

LING 1010



Language and Mind

Prof. Jon Sprouse

03.03.21:

The acquisition of phonemes

Nativism vs Empiricism

The logical problem of language acquisition tends to lead to a debate between two positions. Both of which provide a solution to the problem:

Modern Nativism



Substantial innate knowledge



Input/experience still plays a role, but less than the role it plays in empiricism.

In short, children come to the problem with a lot of genetic help, and then use experience to hone in on the correct answer.

Modern Empiricism



Minimal innate knowledge



Input/experience plays the largest role in learning

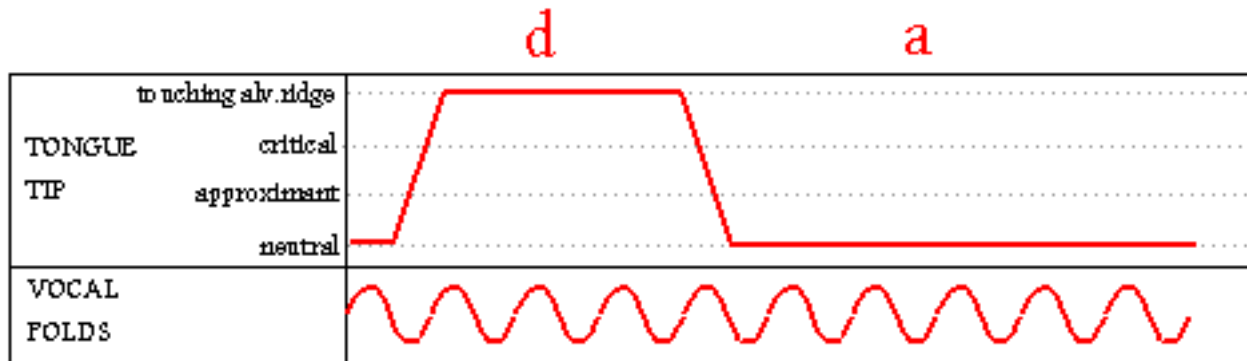
In short, children come to the table with the ability to learn from experience, and use experience to build up all of the complexity of language.

Step 1: Let's look at what
needs to be learned

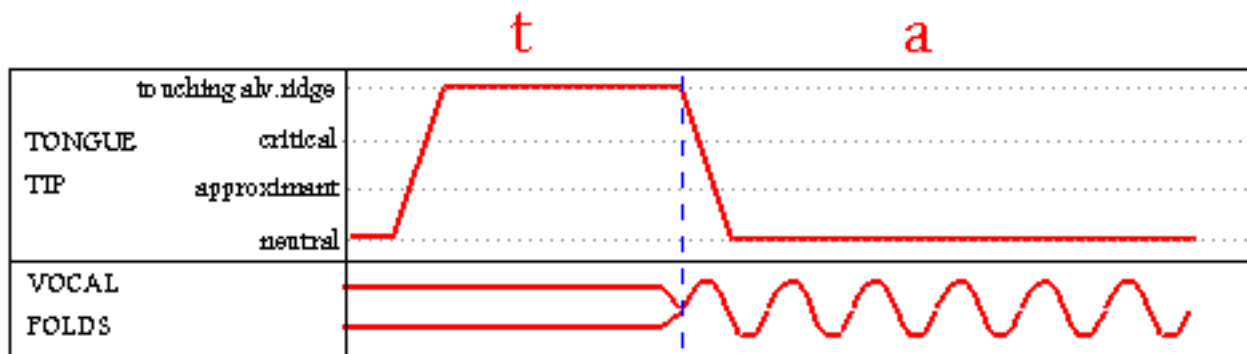
Sounds form a continuum,
and children must learn how
to break it into categories

The continuum of voice onset time

Voiced stop: the vocal folds start vibrating at the same time the air constriction begins



