Unit III: Language Acquisition
Part 1: Acquiring Words and Grammatical Categories

Introduction
It's a remarkable fact that normal human children effortlessly – and without explicit instruction – acquire at an early age a mastery of a systematic and productive language.

This is remarkable for a couple of reasons. First, it sets human children apart from all other creatures that we know of. Second, as anyone who tries to learn a new language as an adult is painfully aware, acquiring a second language takes a very long time, and requires considerable effort and concentration. But children seem to be capable of achieving a much greater fluency in a much shorter time, with minimal effort.

Big Question: How is this possible? How do children manage to learn a language?

This question has attracted a good deal of attention from linguists, psychologists, and philosophers. Two main positions:

- *Nativists* about language claim that at least some important aspects of our knowledge of language is innate.
- *Empiricists* about language claim that our knowledge of language comes from experience; it is not innate.

No one would claim that *all* knowledge of language is innate. Clearly, no one is born knowing specific English words such as "dog" or "epistemology". Rather, the debate is whether language learning can be explained entirely in terms of *domain-general* learning mechanisms, or whether language learning requires *domain-specific* learning mechanisms. Nativists claim that knowledge of language requires the latter: we can only explain how a child learns a language by appealing to language-specific knowledge/learning biases. Empiricists, by contrast, claim that we can explain language acquisition in terms of very general knowledge/learning strategies (e.g., pattern recognition, statistical learning) that apply to a wide variety of areas, not just language.

Components of Language Learning
To answer the Big Question, it is helpful to break down language learning into components. When a child learns a language, it seems they need to learn the following:

1) The words in their language
2) Which grammatical categories these words belong to (e.g., learning that "house" is a noun, "kick" is a verb, etc.)
3) Syntactic rules governing how sentences can be composed out of these grammatical categories (e.g., learning that "The woman swam in the ocean" is a well-formed sentence in English, but "The woman ocean swam swam" is not).

Today, we'll focus on the first task (how a child manages to learn the words in their language), and briefly touch on the second (how a child manages to assign these words to grammatical categories). Next time we'll focus on the third task (how a child manages to learn syntactic rules governing the construction of well-formed sentences).

Acquiring Words

One major task facing the infant language learner is carving the speech stream into individual words.

- It takes a bit of work for adults to appreciate the magnitude of this task. We often think there are "spaces" or "gaps" between words in a sentence. But as a matter of fact the gaps that we "hear" are largely an illusion: in everyday space, acoustic breaks do not reliably signal word boundaries. To get a sense for this, try listening to a conversation in a language with which you are totally unfamiliar and try to identify what the different words in the sentence are.

How do infants manage to do this? Here we will look at three different clues that infants seem to track: (i) phonotactic constraints, (ii) stress patterns, (iii) transitional probabilities.

Phonotactic Constraints

- **Phonotactic constraints** are language-specific constraints that determine how sounds of the language may be combined to form words or syllables.

For example, consider the following nonsense string:

Banriptangbowpkesternladfloop.

Exercise: Try to segment this sentence in an "English-like" way.

You may find some discrepancies with your classmates. But chances are, you did not propose a segmentation like this:

Ba-nri-ptangbow-pkester-nladfl-oop.

This shows that some ways of segmenting a wordstream don't result in what we would consider "possible" words of English. But which segmentations are considered "possible" vary across languages. For example, try segmenting the following stream of Swahili into three words:

nipemkatenzuri

You probably wouldn't guess the correct segmentation, which is:

nipe-mkate-nzuri

While a number of segmentations are "impossible" in English, English is actually relatively loose/permissive when it comes to its phonotactic template. For example, letting "C" represent a consonant and "V" represent a verb, here are some phonotactic templates one finds in other languages:

Hebrew	Hawaiian	Indonesian
CV	V	V
CVC	CV	VC
CVCC		CV
		CVC

Based on head-turn paradigm experiments, it appears that by 9 months of age babies have some knowledge of the templates for proper words in their language. For example, Jusczyk et al. (1993) found that American babies orient longer towards strings of sounds that are legal words in English (e.g., *cubeb, dudgeon*) than they do to sequences that are legal words in Dutch but not in English (e.g., *zampljes, vlatke*). This suggests an important tool that babies use to identify the words in their language: they are sensitive to which words are possible or impossible according to their language's phonotactic template.

Stress Patterns

In addition to phonotactic differences, languages also differ from one another in their characteristic stress patterns. For example, English words stress tends to alternate, giving rise to one of two stress patterns:

- Trochaic stress pattern: the first syllable is stressed, the second is not (e.g., *BLACKmail*)
- Iambic stress pattern: First syllable is unstressed, second is stress (e.g., r*eTURN*)

(In some languages, such as French, syllables are stressed more or less evenly.)

But as a matter of fact, English contains far more trochaic words than iambic words (on the order of 9:1 by some estimates). There is evidence that babies are sensitive to this fact: by 7.5 months, babies in an English-speaking environment have no trouble carving up words with a trochaic stress pattern, but they don't recognize iambic words. This suggests another clue that babies may use to identify words: they make educated guesses about which segments of the speech stream constitute words by recognizing the most common stress patterns in their language.

> **Important Question:** It seems that in order to apply either of these techniques – using phonotactic constraints, or exploiting stress patterns – children would need to have some antecedent conception of words in their language. But how do they get this initial conception in the first place?

