Psych 156A/ Ling 150: Acquisition of Language II

Lecture 6

Speech segmentation I

Announcements

HW1 due today by the end of class

HW2 now available (not due till after midterm: 5/5/16)

Review questions on speech segmentation now available

Midterm review: in class on 4/28/16

Midterm: during class on 5/3/16

Computational problem

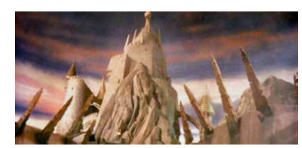
Divide spoken speech into individual words

tuðəkǽsəlbijándðəgáblɪnsíɾi

Computational problem

Divide spoken speech into individual words

tuðəkǽsəlbijándðəgáblɪnsíɾi



tu ðə kǽsəl bijánd ðə gáblɪn síɾi

to the castle beyond the goblin city

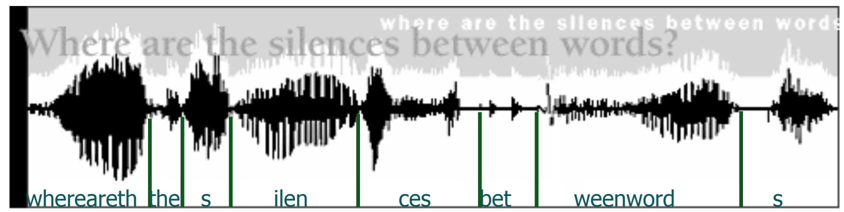
Speech segmentation

"One task faced by all language learners is the segmentation of fluent speech into words. This process is particularly difficult because word boundaries in fluent speech are marked inconsistently by discrete acoustic events such as pauses…it is not clear what information is used by infants to discover word boundaries…there is no invariant cue to word boundaries present in all languages."

- Saffran, Aslin, & Newport 1996

Pauses between words don't really happen

Word boundaries are not necessarily evident in the acoustic waveform



Pauses between words don't really happen

It's harder than you think when you don't know the language!

http://sites.sinauer.com/languageinmind/wa04.01.html

Audio 7: Mandarin sentence

Audio 8: Mandarin words

Audio 9: Farsi sentence

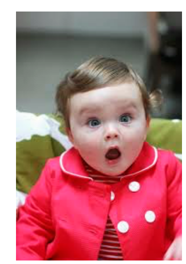
Audio 10: Farsi words

Segmentation mistakes from children

- Two dults
- [Two adults]

- I don't want to go to your ami!
- [I don't want to go to Miami]

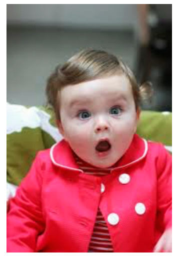
- I am being have!
- [I am behaving!] (in response to "Behave!")



Segmentation mistakes from children

- Oh say can you see by the donzerly light?
- [Oh say can you see by the dawn's early light?]

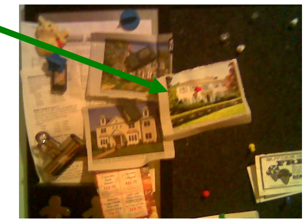
- "A B C D E F G, H I J K, elemenopi…"
- [A B C D E F G, H I J K, L M N O P…



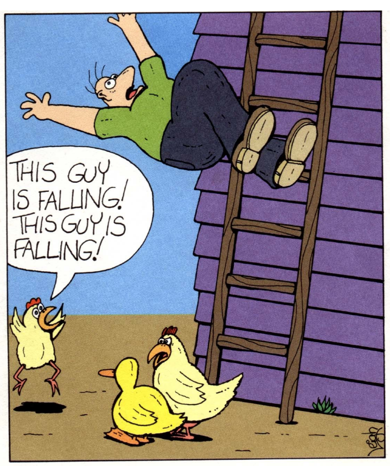
Top-down influence

The white house is under a tack.

The White House is under attack.



Top-down influence



"Boy, he must think we're pretty stupid to fall for that again."