Voiceless stop: the vocal folds start vibrating as the air constriction ends

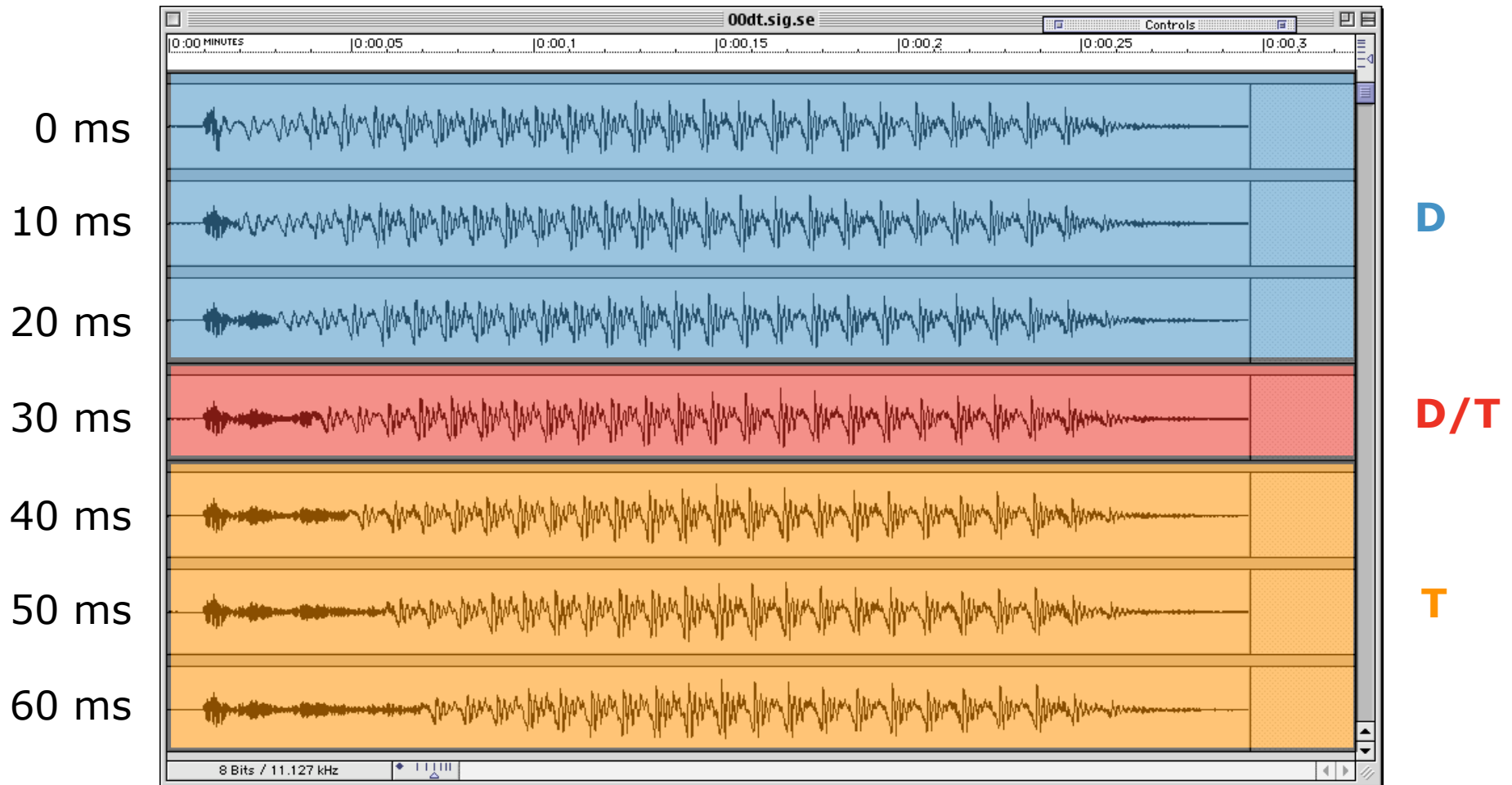


Voice Onset Time is continuous because time is continuous!

Voice Onset Time: A measure of the time between the start of an utterance and the start of the vocal fold vibration

The continuum of voice onset time

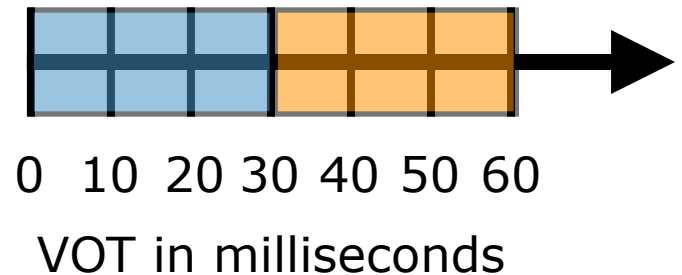
We can create examples of VOT along the continuum. Here are 7 examples, ranging from 0ms VOT to 60ms VOT:



Even though VOT is a continuum, we perceive it as two categories: D and T.

Making categories out of the continuum of voice onset time

When children learn language, they have to learn to break the continuum of VOT into two categories, **D** and **T**. This is the same idea of category that we've seen before - sounds within a category are treated as identical even though they differ in details.



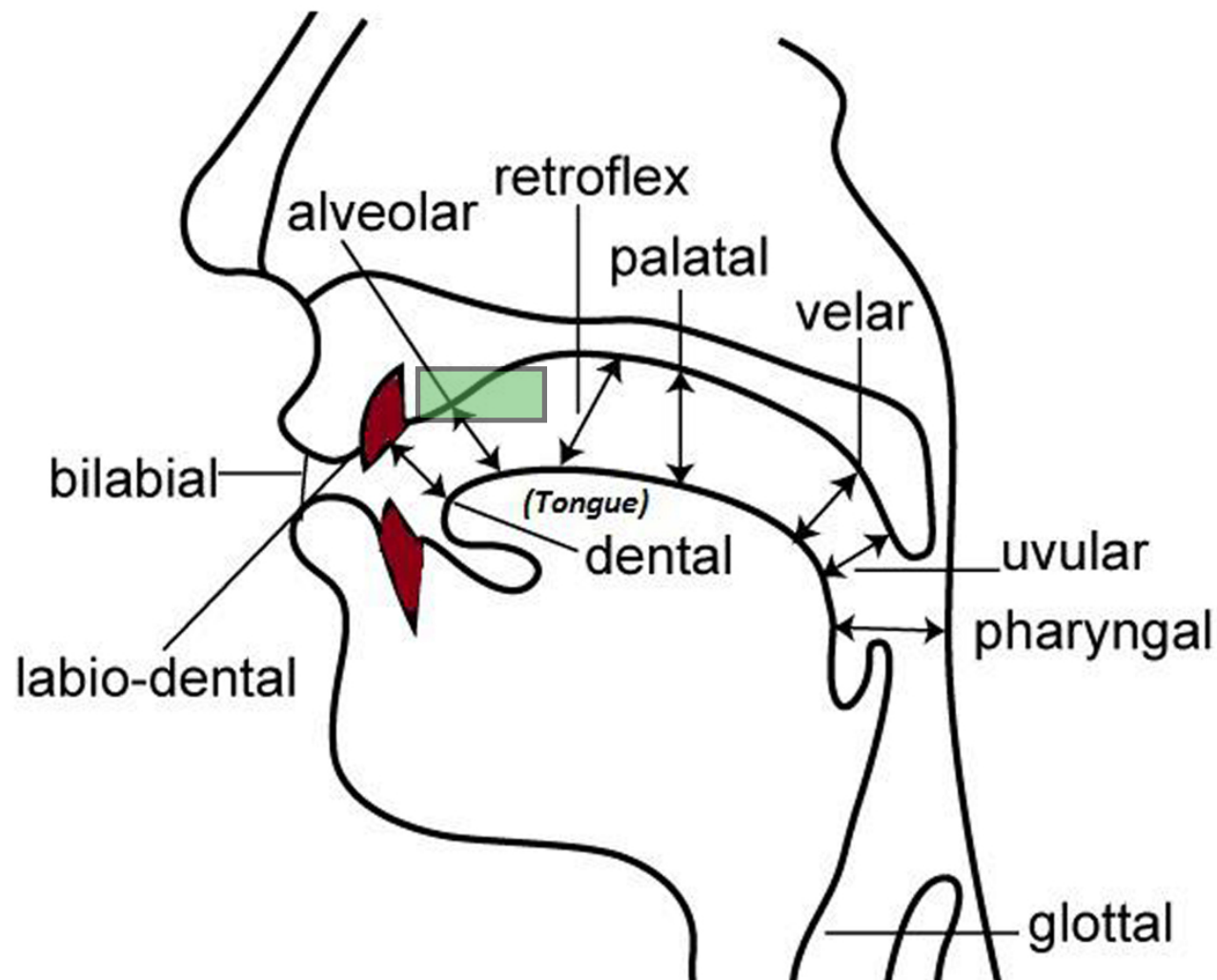
And just to be absolutely clear, the fact that you perceive the low VOT sounds as similar to each other, and the high VOT sounds as similar to each other, has nothing to do with them having similar VOTs. We can show this by creating pairs of sounds that differ by the same VOT, for example, 20ms.

