there.

From the Innateness Hypothesis, we get these other three theories, and they are additions or specifications of this innateness concept, tapping into Universal Grammar and this facility that we have with respect to language.

Active Construction of Grammar Theory is a theory that children actively invent rules as they go along. In other words, they observe and absorb language; they actively and subconsciously create patterns. They think they see and/or hear a pattern, and then they build upon it. It's mostly based off the observation of how children imitate the other children that they're around—not imitating the adults. By the way, this is connected to creolization because, again, we get this concept of peers helping each other out. There have been numerous studies with respect to how children pick up new terms, and most of the time it's not from their adult caregivers. Most of the time is from other children—sometimes siblings or cousins, maybe in the school or daycare situation. It's as if they're taking an input subconsciously, analyzing it subconsciously, hypothesizing what the rule will be, and then apply that to the mental lexicon. We do not have clear evidence of this, like we can't analyze the brain to see that happening. But we definitely do notice that this is part of how children acquire new terminology and new phrasing; they get it mostly from peers. It certainly explains overgeneralization patterns, but frequently people chalked that up to analogy. If you have a child learning a given language—I'll just say English—and they are not around other children, there is a tendency for there to be a slowing of that acquisition. But the more children they're around, the more they acquire and the quicker they acquire it. That being said, it's really hard to observe this actually happening in the brain; this is just based on observation of children in a lab setting and in social settings.

Connectionist Theory is the fact that children learn by creating neural connections and that this is based off of exposure to a variety of stimuli. In this case, that stimulates the new language, and so they basically learn these associations and they build from there—they're making connections. We see this in how they acquire other skills, whether it's walking and moving and motor skills to building with blocks to building a sand castle or anything like that we see them do this with other skills and the thought is okay, maybe they do this with language. There seems to be again some evidence of this; it taps into this overgeneralization aspect, the fact that children are making connections based off of prior knowledge or input or stimulus. That being said, again we have the same problem with active construction, which is we can't observe what's going on between the ears, as it were. It's hard to do that, so we don't really know that that's the case. But we suspect there's something to that.

The last one is **Social Interaction Theory**, which is in some ways a combination of both of the previous two. It's the concept that children acquire language through social interaction; the more they're around peers, in particular—but in this case, you also include the adults—they have more input and they're making more connections. That social interaction also feeds their desire to want to speak more and to communicate in ways that are clear. There's something to be said for this as well, and in a number of cultures, especially European and American cultures, but even in parts of Asia and parts of Africa, we do see that adults will slow down or simplify their language so that it is more comprehensible. The child can interact more because they're not overly intimidated by all these people going a mile a minute with complex phraseology and lexicon. There's something to that, because when you are including people in a dialogue, you want there to be a communication, not just of ideas, but the feeling that you can communicate with me, because I understand you (or want to understand you) and you want to understand me. We slow things down for children, and that's why children can learn from other children a little bit better. Again, we have that same problem of how do you measure that.

There's also the other interesting piece, and I'll use one of my nephews as an example, my older nephew who's now eight. He is fascinated by big words, and has been his entire linguistic life, even as a four-year-old. He did not want to hear so called 'child language' or 'baby language'; he wanted big words. He loved it when Auntie Sarah would start talking; several times, unbeknownst to some of my students, would be sitting here in my room, as I did my online lectures for linguistics, because he wanted to hear the big words. He tried to use the big words—he still tries to use the big words—but it doesn't always work. Now, as an eight-year-old, he's getting stronger with his reading, and he wants the books with big words in them; he wants to try and read

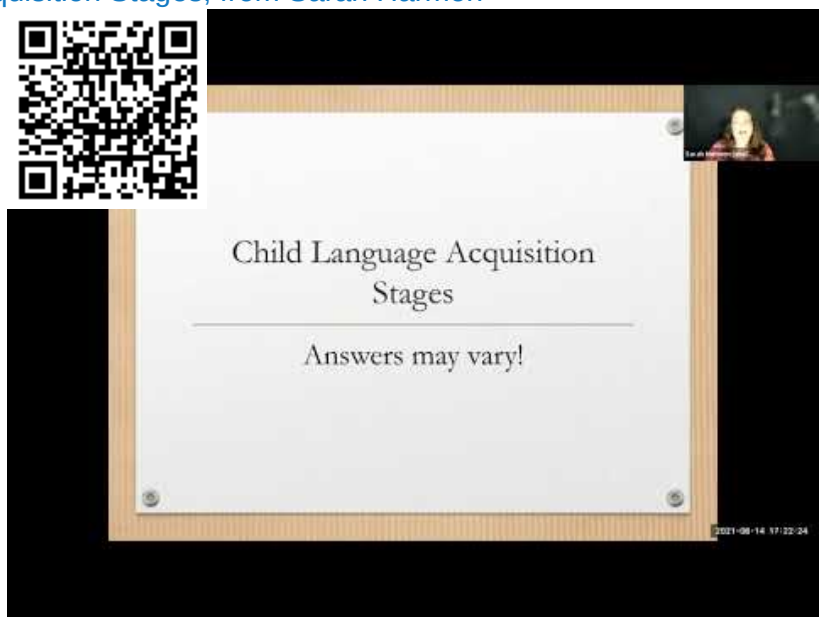
them. There seems to be something with respect to this Social Interaction Theory; if you have a child around adults and children, they're going to want to interact with both at as close to level as possible. All three of those theories are built upon the Innateness Hypothesis.

We don't know for sure, and it will probably be some time before we really can understand, how the brain works, especially with children. But there are certain things that we can observe, and that'll be in the next section, when we talk about the different stages of language acquisition. There is plenty that we can observe just by watching, but hopefully someday we can crack this thing and really get to it.

9.2: Child Language Acquisition Theories is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.3: Child Language Acquisitions Stages

Child Language Acquisition Stages, from Sarah Harmon



Video Script

We're going to go into the stages of child language acquisition. This is based off of about 50 years' worth of data collection, observation, trials, and studies—all ethically above board. That means that at no point was a child was kept in a room, was never talked to, and then we observed how they developed. That has never happened; no one has done that. Ethically, as a researcher, at least to my knowledge. Rather, this is based off of observation. When we bring a child into observation for language acquisition and other aspects of child development, we see them in a lab. We see them play with each other; we give them specific stimuli and we record what they do with it. That being said, know that everything I'm going to cover in this section has an important caveat: Answers may vary. We're going to talk about these stages with respect to certain ages, but understand that answers may vary. No two children will hit the same milestones in the same way, including within a family. I'll use my two nephews as an example; my older nephew is currently eight-years-old and my younger nephew is coming up on six. The two of them developed very differently with respect to language. The older child was running a mile a minute with respect to the language. This kid could talk in complex sentences by the time he was three and a half; when he started kindergarten, verbally he was leaps and bounds ahead of his classmates. Reading for him came a little bit later, but the speaking-wise, he was out the door and running. My younger nephew wasn't the same; in part, it had to do with certain issues with his hearing and with congestion and allergies. As a result, the hearing in his left ear is mildly compromised. Also, he just has a more introverted, shy nature. If he knows you, he will talk to you, quite a bit. But with other folks, he won't talk with them, especially those outside the family. His linguistic development was slower than that of his brother. That's just two personalities, two children.

Answers may vary. Got it?

Let's start off with the **initial stage**, which is pretty much right after birth, the first couple months. We're talking about **babbling**, the first sounds that an infant makes: cooing, crying, laughing, etc. If you've been around infants, you know that set of sounds. If you are a new parent, you're still learning how to differentiate the different cries, that they mean different things; those who are experienced parents or experienced caregivers recognized the differences more readily. In the first few months of life, infants are reacting to stimuli, of all things, not just language. They're reacting in different ways; there is that magic when you see an infant start to smile, especially when they hear the voice of their regular caregivers; they start cooing and laughing when they're around those folks. It is really amazing to hear and watch. What is interesting is that in the first few months, they just react to everything. That makes sense because they're literally brand new; they're still trying to figure things out.

Somewhere around the six-month mark is when things start to change—this is true with respect to all aspects of development, but in this case, we'll focus on language. Those of you who are around young children and infants, you know that right around six months is when the motor skills start coming into effect: they're starting to hold their head up better, if not completely; they're

starting to get this concept of grasping when they want to, not just as a reaction to something hitting their palm. With respect to language, the first aspects of language start to take hold. They're starting to respond to language sounds that they're around regularly. We know this because we we've put infants into rooms with their mother or father or other caregiver, and said caregiver talks in whatever language or languages that are common, and at right around six months the infant is going to start receiving and focusing on those sounds, more so than sounds that are not part of the language at all. We've documented this both with man-made sounds, as far as like machinery, and nature sounds, along with other language sounds. It's really interesting to see this.

I did a minor version of this experiment with my older nephew when he was that young; I would routinely either talk in English to him or Spanish or Italian; my brother and my sister-in-law only speak English and they don't interact much with other languages. Because they saw me work with him and how he would start to react when I spoke in Spanish or Italian, they started exposing him to other languages, both on TV and then in life. This kid, while not multilingual in any sense, constantly wants me to teach him new Spanish and new Italian words. The lingual bug that I put into him at that early stage has stayed with him.

It is also interesting to point out that we even see this in deaf children. This means that it's not just because of what they're hearing, that deaf children right around that six-month point are starting to emit sounds; they are putting things together. How they're doing this, we're still not sure. But we do know that they're trying to make sounds, even if they are in an environment where they cannot hear. The other part of that discussion is CODAs, or Children of Deaf Adults. They do the same thing, even if they are hearing. If one or both of their caregivers is deaf, then primary sign language is going to be used, but the child will also start trying to speak at that six-month mark, and have some of the same patterns that a normal scenario would also include. It's really interesting to see how this plays out.

From here let's talk about the **holophrastic stage**, when we're talking about one word phrasing. That means that one word carries an entire phrase. There is lots of repetition. This stage is right around the one-year mark; if you've been around a one-year-old, you know what I'm talking about. This thing is a bottle, and think of a one-year-old's version of it, probably something close to [baw] or [baba]. Intonation plays a heavy role, so "[bába]!" (😄) is a different phrase than "[bába]?" (🤔) and different than "[bàba]" (😞). Those are three different phrasings for that child based off of what they're trying to communicate; "[bába]!" ( ) could be something like 'I want my bottle!' or 'where is my bottle?!', kind of emphatic. "[bába]?" (🤔) is questioning, and might mean 'where is my bottle?', 'is that my bottle?', something along those lines. "[bàba]" (😞) with the down intonation might imply something like 'I don't have my bottle' or 'I kind of wanted, but I don't know where it is' or 'my bottle is empty'. There's usually a lot of gesturing along with holophrastic phrasing. It's not surprising that around this age is where we get the most use out of baby signs, which are home sign languages. This is around the one-year mark, and it doesn't stop at one year; it's a continuum.

From here, we go to the **two-word stage**; we're talking 18 months to two years, somewhere in there. It could be earlier or later; again, my older nephew was out of the two-word stage definitely by two, but my younger nephew continued on into almost three-years of age; different child, different setup. When we're talking about this two-word stage, **pivot words** are crucial; they're the structure upon which the entire phrase pivots. They have a lot of roles, and frequently involved adpositions (preposition in English). Think of phrases like *allgone*: *allgone shoe; allgone milk, allgone lettuce, allgone outside*. It can mean anything, so *all gone* is the pivot word and then whatever is with it is 'gone' or 'over. Another example is *on*: *blanket on, fix on, bandage on, shoe on*. It may actually work—*shoe on* works, *all gone milk* works—but sometimes it might take a mental leap. For example, *fix on* may not seem like it will work, except what the child is probably saying is, 'can you fix this', 'can you fix this'. Pivot words are crucial, and can range from a few to hundreds. They are frequently saying more than two words, however, the main structure is usually 2-3 words at most.

There was a reason studying about three four years ago, out of the University of Toronto at Mississauga, where Elizabeth Johnson (the main researcher for this project) was looking at 30- to 36-year-olds. They were recognizing adult speech, even when none of their family members were there. This means that they were picking up this concept of by the age of two; they're picking up language from all over. What is really important is that, when they were around other toddlers, if the toddler did not use the same pivot structure, they didn't understand. Let's create an example with 'Jackie' and 'Tracy'. They are with each other in a room. If Jackie uses *allgone* as a pivot, but the Tracy does not, those two children may not always communicate with one another.

Once we get to the two- to three-year range is when we start getting phrases. They are not usually super complex but somewhere in the two- to five-word range. You start to get more syntax and more morphology This is where the creation of different derivations comes into play, where we have correct versions and sometimes incorrect versions. The version I love his *foot* versus *foots* versus *feet* versus *feets*. Every child has some version of that—by the way, not just in English, but this occurs in all languages. The toddlers are going to have more **telegraphic speech,** which is when there is very basic nouns, verbs, and adjectives. They won't

have very much else; maybe prepositions sometimes, certain basic prepositions like *on* or *off*. However, they don't produce embedded clauses or complex derivation. They do produce inflection at a basic level, and it gets better with time.

Speaking of inflection, while English is not a very good example of this, in other Indo-European languages, Niger-Congo languages, and languages around the world, at this telegraphic speech point, the main content words are there. You do get inflection for number—singular versus plural definitely, and if there's a dual sometimes that comes in—and grammatical gender marking, that definitely starts coming into use. Case marking comes later, but it does come. It comes in stages, not all at once. If you have a child, trying to learn Japanese as its first language, that child will produce case marketing eventually, but they will accurately produce plural marking much sooner.

Again, answers may vary. When we talk about these stages, answers may vary; what one child does may not hold true for the other child. These are tendencies, especially the more we understand about autism spectrum disorder and other abilities and issues, we can start diagnosing even in young children. Language production is a part of that diagnosis; we'll come back to that in the next chapter when we talk about the neural processing of language. One of the areas that new parents in particular start freaking out on is whether their child has hit a given benchmark, and language is a really big one, for a number of reasons, and many of them are valid reasons and concerns. I always tell parents who asked about these stages and whether their child is hitting the benchmark, or when they're worried that their child is behind the benchmark or below the benchmark, I keep reminding parents: Answers may vary. Every child is different, and the more you are around young children, and you see their development, you understand that this is at an individual pace. Every child is going to be different. Even within one family, every child is going to be different; just because at two years of age, your child is still only doing one or two words, at a time does not mean that the child has something wrong. Just give it a little more time; have the child be around more people, both of their own age as well as in general. They will catch on.

Answers may vary.

9.3: Child Language Acquisitions Stages is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.4: Child Language Acquisition Linguistics

Child Language Acquisition Linguistics, from Sarah Harmon



Video Script

Let's break down those stages of child language acquisition a little bit more, and specifically the patterns that we observed of how children learn their first language or languages. I keep phrasing it this way, because not all children grow up in a monolingual society; many children grow up in a at least a bilingual society, if not a multilingual society. What we can say is that, regardless of how many linguistic inputs they are receiving with respect to individual languages, we see the following patterns.

Phonology

With phonology, there is a progression on the sounds that children tend to learn first. Whatever the language or languages that they're learning, the progression tends to be the same. When we're talking about manner of articulation, nasal and glides are first, because there's the least amount of control needed to produce those sounds. You don't have to move your articulators so much. Then come stops, then liquids, then fricatives, then affricates.

Let's think about the very young children that you are around. If you think about how they talk at their youngest stages, the first words that they come up with all have lots of nasals, some glides and some stops. If it's a language that's heavy on the fricatives and affricates, those sounds aren't produced very early on. Think of the French-German example that we saw in very first section of this chapter; the child having the worst time with not one, but two different consonant clusters. One of the consonant clusters was an actual affricate, the [pf]; the [kn] is just a consonant cluster. In either case, the child couldn't do it, and simplified everything; he used a basic nasal and a basic stop in their places.

The same is true with place of articulation: labials and velars are the early sounds, and then the sounds in the middle come later, with more precision of the articulators. Palatals are universally the last sounds that the child acquires and articulate well.

I really want to focus on this articulation aspect, that it's the performance that comes late, not the competence. They understand and recognize those sounds, even as early as six months of age. However, producing the sounds that is a different story. The distinction of voiced versus voiceless is an articulation that comes pretty early. Children produce more voiced sounds to start off, and then voiceless sounds come a little bit later.

Think about early words that most children acquire, frequently having to do with mom or dad, because those are the folks that they're mostly around. Those sounds tend to have a lot of nasals and a lot of labials. There early errors in pronunciation, but they're always rule governed. Remember what we said about language change: it's always rule governed. Acquisition is no different.

There is a very strong tendency for children to simplify consonant clusters; we saw this with the French-German example, and you know plenty of other examples. The voicing of final consonants, not just with English, but in general, is also a strong tendency, and

it makes sense. When we are trying to articulate the end of a lexicon, getting that pronunciation at the end is very crucial. The more consonants you have in the coda of the last syllable, the harder it gets to articulate them, and young children tend to chop off the end of the lexicon, and especially the last coda. There's always this consistent voicing of initial consonants and consonant harmony; again, creating patterns and using them in the production of the language.

We see these patterns and tendencies in language after language, in child after child; we see this both in monolingual and multilingual situations. For example, if you have a child who grows up in a bilingual house with bilingual parents, they're going to go through this process with both languages. Remember, competence before performance; we know these children, even at six months, are starting to differentiate the sounds of the language or languages around them. By one year of age, they understand the sounds of whatever languages are around them. They may not be able to produce the sounds correctly, but they understand and recognize them.

Morphology

With respect to morphology, there are some pretty basic changes that we observe. **Overgeneralization**—we've heard that one before. An example is the *foot, foots* or *feet, feets*, along those lines. We see this in pluralization and in all aspects of language; anytime there's any kind of irregularity—and every language has irregularity—then there's going to be some overgeneralization always.

We're still trying to understand why it is the case that children pick up and produce standard pluralization very quickly, whatever language. This is also true of other basic inflection and morpho-syntactic structures. Think back to syntax with respect to subject-verb agreement and the Linear Agreement Rule; that's something that's learned very quickly. The same is true with noun-adjective agreement; if the noun has a certain gender and/or number classification, and it gets inflected as well on the adjective, then that is learned very quickly. The basic derivational rules are also learned very early on; that's why children are able to overgeneralize, because they have learned the basic derivational rules.

One of my favorite activities to do with young children, even as young as two or three, once they start putting those two- to five-word sentences together, I have them tell me stories. I love hearing the combinations that they put together with respect to lexicon; they come up with the best ones.

Syntax, Semantics and Pragmatics

When we talk about syntax, semantics and pragmatics, let's be clear: syntax is kind of on one end of the learning spectrum, and semantics and pragmatics is the other. Basic syntax is learned pretty early on; as the child acquires more is able to produce more. But semantic elements are harder to measure. It's difficult to measure, outside of asking the child, “What does this?”, “What do you call this thing?” We can do that kind of lexical analysis as to how many words the child can produce, as well as how many they can understand. Studies where you have a child in a lab, and the researcher holds this thing (🍼) up using the incorrect name for the item; the babies and toddlers know it's not the correct name. They just may not be able to say ‘bottle’ or [baba]. For example, if I say, “it's zork,” they're going to say, “No, that's not a zork. It’s a bottle.”