The sky is falling!

or

This guy is falling!

- **Adults** can use top-down information (knowledge of words and the world) to help them with word segmentation.
- What about **infants** who have none or few words in their vocabulary?

Statistical information available

Maybe infants are sensitive to the statistical patterns contained in sequences of sounds.

"Over a corpus of speech there are measurable statistical regularities that distinguish recurring sound sequences that comprise words from the more accidental sound sequences that occur across word boundaries." - Saffran, Aslin, & Newport 1996

to the castle beyond the goblin city

Statistical information available

Maybe infants are sensitive to the statistical patterns contained in sequences of sounds.

"Over a corpus of speech there are measurable statistical regularities that distinguish recurring sound sequences that comprise words from the more accidental sound sequences that occur across word boundaries." - Saffran, Aslin, & Newport 1996

Statistical regularity: *ca + stle* is a common sound sequence

to the castle beyond the goblin city

Statistical information available

Maybe infants are sensitive to the statistical patterns contained in sequences of sounds.

"Over a corpus of speech there are measurable statistical regularities that distinguish recurring sound sequences that comprise words from the more accidental sound sequences that occur across word boundaries." - Saffran, Aslin, & Newport 1996

No regularity: *stle + be* is an accidental sound sequence

to the castle beyond the goblin city

word boundary

Transitional probability

"Within a language, the transitional probability from one sound to the next will generally be highest when the two sounds follow one another in a word, whereas transitional probabilities spanning a word boundary will be relatively low." - Saffran, Aslin, & Newport 1996

Transitional Probability = Conditional Probability

TrProb(AB) = Prob(B | A)

Transitional probability of sequence AB is the conditional probability of B, given that A has been encountered.

TrProb("go" "blin") = Prob("blin" | "go")

Read as "the probability of 'blin', given that 'go' has just been encountered"

Transitional probability

“Within a language, the transitional probability from one sound to the next will generally be highest when the two sounds follow one another in a word, whereas transitional probabilities spanning a word boundary will be relatively low.” - Saffran, Aslin, & Newport 1996

Transitional Probability = Conditional Probability

TrProb(“go” ”blin”) = Prob(“blin” | “go”)

Example of how to calculate TrProb:
go…
…bble, …bbler, …bbledygook, …blet, …blin, …tcha
(6 options for what could follow “go”)

TrProb(“go” “blin”) = Prob(“blin” | “go”) = 1/6

Transitional probability

“Within a language, the transitional probability from one sound to the next will generally be highest when the two sounds follow one another in a word, whereas transitional probabilities spanning a word boundary will be relatively low.” - Saffran, Aslin, & Newport 1996

Idea: Prob(“stle” | ”ca”) = high
Why? “ca” is usually followed by “stle”

to the castle beyond the goblin city

Transitional probability

“Within a language, the transitional probability from one sound to the next will generally be highest when the two sounds follow one another in a word, whereas transitional probabilities spanning a word boundary will be relatively low.” - Saffran, Aslin, & Newport 1996

Idea: Prob(“be” | ”stle”) = lower
Why? “stle” is not usually followed by “be”

to the castle beyond the goblin city

word boundary

Transitional probability

“Within a language, the transitional probability from one sound to the next will generally be highest when the two sounds follow one another in a word, whereas transitional probabilities spanning a word boundary will be relatively low.” - Saffran, Aslin, & Newport 1996

Prob(“yond” | ”be”) = higher

Why? “be” is commonly followed by “yond”, among other options

to the castle beyond the goblin city

Transitional probability

"Within a language, the transitional probability from one sound to the next will generally be highest when the two sounds follow one another in a word, whereas transitional probabilities spanning a word boundary will be relatively low." - Saffran, Aslin, & Newport 1996