- | | | |
|-----------------------|------------------|-----------|
| 1. 0 ms versus 20 ms | 20 ms difference | SAME |
| 2. 20 ms versus 40 ms | 20 ms difference | DIFFERENT |
| 3. 40 ms versus 60 ms | 20 ms difference | SAME |

The continuum of place of articulation

Just like voicing is continuous because time is continuous, place of articulation is continuous because physical space is continuous.

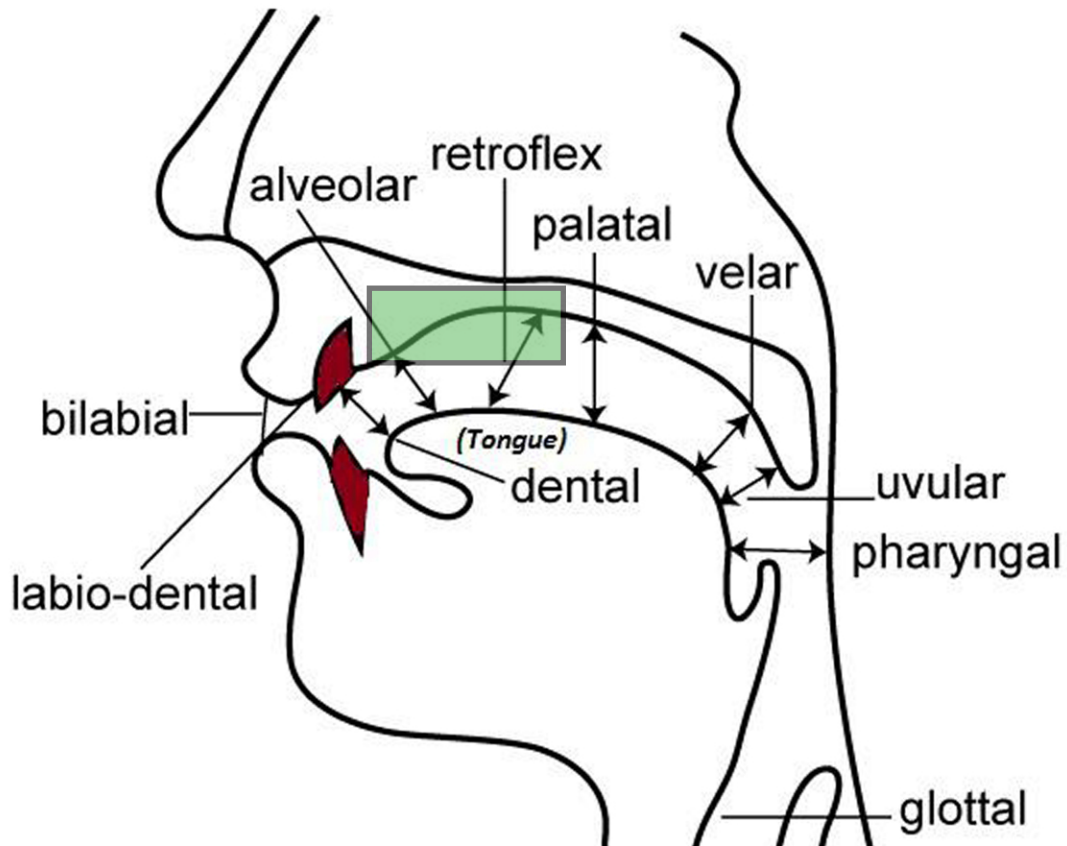
We talk about locations like “alveolar”, but really it is a **region of space**.



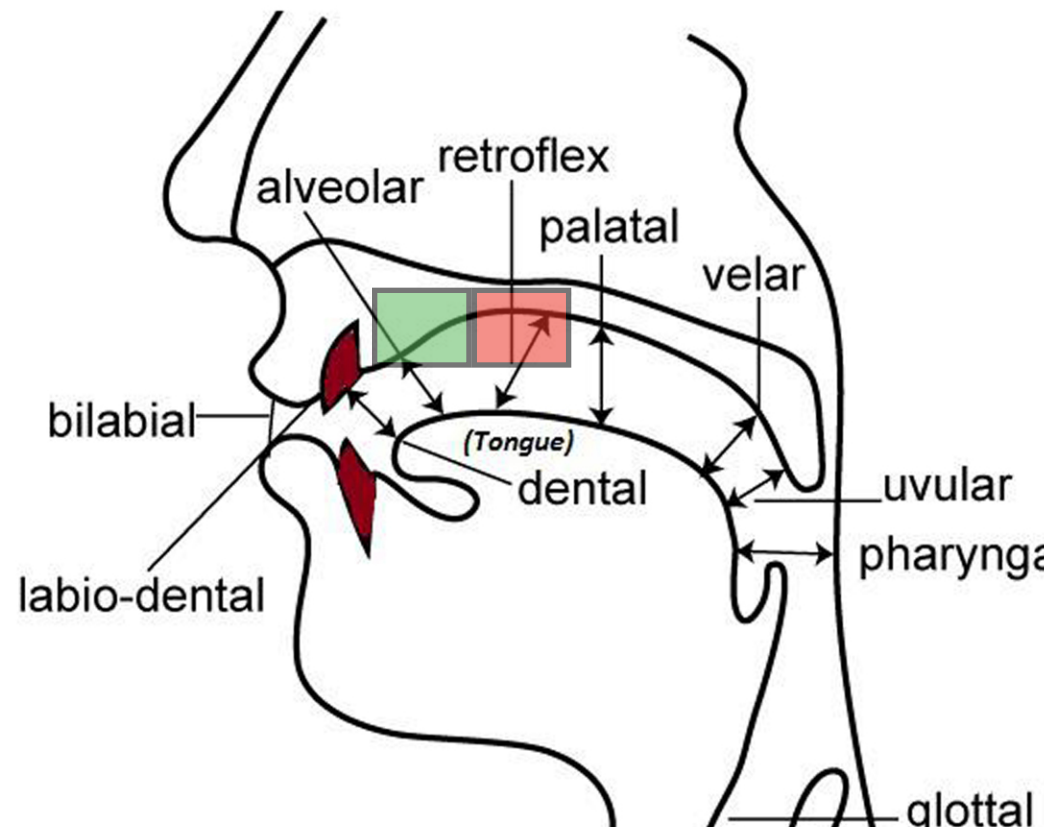
The continuum of place of articulation

And just like voicing, children learning language must learn to break the continuum into categories.