Pragmatic elements are even more difficult to measure. That's true just because of the development of a child's brain; these concepts of politeness, of reading between the lines, of implicature and presupposition, the Maxims of the Cooperation Principles, these are things that a child does not acquire until well into childhood. Most commonly, children don’t understand these concepts well until between the years of seven or six, perhaps not until 9-12, and not until teenage years in some cases. Much of this is culturally simulated, as you would expect, and that's why young children don't understand this. The phrase, “Honesty out of the mouths of babes,” or something along those lines, is seemingly ubiquitous because children don't know when not to say something or how to rephrase it.

That all being said, syntax-semantic relationships, like auxiliaries (like *be*, *have*, *will* or would) and modals (like *could*, *should*, *must*, or *do*) are learned not too early, but not too late; usually somewhere in that three- to four-year range. The more stimulus the babies and toddlers are around, the quicker they pick things up, but also the more complex they become earlier on.

That being said, that doesn't mean all of their input has to be super complex for a number of years. It honestly seems to be cyclical; there have been many times, where the prevailing theory is that you have to surround your children with high level language from the earliest of times. That's not necessarily going to ensure that your child understands, let alone produces, those terms. It is important to understand, though, that just because they hear Cookie Monster all the time, that doesn't mean they're going to talk like Cookie Monster.



This is something that Frank Oz, who created the character of Cookie Monster, as well as voiced him for many decades. I love this; I guess this has happened throughout his career, that the parents were concerned about Cookie Monster's vocabulary and syntax. Somebody asked him if he thought that Cookie's way of speaking could corrupt the children that were watching *Sesame Street*. He said that he didn't foresee a child growing up to become a lawyer and saying, "Me want to represent you."

I think Frank is on to something.

9.4: Child Language Acquisition Linguistics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.5: Child Language Acquisition Bilingualism

Child Language Acquisition Bilingualism, from Sarah Harmon



Video Script

I have referenced several times that most children do not grow up in a monolingual society, that they grew up in at least a multilingual society. In many cases, they grow up in a multilingual household where at least two, maybe more, languages are spoken, and that's not even counting the different dialects. I thought it would be important to have a section talking about bilingualism in children.

When we're talking about what is considered to be **native or quasi-native bilingualism**, we're describing a situation where a child is multilingual from birth, that would be native, or within the first 2-3 years for a quasi-native situation. Some of you are native bilinguals; you grew up in households where more than one language was spoken, or in your society more than one language was spoken. Many of you are quasi-native bilinguals, meaning you spoke one language in the house, and then at a certain point—preschool or kindergarten is usually the start—things change, and then you have to speak a different language.

Let's talk about that process, a little bit. Historically, in the last 40 years, there have been two main hypotheses with respect to childhood bilingualism. These hypotheses are meant to describe how the child is constructing these grammars in their head and their mental lexicon. **Unitary System Hypothesis** states that the child initially only constructs one lexicon and one grammar, and then has a word associated with a term from one language or another, but not necessarily both. The evidence for that suggests that young children will only have one term for a given item, so one reference for a given meaning. For example, if a child is a Spanish-English bilingual, and they see this thing (🍾), they're only going to either say 'bottle' or 'botella'; they're not going to say both. There's some evidence for this. The other hypothesis is **Separate Systems Hypothesis**, which states that the child is constructing one lexicon and one grammar for each language that they're trying to acquire. In this hypothesis, children make the connections along the way. While at times you could argue for either of these hypotheses to explain what happens with native bilingual children, most of the time Separate Systems Hypothesis has more data supporting it.

The real example of this has to do with children when they are producing language, and specifically as they mix their languages. I'll use English and Spanish as the examples. If you take a Spanish-English bilingual child, and you listen to them talk, you'll hear them mix, their languages. There's a system to it; they never use the incorrect syntax. They never try to use morphology from one language on lexicon from a different language. You would hear a child say, "I want to go to la playa." (*La playa* in Spanish means 'the beach'.) They do not say, "I want to voy a la playa." They don't try to mix the morphology and syntax from Spanish in English; they will just use lexicon in the appropriate spaces. If they're talking about 'the brown dog', they will say 'the brown dog'; they will not say 'the dog brown', the order of the noun and the adjective in English is the opposite of what it is in Spanish. The child won't mix the order of the of the noun phrase, they will stay in whatever structure is set up for the given language and they do

this early on. Whether it is Unitary or Separate Systems Hypothesis that best describes native childhood bilingualism, that we're still not sure of. We tend to lean more on Separate Systems Hypothesis, but there is evidence for both.

Let's talk about **code-switching**. In many immigrant societies, code-switching is considered horrible and something to be avoided. Prescriptively, when a child or an adult comes into a formal education or formal setting., they're told they can't code-switch; they have to say in one language or the other but they cannot mix. Here's the reality: code-switching is the best thing that a language learner can do, whether we're talking about a child or an adult. This is particularly true for child language acquisition in multilingual situations. Code-switching shows the amount that the child has acquired in both languages.

I'll give you a great example. We have a neighbor growing up that, while the parents spoke English, the parents were teaching their children Mandarin because they were Taiwanese. The mother was also a teacher of Cantonese, so she was fluent in Cantonese as well as Mandarin. She was teaching the children Cantonese and predominantly Mandarin as their first language; English was going to come later and they just accepted that. When their daughter was 2-4 years old, she would come toddling down the hill with her family for visit; she would start talking to us in her language. I say ‘her language’ because it was a code-switch of Mandarin and Cantonese, with an occasional English word thrown in there for fun. Of course, my parents and I didn't understand the child, because we speak neither Mandarin or Cantonese, so the mother would have to translate. It was so cute to see how she was putting things together, and how the mother, in particular, would have this question mark on her face, when she could not put the pattern together initially.

Children will subconsciously code-switch; they're doing as part of spitting out their thoughts. There's no regard for audience; they haven't learned the contextual clues about not speaking a language in front of somebody if they don't understand. As a result, they just blurt stuff out. There are combinations, but those combinations tell us how much they have learned about the respective languages. Adults are different; adults will consciously code-switch, meaning that they do it with intention. They will only preferably do this with other bilingual because they know the stigma of speaking a language that the other person you're talking to does not understand. it's really interesting to see how code-switching develops in a bilingual child; around the age of eight is when they start understanding when and why to code-switch. If you think about a child's development, that is about that time they're starting to learn social mores and norms. Politeness starts really factoring in around that time, so they're starting to put a lot of pieces together.

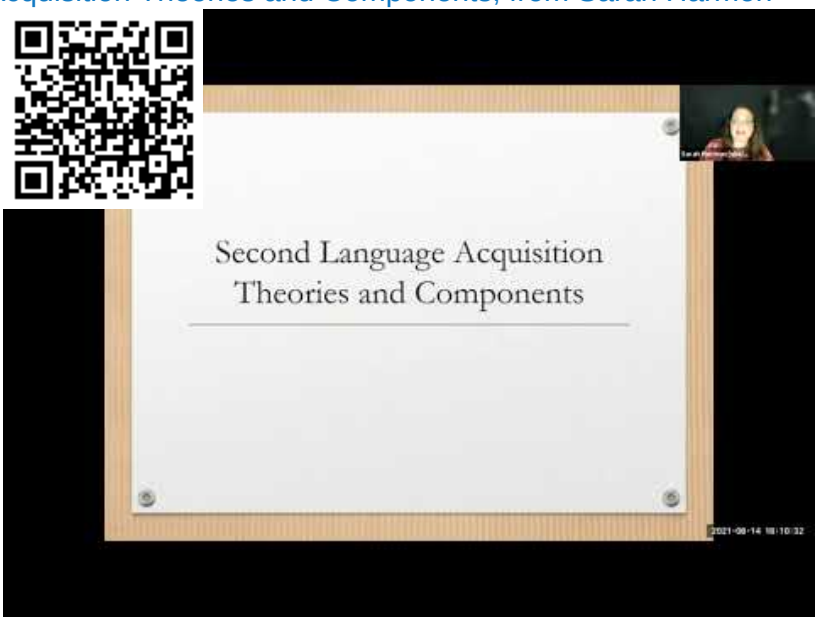
The other important aspect to code-switching is consistent input in both languages. If a child has consistent input in both or all of the languages that they speak, they will continue to code-switch and become more proficient at it. But if the input of one of the languages stops, the code-switching also stops, and they start become more monolingual. It can come back, if they go to learn their heritage language. This gets back to contextual knowledge, and this is where semantics and pragmatics come in; a young child isn't going to have that knowledge, because they haven't developed that mentality, nor have they developed neurologically to that point. As they age, things improve. This is all led to the **Contextual Knowledge Hypothesis**, which is really becoming more plausible when we're talking about bilingualism both in children and adults. It states that as a child reaches that age of about eight, they start intentionally code-switching. It explains a bit about how adults switch between languages—because they're using the contextual knowledge, the environmental cues to know when they can switch. It also explains how listeners go between languages; if you are an English-Spanish bilingual, and I say [si], depending on the context and which language I’m using, that will lead you to understand that syllable, that lexicon differently. If I’m speaking English, and/or if I’m referring to the ocean or the sea, then if I say [si], you're going to think of body of water. If I'm not, and especially if I’m speaking Spanish, then you're going to think of it as ‘yes’, because that is how you say ‘yes’ in Spanish. (It's also how you say ‘if’, but that's another story.) You're using the contextual clues to understand when someone is code-switching; this is both passive and active in its processing, both in receiving/listening to language and when you're producing language. We use **parsing** to help us with this; we’ll come back to parsing in neurolinguistics in the next chapter.

It's really important to understand that this is part of what a child acquires if they're in a multilingual environment. It's not just learning the actual lexicon and the morphology and the syntax and everything else. It also ties into the context clues.

9.5: Child Language Acquisition Bilingualism is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.6: Second Language Acquisitions Theories and Components

Second Language Acquisition Theories and Components, from Sarah Harmon



Video Script

Every single one of you has attempted to learn a second language; in some cases, you may have been successful, while in other cases, not so much so. As we go through these next few sections about second language acquisition, keep in mind your own experiences. I'm sure some of them, if not all of them, will be reflected in what we talk about.

To start things off, this was a hilarious meme that came through Facebook years ago.



Having learned or attempted to learn a couple of these languages, I can relate. For a bit more context:

- Latin: The fact that Latin doesn't have specific lexicon for 'yes' and 'no', but there seems to be 3,000 different ways to talk about killing people.
- Ancient Greek: There's so many different derivations of verbs that you never knew existed until you look at Ancient Greek.
- Ancient Egyptian: The hieroglyphs are the connection.
- French: Certainly, if you tried to learn or if you have learned French, it seems that half of what you see written doesn't get pronounced
- German: If you go back to morphology, we looked at German compounding; look up the word for 'research' in German, and that should tell you some things.
- Mandarin: Remember, it's an isolating language.
- Spanish: Remember, it's a fusional language and a synthetic language.
- English: Okay, maybe not maybe not hell, maybe not the easiest language to learn.

There are certain impressions about learning a different language., that some languages seem easier and others are harder. None of that's really true; they're all difficult, and don't think otherwise. Everybody has their path in learning a language.

Let's talk about some of these theories of language acquisition. With respect to adults, mind you when I say 'adults', I'm saying anybody who has hit puberty and beyond. That has to do with the **Critical Age Hypothesis**; this is a hypothesis that is up for debate in some aspects, but not in others. It does seem to be the case that somewhere around puberty is when our brains have stopped absorbing like sponges. The elasticity and the neural plasticity are not the same when we are teenagers and adults as it is when we are children. Critical Age Hypothesis states that our brains don't seem to be able to pick up various skills as easily, including and especially language. It must be taught via logic, and if you think about a language course, that is really what you're learning: you're learning the logic of the language. Critical Age Hypothesis has its detractors; there are some that suggests that there is no such thing as waning elasticity of the neurons, that any person can learn any skill at any time, and there's no change or difference. That does not seem to be the case at least with respect to this concept of the **Fundamental Difference Hypothesis**, which states that learning a language as an adult is not at all like that of learning a language as a child. Regardless of the detractors, there is a large amount of data that show that with any major skill, children learn these skills as if through osmosis; they just watch,

they practice, and they do, and it seems relatively effortless. The reality may not be the same, but that seems to be the case. As adults, the only time that is true is that if we can build upon a skill we already have; take language.

For example, I can learn a Romance language pretty easily, because I already have acquired and mastered two; based off of those two, I can use those skills to use to learn another Romance language. I will tell you that Germanic languages for me are very difficult; I'm terrible at learning Germanic languages. I just don't get the logic, despite the fact that I natively speak a Germanic language in English. It's because growing up, I didn't hear much in the way of Germanic languages; there was a Swedish family down the street, but they spoke English, and there was a Norwegian woman that one of my cousins married, but she spoke English, so realistically I didn't have much. I've tried to learn Japanese, and there's aspects of it that I get, but I heard Japanese as a kid by a few neighbors. Even though tonal languages seem to elude me as far as getting the tones down, I do see some of the patterns in them, but growing up, I had a number of friends who whose families were from Hong Kong, Shanghai, Beijing, or Taiwan. These are various places that speak different Sino-Tibetan languages, so for me, I start to see those patterns easier because I've been around them and have been exposed to them.

Fundamental Difference Hypothesis underlies mostly language learning, but, as I said, can be extended to other skill sets, especially motor skills. There's this concept with childhood development that when you expose a young child to as many inputs as possible—without overloading but certainly a number of them—and give them consistent access to those inputs, they pick up those talents and skills readily. That it sets up an affinity for those similar skills in their later lives. It's why we try to expose children to music, both listening and playing music from an early age, because if children are exposed to those skills early, they will continue to at least attempt to use them.

These two hypotheses are just that: they're hypotheses. We cannot officially make them into laws or universals, because there are exceptions always. In order to test them properly, we would need some very unethical training and activities, and we're not going to do that. We're not going to deprive a child linguistic input to see what can happen. We're not going to force them to do all sorts of activities that they may not be able to do physically. We're not going to do those things. We do have a goodly amount of evidence that shows that Critical Age Hypothesis does not say that you cannot learn a language once you hit puberty; it says that it must be taught as language acquisition based on logic. It can't be through osmosis.

There are certain common errors and processes that all adult language learners do. Whether they stay in this error-laden stage, or they get past these errors, this depends on a number of things. **Fossilization** is the big one; this is the situation of a language learner who keeps producing incorrect language in certain areas, but it's so ingrained in their mind that they just think it's the right way to do it. A couple of definitions: L2 and L1. 'L' refers to 'language; 'L2' is the second language, while 'L1' is the first or native language. An example of a fossilization error: I teach Spanish, and one of the huge ones that I still get from intermediate students and beyond is, "Me llamo es." Spanish speakers, I apologize; I know how cringe worthy that is. "Me llamo es" is wrong in every way. "Me llamo" is 'I call myself'; it's what you say, when you say, "Oh, my name is Sarah," you say "Me llamo Sarah;" I call myself Sarah. "Es" is a verb; it means 'is'. The problem is "llamo" is also a verb; it means 'I call', so "Me llamo es" is never going to work because you have two conjugated verbs that don't go together. It is a big error.

Interlanguage and interference are connected. **Interlanguage** is what we do as we're learning a language, when we combine elements to try and succeed in producing more of the target language. Think of telegraphic speech in child language acquisition, and this is kind of the adult version. Interlanguage is described as having more basic sentences that eventually lead to more complex sentences and phrases. We use interlanguage as we try to put our pieces together. Sometimes as we do that, we end up having **interference** or sometimes it's called **transfer grammar**. When your L1 interferes with your L2, for example as a native English speaker if I go to learn Russian, and I keep trying to use English phrasing, words, and pronunciation, instead of staying in Russian and learning the Russian, that's interference or transfer grammar. This is really common of all folks who learn a second language; we also see them children, but mostly we see this in adults.

Any language learner can get through all of this. In part, this is done through good instruction, which involves aspects that we're going to talk about soon: encouraging and open environments, like allowing students to feel like they can take a risk. Part of this is also internal motivation; we'll talk about learner affect soon enough. But there are always going to be little bits that keep you from being a native speaker; we can be a near-native speaker, and I'll explain more that in a minute, but some of those near-native elements are directly related to interference and the related to fossilization. But you can get around it.

I would be remiss to not talk about heritage learners. A **heritage learner** is somebody who is trying to learn the language of their heritage or ancestors; frequently it's the language of their parents or grandparents. Think of a case where you have a family that integrates to a new country, a new linguistic community. Their children and grandchildren may learn that new language very quickly, but then may drop the first language, the heritage language, along the way. Many folks, including maybe some of you, try

to learn that heritage language in puberty or adulthood. This is a different process; it's not quite the same as learning from scratch. There are various levels of heritage learners. A low-level heritage learner knows the sounds and some words, but isn't really able to construct the basic sentence in that language. A mid-level or high-level heritage speaker can produce much more. There's always conflicting or competing issues going on with heritage learners, and they have to be addressed in the language and learning environment. There is frequently a desire to 'speak properly'; I can't count the number of heritage Spanish speakers who sign up for first semester Spanish because they want to learn the language 'properly', thinking that what they know is not proper or accurate and somehow substandard. Frequently, that's not the case; in the vast majority of cases, that person is marginally, if not fully, fluent in that language. They just may need to boost their literacy, or learn more complex structures so that they can speak at a higher level, but they are fluent, at least at a basic level. There's also **language attrition**, which we talked a little bit about in child bilingualism. In many cases, once the heritage learner started their formal education in preschool or kindergarten, that is when they had to stop speaking their heritage language; they were forced into speaking the new language. As a result of societal pressures and everything else, they just 'lose' the first language; it's attrition. Just like a muscle that you stopped using and it atrophies, the same thing can be said for a 'home language' or heritage language.

Finally, there's this thing that I like to call 'heritage speaker guilt'. As far as I know, this is a Sarah-ism; I don't know if this is an official term, but it's a specific kind of learner affect. When we are learning a new skill, including a new language, we suffer from some element of **learner affect**. That is the emotion that bubbles up inside of us and says that we can't do it, or we're scared to do it, or we're not confident in our abilities and we don't want to be made to feel the fool. When we're learning a language that does not have our heritage, you have learner affect in varying forms. But when it's part of your heritage, it's present at even higher levels, and it's a very specific one. There are a few flavors of it; there's the 'why aren't I doing this correctly?' or 'why can't I get this right, why aren't I speaking properly?' flavor. There's also the 'why aren't I remembering this?', 'this was something I used to speak or hear when I was a child, why the heck can't I do it as an adult?' flavor. (Frequently stronger language is used. 😉) Learner affect as a whole is an emotional block, and it takes some psychology to go past it, but it doubles down when it is your heritage language. I didn't have this problem as much when I learned Italian, because it was my great-grandparents that immigrated here and spoke it. But my mother was a different story; my mother, to this day, cannot learn Italian. It's not to say she can't really learn it; she actually speaks Spanish, which she has not learned since high school many moons ago. If she goes to Mexico, she can give basic instructions to the taxi driver. That's Spanish, though; that's not Italian. She's got such a block on Italian. She has tried multiple times; first, she wanted me to teach her, which was a bad idea. (We have an amazing relationship, but learning from your child is a little bit different.) I suggested that she take classes at the community college or that she goes to the Italian Cultural Center in San Francisco. She tried a number of ways, and she's never been able to get over that heritage speaker guilt. In her mind, this should be instantaneous, and it's not.