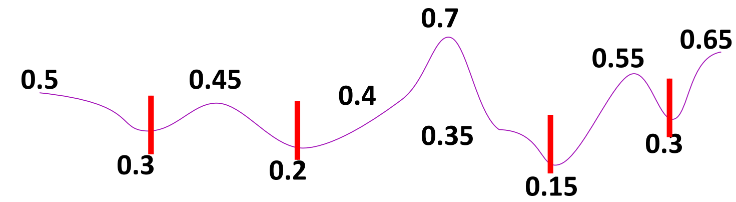
Prob("be" | "stle") < Prob("stle" | "ca")
Prob("be" | "stle") < Prob("yond" | "be")

to the castle beyond the goblin city

TrProb learner posits word boundary here,
at the minimum of the transitional probabilities

Important: doesn't matter what the probability actually is, so long as it's a minimum when compared to the probabilities surrounding it

Transitional probability minima



Transitional probability can be thought of like a landscape.

Every time the transitional probability has a valley (which is a minimum, compared to the "landscape" around it), we put a boundary.

Transitional probability example

un → der → stand → my → po → si → tion

0.9 0.5 0.1 0.3 0.9 0.8

0.5 > 0.1 **0.1 < 0.3**

0.1 = Transitional probability minimum, compared with surrounding transitional probabilities (0.5, 0.3)

Word boundary is here

Another transitional probability example

un → der → stand → my → po → si → tion

0.9 0.8 0.7 0.9 0.9 0.8

0.8 > 0.7 **0.7 < 0.9**

0.7 = Transitional probability minimum, compared with surrounding transitional probabilities (0.8, 0.9)

Word boundary is here

A generic transitional probability example

syl1 → syl2 → syl3 → syl4 → syl5 → syl6 → syl7 → syl8

0.9 0.7 0.8 0.4 0.2 0.3 0.1

0.9 > 0.7 | 0.7 < 0.8 0.4 > 0.2 | 0.2 < 0.3

0.7 and 0.2 = Transitional probability minimum, compared with surrounding transitional probabilities.

Word boundaries are there

8-month-old statistical learning

Saffran, Aslin, & Newport 1996

Familiarization-Preference Procedure (Jusczyk & Aslin 1995)

Measure of infants' response:

Infants control duration of each test trial by their sustained visual fixation on a blinking light.

Idea: If infants have extracted information (based on transitional probabilities) during the habituation trials, then they will have different looking times for the different test stimuli.

8-month-old statistical learning

Saffran, Aslin, & Newport 1996

Familiarization-Preference Procedure (Jusczyk & Aslin 1995)

Habituation:

Infants exposed to auditory material that serves as potential learning experience

Test stimuli (tested immediately after familiarization):

(familiar) Items contained within auditory material

(novel) Items not contained within auditory material, but which are nonetheless highly similar to that material

Artificial language

Saffran, Aslin, & Newport 1996

4 made-up words with 3 syllables each

Condition A:

tupiro, golabu, bidaku, padoti

Condition B:

dapiku, tilado, burobi, pagotu

Artificial language

Saffran, Aslin, & Newport 1996

Infants were familiarized with a sequence of these words generated by speech synthesizer for 2 minutes. Speaker's voice was female and the intonation was monotone. There were no acoustic indicators of word boundaries.

Sample monotone speech:

http://whyfiles.org/058language/images/baby_stream.aiff

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Artificial language

Saffran, Aslin, & Newport 1996

The only cues to word boundaries were the transitional probabilities between syllables.

Within words, transitional probability of syllables = 1.0

Across word boundaries, transitional probability of syllables = 0.33

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Artificial language

Saffran, Aslin, & Newport 1996

The only cues to word boundaries were the transitional probabilities between syllables.

Within words, transitional probability of syllables = 1.0

Across word boundaries, transitional probability of syllables = 0.33

TrProb("tu" "pi") = 1.0

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Artificial language

Saffran, Aslin, & Newport 1996

The only cues to word boundaries were the transitional probabilities between syllables.