In English, children learn to make no distinction between alveolar and retroflex places of articulation. For example, every voiced stop in this region is perceived as **d**.



In Hindi, children learn to make a distinction between alveolar and retroflex places of articulation. Hindi speakers have both **alveolar d** and **retroflex d**.

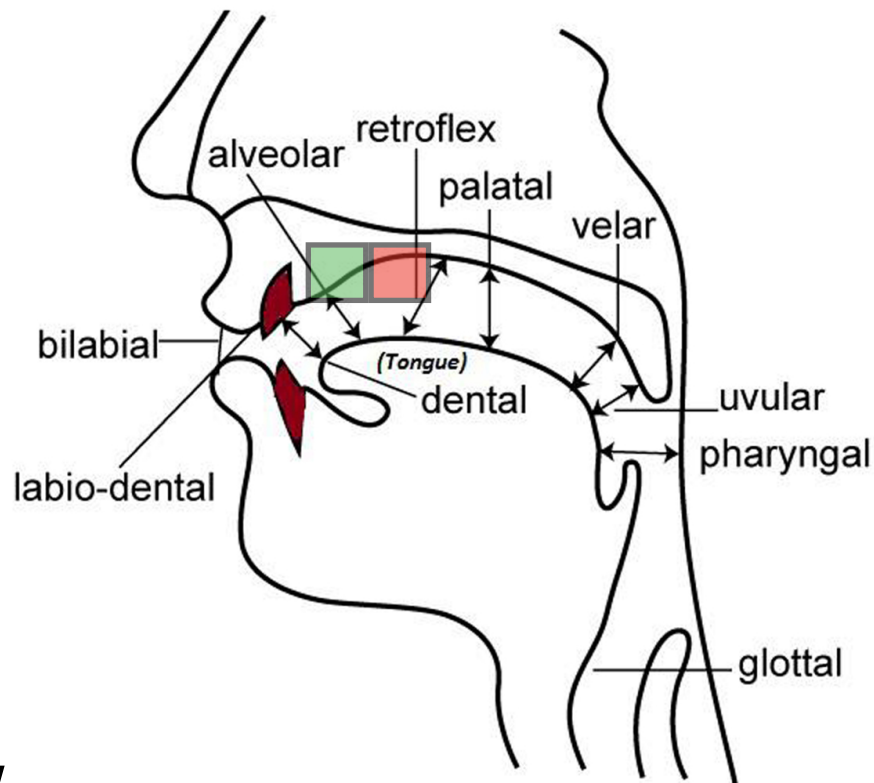


The continuum of place of articulation

Here are eight examples that span the continuum from pure **alveolar d** to pure **retroflex d**.

/d/

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.



/d/

For comparison, here is **d** and **t** again.

/d/

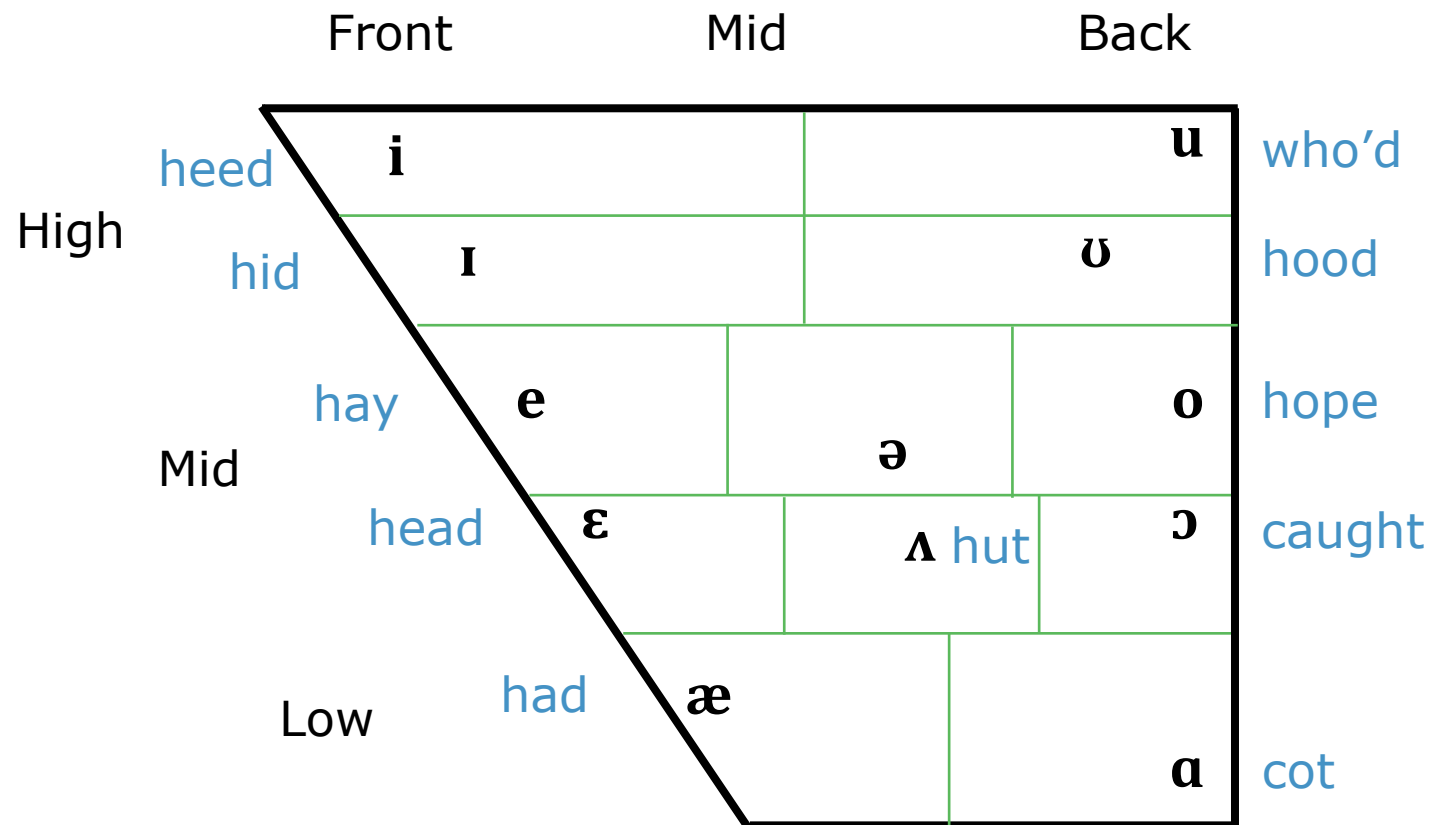
- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.