How do you include heritage speakers? The answer is: whenever possible, especially the higher-level heritage speakers. We try to separate them out into their own class, and that makes sense; they're going to learn the language in a way that's different from a complete beginner. But the big one is support, even if you have a separate class, but especially if they're mixed in with the general population, as it were. You have to give them that emotional support, so they feel like they can take that risk, but also that they can feel pride in what they already know. You include their experiences, especially when discussing elements of culture. One of my favorite lessons when we do intermediate Spanish is when we talk about Spanish in the United States. I love this section, because we start picking apart this concept of what it means to be bicultural. There are so many folks in the class that are bicultural with different cultures, that it brings this really rich conversation. You start seeing the pride well up in them; that starts breaking down the learner affect, the heritage speaker guilt. Alongside that is to expand their knowledge. There are many people who say, "well, I don't need to learn the language; I'm already fluent in it, I already know everything about it." Learning a language, especially a language that has multiple cultures and dialects—English, Spanish, and Mandarin are just three that come to mind—when you start expanding their horizons and cultural knowledge, they start feeling more of that pride. I always grew up with this pride of being an Italian-American. Even though I could count to 10 in our dialect and I knew words for certain food items, and certain not-so-nice things to say to people 😜. But when I was a senior in high school, I started taking Italian at a local community college. It expanded my horizons. Suddenly, I was learning about all sorts of things with respect to Italian culture, and not just my little neck of northwestern Italy, where my family is from. I suddenly was starting to learn about all these different concepts and different ways to phrase things. To this day, I still can't say *focaccia*, the bread, without visualizing the term; it still is always going to come out for me as *fogassa*, which is what we say in Lombardy. But I do know there's a difference, and I know that difference, and I express that difference. That all comes from learning about my heritage language and the cultures and the dialects that are associated with it.

9.6: Second Language Acquisitions Theories and Components is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.7: Second Language Acquisition Teaching Methods

Second Language Acquisition Teaching Methods, from Sarah Harmon



Video Script

I've been fortunate to teach foreign language for over 20 years, and it's been an interesting ride in those 20 years plus. The additional 10 or so beforehand, when I was learning a foreign language, these teaching methods changed just in my own lifetime, very drastically. While most of what you're going to see on this next slide is not in use anymore (thankfully), it certainly represents what most people had to suffer with respect to foreign language learning. I do use the term *suffer* because when we're talking about **grammar translation** and **audio call and response**, these are techniques that have been used for millennia, and with respect to language learning and they don't work. For a very small portion of the population, they actually do help, but for the most part we have left these methods behind. Grammar translation is when you have a term in language one, and then you have the term in language two, and you are supposed to repeat. It's usually a call-and-response type of scenario; I say 'this is not a hat'; 'este no es un sombrero'. You, the student, are supposed to say 'este no es un sombrero'. The lesson would go from there. It doesn't work because it does not actively engage the synapses so that you actually remember. What you're trying to say and what everything means. Remember that first rule of language: everything is arbitrary. If you do not see the pattern, you're not going to see how and why you should remember something, and it just won't stick.

For the last 40 years, we have moved away from that grammar translation, call-and-response mentality, and have been in the content-based instruction mode. There are two main flavors of this. The first is **immersion**, when you immerse yourself in the language 100% of the time. The usual case is that the person moves to a location that speaks that language, and then lives within the population and learns—not through osmosis, might I add—but the input. They're combining what they're hearing around them with classes. It must be said, immersion is all about motivation. For example, if you are forced to move to a new area that speaks a completely different language, you may not have a lot of motivation to learn that new language, you may feel a bunch of resentments, and therefore try to not learn that language. Or, you may be highly motivated because you need a job, so that you can earn money, and therefore be able to afford things. Everybody's motivation is going to be different, and an immersion scenario usually accelerates the motivation in one path or the other.

What we tend to do in a classroom is much more communicative; let me explain what that is. When we talk about the **communicative approach**, mostly we're talking about the work of Stephen Krashen, Bill Van Patten, and a number of others, starting in the late 70s and through to today. Recently, Bill van Patten published a new book about three years ago. The communicative approach is just that: acknowledging that language is a tool, and that we use it to communicate with others. There have been multiple iterations of the communicative approach over the last 40 years, and it is all based off of information that we know to be true, based off of observation and, in the last five years, what we've been able to show as far as how the brain processes language. (More on that in the next chapter.) Suffice to say that so many of the theories that we've come up with over the last 40

years as to how we learn language, how we process language, and how to improve the language learning environment, we're starting to see evidence of that with respect to the neural processing. It is a fascinating time right now to start understanding how language and the brain work.

For now, let's focus on the communicative approach. It's an active learning environment, so you really focus on getting the students to practice the language in multiple ways. The five Cs are the core: communication, culture, comparisons, connections, and communities. When you learn a language in a communicative approach, you are not simply learning the grammar and the vocab; you're learning the culture behind it, or cultures in the case of multicultural languages. You're making comparisons between your first language and this one, and perhaps any others you've learned along the way. You're making connections in a variety of ways: neural connections, the literal connections with respect to the language learning, but also connections with other peoples and then communities, which is you're communicating with other communities, not just locally, but around the world. It is important to remember that, with this communicative approach, we involve technology; this communicative approach starts off just as we were starting to have personal computers, and it really takes off in the 1990s, when we start to have Internet in all homes. Not surprisingly, with the Internet, you can then communicate with the world, literally. Harnessing that power of communication in the language classroom is incredibly important; it can be a hindrance, or it can be a help.

Personally, in my practice, I have always included communications with technology, even at its earliest stages. I was using the Internet to learn languages myself through online bulletin boards or chat. By talking and typing in the languages, I was learning to facilitate; at one point in my life, I could chat English, Spanish and Italian pretty much at the same time. (Those days are long gone, by the way; now, it's two at a time, at most.) That being said, if you encourage students to use the technology to push them forward, as a way to apply that knowledge—to read about something in that new language they're learning and then report on it, to talk to somebody whether it is live or it is email, to investigate a little bit more about that culture and then report back on it—these are ways that technology can help.

But technology in the language learning classroom can be a hindrance. The worst thing that a person can do is try to use some kind of app or website to translate everything. For one, it's not going to work and it's not going to work for some time yet; processors just do not function at that speed with all the permutations possible to give the correct translation. Most trained multilinguals can spot it every single time. The other reason that it does not work is because it is passive. One of the things I tell my students, regardless of the level of Spanish they're in, is I require them to be active in their learning. If you are looking up every single word, that's passive; you're not going to remember any of the vocabulary that you looked up. It's the same as the grammar translation; that never works. If you use that term that you look up, by using it in your speaking in your writing, and then later you hear it, and process it—that is what works. This active learning process gets ingrained into your mental lexicon for that language. That is how technology can be used as a tool.

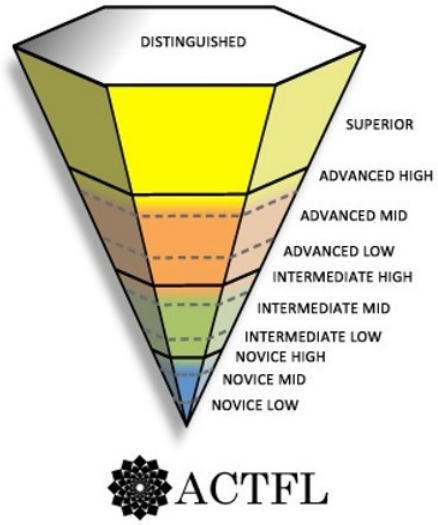
(And no, Apple Translate and Google Translate should not exist; they need to be abolished. 😝)

Flipped classrooms are a part of this concept, where you put the onus of the learning on the student. While the foreign languages have been doing this for some time, we got the idea from a lot of vocational or career technical programs. This scenario is where the student needs to read up and learn about certain processes in that course ahead of the class meeting, and then they come to the lab or they come to the classroom and apply it. If you're in a culinary program, you read about different culinary techniques, and then you go into the kitchens on campus and you apply them. If you're in an automotive program, you learn about a specific system, the transmission or the exhaust system or what have you, and then you go into the mechanic shop and you practice it. Foreign language classrooms have been moving to a flip model for some time; I've been teaching in a flipped model for about 15 years. It really does put the onus of the learning on the student; the student has to take that responsibility. But that's also how you foster two aspects that are really important in combating learner affect: If you put the responsibility for learning on the students, they find the motivation, they want to come to class prepared, or at least with the questions that they need answered. You also break down that learner affect because the student will take that skill that they read about at home. In the classroom, we're going to apply it; I'm going to have you read something and then I want you to tell me about it. Or, I'm going to have you write something to a person or colleague, or to a random person that doesn't exist. I'm going to have you tell a story or write a story; I'm going to have you give your version of Little Red Riding Hood or Sleeping Beauty or whatever folk tale you wish to use. When you use a flipped classroom with active learning, at its core language learning isn't so terrible. When you chunk the learning—meaning you break it down to little pieces and present those chunks in ways that students can acquire on their own—and then apply it in the classroom, suddenly the process isn't quite so onerous.

The last real thing that we're going to talk about with respect to language learning and teaching methodologies has to do with proficiency versus fluency. It is interesting when somebody tells me that they are fluent in a language, and especially if it's a

language I happen to share, and then I will see what that fluency really is about. The popular concept of fluency is something along the lines of, “I can make a basic conversation. I can maybe use some past tense, but I mostly stay in the present tense, because what else do I really need. I can go travel to a place, I can get food or water, do some kind of activity. In my eyes, I'm fluent in the language.”

The problem is when it comes to teaching methods, that's not fluency at all. What you see over here is the ACTFL pyramid; ACTFL is the American Council for the Teaching of Foreign Language. They're one of the preeminent academies, if you will, or associations with respect to language learning.



This inverted pyramid describes what you can do when you first start out, all the way up to the most distinguished and elevated speaker of that language. It doesn't matter what language we're talking about. When you start off, you're at a Novice level; you can make a list of things that you like or dislike, or attributes of something. You're only speaking in the present tense, and your vocabulary is very limited. When you're an Intermediate level, you're starting to branch into a couple of different tenses—not drastic, but a little bit more. You have a more robust vocabulary, but there's still plenty of holes. This is where most people's concept of being fluent is because, in truth, they're not really that fluid. They're really only speaking in the present tense and that's about it. For a linguist or a foreign language professional, in order to be fluent, you need to be in the Advanced level. It means that you modulate your verb structures, so you're using tense, aspect, and mood, those aspects we talked about deixis, semantics and pragmatics. If you recall, many languages have two, if not three tenses. Frequently, there's multiple aspects and there's different moods. If you're an advanced speaker, you're able to use those pieces, and you're able to convey so much more. A novice speaker really is only able to talk about themselves, and maybe immediate family. Intermediate speakers can talk about their immediate community, like their neighborhood. Advanced speakers can talk about their larger neighborhood. Superior is when you are fully fluent; you are a native or near-native, fully fluent speaker. You're able to talk hypotheticals, you're able to talk about potentials, and you're able to talk about massive complex sentence structures that usually involve a goodly amount of discussion.

What is really important with respect to this, is that language is a tool. We use the tool as a skill; I've used that metaphor throughout the course. This pyramid shows us that we always have room to grow.

The very top level is Distinguished, and there are very few Distinguished speakers of any language. For example, I’m a native speaker of English, I have a PhD in Linguistics—yet I am not a Distinguished speaker of English; I am Superior. I can do all of these things: talk all of these complex causal structures, and hypotheticals, and everything that you want. I can write an academic language. But to be a Distinguished person, that would be the highest level of legal language that you can think of. I do not speak legal-ese, as it were, so I’m not there. Most people who have a full formal education, usually through the concept of high school and maybe even a little bit of college, you're going to be in the Superior, or at least Advanced-High level. It's not to say that if you do not have formal education, you can't be up here in the Superior level; it just means you have to get your education from elsewhere. You can be a Superior level speaker in your native language and not have formal education. If you're consistently using hypotheticals, thinking about more than just your immediate society, talking about national or global society—when you start expanding with your language, that is when you come into those higher levels.

That's an important thing to bring up for everyone with respect to language: Push yourself, even in your native language or languages. Keep pushing yourself; keep reading; keep talking; keep expanding what you can do with that tool. You don't just learn to boil water in a cooking class; you learn how to cook, how to make a sandwich, or make a soup. You make some pretty important things; language is no different.

9.7: Second Language Acquisition Teaching Methods is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.8: References

Resources and References for the Chapter

On top of the usual references of Rodman, Fromkin, and Hyams (2019) and *Language Files* (2019), the following resources have been used for this section, and should be consulted:

- Canagarajah, A. Suresh. 2012. "Teacher Development in a Global Profession: An Autoethnography." In *TESOL Quarterly*, 46 (2): 258-279.
- Cenoz, Jasone, and Durk Gorter. 2011. "A Holistic Approach to Multilingual Education: Introduction." In *The Modern Language Journal*, 95 (3): 339-343.
- Dumitrescu, Domnita. 2014. "English-Spanish Code-switching in Literary Texts: Is It Still Spanglish as We Know It?" In *Hispania*, 97 (3): 357-359.
- Huguet, Ángel. 2006. "Attitudes and Motivation Versus Language Achievement in Cross-lingustic Settings: What Is Cause and What Effect?" In *Journal of Multilingual and Multicultural Development*, 27 (5): 413-429.
- Pham, Giang and Timothy Tipton. 2018. "Internal and External Factors that Support Chldren's Minority First Language and English." In *Language, Speech, and hearing Services in Schools*, 49:595-606.

9.8: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

10: Language Processing and the Human Brain- Cognitive Linguistics

Topic hierarchy

10.1: Human Language versus Animal Communication
10.2: Evolution of Human Language
10.3: Language Processing and Parsing
10.4: Language Processing in the Brain
10.5: Aphasia
10.6: Autonomy of Language
10.7: References

10: Language Processing and the Human Brain- Cognitive Linguistics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.1: Human Language versus Animal Communication

10.1.1: Human Language versus Animal Communication, from Sarah Harmon



10.1.1.1: Video Script

part of it. I know a number of you have felt it was difficult throughout this whole process; I'm not going to argue. But when we talk about the actual neural components to language—how the brain processes language—this is where things get really technical.

For this section on animal vs human communication, this is evidence that we have up to this point. It's still evolving and we're still learning, but I can be more confident with what we're saying and what we're presenting here. There is so much that we're still learning about this thing between our ears and how it processes everything, including language, that this is constantly evolving. Just for a frame of reference, I am recording this mid-August of 2021; by the time you watch this, there could be some significant advances, and I won't know it until we get there. Just know that everything I'm presenting is with a grain of salt, in the sense of I am presenting the latest that we have up to this point. There is much more research to come, there is also more research underway and, additionally, our understanding of everything is changing. I'm confident and what I'm going to present for this chapter, but understand that things are changing.

Let's start off with a discussion that we had back in Chapter 1 when we talked about animal communication and human language. Let's refresh our memories a bit; let's go back to those **attributes or hallmarks of human language**. We understand that human beings have something that is potentially unique, if not very rare, that this form of communication—we talk with arbitrary signs and signals, we talk about things in front of us and not in front of us., we transmit aspects of our culture of our lives, we make an infinite use of finite means, we only have a certain number of vocabulary or lexicon, we only have a certain number of ways that we can put them into a phrase, and yet we can say anything that comes to mind, that we are productive and creative and constantly changing and adapting with our language, we can talk about all these things and ourselves, we can create, we can express and we can talk about things that are not just here in now, not just immediate needs, we go well beyond that. Just think of this class alone!

There are some similar aspects that we see in other forms of animal communication, but limitedly. Let me explain a little bit. There's no question that all animals have some form of communication to express their needs and basic desires; that's not in-argument. The question is: Are they be able to talk about things, not in front of themselves? Are they able to talk about hypotheticals, or make suggestions? Are they able to talk about things that happened in the past, or what may happen in the future?

Maybe. This is why I'm going to say maybe. We know, for example, that most species of birds are able to describe where food or mates are when they're not in front of them—as if to say, "Oh yeah, that flower, that is about five miles that way. That's a really good place to go get nectar." We know that birds and bees tend to do this, and we see other mammals, especially other primates, have aspects to this in their forms of communication. However, we're still not sure what they actually do among their species.

We are not talking about mimicking human language, and that's a really crucial piece of this argument. We are not talking about when we try to teach a parrot how to talk, or when we try and force another primate to learn a primary sign language. That is not how they communicate with each other, and so we have to abolish that concept entirely.

Because we understand that what a different species used to communicate with its peers is going to be different than what humans do, the research that I'm referring to requires analyzing, observing and being descriptive about what species do amongst their own kind, and to a lesser extent to other animals in the region.

When we talk about animal communication, I love this old *The Far Side* comic—I'm sorry, I'm a Gen X and *The Far Side* was part of my upbringing.



It encapsulates everything that we think of about animal communication versus human language, as far as what we say to them versus what they hear and understand. I absolutely love and adore *The Far Side*, especially this comic but I'll give you an example in real life. I have a cat her name is Bella and she is 16. When I think about my cat, and I've had her since she was a kitten since she was three months old, there's a ton of things that we communicate to each other through voice and through body language. She tells me when she needs attention and love, and when she thinks I need attention and love. She tells me very clearly when she has no food in her bowl or its old food and it's not acceptable anymore. She plays well, not so much anymore, but certainly when she was younger, and she definitely communicates that when I go away, she doesn't like that, and when I come home, she makes that very well known. I communicate to her when she does it behavior that I do not approve of, like if she were to scratch the furniture —which she's never done save for once, and it was to get my attention because I forgot to feed her, so she knows how to get my attention. We have a form of communication. When I am at a low point, she's one of those folks I confide in; I cry and she's there. She cuddles me, and I cuddle her; I tell her all my hopes, fears, desires and wishes. and she purrs.

