03.06.19: The acquisition of morphology
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 puzzles in learning words/morphemes

There are tons of open research questions when it comes to learning words. Today we are going to talk about three of the most basic questions.

1. How do children segment the speech stream into words?
2. How do children learn what a given word means?
3. How do children learn morphological rules?

(sound)  [k æ t]  (meaning)  [m]ah[tummy]  has a [happy]

lexical entries

morphological rules

un  ADJ

lock  able
The timeline we are talking about

Just to orient today’s discussion relative to last time, we can look at the timeline of word/morphology acquisition:

- Babbling begins at 6mo, becomes variable and language-specific by 12 mo.
- First words are produced between 10-15 mo.
- For many children, word learning accelerates dramatically around 18 mo. This is called the **vocabulary explosion**.
- Complex morphology appears on words.
- Two word utterances, function words and longer utterances.
Word Segmentation
The word segmentation problem

You may recall from the first section of class that we learned that there is no obvious way to identify individual speech sounds in a stream of speech.

Well, this problem scales up to words too. The stream of speech is a continuous modulation of amplitude and frequency. There are no obvious breaks in the physical signal that correspond to breaks between words.

The word segmentation problem is the fact that children must somehow decide where the breaks are between words in the speech stream, despite the fact that there are no physical breaks in the stream (i.e., they must segment the speech stream into words)
Don’t be fooled by your adult processing ability!

When you hear speech, you feel like there are individual words in the speech stream. Don’t be fooled by this!

Adults have already learned words. So you can use this knowledge to help you segment novel speech streams.

To get an idea of the problem that children face, try listing to a language that you don’t know, and figuring out where the word boundaries are!

Or just look at a waveform without the words superimposed, and try to guess where the word boundaries are:
There are segmentation errors

We can also see that the word segmentation problem is real by the fact that people make word segmentation errors:

**Examples from daily speech:**

The sky is falling.
This guy is falling.

The white house is under attack.
The white house is under a tack.

**Examples from history:**

a napron → an apron

an other → a nother

The n is reanalyzed as being part of the determiner!

That is a whole nother thing!
Transitional Probabilities

One way children might solve the problem is to track how often each sound follows other sounds.

We call this the **transitional probability** - it is the probability of transitioning from one specific sound (e.g. s) to the letter that comes after it (e.g. j).

It is really easy to calculate (on a computer). You simply find every instance of a sound (e.g. s) in a corpus, and then look at the sound that comes after it each time. You then pick one sound (e.g. j), and divide the number of times you see j after s by the number of times s appears:

\[
\text{transitional probability (s j)} = \frac{\text{# of j’s following s’s}}{\text{# of s’s}}
\]

(For the mathematically minded, this is the same thing as a conditional probability. It is the probability of j given that you already saw a s, or \(p(j|s)\).)
How do transitional probabilities help?

The idea behind transitional probabilities is that sounds that appear next to each other inside of a word will be more frequent than sounds that do not appear next to each other in a word.

The reason this will be true is that the sequence of sounds inside a word will be spoken every time the word is spoken. Sequences of sounds outside of words will only be spoken when those two words are randomly stuck next to each other.

So, children can use the transitional probabilities of sounds to guess word boundaries:

- Word internal sequences will have high transitional probability
- Word boundaries will have low transitional probability
Children need more

It turns out that transitional probabilities alone are not quite enough to solve the word segmentation problem.

This is an open area of research, but current research suggests that children use multiple sources of information to identify word breaks:

1. Children may use the transitional probability between phonemes.

   This is just a schematic

2. Children may also use the transitional probability between syllables because most low frequency transitions happen across word boundaries:

   This is just a schematic

3. Children may use the fact that words tend to have one primary stress as a way to help identify separations between words:

   permit permit
Word Segmentation: Nativism vs Empiricism

So now we can look at the knowledge that children must have in order to solve the word segmentation problem, and ask which theory it fits with:

**Modern Nativism**
- The ability to track transitional probabilities
- Knowledge of syllable boundaries
- Knowledge that words only have one primary stress

**Modern Empiricism**
- The ability to track transitional probabilities
Learning word meanings
Nouns first

Very young children (<18 mo) are often described as having a noun bias. The first 50 or so words that they learn tend to be nouns: names for people they are around (mama, dada, etc), the food they eat, body parts, clothing, animals/pets, toys, etc.

Here is a wordle that a mother made of her child’s first words. Size indicates the order in which they were learned (based on her hearing the child produce them).