/t/

The continuum of vowel space

As we learned, vowels are determined by the physical location of the tongue in the mouth. This space is continuous.

When children learn the vowels in their language, they need to learn the **breaks** to make in the space.

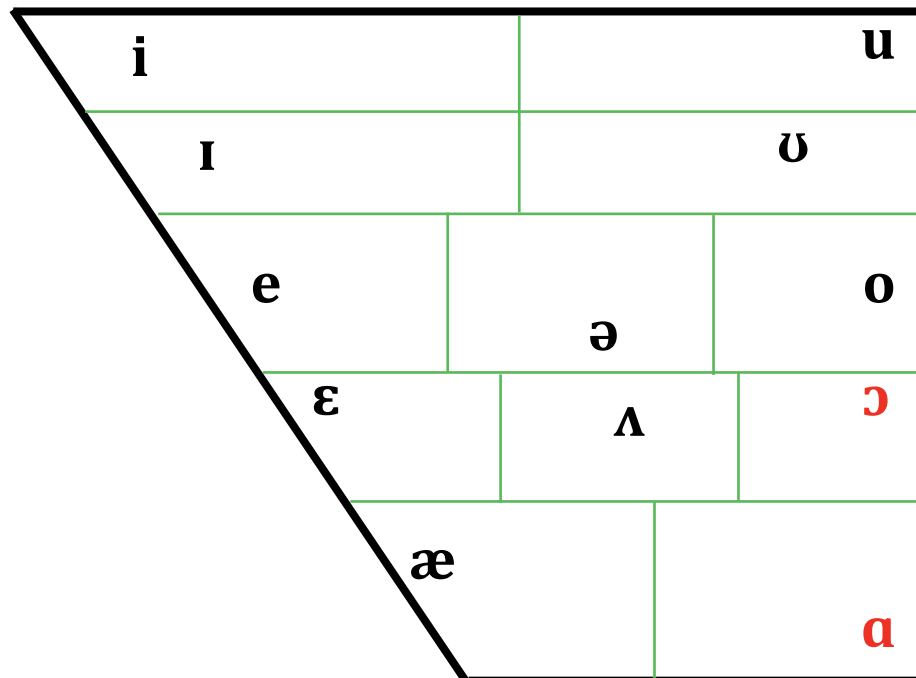


A cool example: the cot-caught merger

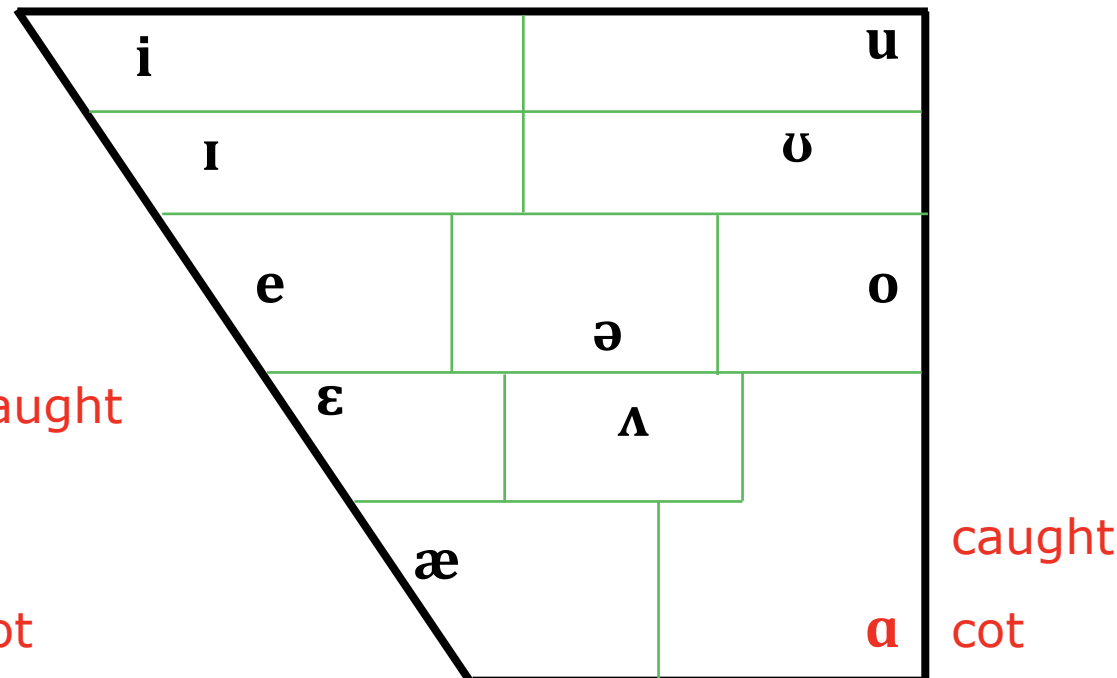
Languages change over time. One way in which they change is by adding or eliminating vowels... in spatial terms, this means adding/eliminating boundaries between areas of the vowel space.