Now, does she understand what I'm saying? Or, is she like *The Far Side* comic where she just hears noise and she doesn't know what it is? I don't speak cat and she doesn't speak human, so I don't know what really is able to be communicated as far as displacement, as far as productivity or creativity. We certainly have arbitrary sounds and meanings for those sounds. It has often

been said, the cats probably learn to meow because of humans their interactions with humans. When they meow in certain ways, humans do certain activities, and it's a symbiotic relationship. I think there's some of that that's true. But she's not able to tell me her deepest hopes, wishes and desires; she's not able to tell me what she thinks might happen in the future, or what did happen in the past. I don't know what she thinks really, although as I'm saying this, she's walking around my feet right now, because she's clearly telling me she doesn't want me talking like this, she doesn't want the camera and she doesn't want the lights. She wants me on the bed right now; she's able to communicate her needs and basic desires. But not much more. Is that to say that she can do that with a different cat? Who's to say?

Where we have been starting to observe a few pieces with respect to animal communication and whether they might have a language, primarily, has to do with our primate cousins. We know certain things to be true. First of all, their vocal tracks are not like human vocal tracks; they are well more primitive, to the point that they cannot produce the sounds that we can produce. We know that part that goes out the window. Yes, it is true that, certainly for other great apes like chimpanzees and gorillas, some have been taught American Sign Language, in particular, and a few other primary sign languages. But—and this is a huge 'but'—their learning is very slow and formulaic, and they basically get stuck at the level of a three-year-old. If you remember telegraphic speech from child language acquisition in the previous chapter, they're not able to do much more than that, at least not in ASL. They're also not able to create with ASL very well at all. Therefore, I would argue that you can throw out using any kind of human language with a primate; it's probably not going to work.

All that being said, there is quite a bit of evidence to suggest that they *might* have something primitive. I'm saying the term 'primitive', but I do not want you to think that this is a prescriptive use of it. It's saying that this is a very early stage, and maybe in a millennium or several they might have the capability to use a language, much like a human language. At this stage, we don't know. What do we know is that chimpanzees and other great apes are able to teach each other tools. Chimpanzees are particularly good at this, but even we see this in gorillas and some other great apes. We also know that other primates use sounds to communicate things beyond basic needs and desires, not just a warning system, not just to say, "Hey, I need food" or "Hey, I need sex." You do observe them using the sounds in more arbitrary ways. But—and this is an incredibly important point—we are still trying to decipher what those calls and sounds mean. When we observe our primate cousins teaching each other how to use tools, they are not necessarily using a vocal communication to do it. There is some kind of gesturing. I don't really want to call a sign language yet, because I think it's too early to say that, but our colleagues and primatologist are showing us that our primate cousins, especially the great apes are able to use some kind of communication that's at a higher level than what most other animals do.

Primates are an interesting discussion. What is also interesting, and this is in the video below is Zipf's Law, and the video is going to go a little more into that. Here's the interesting thing: It could be that dolphins in particular might have a language. You may have heard of studies on dolphin communication before, and this is an area that continuously evolves. Suffice it to say we are very much at the precipice of understanding what other animals do when they need to talk to each other, when they need to communicate to one another, beyond their basic needs, hopes and desires. We are still learning so much about what our own brains do, let alone what other brains of other animals do. So, we'll come back to this—maybe not in this class, and maybe not in the next year, but certainly in the future, so keep an eye on this.

10.1.1.2: 10.1.2 More on Zipf's Law, from NOVA Wonders (PBS)

To finish things off, watch the video below about Zipf's Law, and why we still have so much more to learn about other animals and their methods of communication. (The video is captioned.)



10.1: Human Language versus Animal Communication is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.2: Evolution of Human Language

Evolution of Human Language, from Sarah Harmon



Video Script

When it comes to the evolution of human language, we still have quite a few missing pieces, but there is more that we can say with certainty at least up to this point. As far as when the first human language came into existence, let me dispel something really quickly: we don't know. We know that *homo sapiens sapiens*, our current version of the species, seems to have a language capacity from the start> For reasons that I'll come to in a minute, we are starting to understand that maybe other hominids also had if not as robust language as we have now, certainly something pretty close. There's so much more we have to know; there's so much that we probably will never know. But let's start off with some basics.

Based off of data—not just linguistic data, but biology, paleontology, anthropology and archaeology, all of the studies of the human existence, we know the following:

- The formation of our vocal tract as well as our ear is unique. Even amongst primates, it is unique. That may have something to do with this concept; it could be that we evolved to have these apparatuses so that we could communicate via language. As I said, this is something unique amongst even the primates.
- We also know that there are a couple of areas of the brain, in particular, that in humans seem to be more developed than any other species. Specifically, that's **Broca's Area** and **Wernicke's Area**; in the next couple of sections, you'll learn more about them.
- We also know that humans have a very specific genetic mutation, which I'm going to talk about that in a minute.

All of this aside, we still have folks who are deaf from birth, yet they seem to be able to communicate with a language, if not more than one language just fine—so maybe the vocal tract and ear may not be exactly connected to this linguistic faculty; it may, but we're not sure.

As far as theories about how languages evolved in the first place, a couple of them have come about that should be debunked, but others have a little bit more merit. The first one is that our linguistic faculty comes from **gestural origins**. I like thinking of the caveman metaphor, as it were, that early hominids or early humans gestured and grunted a lot, and then they decided to become more fluent and more expressive. This pretty much has been thrown out; we are pretty confident in saying that that probably is not the case. While there certainly was an evolution of language, and the first language that was spoken was not as robust as modern English or Sanskrit or Old Chinese or Sumerian, but evolved to that point. The other hypothesis comes out of primatology. The **grooming hypothesis** is the concept that if you watch other primates, in particular, and some other mammals, we groom each other. Let's face it, folks; what happens when we go get our hair done? What happens when we get our hair cut in a hair salon, whether we go to the same person over a number of years or it's a brand-new person? We talk a lot. Therefore, the thought is perhaps that's

part of the grooming experience, that you talk and exchange information. You get to know each other a little bit better, if you don't already, and life goes on. However, linguists by and large don't tend to subscribe to this hypothesis. It's circumspect; there are plenty of people who don't say anything when they are being groomed. There's also plenty of times when we communicate and it has nothing to do with grooming.

Genetics might hold a key, and this is tied to **Universal Grammar**, the fact that we have this template or this faculty for language. That was the hypothesis of Noam Chomsky back in the late 50s-early 60s, and we more or less continue with a version of that to this day. Interesting data have come about and, specifically, it has to do with the specific gene called the FoxP2 gene; if you take an anthropology course, you may hear about it. The FoxP2 gene in humans has a very specific mutation that you do not see in any other primates or any other mammal or any other animal. There are people who have a different mutation of the FoxP2 gene, and they do not have as strong or robust a faculty for language; they are not as expressive linguistically. They are more basic in their needs, and their mutation is very similar to what we see in other primates. We don't fully understand the role of the FoxP2 gene; it seems to have something to do with the motor cortex, which is part of the left hemisphere that hits right about where Broca's area is and has something to do with motor skills, along with muscles in the face. Maybe it has something to do with language. And this is a maybe. We have no proof yet; we're still collecting data and trying to understand the brain and everything about it. But with respect to that one gene, perhaps that is the case.

The final one to talk about the **social cognition**, which again tapping into this concept that human beings are social creatures. We crave being in a social environment and, if you think back to the previous chapter when we talked about theories of how children learn a language, a couple of them had social components to them. That's because we can observe how children interact both amongst other children as with adults and how they acquire language. One thing we suspect—I won't go stronger than 'suspect'—is that if we see this in modern humans now, this might tie into how humans acquired language originally. This is a very questionable idea; we do not know whether this is true, mostly because we're lacking a few things. For example, a time machine would be helpful. 😉 It's a nice idea to think about. What we have been able to show is that many of the same patterns we observe with children and how they acquire a language come up when we talk about creolization and second language acquisition as an adult. Maybe there is something to this concept that as social creatures, we feel the need to express more. But there are other animals, especially other mammals, that are equally social, yet they don't seem to have a language like ours, at least not something we can identify. Therefore, this is a 'perhaps'.

Truth be told, what we can say is that there might be a little bit of all those reasons; maybe not so much on the grunting and the gesturing, and maybe less so on the grooming hypothesis. With the latter two, they seem to have some more plausibility. But here's the thing: there could be other aspects, with respect to the development and the evolution of language, that we don't know. What we really need is that time machine. So, get on it, folks; create that time machine, so that we historical linguists can go back in time and we can go find out the answer. 😉

10.2: Evolution of Human Language is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.3: Language Processing and Parsing

10.3.1 Language Processing and Parsing, from Sarah Harmon



Video Script

10.3.1.1 Language Processing and Parsing

Let's go a little more into language processing and parsing. This is what we know so far, what we have data to show, and not just hypotheses but actual brain scans. We're able to start parsing out how a human processes language. This is also a very nascent area of linguistics; we have had theories on how language is parsed for a number of years, really for decades. However, we're only at the starting point of understanding how the brain works; the technology is only starting to allow us of view into how this is happening. A lot of what is presented here, hopefully, will either be supported in the future, or shown to be different; we will see.

Let's start with the specific processing of language. There are two views of this: top-down and bottom-up, and that has everything to do with the input versus the output. When we're talking about processing, we're talking about when you hear language and you understand it—competence, in this case. **Bottom-up processing** would be that you get your input from your ear (or if you're signing, from your eyes), then you process the small to the big. You start with the sounds and the phonemes, then you go to the puzzle pieces and the morphemes. From there, you start processing how they combine in that phrase, and then the semantic or pragmatic interpretations. That's one way to view processing. The other view would be **top-down processing**, which would mean that you start with the overall meaning—concepts, context, pragmatic aspects, and the phrasal aspects—and then break it down to the morphology, then phonology and phonemes.

Both of those theories of processing are highly abstract, and we actually have a little bit of evidence for both. First, think about if you are in a really noisy environment and you're trying to have a conversation with somebody. You cannot hear everything that they say, and yet you follow along; you get the gist of what they're saying. That's an example of top-down processing; you're getting the overall context and then you're filling in the blanks, as it were, for the pieces that you did not hear. There have been a number of studies that showcase that if you get folks trying to pronounce isolated words, and then have them use those words in a sentence, they will make more mistakes if they pronounce the word in isolation. Interestingly enough, if you create a gibberish sentence—a sentence that fails compositionality and/or grammaticality, it could be either one—that there are errors in different ways. That's an example of both bottom-up and top-down parsing; bottom-up because you're focusing on that lexicon in isolation and therefore, you're focusing on the phonemes and then going up from there. Top-down because in a sentence you're able to use the context clues to fill in the gaps.

Speaking of gaps filling in gaps, I think of this kind of like Mad Libs—the game where one person has a story or a paragraph with certain words out, and they ask their friends or colleagues or audience for a specific kind of word a noun, a color, a verb, an action, and then you fill in the gaps. Filling-in-the-gap experiment is pretty much the same thing; you leave out certain key information

and the audience, the person you're talking to or writing to, fills in the gap and frequently will use the context clues to fill in the gap. Here's the interesting piece: that's mostly a top-down parsing, but there's also a bottom-up parsing, if they're given a partial clue. Let's say they were given a sentence and a key verb was missing. If they only heard one sound initial sound of that verb, they might fill in the gap with the appropriate term, thereby building up from the bottom, from the sound, and building up to the context. We see a number of cases of both bottom-up and top-down processing. There isn't one way that we do this; it's both.

The final example of this has to do with segmentation; I'll give you a great example. You see in IPA this term here: [grede]. Unless you know the context, that could be two possible meanings. It could be *grade A* as an something was given the grade of A, so there's a syllable break, as it were, between the [d] and the [e]. Or, it can be a *grey day*, so the syllable break would be between the first [e] and the [d]. Unless you have the context, you don't know. This is where arbitrariness really comes into play; I keep bringing this term up throughout the course, and I told you it was the most important of all the hallmarks of human language. This is an example; we cannot parse something unless we understand the context. This is just English, but this is true for any language, where you have any level of homophony.

With respect to processing, we also have to factor in anticipation and priming; you'll see a little bit more about priming in the next section that Catherine Anderson includes, but she mixes up a few things. That's why I'm not going to do it here. **Anticipation**, especially with respect to phonology, is a signal that you are fluent in a language. If you hearken back to the very last section of language acquisition, the previous chapter, we talked about what fluency really meant and that the average person's definition of fluency is lower, and does not satisfy the definition that a linguist or a foreign language professional would count as fluency in a language. Anticipation is one aspect; if I'm fluent in a language, I can anticipate based off of what the person is saying. The sound and the context will help me parse out everything else. If I'm not fluent in that language, I have a harder time parsing that information and anticipating what will happen. The same is true with **priming**; that's focused more semantic and morphologic input, where certain cues will lead you to think about other things. If I'm talking about a picnic that I'm organizing, and I'm trying to figure out who should bring what, and I say, "Well, first of all, I kind of would like to have a hot dog; somebody have hot dogs at the at the picnic." That might lead other people to say, "Well, we should bring burns. Maybe we could do hamburgers or other meats or meat substitutes. Oh, we need condiments like mustard, ketchup, mayo and pickles." One seed, as it were, grows into all sorts of connections, and that is what priming does.

What is really cool is to see how categorical language is. The video below talks about the brain dictionary, and it shows the hypothesis that we have been working with for some time about a mental lexicon. It seems that we do actually see in different parts of the brain categories of lexicon, and they seem to pop up in different areas of the brain.



10.3.1.2 Failure to Parse

If we talk about parsing and processing, we also have to talk about what happens when we don't parse correctly, when there's something going on. If we talk about arbitrariness, we also have to talk about ambiguity. We talked a little bit about ambiguity with respect and syntax and semantics. There's more to it; there are two specific principles—The Principle of Minimal Attachment and

The Principle of Late closure—that factor heavily into ambiguity. The **Principle of Minimal Attachment** is that you build the simplest structure consistent with the grammar of the language. If something doesn't make sense, it could be in violation of the Principle of Minimal Attachment, because you got too complicated; you need to keep it simple and straightforward. Don't try and attach extra pieces. The **Principle of Late Closure** says that we attach incoming material to a phrase that is currently being processed, we anticipate and prime ourselves; if what we're priming and anticipating doesn't match what happens, there's ambiguity. Both of these are factors when we say we don't understand something, and you can also include the Principle of Compositionality, which we talked about in the Meaning chapter. They all work together; if there's ambiguity, something is breaking down. Either the structure doesn't follow what we expect—that's Minimal Attachment—or the semantics and the content doesn't match what we expect—that's Late Closure.

What we can do sometimes is **backtrack** and use a **garden path**, as it were, to backtrack and figure out what the person was intending to say. This, especially happens with respect to presupposition and implicature. If we didn't understand the context that the speaker was intending, then we'll backtrack or use the garden path, as it were, to go back to where we were to try and piece it together. You can see, this all the time in a conversation. If you have a conversation with somebody, and you aren't quite picking up what they're saying, or you're not quite understanding it, it means that you're not able to parse it. It's ambiguous. You usually have some kind of way of stopping the conversation and go back to the part that you last understood. It could be as simple as saying, “Wait just a second, let me go back,” or some version of that phrasing. Or you can say, “Hold up; say that again,” and that's another way to cue your speaker that they're not making sense. This is something called **dischordia**. We signal dischordia to explain that scenario when the person you’re talking/signing to says something that you are able to parse. This definitely can lead to **re-analysis**. By the way, if you think about historical linguistics, when we talked about language change and specifically re-analysis, that *an apron* really started off in life as *a napron*, so there was a question as to priming and parsing, probably a Late Closure error. You can also see it in acquisition of lexicon from other languages, when we borrow something and then add or subtract from it. We see this in a variety of ways, and this all has to do with parsing.

We can even talk about bilingual parsing, when we are talking about aspects of language that are not native to us, and the use of context clues. We talked about how if I said [si] to a Spanish-English bilingual, depending on the context depends on whether it is a body of water or an affirmative comment, and that has everything to do with context clues. It also ties into knowing your audience, whether you know the person that you're talking to speaks the language that you're intending to use. Implicature and presupposition is also part of this. We build as we learn a language, whether it is our native language or one we have acquired, whether it is our first language or our 10th language we constantly build all of these pieces in such that parsing becomes a high level of processing language.

Of course, there are errors in parsing. When we have **slips of the tongue**—I meant to say *bridge* and I said *bork*—or **word substitutions**—I can't think of *sofa*, so I’m going to say *couch*, but for whatever reason that semi-synonym may not work for the context. There are **blending errors**, which is when instead of combining lexicon that work together, like *smog* from *smoke* and *fog*, I say *skog* instead and there's an error in the blending. **Spoonerisms** are when you swap out full syllables and swap them, and then create a metastasis that may create a whole different word. **Metaphor**, when it's not shared meaning, and the other person doesn't understand the context, all of these are errors in parsing, in comprehension.

What is also important to remember is that every single human being innately and subconsciously shows that they have an error in parsing and comprehension, and it can be followed by a conscious thought and even statement. When we talk about animal communication, that is something that we don't see much of; we don't see errors in parsing. You don't see a cat telling another cat, “Hey, I didn't understand you. Would you say it again?” We don't (so far) know that they do that.

