24.904
Language Acquisition

Class 4: Word Segmentation

Word learning

“Words”

- **Intuitive notion:** a sign that stands in a one-to-one relation to a meaning (e.g. a sound form in spoken language, a group of letters with spaces on either side in written language)

ENG: cat

SPA: gato

MAL: puucca

Image courtesy of Jan Miller on Flickr. License: CC BY-NC-SA.

“Words”

- Intuitive notion is not quite right…
  - A single sign representing multiple pieces of meaning

    went = go + past.tense

    matuisaaliqqauviuk? [Inuktitut]
    open.early.rec.pst.interr.2s.S-3s.O
    ‘Did you open it early?’

    [Verb] Adv] Tns] Mood] AgrS] AgrO]

“Words”

- Intuitive notion is not quite right…
  - A single sign representing multiple pieces of meaning
  - A single meaning expressed across multiple signs

kick the bucket = die

“Words”

- Intuitive notion is not quite right…
  - A single sign representing multiple pieces of meaning
  - A single meaning expressed across multiple signs
  - A single meaning expressed by a sub-part of a sign

    un- (as in *unhappy*) = not

“Words”

- **Listemes:** a set of “listed” units of language; not predictable and therefore must be learned and stored in memory

- **Phonological words:** a unit that can, for instance, be preceded and followed by a pause in spoken language.
  - *cat, cats, they’d, whatcha* (as in ‘whatcha doing?’)

“Words”

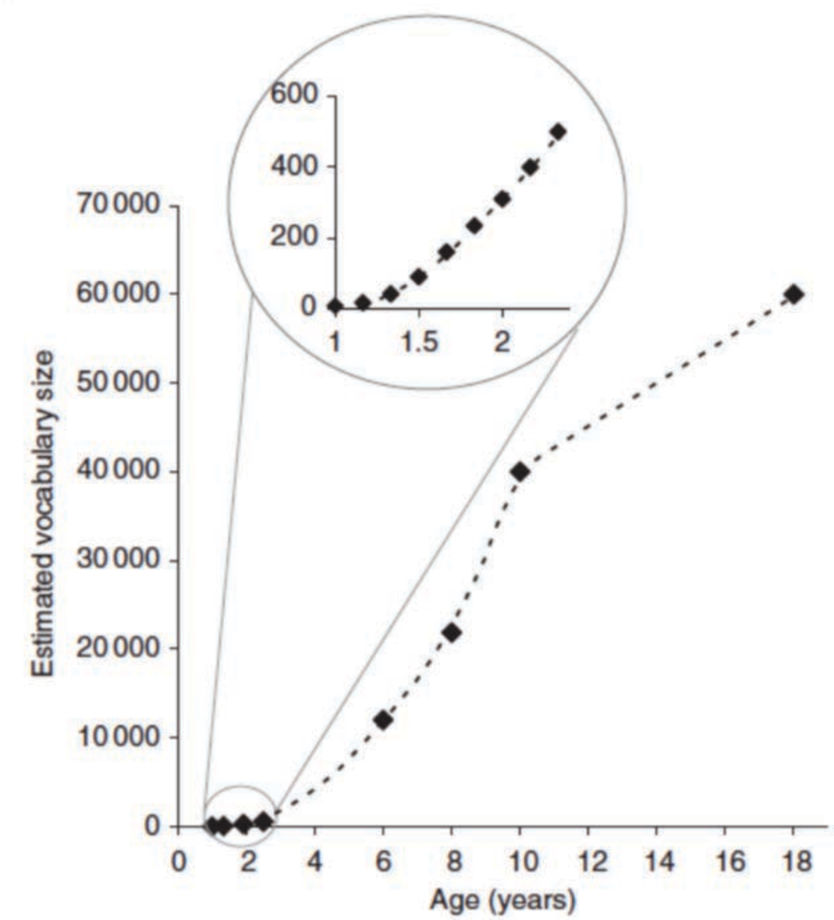
- **Listemes:** a set of “listed” units of language; not predictable and therefore must be learned and stored in memory
- **Phonological words:** a unit that can, for instance, be preceded and followed by a pause in spoken language.
  - *cat, cats, they’d, whatcha* (as in ‘whatcha doing?’)
- **“word”, a shorthand**: “listeme that corresponds to a phonological word”