Within words, transitional probability of syllables = 1.0

Across word boundaries, transitional probability of syllables = 0.33

TrProb("tu" "pi") = 1.0 = TrProb("go" "la"), TrProb("pa" "do")

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Artificial language

Saffran, Aslin, & Newport 1996

The only cues to word boundaries were the transitional probabilities between syllables.

Within words, transitional probability of syllables = 1.0

Across word boundaries, transitional probability of syllables = 0.33

TrProb("ro" "go") < 1.0 (0.3333...)

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti...

Artificial language

Saffran, Aslin, & Newport 1996

The only cues to word boundaries were the transitional probabilities between syllables.

Within words, transitional probability of syllables = 1.0

Across word boundaries, transitional probability of syllables = 0.33

TrProb("ro" "go"), TrProb("ro" "pa") = 0.3333... <
1.0 = TrPrb("pi" ro"), TrProb ("go" "la"), TrProb("pa" "do")

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti...

word boundary

word boundary

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 1, test trial:

Each infant presented with repetitions of 1 of 4 words

2 were "real" words

(ex: *tupiro, golabu*)

2 were "fake" words whose syllables were jumbled up

(ex: *ropitu, bulago*)

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti...

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 1, test trial:

Each infant presented with repetitions of 1 of 4 words

2 were "real" words

(ex: *tupiro, golabu*)

2 were "fake" words whose syllables were jumbled up

(ex: *ropitu, bulago*)

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti...

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 1, results:

Infants listened longer to novel items (non-words)

(7.97 seconds for real words, 8.85 seconds for non-words)

Implication: Infants noticed the difference between real words and non-words from the artificial language after only 2 minutes of listening time!

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 1, results:

Infants listened longer to novel items (non-words)

(7.97 seconds for real words, 8.85 seconds for non-words)

Implication: Infants noticed the difference between real words and non-words from the artificial language after only 2 minutes of listening time!

But why?

Could be that they just noticed a familiar sequence of sounds ("tupiro" familiar while "ropitu" never appeared), and didn't notice the differences in transitional probabilities.

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 2, test trial:

Each infant presented with repetitions of 1 of 4 words

2 were "real" words

(ex: *tupiro, golabu*)

2 were "part" words whose syllables came from two different words in order

(ex: *pirogo, bubida*)

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 2, test trial:

Each infant presented with repetitions of 1 of 4 words

2 were "real" words

(ex: *tupiro, golabu*)

2 were "part" words whose syllables came from two different words in order

(ex: *pirogo, bubida*)

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Testing infant sensitivity

Saffran, Aslin, & Newport 1996

Expt 2, test trial:

Each infant presented with repetitions of 1 of 4 words

2 were "real" words

(ex: *tupiro, golabu*)

2 were "part" words whose syllables came from two different words in order

(ex: *pirogo, bubida*)

tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti…

Testing infant sensitivity

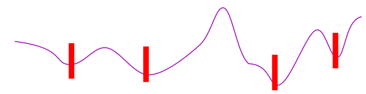
Saffran, Aslin, & Newport 1996

Expt 2, results:

Infants listened longer to novel items (part-words)

(6.77 seconds for real words, 7.60 seconds for part-words)

Implication: Infants noticed the difference between real words and part-words from the artificial language after only 2 minutes of listening time! They are sensitive to the transitional probability information.



Testing infant sensitivity

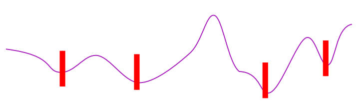
Saffran, Aslin, & Newport 1996

Getting a feel for what infants were able to do.

http://sites.sinauer.com/languageinmind/wa04.03.html

Audio 2

Speech segmentation test



Recap: Saffran, Aslin, & Newport 1996

Experimental evidence suggests that 8-month-old infants can track statistical information such as the transitional probability between syllables. This can help them solve the task of word segmentation.

Evidence comes from testing children in an artificial language paradigm, with very short exposure time.