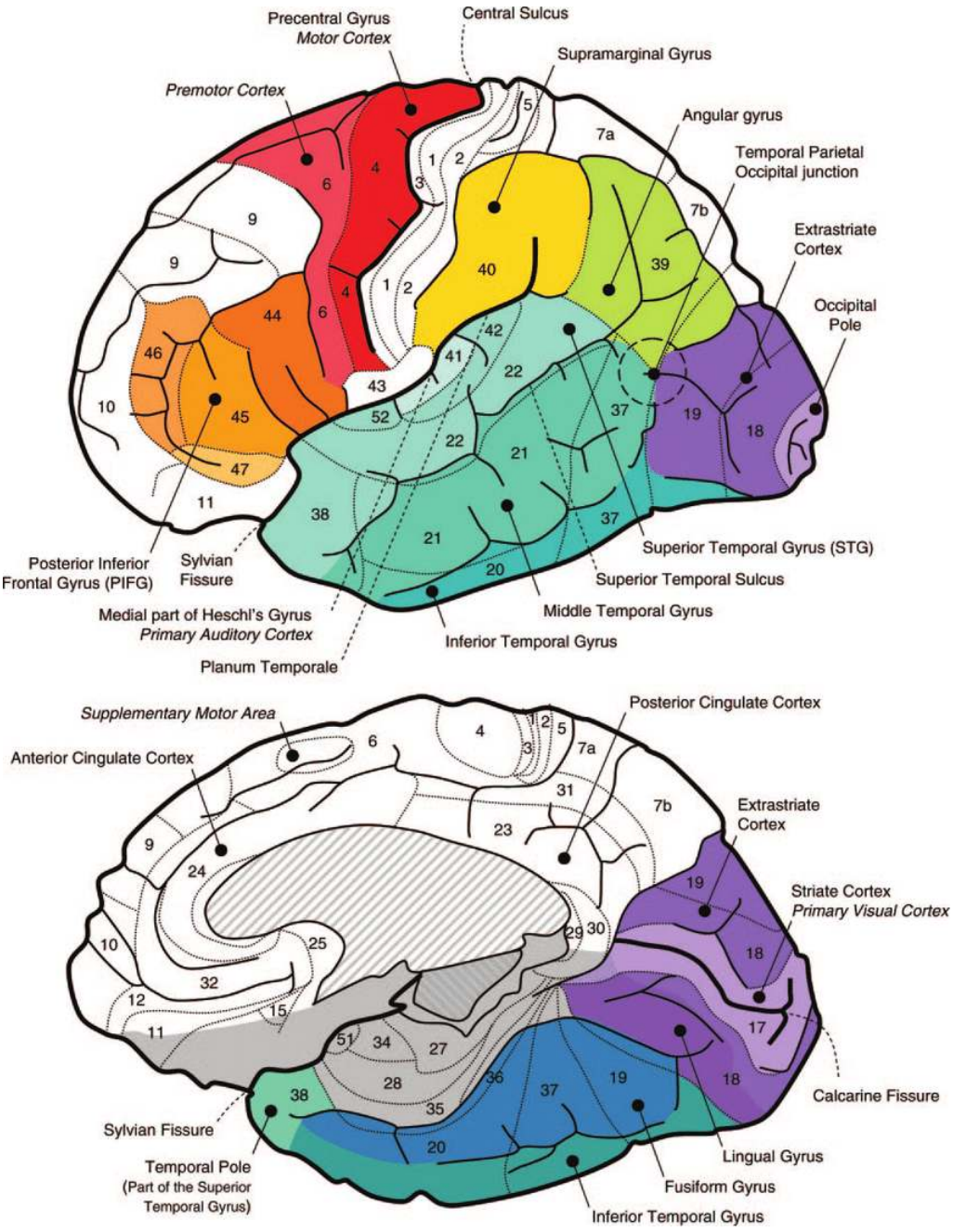
10.3: Language Processing and Parsing is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.4: Language Processing in the Brain

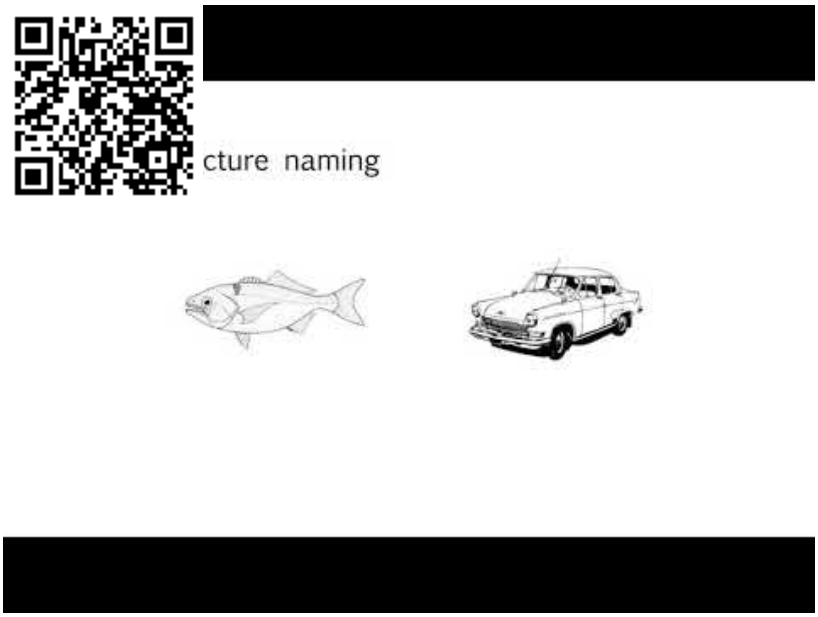
The brain and language processing

Below are some updated graphics of the brain with respect to processing. They may be useful to you as you read the subsequent material.

The anatoy of the functional areas of the cerebral cortex



Neurolinguistics: Syntactic Category Differences in the Brain, in Anderson's *Essentials of Linguistics*



Video Script

We've seen that we can group words into categories according to how they behave. We know that words within a particular **syntactic category** behave similarly to the other words in that category. They're similar in their inflectional morphology, and in their syntactic distribution, that is, what positions they can occupy in a sentence. That's some linguistic evidence that syntactic categories are real. There's also **neurolinguistic** evidence that our brains respond differently to words from different categories.

Lorraine Tyler and her colleagues used **functional magnetic resonance imaging (fMRI)** to measure blood flow to different regions of the brain. The idea behind fMRI is that brain activity consumes oxygen, and when a particular area of the brain is active, then more blood flows there to bring it more oxygen. The researchers asked people to do an easy task. They showed them lists of three words and asked them to decide if the third word, the one in all caps, was related to the first two. So in this example, *sparrows, thrushes, and wrens* are all kinds of birds, so the participants would respond Yes. In this next example, *hammer, wrench, banana*, the first two are tools but the third one is a fruit, so it's not related to the first two, so the participants would answer No.

Some of the words in this task were nouns, like the lists we just saw, and some were verbs, like these ones: *eating, grazing, dining*. All of those words are related to eating, so the participant would decide Yes. This is a pretty simple task, but what the researchers found in the fMRI was that there were several areas of the brain that showed greater blood flow for verbs than for nouns! Apparently, the brain was reacting differently to words from these two different syntactic categories, even though the task was the same for both categories.

We also see differences between nouns and verbs in the brains of people with **aphasia**. Aphasia is the name for any kind of language disorder that results from an injury to the brain, such as a stroke or a tumour. There are different kinds of aphasia that have different kinds of symptoms. Louise Zingeser and Rita Sloan Berndt found a **dissociation** between nouns and verbs in the speech of two different groups of people with aphasia.

The researchers asked their participants to do a few simple tasks. One was a picture naming tasks, where the researcher would show a line drawing and ask the participant to say what it was, like a *fish* or a *car*. In another task, they asked participants to describe how they would go about a particular action, such as making a birthday cake or attending a concert. And in the last task, they gave participants a picture book that depicted a well-known fairy tale but didn't have any words in it and asked them to tell the story. So from all these tasks, they had a good collection of each person's speech. For each person, they calculated the ratio of nouns to verbs.

It turns out that in people without any brain injury, the ratio of nouns to verbs is pretty close to one. That means there are about the same number of nouns as there are verbs in the average person's speech for these tasks. But for people with **agrammatic aphasia**, verbs are very difficult to produce. These people had more than twice as many nouns as verbs. And for people with **anomic aphasia**, nouns are quite difficult. This group of people had fewer nouns than verbs.

Aphasia researchers call this kind of pattern a **dissociation**. We say that nouns are verbs are dissociated from each other because it's possible to have verb production impaired while noun production is ok, or vice versa. This dissociation is consistent with the idea that verbs and nouns are processed differently in the brain.

So what does all this mean? We've seen that, in typical brains, a simple task with verbs involves greater brain activity than the same task with nouns. And in people with brain injuries, some people have an impairment of verbs but not nouns, while others have an impairment of nouns but not verbs.

All this suggests that our brains treat words differently depending on what category they're in. Or in other words, different syntactic categories exist not just in language, but also in the brain!

Check Yourself

? Exercise 10.4.1

Comparing the following sets of words, which would you predict would lead to greater blood flow in more areas of the brain?

- Humming, singing, whistling.
- Piano, flute, guitar.
- Gloves, scarf, hat.

Answer

"Humming, singing, whistling."

Hint: Think about the categories of each group of lexicon, and remember which took more bloodflow to recall.

? Exercise 10.4.2

When shown a picture of a pair of tongs, a patient describes the picture, "You pick up things with it". Which type of aphasia is this response more typical of?

- Anomic aphasia.
- Agrammatic aphasia.

Answer

"Anomic aphasia."

Hint: Think about the word that the patient can't recall, and then think about its category.

? Exercise 10.4.3

When describing an injury to his knees, a patient says, "no good uh ache and uh uh uh knees and ankles uh home doctor and legs". Which type of aphasia is this response more typical of?

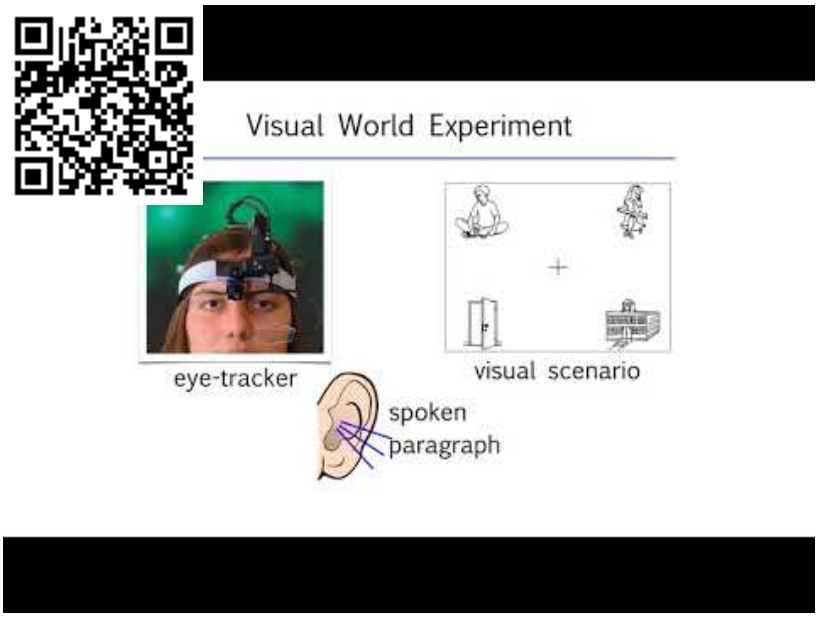
- Anomic aphasia.
- Agrammatic aphasia.

Answer

"Agrammatic aphasia."

Hint: Think about the word that the patient can't recall, and then think about its category.

Psycholinguistics: Traces in the Mind, in Anderson's *Essentials of Linguistics*



Video Script

We've been working with a theory that says that the operation MERGE generates a Deep Structure. For this *wh*-question, *Who did Lucy invite to wedding*, the Deep Structure looks like this. This *wh*-pronoun *who* refers to whoever it is that Lucy invited, and it is generated in this position in the complement of the Verb head, which is exactly where the noun phrase complement would be if we know who it was that Lucy invited. The preposition phrase gives us more information about the event of inviting, and it's adjoined at the V-bar level. Because this is a *wh*-question, there's both a [+Q] feature and a *wh*-feature in the C-head position.

Then the operation MOVE does its work.The *wh*-phrase moves up to the Specifier of CP, where it can support the *wh*-feature in C. Then *do* comes into the T-head in its past-tense form, *did*, then moves up to the C-head position.

One element of this theory that we've been taking for granted so far has to do with the **trace** that's left behind when something moves. When we speak a sentence, we pronounce words in their Surface Structure positions, but we don't pronounce anything in the Deep Structure position. But when we draw the tree, we show the deleted copy in that Deep Structure position, to suggest that, in the underlying representation, in our mental grammar, there's something unspoken occupying that position.

There is some linguistic evidence for the existence of traces in our mental grammar. We're claiming that there's a trace in this position in the complement of *invite*. Notice that it's not possible for any other phrase to occupy that position: if we try to put another noun phrase in the complement position, we can observe that each attempt is ungrammatical.

There's also some psycholinguistic evidence for the existence of traces. The evidence comes from what's called a **visual world experiment**. In this kind of experiment, a person's eye-movements are measured using a device called an **eye-tracker**. The eye-tracker records where their eyes move while they listen to a spoken paragraph and look at a visual scene. The spoken paragraph goes like this:

> *This story is about a boy and a girl. One day they were at school. The girl was pretty, so the boy kissed the girl. They were both embarrassed after the kiss.*

The idea behind a visual world experiment is that you look at what's being mentioned. So when you hear *the boy*, your eyes move to the picture of the boy, and when you hear *the girl*, your eyes move to the picture of a girl.

At the end of this paragraph, one group of listeners heard a *wh*-question, *Who did the boy kiss that day at school?* A different group of listeners heard the same paragraph, but followed by a *yes-no* question, *Did the boy kiss the girl that day at school?*

These two sentences are similar in their structure, but they have a crucial difference. In the complement of the verb position, the *yes-no* question has an overt noun phrase, *the girl*. In that same position of the *wh*-question, there's a trace, a deleted, unpronounced copy of the moved *wh*-word.

The researchers focused on this exact position in the spoken question: they observed where the participants' eyes moved after the verb *kiss*. They compared how often the participants looked at the girl vs. how often they looked at the boy. In the *yes-no* question, when they heard the verb *kiss*, people looked at the boy 11% more often than to the girl, maybe because the boy is the one doing the kissing. But in the *wh*-question, when they heard the verb *kiss*, people looked to the girl 21% more than to the boy.

In both scenarios, the boy kissed the girl. But people's eye movements differed in the two conditions. We know from previous studies that eye movements are quite closely synchronized to what's being mentioned in the discourse. In this study, people's eyes move to the girl not just when the sentence refers to her overtly, but also when the deep structure contains a trace that refers to her. The evidence from this eye-tracking experiment suggests that traces don't just exist in tree diagrams, but also in our minds.

Check Yourself

Exercise 10.4.4

Which of the following illustrates the position of the trace in the wh-question *What did Christina order at Chipotle?*

- What did Christina order at Chipotle what.
- What did Christina order what at Chipotle.

Answer

"What did Christina order what at Chipotle."

Hint: Recall what we studied earlier in Syntax with respect to movement, and then think about where that 'what' would have originated from.

Exercise 10.4.5

Which of the following ungrammatical sentences gives evidence that unpronounced traces exist in our mental representations of sentences?

- *Did you ate what for lunch?
- *What did you eat sandwiches for lunch?

Answer

"*What did you eat sandwiches for lunch?"

Hint: Recall what we studied earlier in Syntax with respect to movement, and then think about where that 'what' would have originated from.

Exercise 10.4.6

Predict which sentence would lead to more eye movements to a picture of a rabbit after the verb *chase*:

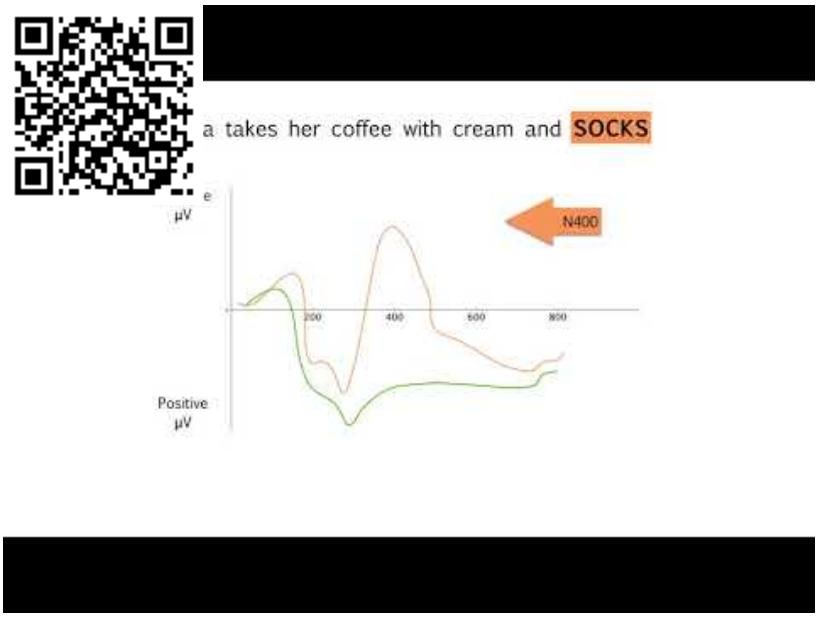
- What did the fox chase ^ into the hedge?
- Did the fox chase ^ the rabbit into the hedge?

Answer

"What did the fox chase ^ into the hedge?"

Hint: Think about the role that a Prepositional Phrase plays, which includes movement or location.

Neurolinguistics: Using EEG to Investigate Syntax and Semantics, in Anderson's *Essentials of Linguistics*



Video Script

When we started talking about semantics, we observed that a sentence's syntax influences its semantics, because of the principle of **compositionality**. For example, we saw that a given string of words can have two different meanings if it has two different grammatical syntactic structures. And yet, we also observed that syntax is independent of syntax. A noun phrase that has the semantic thematic role of Agent often occupies the syntactic position of Subject, but not all Agents are Subjects, and not all Subjects are Agents!

The division of labour between thematic roles and grammatical roles is some evidence that syntax and semantics are represented differently in our minds. There's also evidence from **neural imaging** to show that our brains process semantic information differently from syntactic information. This evidence comes from **electroencephalography** or **EEG**. Electroencephalography uses electrodes to measure electrical activity on a person's scalp from which scientists can draw conclusions about the person's neural activity. The particular EEG technique that gets used in neurolinguistics is **ERPs** or **event-related potentials**, which measure the timing of the neural response to a particular event, like a sound or a word.

When we're observing ERPs, we always do so by comparing responses to different kinds of events, and the usual comparison is between events that are expected and events that are unexpected. For example, a sentence like, "She takes her coffee with cream and …" sets up a very strong expectation in your mind of what the next word will be. If the next word that arrives in the sentence matches your mind's expectation, then the electrical response at your scalp will look something like this: the baseline condition. But if the next word that shows up violates your mind's expectation, then compare your brain's response: We observe a spike in negative voltage about 400 milliseconds after that unexpected word appears.

This response is called an **N400**. The N in N400 stands for a negative voltage, and the 400 indicates that this spike in negative voltage shows up, on average, about 400 milliseconds after the event. The N400 was first observed in 1980 by Kutas & Hillyard and has been replicated hundreds of times since then. It's clear from all these studies that the particular kind of event that leads to an N400 response is a word that is unexpected in the semantic context.

The N400 is the brain's response to an unexpected or surprising event, but not every kind of surprise will produce an N400. In other words, we have expectations about things besides the meanings of sentences. Think about a simple sentence like, "The bread was…" If that sentence finishes with "eaten", that fits our mind's expectation, and this is the baseline brain response. Now, what expectation do you have for this sentence, "The ice cream was in the.."? You probably expect a noun to come next, to follow the preposition and determiner. But if what comes next is not a noun but a verb participle, this violates your mind's expectation. Notice that the word *eaten* is semantically consistent with ice cream, but is not consistent with the syntax of the sentence: determiners are followed by nouns, not verbs. So the brain's response is a positive voltage about 600 milliseconds after that unexpected word: a **P600**.

When we're using language in real time — either reading or listening — our mind sets up expectations about what's going to happen next. If what happens next violates our semantic expectations, the brain's response is an N400. And if what happens next violates our syntactic expectations, the brain's response is a P600.

These two different brain responses give us further evidence that syntax is independent of semantics in our brains!

Check Yourself

Exercise 10.4.7

What kind of ERP is likely to be observed in response to this sentence?

The children's feet have grown so I bought them some new *spoons*.

- N400.
- P600.

Answer

"N400."

Hint: Think about what is wrong about the indicated word, and the types of ERPs described above.

Exercise 10.4.8

What kind of ERP is likely to be observed in response to this sentence?

Elias told the coach that he wanted to learn to *swam*.