Transitional Probabilities

In recent years, a number of researchers have tried to answer this residual question by appealing to the idea that babies are capable of tracking *transitional probabilities* between syllables. To illustrate this idea, consider the following syllables:

ti – pre – bay

Which of these would you most likely hear following the other?

One thing to notice here is that in English, pre – ti makes a word ("pretty"), whereas ti – pre, pre – bay, bay – pre, and ti – bay do not. Now, you might still hear some of these combinations within the speech stream, but it would only be at the boundaries between words. As a result, the probability of hearing ti followed by bay is pretty low – it would only happen if you heard two words in succession (e.g., "pretty bay"). By contrast, the probability of hearing pre followed by ti is much higher.

To make this a bit more precise, we can talk about the transitional probability of a syllable Y given a syllable X (written: "T(Y|X)") – that is, the probability that syllable Y will occur given that syllable X has just occurred. We can calculate this by dividing the frequency of the string XY (that is, X followed immediately by Y) by the frequency of X:

$$T(Y|X) = \text{frequency }(XY)/\text{frequency }(X)$$

Exercise: using this formula, calculate T(ti|pre) and T(ti|bay) for the following sentence:

"We went to a bakery and watched the pretty boats on the pretty bay."

- *Hypothesis:* babies pick up on these transitional probabilities, and use them to break up the speech stream into word-size chunks.

An important study by Saffran et al. (1996) seems to bear out this hypothesis:

- In the familiarization phase, eight month old infants were exposed to a two-minute speech stream made up of three syllable nonsense "words" – e.g., *bidaku, golabu, padoti, dutaba.* The only cues to word boundaries were the transitional probabilities between syllables, which were higher within words than at word boundaries.
- In the test phase, researchers looked at how much babies pay attention to the "words" they heard in the familiarization phase, and compared this to the extent to which they paid attention to three-syllable nonsense "words" that straddled the word boundaries (e.g., *dakupa*). Babies paid more attention to the latter, arguably because they were bored by the frequent repetition of the word units.

There is a further question as whether the statistical learning mechanisms that babies use to segment words are unique to humans. Some evidence that it is not:

- Hauser and colleagues replicated Saffran et al.'s experiments on tamarin monkeys. In the familiarization phase, the monkeys heard 20 minutes of a speech stream in a language of made up words. In the test phase, they heard some words from the familiarization phase, as well as some "words" that straddled word boundaries. The monkeys seemed to distinguish between them, staring at the researchers more when they heard one of the unfamiliar "words".
- Toro and Trobalón (2005) also reproduced the results on rats, using the same artificial language used on human infants and monkeys. They found the rats were also able to distinguish between the words in the original speech stream and the unfamiliar "words" straddling word boundaries.[1]

Learning Grammatical Categories

Part of what a child learns when she/he learns English is that sentence (1) is well-formed, but sentence (2) is not:

(1) He admires her.
(2) * He her admires.
(3) *admires he her.

This reveals a familiar fact: within a language, only certain ways of ordering words are permissible.

The permissible ordering of words varies by language. For example, English and Chinese are SVO languages – the Subject comes first, then the Verb, then the Object. Not all languages are like this. For example, Japanese and Hindi are SOV languages. In these languages, (2) is well-formed, but (1) and (3) are not.

[1] These studies only show so much: they show that non-humans are capable of tracking transitional probabilities between words. It remains an open question whether human infants are also capable of tracking more subtle sorts of statistical regularities – regularities that non-human animals cannot track.

Note that while there is considerable variation in permissible word order between languages, at least one thing is constant: *all languages seem to care about grammatical/syntactic categories*. For example, all languages seem to be sensitive to the categories of "subject" and "verb". This raises a question:

How do children learn grammatical categories in the first place?

One hypothesis is that children are sensitive to *distributional evidence*: they pay attention to where certain sorts of words tend to appear in the sentence. For example, a child might realize that certain sorts of words – call them "Class A" words – tend to appear at the beginning of a sentence. They might then notice that a different group of words – call them "Class B" words – tend to appear immediately after Class A words. They might then notice that whereas Class A words tend to refer to actions, Class B words tend to refer to people, places, and things. This might then allow them to acquire a concept of *subject* and a concept of *verb*, and also enable them to assign specific words to these categories.

Some have doubted whether distributional evidence is enough to acquire grammatical categories. For example, Pinker notes that distributional evidence could lead us to make many faulty generalizations. For example, one might be led from the observation that "rabbits" and "fish" occur in the same environments in (4) and (5) to the conclusion that (7) is grammatical:

(4) John ate fish.
(5) John ate rabbits.
(6) John can fish.
(7) *John can rabbits.

These doubts have led some researchers to hypothesize that at least some of our knowledge of grammatical categories is *innate* – it can't be gleaned from distributional evidence alone.

In response to this nativist hypothesis, some have insisted that we need to look at the right sort of distributional evidence. An important area of ongoing research contains lexical co-occurrence patterns. For example, consider a two-word sequence (or "bigram"):

The ____

Chances are you would fill in a noun or a verb after "the". You wouldn't fill in a verb, for example. And if we look at a three-word sequence (or "trigram") your choices become even more restricted:

The ___ is

While either an adjective or a noun can occur after "the" on its own, only a noun can occur between "the" and "is".

Researchers have found that just based on three-word sequences like this, it is possible to predict grammatical categories with a high degree of accuracy. Perhaps then, just as infants are able to track statistical patterns to make conjectures about word boundaries, they are also able to track statistical patterns to make conjectures about grammatical categories.