Currently, various types of English are undergoing (or have completed) a merger between the vowels in **cot** and **caught**.

unmerged dialect

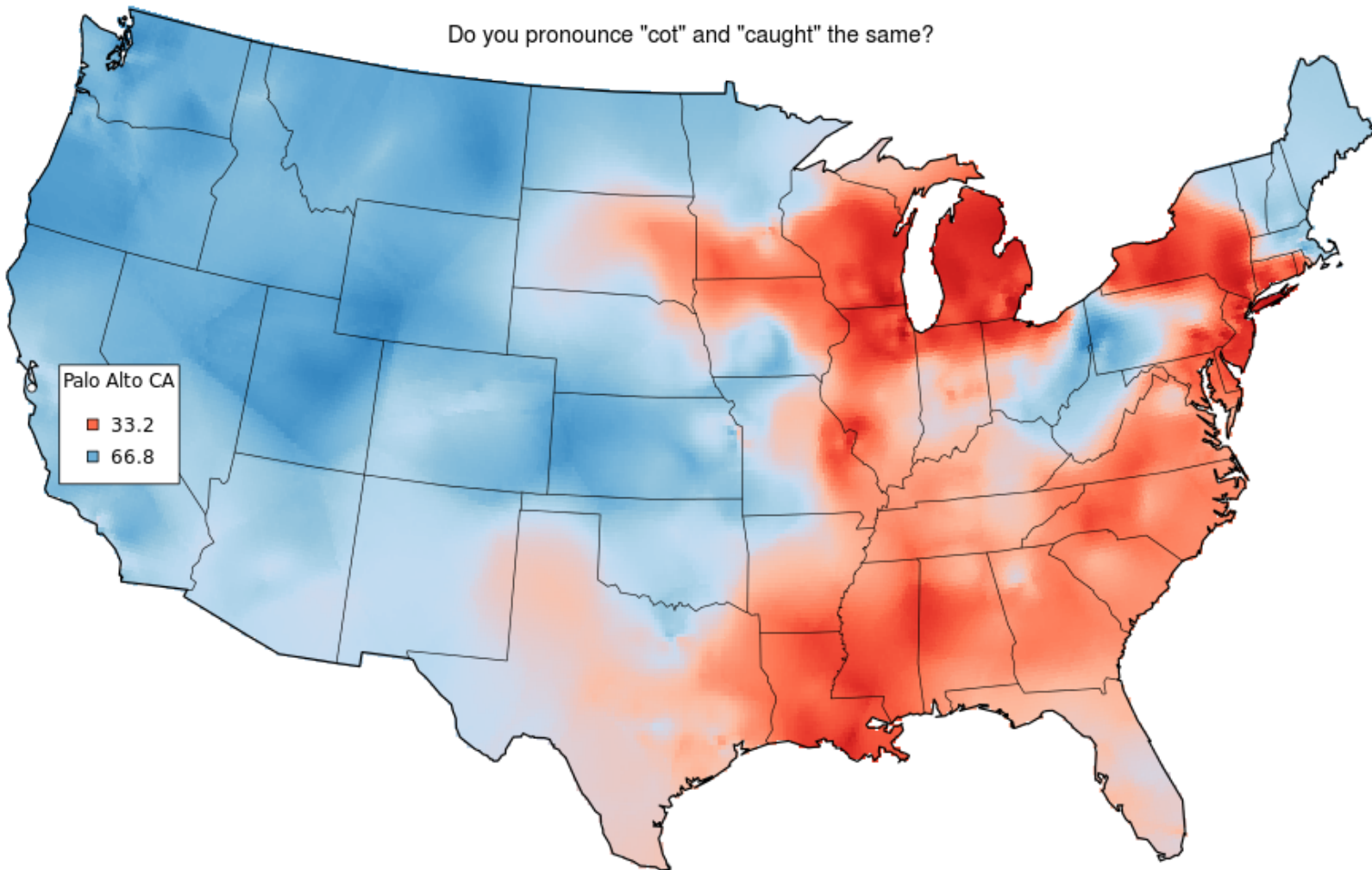


merged dialect



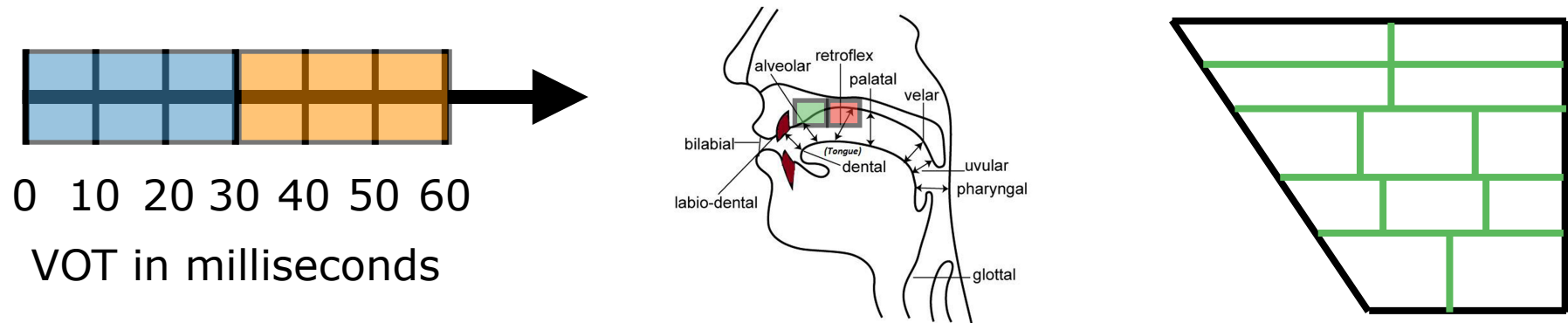
A cool example: the cot-caught merger

As you can see in the map below, Northern New England has already **merged**. Connecticut speakers are still maintaining the **distinction**, but some speakers are in transition...



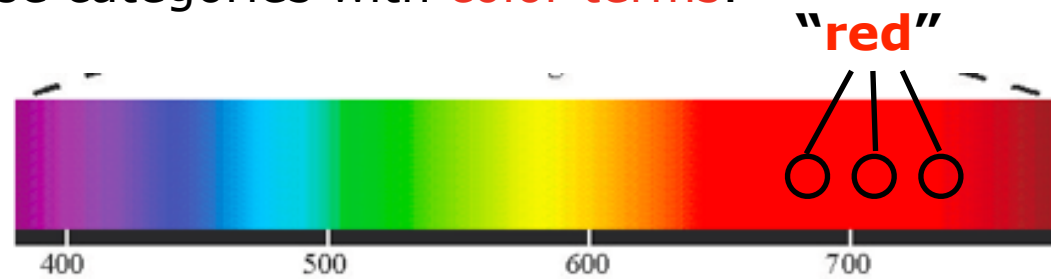
Categorical Perception

Whether it is VOT, place of articulation, or vowel height/backness, the story is the same: we take a continuum, and carve it into categories.



We call this **categorical perception**. It is the act of perceiving physically distinct stimuli as identical for a specific purpose.