- N400.
- P600.

Answer

"P600."

Hint: Think about what is wrong about the indicated word, and the types of ERPs described above.

Exercise 10.4.9

What kind of ERP is likely to be observed in response to this sentence?

Before the dinner party, Isla spent a long time in the *cooked*.

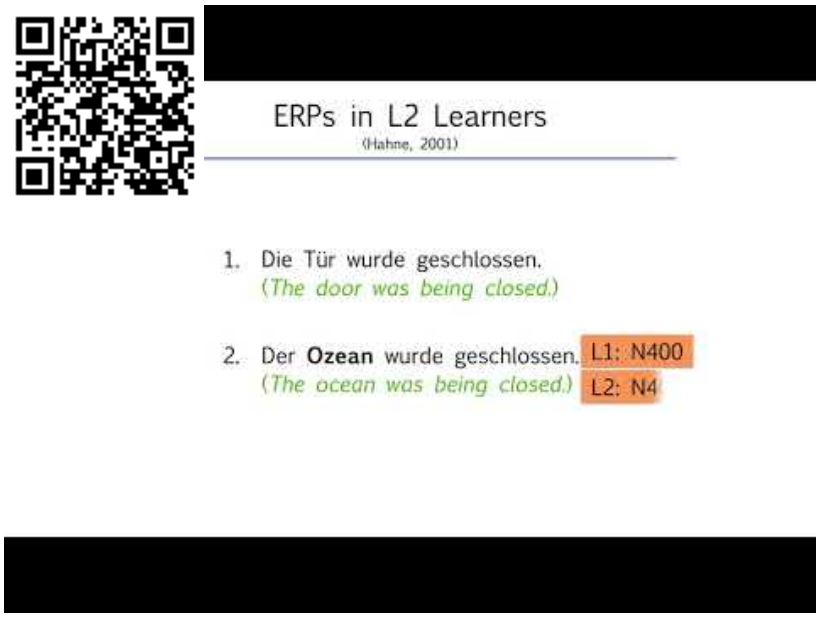
- N400.
- P600.

Answer

"P600."

Hint: Think about what is wrong about the indicated word, and the types of ERPs described above.

Neurolinguistics and Second Language Learning, in Anderson's *Essentials of Linguistics*



Video Script

As we've been talking about mental grammar, we've concentrated almost entirely on the mental grammar of your **native language** — the language you learned to speak in childhood, in your home. Linguists refer to your native language, your first language, as your **L1**. But many people in the world speak more than one language, and many of those people learned a second or third language in a different way from their L1. Any language that you learned after childhood, whether you learned it in school, using software, by travelling or immigrating somewhere, is called an **L2** (even if it's really your third or fourth language).

Learning an L2 is different from learning an L1 for a couple of different reasons. One is that, obviously, the language learner is not a child, so their cognitive processes might be different from those of a child. L1 learning happens by being immersed in a language environment, and most of the learning is unconscious, without overt teaching. L2 learning often happens with a lot of conscious effort: studying and memorizing and practicing.

But of course, the biggest difference between L1 learning and L2 learning is that when you start learning an L2, you already know at least one other language. The mental grammar of your L1 can influence the mental grammar that you're developing for your L2: this is called **transfer**. Transfer can be helpful in L2 learning or it can pose a challenge. If your L1 includes a structure that's similar to a structure in the L2, then you might experience **positive transfer**, which facilitates learning the L2: you can transfer what you know from L1 and apply it to the L2. But if the structures that you're learning in L2 are different from those in your L1, then you might experience **negative transfer**: the knowledge from your L1 could make it more difficult to learn the new structures in the L2. And of course, you might experience both positive and negative transfer from your L1 to different parts of the grammar of your L2.

One theory of second language acquisition predicts that we would observe differences between native speakers and beginner L2 learners, but as the L2 learners become more proficient, their mental processes should become more and more native-like — that is the mental grammar of a fluent L2 speaker should look very similar to the mental grammar of an L1 speaker of that language. We can use the tools of psycholinguistics and neuroscience to learn about the mental grammars of L2 learners. Let's take a look at some of the evidence.

Several studies have compared N400 effects in L1 and L2 speakers of a language. Remember that the N400 is an electrophysiological response that our brains show when a word is semantically unexpected in a given context. The brains of native speakers of English show a negative voltage about 400 milliseconds after a semantically unexpected word (*socks*), compared to an expected word. But what do the brains of non-native speakers show? What do we see in L2 learners?

A 2001 paper by Anja Hahne compared L1 speakers of German with L1 speakers of Russian who had moved to Germany in their 20s and had been living there and studying German for an average of six years. The experiment used fairly simple German sentences like these ones:

> Die Tür wurde geschlossen / The door was being closed.

Der Ozean wurde geschlossen / The ocean was being closed.

Obviously, the word *closed* is a reasonable way for the first sentence to end but is a pretty unexpected way for the second sentence to end. So it's not surprising that the native speakers of German showed an N400 in response to sentence 2 compared to sentence 1. The L2 speakers, the ones who had started learning German in their 20s, also showed an N400 to sentence 2. Hahne concluded that words that are semantically unexpected cause "essentially similar semantic integration problems in native participants and second-language learners" (Hahne, 2001: 263). In other words, the evidence from the N400 suggests that the lexical semantic component of an L2 learner's mental grammar is not too different from that of an L1 grammar.

Now, we know that there's a whole lot more to mental grammar than just the meanings of words. What can ERPs tells us about morphology and syntax? Remember that native speakers' brains often show a P600 response to sentences that are syntactically unexpected. For many years, studies that looked at the P600 in L2 learners seemed to suggest that adult language learners never really approached native-like proficiency in their L2 morphosyntax: the P600 response to syntactic violations was significantly delayed or not there at all in these late learners. But some more recent research has suggested that maybe those earlier studies just didn't give the learners enough time to learn their L2 — of course, their mental grammar wasn't native-like if they hadn't been learning the language for very long.

A 2013 study by Harriet Bowden and her colleagues looked at L1 English speakers who started learning L2 Spanish in university. They compared learners who had completed first-year Spanish to learners who had completed more than three years of university Spanish and had spent a year abroad. And they included a control group of L1 speakers of Spanish. The researchers presented Spanish sentences that violated syntactic expectations about word order like these ones. Sentence 1, "I have to run many miles this week" has the expected word order, while sentence 2, "I have to miles many run this week" is unexpected in its word order: the quantifier *many* comes after the noun *miles*, and that whole complement phrase comes before the verb *run*. Sentence 2 is ungrammatical in Spanish.

As you'd expect, the native speakers of Spanish showed a P600 in response to the ungrammatical sentence. But so did the advanced L2 learners: their ERP response was the same as that of the L1 Spanish speakers. It was only the beginning learners, the ones who had had only a year of Spanish, who showed an atypical P600: it was a smaller response and more diffuse. The researchers concluded that "University foreign-language learners who take L2 classes through much of college and also study abroad for one or two semesters …show evidence of native-like brain processing of syntax" (Bowden et al., 2013: 2508).

So this study suggests that one year of studying a language maybe isn't enough to achieve native-like fluency, and three years of study including a year of immersion allows a learner to approach native proficiency, but the researchers also wondered whether the kind of language-learning makes a difference to learners. If you're learning a language in university, you probably spend three or four hours a week in the classroom, and maybe two or three more hours each week studying. But that's not the only way to learn a language.

A study in Montreal looked at university students who were L1 speakers of Korean and Chinese who were enrolled in a nine-week intensive English L2 course. These learners were studying, practising, using English at least 8-10 hours a day, five days a week, for nine weeks. The researchers tested the learners on sentences with morphosyntactic violations in the tense features on the verb, like these ones:

1a. The teacher did not start the lesson / 1b. The teacher did not started the lesson.

2a. The teacher had not started the lesson / 2b. The teacher had not start the lesson.

Notice that in 1b and 2b, the verb has unexpected morphology on it. A native speaker of English would show a P600 response to 1b and 2b in comparison to 1a and 2a. In this study, the researchers measured learners' ERP responses at the beginning of the course and after the nine weeks, and they also asked the learners to judge whether the sentences were grammatical. At the beginning of the course, none of the learners showed P600s in response to the syntactically unexpected sentences, and they also weren't very successful at deciding whether sentences were grammatical or ungrammatical. After the nine-week course, all of the learners showed P600 responses to the syntactically unexpected sentences, and the learners who scored highest on the grammatical judgments showed the largest P600s. This study suggests that even short-term, intensive L2 learning can help a learner develop a mental grammar that approaches that of a native L1 speaker.

And the results of all of these studies tell us that L2 language learners can achieve fluency that compares to that of a native speaker; it just takes lots of training to get there!

Check Yourself

Exercise 10.4.10

French uses morphology indicate whether nouns, adjectives and determiners are masculine or feminine. If an L1 speaker of English is learning French, what kind of transfer are they likely to experience in learning this property of French grammar?

- Positive transfer.
- Negative transfer.

Answer

"Negative transfer."

Hint: Think about the structures of French and English, and think about whether they are more similar or more different.

Exercise 10.4.11

Russian does not have definite or indefinite determiners like English *a* and *the*. If an L1 speaker of Russian is learning English, what kind of transfer are they likely to experience in learning this property of English grammar?

- Positive transfer.
- Negative transfer.

Answer

"Negative transfer."

Hint: Think about the structures of Russian and English, and think about whether they are more similar or more different.

Exercise 10.4.12

Russian groups nouns by their grammatical gender, either masculine, feminine or neuter. Look again at the facts about French presented in Question 1. If an L1 speaker of Russian is learning French, what kind of transfer are they likely to experience in learning this property of French grammar?

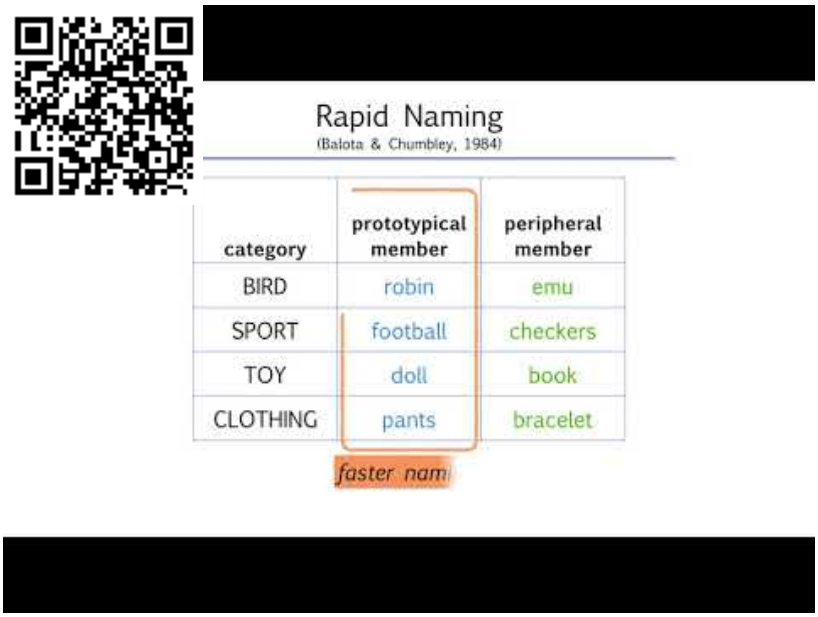
- Positive transfer.
- Negative transfer.

Answer

"Positive transfer."

Hint: Think about the structures of Russian and French, and think about whether they are more similar or more different.

Psycholinguistics of Word Meanings, in Anderson's *Essentials of Linguistics*



Video Script

In the last unit, we suggested that **intensions** for word meanings might be organized in our minds in **fuzzy categories**. Our minds construct categories of things based on our experience in the world: each time we encounter an extension of a word, we count it as an **exemplar** in that fuzzy category. There is some evidence from psychology and psycholinguistics that our mind really does represent a difference between **prototypical** category members and **peripheral** members. For lots of categories, we have some instincts about what kinds of exemplars are prototypical and what kinds are peripheral. When we give somebody the name of a category and ask them to name an exemplar, people from a given language community are remarkably alike in the first things they name as exemplars. If your mental grammar for English is like mine, then perhaps your prototypical bird is a robin, your prototypical fruit is an apple, and your prototypical tool is a hammer.

In a behavioural study of word recognition, participants saw a word appear on a screen and had to say the word out loud. This is called a **rapid naming** task. Some of the words referred to prototypical exemplars of their particular category and some of them referred to peripheral exemplars. The prototypical and peripheral exemplars were all mixed up in the experiment, but when the researchers measured how fast people had been able to name the word that they saw, the found that people were faster to name the prototypes than the peripheral exemplars.

The same researchers used these words in a lexical decision task. In this kind of task, a word appears briefly on a screen, and the person's job is just to decide whether it's a word or not, and say Yes or No. So if the word *pants* appears on the screen, you would say "Yes", because it's a real word in English. But if *pfonc* appears, you say "no", because that's not a word of English. What the researchers found in the lexical decision experiment was, again, that people are fast to make a decision about a word if it refers to a prototypical category member, and slower to make the decision if the word refers to a peripheral member.

These findings indicate that the process of recognizing a word is easier and faster if that word refers to a prototype. We can interpret these findings to mean that our intensions for categories are made up of exemplars and that prototypical exemplars have a privileged position in our intensions.

So that's a couple of examples of psycholinguistic tasks we can use to observe how words are processed in our minds: a simple naming task, and a lexical decision task. There's an additional task that we can combine with each of these, to allow us to investigate relationships between different words. That task is called **priming**. A **primed lexical decision** task works like this: First, a word appears on the screen for a very short length of time: that word is called the **prime**. The prime disappears, and then a second word appears on the screen. This word is the **target**, and the participant makes a lexical decision about the target. The prime word can have an influence on how quickly people make their lexical decision about the target word.

For example, in one condition, the prime might be *doctor* and the target *nurse*. In another condition, the prime could be *apple* and the target *nurse*. As you might expect, people are faster to make their lexical decision to *nurse* when it's primed by *doctor* than

when it's primed by *apple*. When we observe this faster lexical decision, we interpret that to mean that these two words are connected to each other in our minds.

Over the years, psychologists and psycholinguists have conducted thousands of experiments on priming, and the results of these experiments show us how words are related to each other in our minds. The scientific literature has shown priming between words that are members of the same category, for words that are synonyms, antonyms, and even for words that describe things that share attributes. For example, an orange and a baseball aren't members of the same category, but they're both spheres, so they can prime each other.

Looking at all these and many other priming effects, we can conclude that those semantic relationships play an important role in how the meanings of a word are organized in our minds.

Check Yourself

? Exercise 10.4.13

Which of the following words would we expect to prime a target word *carrot*?

- Power.
- Broccoli.
- Cattle.
- Car.

Answer

"Broccoli."

Hint: Think about the meaning of *carrot*, and the meaning of *broccoli*, and what they share in common.

? Exercise 10.4.14

Which of the following words would we expect to prime a target word *happy*?

- Rabbit.
- Umbrella.
- Sad.
- Zoo.

Answer

"Sad."

Hint: Think about the meaning of *happy*, and the meaning of *sad*, and what they share in common.

? Exercise 10.4.15

Which of the following words would we expect to prime a target word *week*?

- Month.
- Actor.
- Goal.
- Unit.

Answer

"Month."

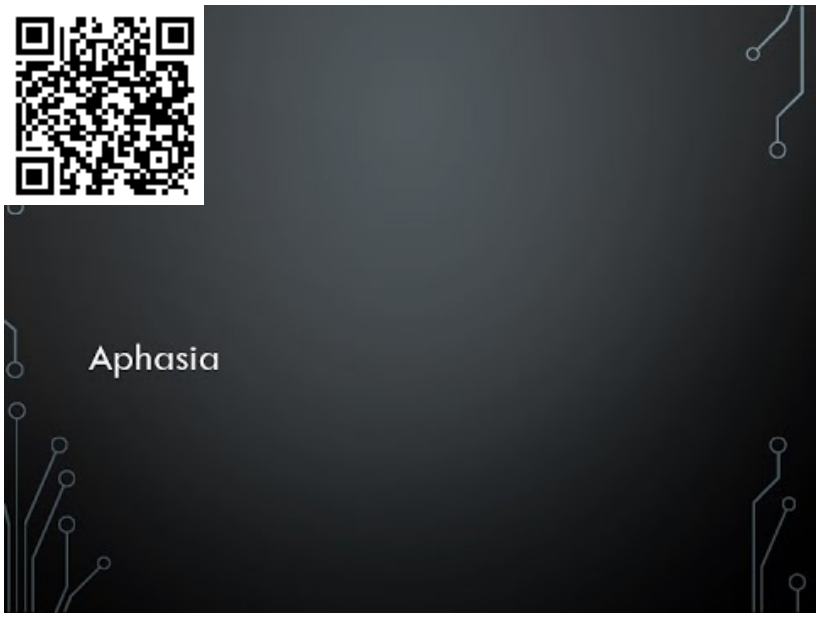
Hint: Think about the meaning of *happy*, and the meaning of *sad*, and what they share in common.

10.4: Language Processing in the Brain is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **7.7: Neurolinguistics- Syntactic Category Differences in the Brain** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **8.3: Psycholinguistics- Traces in the Mind** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **9.4: Neurolinguistics- Using EEG to Investigate Syntax and Semantics** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **9.6: Neurolinguistics and Second Language Learning** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.
- **10.3: Psycholinguistics of Word Meanings** by Catherine Anderson is licensed CC BY-SA 4.0. Original source: https://ecampusontario.pressbooks.pub/essentialsoflinguistics.

10.5: Aphasia

Aphasia, from Sarah Harmon



Video Script

In the previous section Catherine Anderson described a little bit about aphasia. But this is an area that I think deserves much more discussion, so I'm going to create a whole mini-lecture on this. It also is something that affects me personally; I do not have aphasia but my husband does. He suffered a stroke about eight years ago, and while he is as humanly recovered as possible, he's still showing effects of aphasia. Not all the time, but occasionally. This is not uncommon; as more and more people survive strokes or massive head traumas and severe concussions, there's more evidence for aphasia. I think it's one of those areas that most people do need to understand, at least at a basic level.

Just like with every other section of this chapter, the information I'm going to provide you is as up-to-date as we can get for our understanding of the brain. There is much more that we are learning; in fact, almost weekly, if not monthly, there's a new study on various types of aphasia and how to improve and treat aphasia. We'll see some of this here, but know that this is a constantly evolving area of neurolinguistics and neurology as a whole.