examples from English

- **Nouns:** dog, chair, windowsill, spinach, peninsula, silence, force, ...
- **Verbs:** laugh, smile, grow, eat, swallow, give, send, donate, think, talk, speak, tell, ...
- **Adjectives, Adverbs:** blue, round, tall, full, new, fake, well, former, possibly ...
- **Prepositions:** in, on, over, along, between, about, ...
- **Function words:**
  may, can, should, …
  a, the, every, all, many, ...
  only, even, too, also, ...
  and, or, not, if, ...
- ...

Word learning over time

- Infants produce their first word between ages 10-12 months
- They understand several words by 6 months (Bergelson 2016)
- By 6 years of age, an average (English-speaking) child knows ~13,000 words (Ames 1964)
    - They've only been alive for 2000 days; that's 6-7 words per day



Snedeker (2009)

Learning task

- For any given listeme w with phonological form ϕ_w and meaning μ_w:
  - What are the properties of ϕ_{wi} (boundaries, length, phonemes, co-occurrence constraints, stress, ...)
  - What are the properties of μ_{wi} (semantic features/ concept, conditions on use, ...).

Challenges of word learning

Each task comes with its own set of challenges…

1. Finding words in the speech stream

2. Associating words with meaning

The problem of finding words

The acquisition task

- Divide continuous (fluent) speech into individual units (typically words)
- The problem: word boundaries are not necessarily evident in the acoustic waveform

Kuhl 2004

The acquisition task

- It’s harder than you might think:

  https://learninglink.oup.com/access/content/sedivy-2e-student-resources/sedivy2e-chapter-4-web-activity-1

The acquisition task

Segmentation errors from child speech

- “I don’t want to go to your ami!”
  [Response to: we are going to Miami]

- “I am being have!”
  [Response to: “Behave!”]

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

Some possible solutions: isolated words

- Basic idea: children learn individual words when they hear it in isolation, then use this to identify more words when they hear it in context

Some possible solutions: isolated words

- About 9% of utterances directed at children by their mothers are single word utterances
- How does the child know what is a single word utterance?

Some possible solutions: statistical information

- Basic idea: 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

Some possible solutions: statistical information

- Basic idea: 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

what a pretty kitten

Some possible solutions: statistical information

- Basic idea: 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

what a pretty kitten

Some possible solutions: statistical information

"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

$$TP(XY) = P(Y|X) = \frac{\mathit{Frequency}(XY)}{\mathit{Frequency}(X)}$$

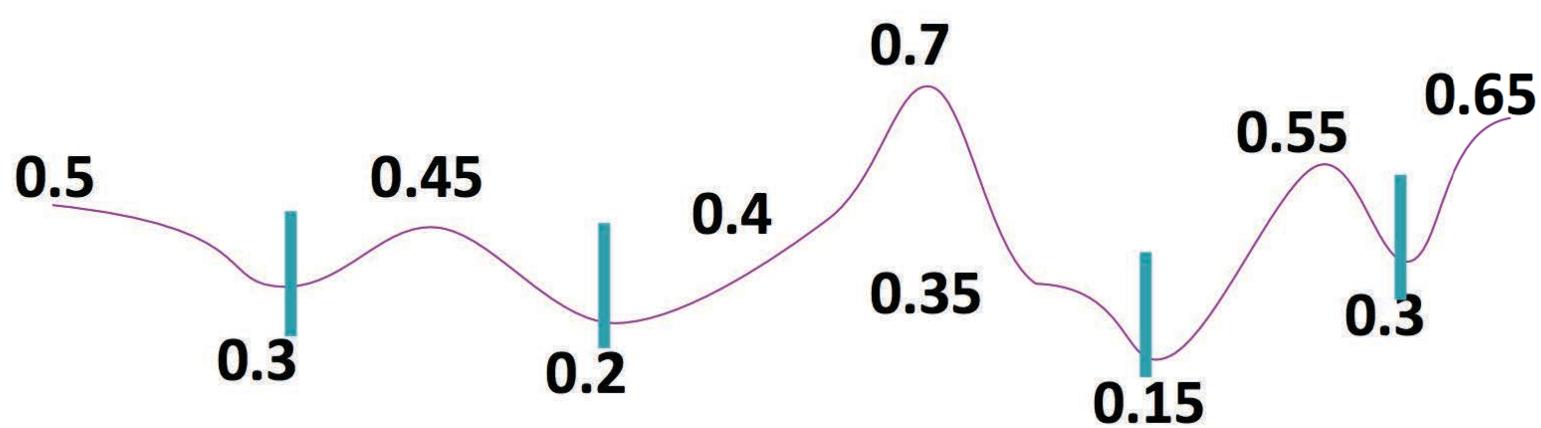
$$TP(ki.ti) = P(ti|ki) = \frac{\mathit{Frequency}(kity)}{\mathit{Frequency}(ki)}$$