The idea of categorical perception isn't new to you. We do it all the time, but probably never talk about. Take **color** as an example. Color is continuum (of wavelengths of light). We know there are infinitely many colors along the continuum, but we treat certain colors as coming from the **same category**. And then we label those categories with **color terms**.



Step 2: Let's look at how
children actually learn it (no
theories yet, just facts)

How children seem to do it

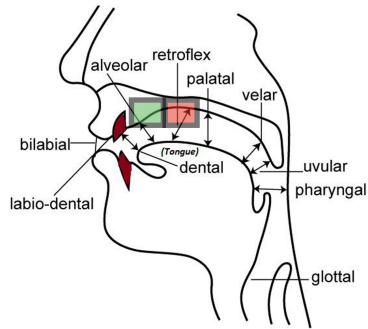
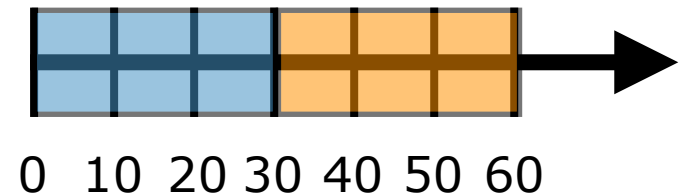
1. Children appear to be born with the ability to discriminate every sound difference found in human languages. They are **universal listeners**.
2. The adults speaking the language around them will **tend to produce sounds that are in the middle** of that language's categories. Children are able to notice this.
3. Over time, children lose the ability to discriminate speech sounds that aren't in the language being spoken around them. Only the category boundaries in their language remains.
4. When this process is finished, the children are just like the adults: they can only discriminate between sounds that are in their language.

This whole process takes about 10-12 months. By the time infants are 10-12 months old, they have learned the speech sounds of their language.

Universal listeners and learning

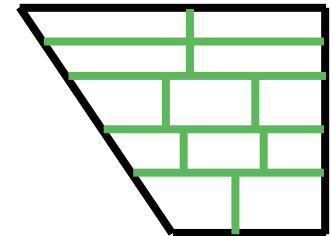
When we say that children are **universal listeners**, what we mean is that they seem to begin with **all possible category boundaries already in place**.

For VOT, this means that they already have the **d/t** boundary in place.



For place of articulation, this means that they already have the **alveolar d** / **retroflex d** boundary in place.

For vowels, this means that they already have the **cot** / **caught** boundary in place.



The act of learning the phonemes of language is the act of **“losing”** the boundaries that the child doesn’t need for their specific language.

How did we figure all of this out?

If you have ever played with an infant from 0-12 months, you know that they aren't really doing much that looks like language. You certainly can't ask them whether they can discriminate different sounds. So how did we figure all of this out?

The answer is something called the **Conditioned Head-Turn Procedure**.

Here is Janet Werker explaining the task. She is the one who first determined that children lose the ability to discriminate sounds by 10-12 months of age!

https://kaltura.uconn.edu/media/infant+speech+perception.mp4/_g291ru68

As you can see, the Conditioned Head-Turn procedure takes advantage of children's desire to see novel fun things, and their ability to turn their head. The idea is that we can train them to expect a novel fun thing after a change in the (boring!) sound being played in the background. If they can hear the difference in the sound being played, they turn their head to look for the fun thing. If they can't hear the difference, they don't turn their head!

The other side of the coin: babbling

Experimental procedures like the conditioned head-turn procedure let us see what children can do during **language comprehension**. But what about **language production**?

It turns out that the two abilities track each other fairly well during the first 12 months of life!

Age

- 6 months:** Babbling begins. Babbling at this age tends to be repetitive (ba ba ba ba ba), and does not necessarily correspond to the language being spoken by adults!
- 6-10 months:** Over time, babbling starts to show variability (ba bi da di do), and slowly starts to take on more and more characteristics of the language being spoken by adults.
- 10-12 months:** The sounds created during babbling only come from the adult language. This is the last babbling stage before true words are spoken (around 12 months).

And here are two twin boys babbling with each other:

https://www.youtube.com/watch?v=_JmA2CIUvUY&feature=related

Step 3: Let's compare
Nativism and Empiricism

Phonemes: Nativism vs Empiricism

Both theories can explain the acquisition of phonemes. This is because both theories allow for innate knowledge (the ability to be a universal listener), and both theories allow for experience to play a role (the loss of universal listening).

Modern Nativism



Substantial innate knowledge



Input/experience still plays a role, but less than the role it plays in empiricism.

Modern Empiricism



Minimal innate knowledge



Input/experience plays the largest role in learning

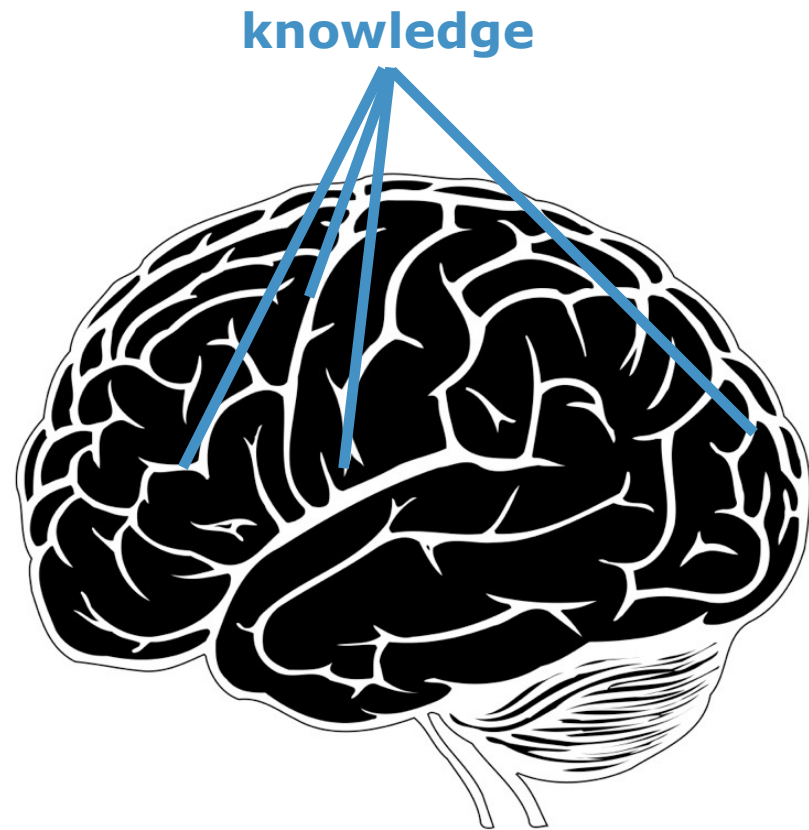
So the question is whether we can make a finer-grained distinction between these two theories, and see which better explains the phoneme facts.