A video of first words:
https://www.youtube.com/watch?v=dlqq4-fRIdY
http://nipitinthebud.wordpress.com/2013/04/19/first-words-wordles/
Why are nouns first?

Most likely this has to do with the salience of nouns in scenes. The actors and objects in a scene are much easier to figure out (without language) than the actions themselves.

This is even true for adults:

Snedeker et al. 1999 showed adults clips of mothers playing with their children. The videos were silent, except for a beep when the mystery word was spoken by the mother.

They asked the adults to guess what the word was that the mother said.

Nouns were much easier to guess than verbs, even for adults!
Verbs second

When the vocabulary explosion hits, children begin to learn concrete verbs (by concrete we mean actions that they observe like kick, as opposed to abstract ideas like think). They also start to put them together into two word utterances.

One interesting aspect of the timing (nouns first, verbs second) is that once nouns are learned, children can use the nouns to help them learn the verbs!

Snedeker et al. 2002 used the same idea as before (silent clips with beeps where the words go), but added nouns to the clips.

They found that, yes, adding the nouns makes adults slightly better at guessing the verb. They still aren’t great, but it suggests that some verbs could be learned this way.
Abstract nouns (idea, death), abstract verbs (think, die), and function words like prepositions (near, behind, etc) all tend to appear after concrete nouns and verbs.

Again, the timing might be telling us something. For example, once nouns and basic verbs are learned, it is possible to use them to learn a basic grammar, and start figuring out what the more abstract words are.

For example, if the word appears in a syntactic verb frame (remember those!), the child would know it is a verb.

If the word shows up after verbs and before nouns, the child might figure out that it is a preposition.

Snedeker et al. 2002 showed that syntactic frames help adults figure out words!
Word learning: Nativism vs Empiricism

So now we can look at the knowledge that children must have in order to solve the word segmentation problem, and ask which theory it fits with:

**Modern Nativism**
- The ability to identify actors/objects
- The ability to identify verbs after actors/objects have been identified
- The ability to use syntactic frames to learn abstract words

**Modern Empiricism**
- The ability to identify actors/objects
- The ability to identify verbs after actors/objects have been identified
Learning morphological rules
As you know, the past tense in English is typically formed by adding the suffix -ed to a verb:

- talk \rightarrow talked
- jump \rightarrow jumped
- kick \rightarrow kicked

This looks like a rule:

\[ V + \text{ed} \rightarrow \text{past} \]

But this isn’t true for all verbs. There are irregular verbs in English that form the past tense some other way:

- sing \rightarrow sang
- go \rightarrow went
- think \rightarrow thought

These look like **exceptions** to the past tense rule above.
Children and rule learning

One of the most interesting aspects about morphological rule learning in children is that their performance for irregulars follows a “u-shaped” pattern:

U-shaped curves like this are incredibly valuable in theories of child development, as they show **two changes in mental ability**: the change from the first part of the U to the minima, then the change to the second part of the U.

**A video with examples:**
https://www.youtube.com/watch?v=2lyNB-HE0yY
So what is going on?

The U-shape suggests a timeline for the learning of the rule:

Rule learning requires the child to **generalize**. They see several verbs all ending in -ed, and all meaning “past”. They must generalize this into a **rule**.

At the red stage, we say that the child is **overgeneralizing**. They are over applying the rule (applying it to words that don’t use it). It takes time for them to learn the exceptions.
Rule learning: Nativism vs Empiricism

So now we can look at the knowledge that children must have in order to solve the word segmentation problem, and ask which theory it fits with:

- **Modern Nativism**
  - Memorizing the (past tense) form of verbs
  - Generalizing to a rule

- **Modern Empiricism**
  - The question is whether there are any aspects of the rule learning that require domain-specific knowledge.
Some Conclusions

The **word segmentation problem** is the fact that children must somehow decide where the breaks are between words in the speech stream.

Transitional probability, syllables, and stress may all be pieces of information that children use.

The **word learning problem** is the fact that children must somehow decide what words mean.

Nouns first, verbs, second, abstract words third. Nouns appear to be easy to learn, and can help children verbs, which can help children learn syntactic frames, which can help children learn abstract words.

Morphological rule learning follows a **u-shaped curve**, suggesting a phase of memorization, followed by generalization, followed by overgeneralization, followed by the learning of exceptions (memorization).

Many of these abilities seem in line with modern empiricism, but some are on the fence. They may end up requiring domain-specific innate knowledge.