One issue with infants using transitional probabilities

In general, it seems that infant statistical segmentation abilities (and the forms segmented) may be fragile for young infants (*see Sondregger 2008 for a thorough review of this*).

Johnson & Tyler 2010, Mersad & Nazzi 2012:

8-month-olds fail at utilizing transitional probabilities when the word forms in the artificial language are different lengths.

Success (all 3 syl):	Failure (some 2 syl, some 3 syl):
tupiro, golabu, padoti	pabi, tibu, golatu, daropi

Cues in combination

Still, infants may be able to utilize multiple types of cues to help. For example: transitional probabilities & familiar words

Mersad & Nazzi 2012:

8-month-olds succeed at segmenting artificial languages with words of different lengths if one of those words is a familiar word and transitional probabilities are informative.

Success (some 2 syl, some 3 syl, one familiar word):

pabi, mama, golatu, daropi

Cues in combination

Hearing words in isolation can also help infants segment streams that contain those words and other words of different lengths. This may help infants to recognize these words as "familiar", even if only briefly.

Lew-Williams, Pelucchi, & Saffran 2011:

English 9-month-olds succeed at segmenting non-native language streams with words of different lengths if one of those words is presented in isolation and the transitional probability within the word is high.

Success (utterance + isolated word with high internal TrProb):

melo, Il picchio si abitua a fare la sua casa in ogni *melo* cavo e alto

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants use the prosody (rhythm) of an utterance to help them identify likely boundaries for words (sequences that cross utterance or clause boundaries are less likely to be words). [Gout et al. 2004; Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; Gerken et al. 1994; Nazzi et al. 2000; Seidl 2007, Millotte et al. 2013]

clause boundary

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

utterance boundary

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants use the prosody (rhythm) of an utterance to help them identify likely boundaries for words (sequences that cross utterance or clause boundaries are less likely to be words). [Gout et al. 2004; Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; Gerken et al. 1994; Nazzi et al. 2000; Seidl 2007, Millotte et al. 2013]

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

{pause}

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants use the prosody (rhythm) of an utterance to help them identify likely boundaries for words (sequences that cross utterance or clause boundaries are less likely to be words). [Gout et al. 2004; Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; Gerken et al. 1994; Nazzi et al. 2000; Seidl 2007, Millotte et al. 2013]

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

Not crossing a clause or utterance boundary - more likely to be a word

{pause}

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants use the prosody (rhythm) of an utterance to help them identify likely boundaries for words (sequences that cross utterance or clause boundaries are less likely to be words). [Gout et al. 2004; Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; Gerken et al. 1994; Nazzi et al. 2000; Seidl 2007, Millotte et al. 2013]

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

Crossing a clause boundary - less likely to be a word

{pause}

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants use the prosody (rhythm) of an utterance to help them identify likely boundaries for words (sequences that cross utterance or clause boundaries are less likely to be words). [Gout et al. 2004; Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; Gerken et al. 1994; Nazzi et al. 2000; Seidl 2007, Millotte et al. 2013]

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

{pause}

Crossing an utterance boundary - less likely to be a word

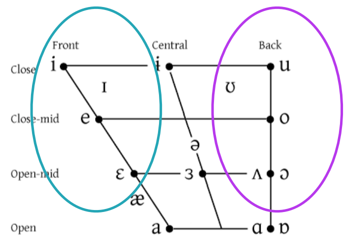
Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Language-specific properties like vowel harmony can signal that syllables belong to the same word in languages that have vowel harmony, like Turkish, Finnish, and Hungarian (Mintz & Walker 2006, van Kampen et al. 2008, Ketrez 2014).

Example: Words contain all front vowels or all back vowels.

From Turkish:
kediler = cats
yolunuz = your road



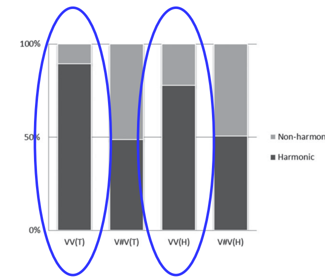
Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Language-specific properties like vowel harmony can signal that syllables belong to the same word in languages that have vowel harmony, like Turkish, Finnish, and Hungarian (Mintz & Walker 2006, van Kampen et al. 2008, Ketrez 2014).