What is aphasia? **Aphasia** is an issue with speech or language; we do see aphasia also in deaf folks who use a primary sign language, so this affects both spoken and signed language users. We know that, as you read before, language processing is predominantly a left-brain activity, while lexicon are stored throughout the brain. Primarily, the processing is in the left side and that the **Broca's Area**, which is here in the front part of the temporal lobe on the left hemisphere, that is where much of the syntax and sentence formation occurs. The morphology and function words tend to reside there, too. **Wernicke's Area**, which is more or less behind it—but that's kind of nebulous—that is controlling of semantics and processing. There are pieces in the middle; **Exner's Area** controls writing, and that is in the motor cortex; the **Articulate Fasciculus** connects Broca's and Wernicke's Area; the **Angular Gyrus** is where a lot of the visual processing is done, and it's more towards the occipital lobe. Therefore, if you have a lesion in Broca's Area, you tend to suffer from Broca's Aphasia; if you have a lesion in the Wernicke's Area, it tends to be Wernicke's Aphasia. The name of the aphasia is tied to the primary area that is being affected.

Let's talk about these different aphasias.

Expressive or **non-fluent aphasia** is one entire category, and this is when the patient knows what they want to say, but they're not able to get it out. There are various levels of this aphasia. The lowest level is something called an **anomia**, where the motor skills of the patient are fine, but occasionally the patient is never able to come up with certain terms. Anomia frequently is called **jargon aphasia**, but jargon aphasia also affects Wernicke's Area, so we'll talk about that in a minute. If the patient has trouble getting out language regularly—everybody suffers a bit when they're tired, and they can't express yourself. But if this is a chronic issue, so when you have the phrases and words that you want to say, but you can't get them out—that feeling that it's like at the tip of your

tongue and they just not coming out—that is called either **motor aphasia** or **dysarthria**. That is what my husband has. If you're talking to him normally, most of the time you won't know that he has anything wrong. But if he's in a lot of pain, if he's very frustrated or upset, or if he's tired, he can't get out what he wants to say. He has to circumnavigate; he has to pivot and trying to think of a different way to say it. Again, all of us suffer from that a little bit when we're tired, or when we're sick, but certainly not at a chronic level; this is chronic. One of the things that dysarthria patients suffer from is extra wordiness, because they can't get out exactly what they're trying to say. They have to pivot, frequently that results in the patient getting very wordy. Frequently, it feels like they're going on a tangent, to a different conversation, and that's not actually the case. They desperately want to stay in-lane, as it were; they just can't because they've had to find other ways to describe their thoughts. Motor aphasia or dysarthria is challenging, and most patients who have dysarthria or motor aphasia also have some kind of coordination issue. In my husband's case, he certainly has a balance issue, and he also has certain dexterity issues, that are related to a stroke. A more aggravated version would be **Broca's Aphasia**. This aphasia can vary greatly, but by and large, if you have Broca's Aphasia, you are not putting sentences together fluently; this is why we call it non-fluent aphasia. Everything comes out in chunks; it may not be the correct term at first, and/or you may have to pivot and think of something else you know. Importantly, you know that you're not communicating properly. Unfortunately, many Broca's aphasics, if they cannot find any fluency, if they are permanently non-fluent, most of them just stop talking, or only talk to one or two very trusted folks. They're very aware that they are not able to communicate.

Below in 10.5.2, you have examples videos of a young lady named Sarah Scott. She is a British woman who, when she was 18 years old, suffered a massive stroke in her English class. As a result, she ended up with Broca's Aphasia. On the YouTube channel that her mother created and now she's a heavy participant of she shows her progress over time. I want you to watch the videos, at least in part. When you see Sarah Scott one year after her stroke, she shows very hesitant speech; stuff just isn't coming out. She doesn't have the right word, or she might know the word but it just is blocked and she can't do it.

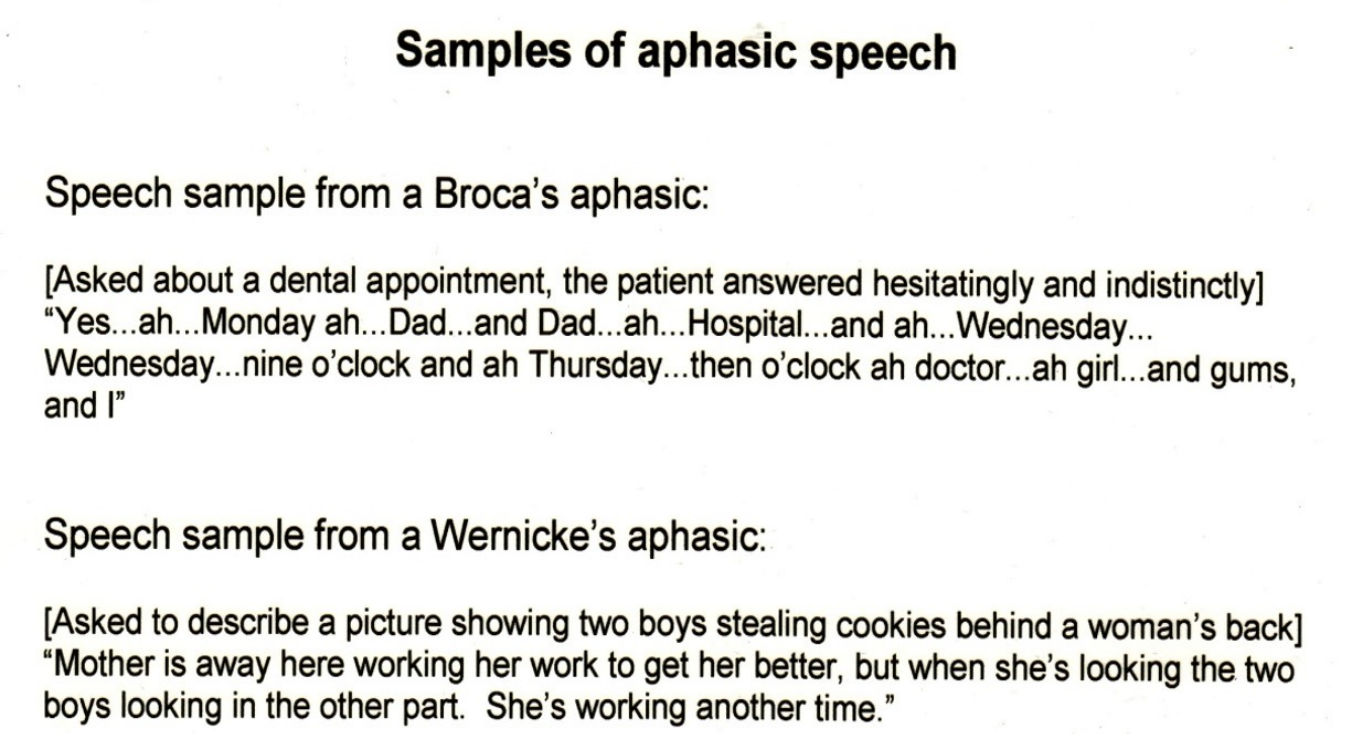
Broca's aphasics, however, do improve, with many, to the point of just having dysarthria. Check out Sarah's progress over the years—and this is just a couple of videos; there are many more on their channel. The same happened with my husband. When he first had his stroke, the first 2-3 months he was in Broca's Aphasia-land; he was able to get out much of what he wanted to say, but there was great hesitation, frequently searching for the word and not quite getting there. Between therapy and genetics, and the fact that he was young—he was only 41 when he had a stroke—that helped him to recover, to the point of dysarthria. As I said, if he's having a good day, with good energy and in good mood, you would never know he ever had a stroke. On a bad day, you suspect something has happened.

When we talk about **receptive aphasias**, frequently called **fluent aphasias**, that's a different story. **Wernicke's Aphasia** is the main one here, and this is when the patient's language is a jumble, that what the person says has no semantics; they're completely gone. Sometimes in extreme cases that can be called **jargon aphasia** because you're saying that the person is not able to come up with the right terms; the syntax is frequently decent, if not intelligible. There's grammaticality in many cases, but the semantics is wrong, the compositionality is wrong, nothing to do with anything else. When the lesion is away from the Wernicke's Area and closer to the Articulate Fasciculus—closer to Broca's Area—we see **conduction aphasia**, which is when the person can hear and understand and perform but can't repeat. They otherwise speak normally, but if you say, "Could you repeat that back? Can you repeat back what I said?" they can't do it. It is interesting to know that while dyslexia as a whole is still being researched, and we still don't really know a lot about how and why it forms, there are types of **acquired dyslexia** that have to do with language, specifically reading. These acquired dyslexia conditions do involve Wernicke's Area, so the person with acquire dyslexia frequently is able to speak perfectly fine, but because of where the lesion is, at the back part of Wernicke's Area affecting reading and processing, they're not able to read and comprehend.

There is important to underscore with respect to aphasia. Aphasia is fairly permanent; it may progress to a state that is better, especially with expressive or non-fluent aphasias, but even with receptive aphasias you can see progress, where the patient does get back some of their linguistic facility. But it never goes away; it's never 100% healed, at least not at this point in time. Certainly, there are factors that are in the favor or against improvement. When a younger person under the age of 60 has some kind of trauma or lesion, the brain is able to adapt; the brain is able to reroute a lot quicker and especially if that person was otherwise decently healthy. Even though a patient is not able to fully recover from any aphasia, there can be progress. As I mentioned earlier, there tends to be more progress with expressive aphasics. However, Wernicke's patients, in particular, are a different story; they very rarely recover. Although it can happen, it is very difficult. Conduction aphasia has much more success, and usually patients are able to recover some of that ability. It is something that takes a lot of time and there's a number of factors that need to be included.

Just to give you a written example of aphasic patients, specifically Broca's Aphasia versus Wernicke's Aphasia. You can see that with Broca's Aphasia, there's a lot of pausing—that's what the ellipses represent—that more or less, you can understand what the

person is trying to say. They were asked about a dental appointment and they described what happened. But notice how many times the patient has to pivot and start saying one thing in order to really say another thing. See the graphics below.



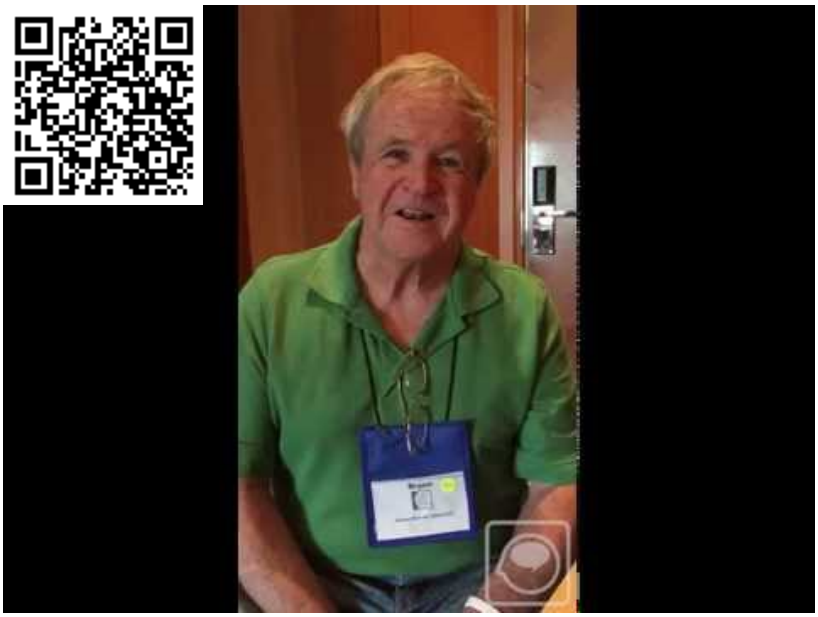
It's very hesitant, and it makes it hard to figure out exactly what he's saying, but you get the idea. This person is trying to convey that they went to a doctor's appointment that was for their teeth—therefore dentist—their father sent them or came with them, and it was probably in the morning. You look at Broca's aphasic, you can get the point. But Wernicke's aphasics are a different matter. In this scenario, we see a Wernicke's aphasic. They have a picture in front of them with two boys stealing cookies behind a woman's back. Now, I realize both examples are very difficult to understand. But I think you would agree with me that, with the Broca's aphasic, you got some of what this person was trying to convey; it was difficult, but you could do it. In the case of the Wernicke's aphasic, it's impossible. Yes, the person says that there's a mother and that there's two boys. But there's nothing about a kitchen; there's nothing about cookies; there's nothing with stealing; there's nothing about anything. This is what makes Wernicke's so difficult to treat is the processing part of it. In most cases Wernicke's aphasics don't necessarily understand that they're not communicating; sometimes they do, but frequently they don't.

All that being said, what aphasia tells us and helps us to understand is that language is independent of intelligence. Even at his worst moments, my husband may not have always been able to communicate exactly what he wanted to say. But he was just as intelligent as he was before the stroke. Now he's able to communicate pretty well. It just goes to show that there's always hope.

Examples of Aphasia

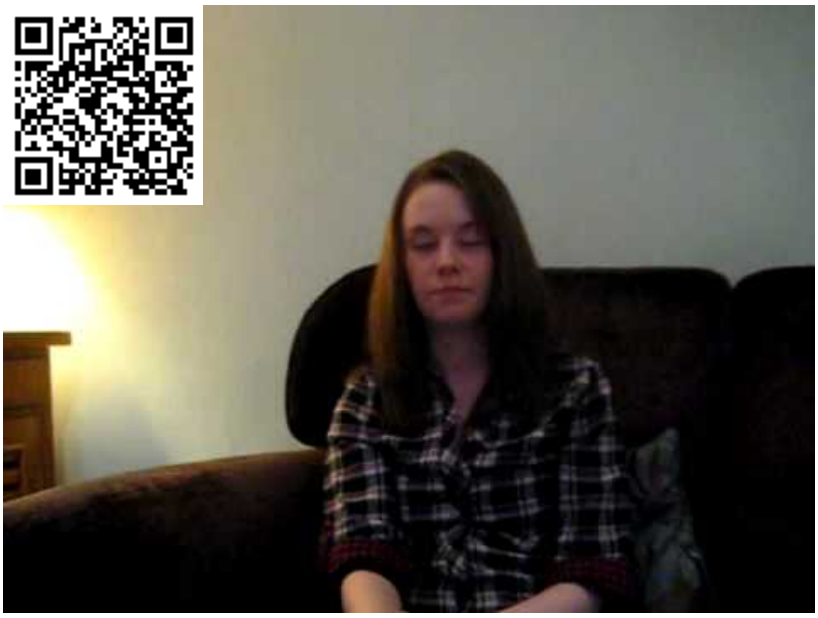
All videos below are captioned.

Wernicke's Aphasia

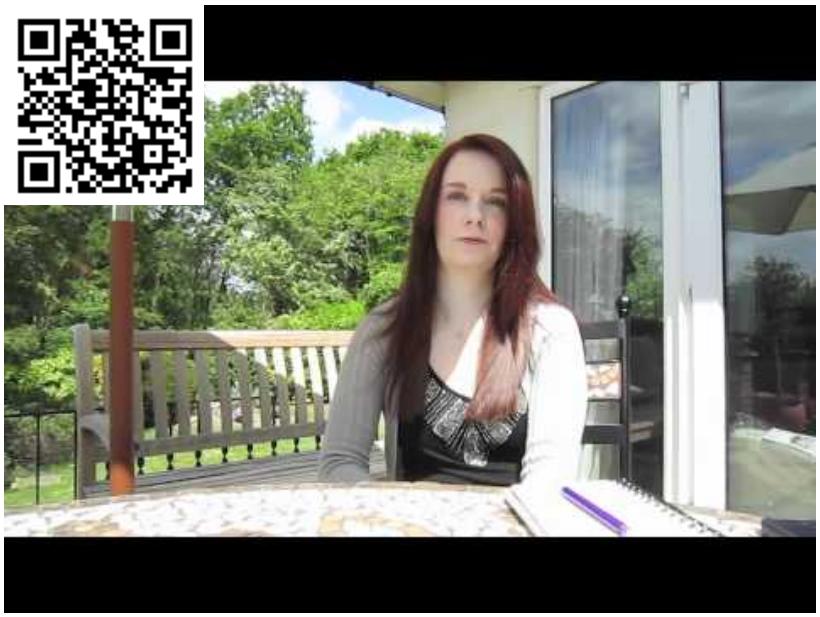


Broca's Aphasia

Sarah Scott, 1 year out from her stroke:



Sarah Scott, 2 years out from her stroke:



Sarah Scott, 5 years out from her stroke:



Bilingual Aphasia



More on Aphasia

Watch this TedEd video in which Susan Wortman-Jutt explains a bit more about the affects of aphasia on the brain. (The video is captioned.)



10.5: Aphasia is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.6: Autonomy of Language

Autonomy of Language, from Sarah Harmon



Video Script

To close out the chapter, the course and the book, let's remind ourselves that language is only one thing that our brain does. Our brain is this multifaceted, amazing, complex combination of neurons and other matter that does so much, and we're still learning about what the brain can do.

For example, through **lateralization studies** and **brain mapping**, we're able to start figuring out exactly how the brain processes various things, including language. Catherine Anderson in an earlier section talked about **PET scans** and **MRIs** and other things, so I'm not going to go into that too much here. Certainly, when we look at **Event-related Brain Potentials** (**ERPs**), we're able to start mapping out how the brain processes language. This is even evolved in the 15 years or so that I've had this slide. Realistically, we use ERPs and many other tests and procedures to describe and capture how the brain processes language and everything else it does. We're able to look at speech versus non-speech stimuli and see how the brain lights up, how it processes, how that light processes through different parts of the brain. We're also able to show pretty distinctively that the left hemisphere is where language is processed.

We should also remind ourselves about **the Critical Age Hypothesis**, and we talked about this with respect the second language acquisition in the previous chapter. We know, and we can show even neurologically, that sometime around puberty is when the elasticity of the brain changes. Now, how it changes is still being understood; we do know that once humans hit right around puberty, they learn by logic; we don't absorb information like a sponge, rather we are processing it based off of previous information that we have learned. Even when you think somebody is learning through osmosis and they're just soaking up information and spitting it out all, they are processing, probably at a faster pace or at a higher rate than we mere mortals, and that they're basing it off of knowledge they've learned previously. We do know that there are cases where, unfortunately, children have been deprived of language; you may read about the case of Genie or Chelsea, or other children who, at birth, their parents or caregivers thought that they were somehow 'not normal'. They therefore deprived these children of everything, including language. These children were stuck in a room and were not talked to, not played with, not interacted with, abused and neglected. In those cases, yes, it is true that the children have exhibited different developmental issues with respect to the language, such that when they are found as teenagers or adults, unfortunately, they do not have the linguistic capabilities that most others do. That being said, their treatment was certainly unethical and inhumane and should never be duplicated. What is more, we don't actually know whether their lack of linguistic inputs affected their linguistic productivity later on, or whether there was something else going on, and most of these cases that we hear about come from prior eras, meaning we didn't have any concept of the neurological processes that we do now, or have had for the last 30 years. What we can say is this: When we get older, we have to rely on making neural connections and learning based off of prior experience. The more prior experience you have, the easier the learning is.