Some possible solutions: statistical information

- **Important:** absolute probability doesn't matter, so long as the relevant TP is a minimum when compared to the TPs surrounding it

- Think of it as a sort of landscape:

  - Possible strategy: every time there is a “valley” (a minimum compared to its surroundings), insert a word boundary



Saffran et al. 1996

- Since different syllables occur next to each other more frequently when they make part of a word than otherwise, if infants can extract and keep track of this information, they can use it to identify words.
- Question in Saffran et al. (1996): can infants use transitional probabilities to extract word forms?

Saffran et al. 1996

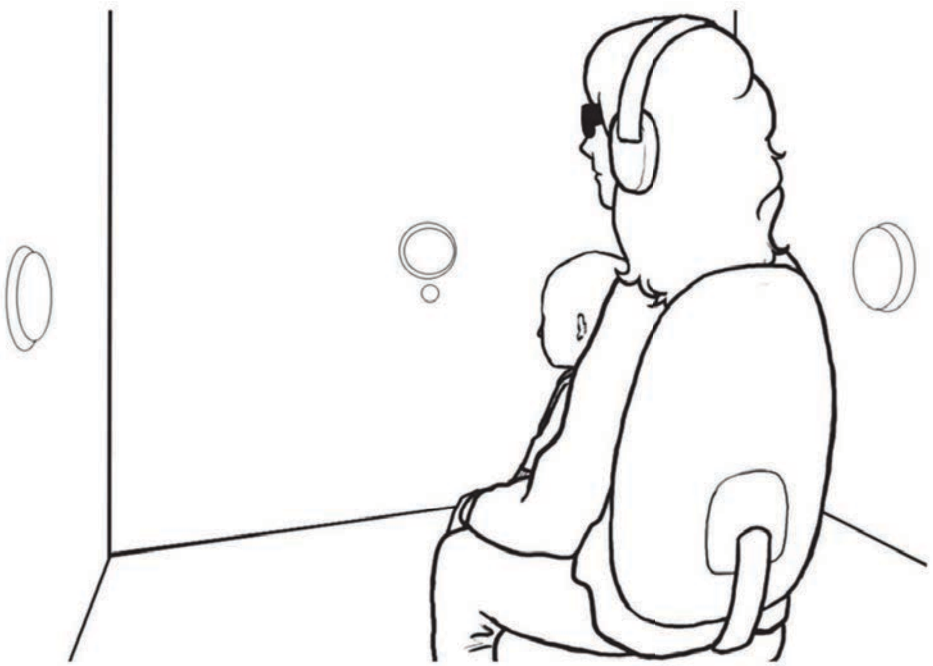
- Artificial language learning
- 3-syllable nonsense words (CVCVCV), e.g. *pabiku*, *tibudo*, *golatu*, *daropi*
- Synthesized with no intonation, then stitched together into a continuous stream of 270 word tokens, e.g *bidakupadotigolabubidaku*…

Saffran et al. 1996

- Head-Turn Preference Paradigm (HTPP)
  - 7.5-month-old infants are exposed to the stream for 2 minutes
  - **Test:** How long do infants listen to “words” in the stream (e.g. pabiku) compared to “part/non-words” (e.g. biku][ti)