Domain-General and Domain-Specific

Domain-general knowledge is knowledge that is used by multiple cognitive abilities.

We can minimize the amount of innate knowledge necessary by relying solely on domain-general knowledge. One piece of knowledge can assist multiple cognitive abilities.

In cases where **Empiricism** needs to posit innate knowledge, like the universal listener aspect of phoneme learning, it tends to posit innate domain-general knowledge so as to minimize the amount of knowledge being postulated.

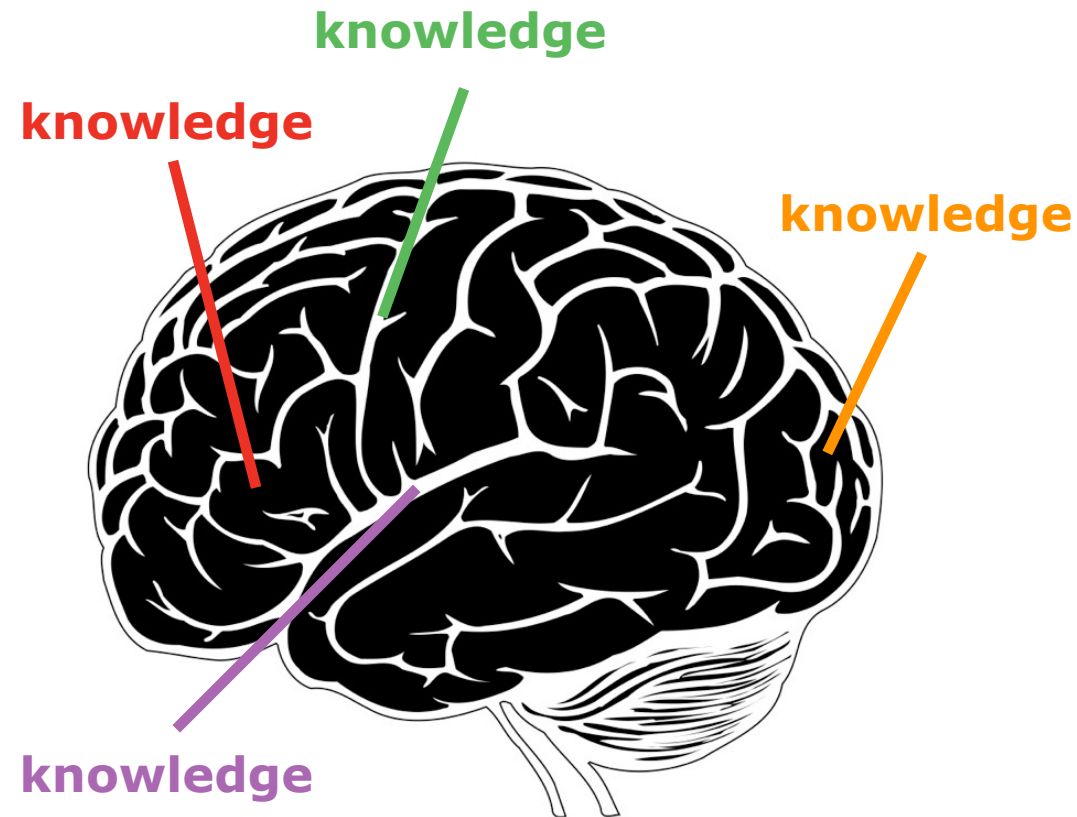


Domain-General and Domain-Specific

Domain-specific knowledge is knowledge that is used by one cognitive ability.

Innate domain-specific knowledge tends to increase the amount of innate knowledge necessary to solve problems, because every cognitive ability will require its own innate knowledge.

Nativists don't mind this state of affairs. They are willing to posit innate domain-specific knowledge if it appears to be necessary to solve the problem.



Nativism vs Empiricism

From this discussion we can update the properties of our two theories to include their positions about the domain-specificity of innate knowledge:

Modern Nativism



Substantial innate knowledge



Input/experience still plays a role, but less than the role it plays in empiricism.



The innate knowledge can be domain-specific.

Modern Empiricism



Minimal innate knowledge



Input/experience plays the largest role in learning



If there is innate knowledge, it is domain-general

An open question about Domain-Specificity

We've already established the fact that adults **can't distinguish sounds** from other languages.

But here is an interesting fact: they can distinguish sounds from other languages **if you tell them that they are listening to something that is not language!**

For example, Janet Werker (Werker and Tees 1984) showed that adults could distinguish two sounds not in their language if you told them that they were listening to water dropping into a bucket. (You ask them to tell you when it sounds like the bucket changed size).

This looks like the change in ability is **domain-specific** (specific to language).



The open question is whether the specificity of the learning change requires domain-specific knowledge or not. We don't know... but it is interesting!

Nativism vs Empiricism

Just to recap, here are the properties of Nativism and Empiricism, including their stances on the type of innate knowledge that is possible.

Modern Nativism



Substantial innate knowledge



Input/experience still plays a role, but less than the role it plays in empiricism.



The innate knowledge can be domain-specific.

Modern Empiricism



Minimal innate knowledge



Input/experience plays the largest role in learning



If there is innate knowledge, it is domain-general

Some Conclusions

The physical properties of phonemes, like their VOT, place of articulation, or vowel height/backness, are **continuous**. This means two things:

Adults must treat the different sounds as categories (**categorical perception**).

Children must **learn the boundaries** that form those categories.

Children appear to be born with the ability to discriminate every sound difference found in human languages. They are **universal listeners**.

By the age of 10-12 months, children **lose the ability to discriminate speech** sounds that aren't in the language being spoken around them. This coincides with changes in their babbling (they only babble sounds in their own language).

Domain-general knowledge is knowledge that is used by multiple cognitive abilities. If empiricists require innate knowledge, it will be domain-general.

Domain-specific knowledge is knowledge that is used by one cognitive ability. Nativists allow for innate domain-specific knowledge because they don't mind multiple types of innate knowledge.