Harmony within words (VV) vs. across words (V#V) in (T)urkish and (H)ungarian (Ketrez 2014):

High within words
At chance across words



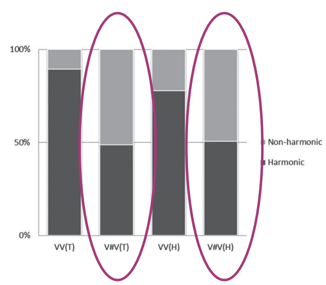
Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Language-specific properties like vowel harmony can signal that syllables belong to the same word in languages that have vowel harmony, like Turkish, Finnish, and Hungarian (Mintz & Walker 2006, van Kampen et al. 2008, Ketrez 2014).

Harmony within words (VV) vs. across words (V#V) in (T)urkish and (H)ungarian (Ketrez 2014):

High within words
At chance across words



Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants distinguish between stressed and unstressed syllables, and they learn language-specific biases. English infants prefer words to begin with stress [≈trochaic] (Jusczyk et al. 1993, Jusczyk et al. 1999) while French infants prefer words to end with stress [≈iambic] (Vihman et al. 1998).

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

{pause}

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants distinguish between stressed and unstressed syllables, and they learn language-specific biases. English infants prefer words to begin with stress [≈trochaic] (Jusczyk et al. 1993, Jusczyk et al. 1999) while French infants prefer words to end with stress [≈iambic] (Vihman et al. 1998).

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

{pause}

Pretty good strategy for English…

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

- Infants distinguish between stressed and unstressed syllables, and they learn language-specific biases. English infants prefer words to begin with stress [≈trochaic] (Jusczyk et al. 1993, Jusczyk et al. 1999) while French infants prefer words to end with stress [≈iambic] (Vihman et al. 1998).

{pause}

"I went to the castle beyond the goblin city, which was very hard to get to. I saw the goblin king."

{pause}

…though it's not perfect

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

But how do infants learn these language-specific stress biases? Swingley (2005) suggests that they arise from the initial words infants extract by using statistical cues. This initial set of words is sometimes called a proto-lexicon.

went	castle	goblin	city
very	hard	get	saw
king			

All words in this English proto-lexicon appear to begin with a stressed syllable.

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

Some evidence that this is the right sequence of events:

Thiessen & Saffran (2003) found that 6-month-olds prefer to segment using statistical cues (like transitional probability), but 9-month-olds prefer to use lexical stress cues. This suggests that infants first rely on statistical cues, and use the proto-lexicon derived from these statistical cues to infer the appropriate lexical stress bias.

Other cues

In additional to statistical information, infants can also use other cues to help them identify words in fluent speech.

Some evidence that this is the right sequence of events:
Thiessen & Saffran (2007) found that 7-month-old English learners can infer from artificial language data with word-final stress that words should end with stress. (Generalization of non-English lexical stress pattern from the artificial input if proto-lexicon provided, despite conflict with TrProb cues → infants are actively using the words they learn.)

stress-final proto-lexicon

padú, topí, gubó, tibí … → ✓ dapú vs. ✗ púdo

da pú do bí bu gó di tí …

low TP High TP low TP

dapú: low TP, stress-final

púdo: High TP, stress-initial

Recap: Other cues

Besides statistical cues to word segmentation, infants are apparently sensitive to familiar words, prosodic cues such as clause and utterance boundaries, and also lexical stress patterns.

It seems that some of the lexical stress cues infants use are language-specific, so these cues are probably not used initially. Instead, these cues may be derived from the proto-lexicons infants have after using statistical cues.

Questions?

You should be able to do up through question 4 on HW2 and up through question 7 on the speech segmentation review questions.