Saffran et al. 1996

- A bit more on the paradigm:
  - capitalizes on habituation vs. (renewed) interest
  - **Habituation:** Infants exposed to auditory material that serves as potential learning experience
  - **Assumption:** infants will listen longer to unfamiliar or unexpected stimuli
  - **Test:**
    - familiar: items contained within auditory material
    - novel: items not contained within auditory material, but which are nonetheless highly similar to that material



Saffran et al. 1996

- Condition A: tupiro, golabu, bidaku, padoti
- Condition B: dapiku, tilado, burobi, pagotu
- The only cues to word boundaries were the transitional probabilities between syllables.
  - within word: 1
  - across words: 0.33

Saffran et al. 1996

- Condition A: tupiro, golabu, bidaku, padoti
- Condition B: dapiku, tilado, burobi, pagotu
- The only cues to word boundaries were the transitional probabilities between syllables.
  - within word: 1
  - across words: 0.33

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

TP = 1 TP = 1

Saffran et al. 1996

- Condition A: tupiro, golabu, bidaku, padoti
- Condition B: dapiku, tilado, burobi, pagotu
- The only cues to word boundaries were the transitional probabilities between syllables.
    - within word: 1
    - across words: 0.33

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

TP = 0.33 **TP = 0.33**

Saffran et al. 1996

- Condition A: tupiro, golabu, bidaku, padoti
- Condition B: dapiku, tilado, burobi, pagotu
- The only cues to word boundaries were the transitional probabilities between syllables.
  - within word: 1
  - across words: 0.33

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

TP = 0.33 **TP = 0.33**

Saffran et al. 1996

- Experiment 1, test trials:
- 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)

Saffran et al. 1996

- Experiment 2, test trials:
- 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)

Saffran et al. 1996

- Results

Table 1. Mean time spent listening to the familiar and novel stimuli for experiment 1 (words versus nonwords) and experiment 2 (words versus part-words) and significance tests comparing the listening times.

Experiment	Mean listening times (s)		Matched-pairs *t* test
	Familiar items	Novel items	
1	7.97 (SE = 0.41)	8.85 (SE = 0.45)	$t(23) = 2.3, P < 0.04$
2	6.77 (SE = 0.44)	7.60 (SE = 0.42)	$t(23) = 2.4, P < 0.03$

Saffran et al. 1996

- **Upshot:** infants around 8-months of age can track statistical information such as the transitional probability between syllables. This can help them solve the task of word segmentation.

Limitations of this strategy

- In general, it seems that infant statistical segmentation abilities may be fragile for young infants (see Sondregger 2008 for review).

- 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: tupiro, golabu, padoti
        - Failure: pabi, golatu, tibu

Cues in combination?

- Infants may be able to utilize multiple types of cues to help.
- E.g. 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 may also help infants segment streams that contain those words and other words of different lengths.
- 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.

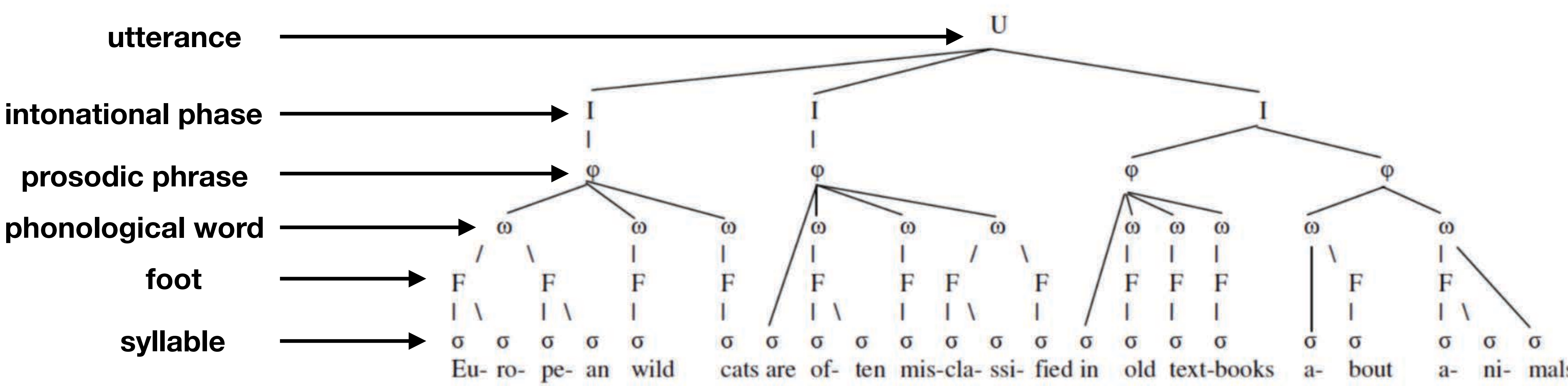
melo…. Il picchio si abitua a fare la sua casa in ogni melo cavo e alto