It is also important to bring up when we have cases of children or young adults who seem either to be on the autism spectrum, or have some other developmental capability that precludes them from talking too much, or talking in ways that are not quite as fluent as you would expect for somebody of their age. We definitely have cases and continue to have cases of children who linguistically seem to be okay, but developmentally are behind their cohorts. We have cases with respect to autism spectrum disorder where linguistic development is not on par with other aspects of the neurological development. However, notice that it doesn't necessarily mean that just because you can't speak well or speak at a certain level that you are not intelligent. There is true autonomy, and we can see this in a variety of cases. Numerous folks who are on the autism spectrum show amazing brilliance with respect to everything to math and science to art. They just may not be able to communicate verbally or through sign language what they're actually thinking. That does not mean that language and intelligence are connected; in fact, it shows pretty conclusively they're not.

Finally, there are a couple of examples to bring up with respect to **Specific Language Impairment (SLI)**. Our data on this are changing, and I mean rapidly. Historically, SLI has been something that runs in families. For example, if a four year old says, “Woofwoof run door,” “Give me apple,” or “Man speak yesterday,” you would expect that kind of language out of a two-year-old. Children are frequently told or diagnosed with specific language impairment right around 4-6 years of age. This is changing rapidly; much of what was previously diagnosed as SLI might actually be autism spectrum disorder. In some cases, it may be nothing at all, that maybe at the age of four they're still talking like this, but come age five or six are significantly more verbal and fluid in their speech. Sometimes it could be environmental factors; if you have a child who was in a family that doesn’t talk very much to the child, and they’re allowed to play very much, their linguistic output is going to be reduced compared to someone else in their same age range.

Our understanding of how we process language is changing almost daily; this new information that we keep getting over and over again shows that language and intelligence and any other aspect of cognitive development are separate. You may be highly proficient in math but not be able to tie your shoelaces, or so the story goes with Albert Einstein. The same is true for language. You could be perfectly highly fluent in language but not be able to solve this simple math problem, or be able to understand a piece of art, or be able to sing a song. Language is autonomous from every other cognitive process, and even if there is damage or lesions or other type of malady that affects the brain, and it can affect both language and other aspects of cognitive processing, that's not to say that language can't come back, or any of these other areas of processing can't come back. With the brain, anything seems to be possible.

10.6: Autonomy of Language is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.7: References

List of References

My first ever trapse into neurolinguistics was Rodman and Fromkin's 6th edition of *Introuction to Language*, and little did I know how important Victoria Fromkin was to this particular area. Her UCLA homepage is a list of starter material to first understand what we know of brain development. Sadly, she passed away in 2000, and the world of neurolinguistics has grown exponentially since then. Below are a list of the resources that I've used, and they are good starter material.

One other point of reference: When my husband suffered his stroke and worked through his recovery, it was his neurologists and speech pathologists at the Palo Alto Medical Foundation who helped me flesh out much more information about dysarthria and other aphasias. That list of folks is fairly long, and most are no longer working at PAMF, but I'm immeasurably grateful for their knowledge and guidance.

- Cataldo, Dana Michelle, Andrea Bamberg Migliano, and Lucio Vinicius. 2018. "Speech, stone tool-making, and the evolution of language." In *PLOS One*.
- Jackendoff, Ray. 2011. "What is the Human Language Faculty? Two Views." In *Language* 87 (3): 586-624.
- McMahon, April and Robert McMahon. 2014. *Evolutionary Linguistics*. Cambridge University Press.
- Mesulam, M-Marsel, Cynthia K. Thompson, Sandra Weintraub, and Emily J. Rogalski. 2015. "The Wernicke Conundrum and the Anatomy of Language Comprehension in Primary Progressive Aphasia." In *Brain: A Journal of Neurology* 138 (8): 2423-2437.
- Pagel, Mark. 2017. "Q&A: What is human language, when did it evolve, and why should we care?" In *BioMed Central* 15.
- Vukovic, Nikola, Matteo Feurra, Anna Shpektor, Andriy Myachykov, and Yury Shtyrov. 2017. "Primary Motor Cortex Functionally Contributes to Language Comprehension: An Online rTMS Study." In *Neuropsychologia* 96: 222-229.

10.7: References is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by LibreTexts.

Index

A

acoustic energy
2.5: Sound Classes and Hierarchy
acoustic phonetics
2.2: Articulators and Airstream Mechanisms
affricate
2.3: Voicing, Resonance and Articulation
all languages are equally valid
1.6: Language universals
all languages have grammar
1.6: Language universals
alveolar
2.3: Voicing, Resonance and Articulation
alveolar ridge
2.2: Articulators and Airstream Mechanisms
approximant
2.3: Voicing, Resonance and Articulation
arbitrariness
1.3: Hallmarks of human language
articulation
2.2: Articulators and Airstream Mechanisms
articulators
2.2: Articulators and Airstream Mechanisms
articulatory phonetics
2.2: Articulators and Airstream Mechanisms
articulatory system
0.1: Linguistics is a science
auditory system
0.1: Linguistics is a science

B

bilabial
2.3: Voicing, Resonance and Articulation
broad transcription
2.4: IPA and Charts

C

cluster
2.6: All About Vowels
coda
2.6: All About Vowels
cognitive science
1.1: What is linguistics?
Competence
1.5: Language and thought
consonants
2.5: Sound Classes and Hierarchy
consonats
2.3: Voicing, Resonance and Articulation
creativity
1.3: Hallmarks of human language
cultural transmission
1.3: Hallmarks of human language

D

descriptive approach
0.1: Linguistics is a science
descriptive grammar
1.2: Linguistic definitions
diacritic
2.6: All About Vowels
dialects
1.2: Linguistic definitions
diphthongs
2.6: All About Vowels
displacement
1.3: Hallmarks of human language
duality
1.3: Hallmarks of human language

E

empirical observations
0.1: Linguistics is a science
every language changes over time
1.6: Language universals

F

feature matrix
2.5: Sound Classes and Hierarchy
fingerspelling
1.7: Signed languages
flap
2.3: Voicing, Resonance and Articulation
fricatives
2.3: Voicing, Resonance and Articulation

G

generative
1.3: Hallmarks of human language
glides
2.4: IPA and Charts
2.5: Sound Classes and Hierarchy
glottal
2.3: Voicing, Resonance and Articulation
glottalic sounds
2.2: Articulators and Airstream Mechanisms
glottis
2.2: Articulators and Airstream Mechanisms

H

handshape
1.7: Signed languages
home sign languages
1.7: Signed languages

I

International Phonetic Alphabet
2.4: IPA and Charts
Intonation
2.6: All About Vowels
IPA
2.4: IPA and Charts

K

knowledge
1.5: Language and thought

L

labiodental
2.3: Voicing, Resonance and Articulation
language
1.2: Linguistic definitions
larynx
2.2: Articulators and Airstream Mechanisms
lateral approximant
2.3: Voicing, Resonance and Articulation
length
2.6: All About Vowels
lengthening of vowels
2.6: All About Vowels
lexicon
1.2: Linguistic definitions
linguistic determinism
1.5: Language and thought
linguistic relativism
1.5: Language and thought
linguistic universals
1.5: Language and thought
lips
2.2: Articulators and Airstream Mechanisms
2.3: Voicing, Resonance and Articulation
liquid
2.6: All About Vowels
location
1.7: Signed languages
loudness
2.6: All About Vowels

M

major class features
2.5: Sound Classes and Hierarchy
manner of articulation
2.3: Voicing, Resonance and Articulation
mental grammar
1.1: What is linguistics?
1.3: Hallmarks of human language
morphology
1.1: What is linguistics?
movement
1.7: Signed languages
mutual intelligibility
1.2: Linguistic definitions
mutually intelligible
1.2: Linguistic definitions

N

narrow transcription
2.4: IPA and Charts
nasal
2.3: Voicing, Resonance and Articulation
2.6: All About Vowels
nasal cavity
2.2: Articulators and Airstream Mechanisms
nasalization
2.6: All About Vowels
natural class
2.5: Sound Classes and Hierarchy
neural system
0.1: Linguistics is a science
nucleus
2.5: Sound Classes and Hierarchy
2.6: All About Vowels

O

obstruction
2.3: Voicing, Resonance and Articulation
obstruents
2.5: Sound Classes and Hierarchy
occlusion
2.5: Sound Classes and Hierarchy

onomatopoeia
1.4: Arbitrariness and ongoing changes
onset
2.6: All About Vowels
oralism
1.7: Signed languages
orientation
1.7: Signed languages

P

palatal
2.3: Voicing, Resonance and Articulation
palate
2.2: Articulators and Airstream Mechanisms
parameters
1.7: Signed languages
performance
1.5: Language and thought
pharyngeal
2.3: Voicing, Resonance and Articulation
phonation
2.2: Articulators and Airstream Mechanisms
phonetic transcription
2.4: IPA and Charts
phonetics
1.1: What is linguistics?
phonology
1.1: What is linguistics?
pitch
2.6: All About Vowels
place of articulation
2.3: Voicing, Resonance and Articulation
plosives
2.3: Voicing, Resonance and Articulation
posody
2.6: All About Vowels
prescriptive grammar
1.2: Linguistic definitions
primary sign language
1.7: Signed languages
productivity
1.3: Hallmarks of human language

R

reflexivity
1.3: Hallmarks of human language
respiration
2.2: Articulators and Airstream Mechanisms
rhotic
2.4: IPA and Charts
rhyme
2.6: All About Vowels

S

schwa
2.4: IPA and Charts
semantics
1.1: What is linguistics?
shared system
1.1: What is linguistics?
simple vowels
2.3: Voicing, Resonance and Articulation
sonority
2.5: Sound Classes and Hierarchy
sound classes
2.5: Sound Classes and Hierarchy
sound hierarchy
2.5: Sound Classes and Hierarchy
sound systems
1.1: What is linguistics?
square brackets
2.4: IPA and Charts
stop
2.3: Voicing, Resonance and Articulation
suprasegmental
2.6: All About Vowels
syllabic consonant
2.6: All About Vowels
syllable
2.5: Sound Classes and Hierarchy
syntax
1.1: What is linguistics?

T

teaching grammar
1.2: Linguistic definitions
teeth
2.2: Articulators and Airstream Mechanisms
term
1.2: Linguistic definitions
terminology
1.2: Linguistic definitions
the science of human language
0.1: Linguistics is a science
the science of language
0.1: Linguistics is a science
Thinking for Speaking
1.5: Language and thought
tone
2.6: All About Vowels
trapezoid
2.3: Voicing, Resonance and Articulation
trill
2.3: Voicing, Resonance and Articulation
typology
1.6: Language universals

U

universal grammar
1.6: Language universals
universal properties
1.6: Language universals
uvula
2.2: Articulators and Airstream Mechanisms
uvular
2.3: Voicing, Resonance and Articulation

V

velar
2.3: Voicing, Resonance and Articulation
velaric sounds
2.2: Articulators and Airstream Mechanisms
velum
2.2: Articulators and Airstream Mechanisms
village sign language
1.7: Signed languages
vocal folds
2.2: Articulators and Airstream Mechanisms
voiced sounds
2.2: Articulators and Airstream Mechanisms
voiceless sounds
2.2: Articulators and Airstream Mechanisms
voicing
2.3: Voicing, Resonance and Articulation
vowels
2.3: Voicing, Resonance and Articulation
2.5: Sound Classes and Hierarchy

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: ENGL LING 200: Introduction to Linguistics

Webpages: 107

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC 4.0: 99.1% (106 pages)
- Undeclared: 0.9% (1 page)

By Page

- ENGL LING 200: Introduction to Linguistics - *CC BY-NC 4.0*
  - 0: What is this course about? - *CC BY-NC 4.0*
    - 0.1: Linguistics is a science - *CC BY-NC 4.0*
    - 0.2: Linguistics as a social science - *CC BY-NC 4.0*
    - 0.3: The resources used for this course - *CC BY-NC 4.0*
  - Front Matter - *CC BY-NC 4.0*
    - TitlePage - *CC BY-NC 4.0*
    - InfoPage - *CC BY-NC 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-NC 4.0*
  - 1: Introduction- What is language? - *CC BY-NC 4.0*
    - 1.1: What is linguistics? - *CC BY-NC 4.0*
    - 1.2: Linguistic definitions - *CC BY-NC 4.0*
    - 1.3: Hallmarks of human language - *CC BY-NC 4.0*
    - 1.4: Arbitrariness and ongoing changes - *CC BY-NC 4.0*
    - 1.5: Language and thought - *CC BY-NC 4.0*
    - 1.6: Language universals - *CC BY-NC 4.0*
    - 1.7: Signed languages - *CC BY-NC 4.0*
    - 1.8: Next steps and references - *CC BY-NC 4.0*
  - 2: Sounds, Part 1- Phonetics - *CC BY-NC 4.0*
    - 2.1: Introduction to Phonetics - *CC BY-NC 4.0*
    - 2.2: Articulators and Airstream Mechanisms - *CC BY-NC 4.0*
    - 2.3: Voicing, Resonance and Articulation - *CC BY-NC 4.0*
    - 2.4: IPA and Charts - *CC BY-NC 4.0*
    - 2.5: Sound Classes and Hierarchy - *CC BY-NC 4.0*
    - 2.6: All About Vowels - *CC BY-NC 4.0*
    - 2.7: References - *CC BY-NC 4.0*
  - 3: Sounds, Part 2- Phonology - *CC BY-NC 4.0*
    - 3.1: Comparing Sounds and Distribution - *CC BY-NC 4.0*
    - 3.2: Assimilation and Dissimilation - *CC BY-NC 4.0*
    - 3.3: Other Phonological Rules - *CC BY-NC 4.0*
    - 3.4: Analyzing Phonological Data - *CC BY-NC 4.0*
    - 3.5: References - *CC BY-NC 4.0*
  - 4: Words- Morphology - *CC BY-NC 4.0*
    - 1: Back to the Arbitrary - *CC BY-NC 4.0*
    - 2: Morphological Definitions - *CC BY-NC 4.0*
    - 3: Morphemes - *CC BY-NC 4.0*
    - 4: Affixation and Other Morphological Processes - *CC BY-NC 4.0*
    - 5: Morpho-phonology - *CC BY-NC 4.0*
    - 6: Parts of Speech and Word Formation - *CC BY-NC 4.0*
    - 7: Creation of New Lexicon - *CC BY-NC 4.0*
    - 8: Morpho-syntax - *CC BY-NC 4.0*
    - 9: Analyzing Morphological Data - *CC BY-NC 4.0*
    - 10: References - *CC BY-NC 4.0*
  - 5: Phrases- Syntax - *CC BY-NC 4.0*
    - 5.1: Syntax Terminology - *CC BY-NC 4.0*
    - 5.2: Word Order and Lexical Categories - *CC BY-NC 4.0*
    - 5.3: Phrase Structure Rules, X-Bar Theory, and Constituency - *CC BY-NC 4.0*
    - 5.4: Analyzing a PS Tree - *CC BY-NC 4.0*
    - 5.5: Grammatical Dependencies - *CC BY-NC 4.0*
    - 5.6: Grammaticality - *CC BY-NC 4.0*
    - 5.7: References - *CC BY-NC 4.0*
  - 6: Meaning- Semantics and Pragmatics - *CC BY-NC 4.0*
    - 6.1: Arbitrariness and Compositionality - *CC BY-NC 4.0*
    - 6.2: Lexical Semantics - *CC BY-NC 4.0*
    - 6.3: Phrasal Semantics - *CC BY-NC 4.0*
    - 6.4: Semantic Features - *CC BY-NC 4.0*
    - 6.5: Pragmatics Definitions - *CC BY-NC 4.0*
    - 6.6: Grice and the Cooperative Principle - *CC BY-NC 4.0*
    - 6.7: Reading Between the Lines - *CC BY-NC 4.0*
    - 6.8: Deixis - *CC BY-NC 4.0*

- 6.9: References - *CC BY-NC 4.0*

- 7: Language and Society- Sociolinguistics - *CC BY-NC 4.0*
  - 1: Sociolinguistics definitions - *CC BY-NC 4.0*
  - 2: Geographic dialects - *CC BY-NC 4.0*
  - 3: Appalachian American English (AE) - *CC BY-NC 4.0*
  - 4: African American English (AAE) - *CC BY-NC 4.0*
  - 5: Other Social Dialects - *CC BY-NC 4.0*
  - 6: Gender and Class Dialects - *CC BY-NC 4.0*
  - 7: Prestige and Politics - *CC BY-NC 4.0*
  - 8: Social Norms - *CC BY-NC 4.0*
  - 9: Languages in Contact - *CC BY-NC 4.0*
  - 10: Indigenous North American Languages (Optional) - *CC BY-NC 4.0*
  - 11: References - *CC BY-NC 4.0*
- 8: Language Change- Historical Linguistics - *CC BY-NC 4.0*
  - 1: Language Change Facts and Definitions - *CC BY-NC 4.0*
  - 2: Reconstructions and Analysis - *CC BY-NC 4.0*
  - 3: Indo-European and Romance Phonological Reconstruction - *CC BY-NC 4.0*
  - 4: Stages of English - *CC BY-NC 4.0*
  - 5: Morphological Stages - *CC BY-NC 4.0*
  - 6: Syntactic and Lexical Changes - *CC BY-NC 4.0*
  - 7: Typology - *CC BY-NC 4.0*
  - 8: Language Families - *CC BY-NC 4.0*
  - 9: Writing Systems - *CC BY-NC 4.0*
  - 10: Arabic and Romance- A Comparison (Optional) - *CC BY-NC 4.0*
  - 11: References - *CC BY-NC 4.0*
- 9: Learning Languages- Language Acquisition - *CC BY-NC 4.0*
  - 9.1: Universals and Definitions - *CC BY-NC 4.0*
  - 9.2: Child Language Acquisition Theories - *CC BY-NC 4.0*
  - 9.3: Child Language Acquisitions Stages - *CC BY-NC 4.0*
  - 9.4: Child Language Acquisition Linguistics - *CC BY-NC 4.0*
  - 9.5: Child Language Acquisition Bilingualism - *CC BY-NC 4.0*
  - 9.6: Second Language Acquisitions Theories and Components - *CC BY-NC 4.0*
  - 9.7: Second Language Acquisition Teaching Methods - *CC BY-NC 4.0*
  - 9.8: References - *CC BY-NC 4.0*
- 10: Language Processing and the Human Brain- Cognitive Linguistics - *CC BY-NC 4.0*
  - 10.1: Human Language versus Animal Communication - *CC BY-NC 4.0*
  - 10.2: Evolution of Human Language - *CC BY-NC 4.0*
  - 10.3: Language Processing and Parsing - *CC BY-NC 4.0*
  - 10.4: Language Processing in the Brain - *CC BY-NC 4.0*
  - 10.5: Aphasia - *CC BY-NC 4.0*
  - 10.6: Autonomy of Language - *CC BY-NC 4.0*
  - 10.7: References - *CC BY-NC 4.0*
- Back Matter - *CC BY-NC 4.0*
  - Index - *CC BY-NC 4.0*
  - Glossary - *CC BY-NC 4.0*
  - Detailed Licensing - *CC BY-NC 4.0*