Other cues?

- There may be cues beyond statistical regularities that help infants
- e.g. prosody (rhythm)

Prosodic domains

- In an utterance, words tend to be grouped into prosodically cohesive units
- These units tend to correspond roughly to syntactic clauses and phrases (Selkirk 1984 *et seq.*)



(1) While I was sleeping, someone broke into my house.

Prosodic domains

- Intonational phrase
  - typically corresponds to clause
  - generally delimited by final lengthening and a pause.
- Phonological phrase
  - Boundaries typically coincide with syntactic phrase boundaries
  - exhibit final lengthening
  - exhibit a single pitch contour
  - exhibit greater initial strengthening (first phoneme of the phrase typically stronger articulated)

Prosodic domains

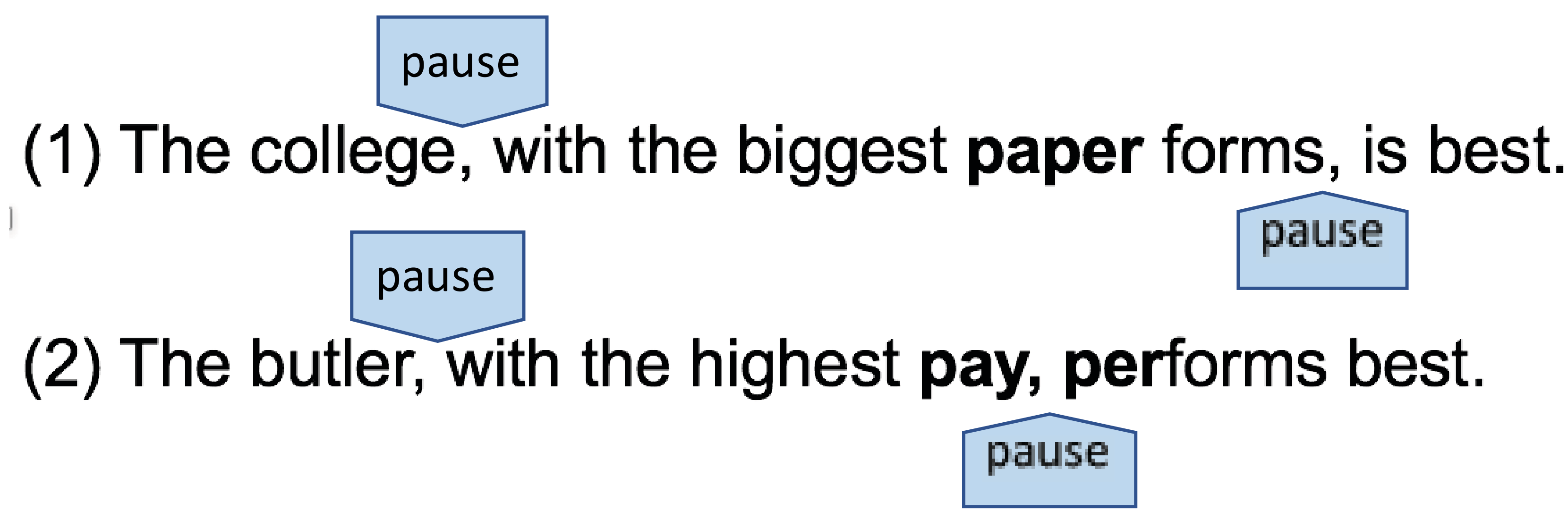
- Further evidence for these domains comes from phonological processes that reference these domains.
- Example of phonological phrase (ϕ) bound process in English:

(1) a. The barriers boxed in the crowd
b. (The barriers)$_{\phi}$ (boxed [i]n)$_{\phi}$ (the crowd)$_{\phi}$

(2) a. The sluggers boxed in the crowd
b. (The sluggers)$_{\phi}$ (boxed)$_{\phi}$ ([ə]n the crowd)$_{\phi}$

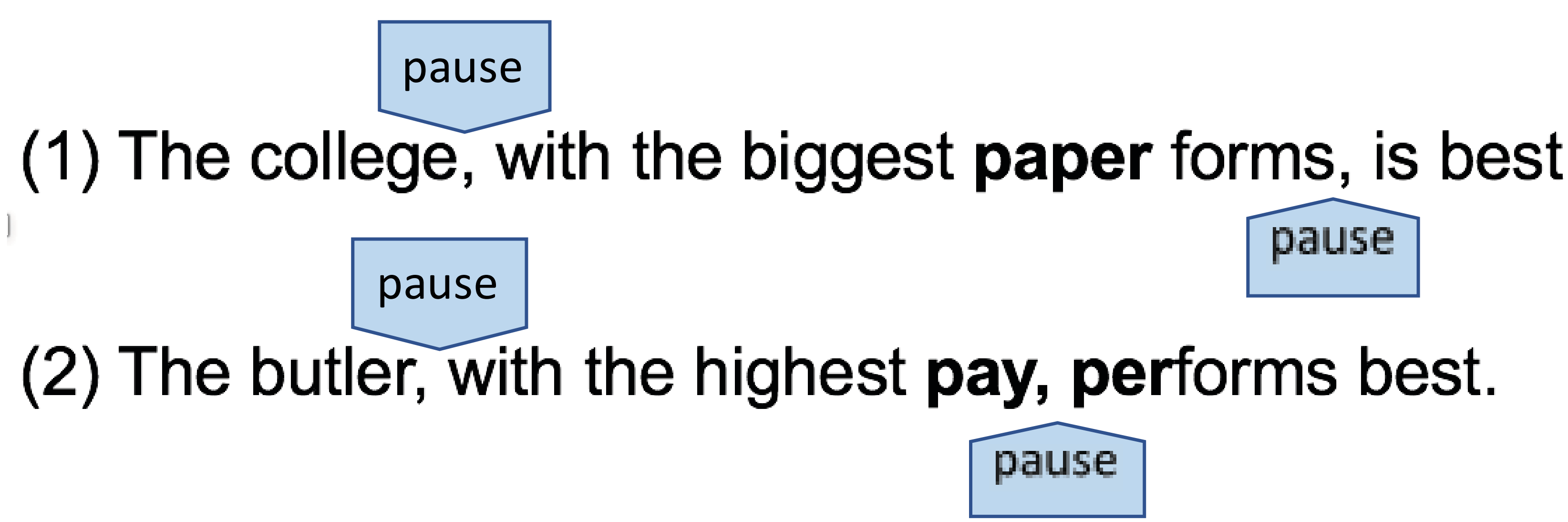
Gout et al. 2004

- Question: do infants use phonological phrase boundaries to find words?



Gout et al. 2004

- The critical bi-syllabic sequence has very high TP in experimental corpus
- If crossing a ϕ-boundary impedes recognition, ϕ-boundaries play an important role in word segmentation



Gout et al. 2004

- Experiments 1 & 2
  - Paradigm: Head-Turn Preference Procedure (HTPP)
  - Familiarized on isolated tokens of 2 bi-syllabic words (paper, beacon; accumulated listening time of 30s)
  - Tested on passages with two contained one of the familiar words and two contained the syllables of the familiar words across a ϕ-boundary

(1) a. [The scandalous paper] [sways him] [to tell the truth].
b. [The outstanding pay] [persuades him] [to go to France]

(2) a. [The owner of the beacon] [founded the association]
b. [The color of the bee] [confounded the new beekeeper]

Gout et al. 2004

- Experiments 1 & 2
  - 13-mos listened longer to passages where the familiar sequence aligned with the ϕ-boundary
  - 10-mos did not

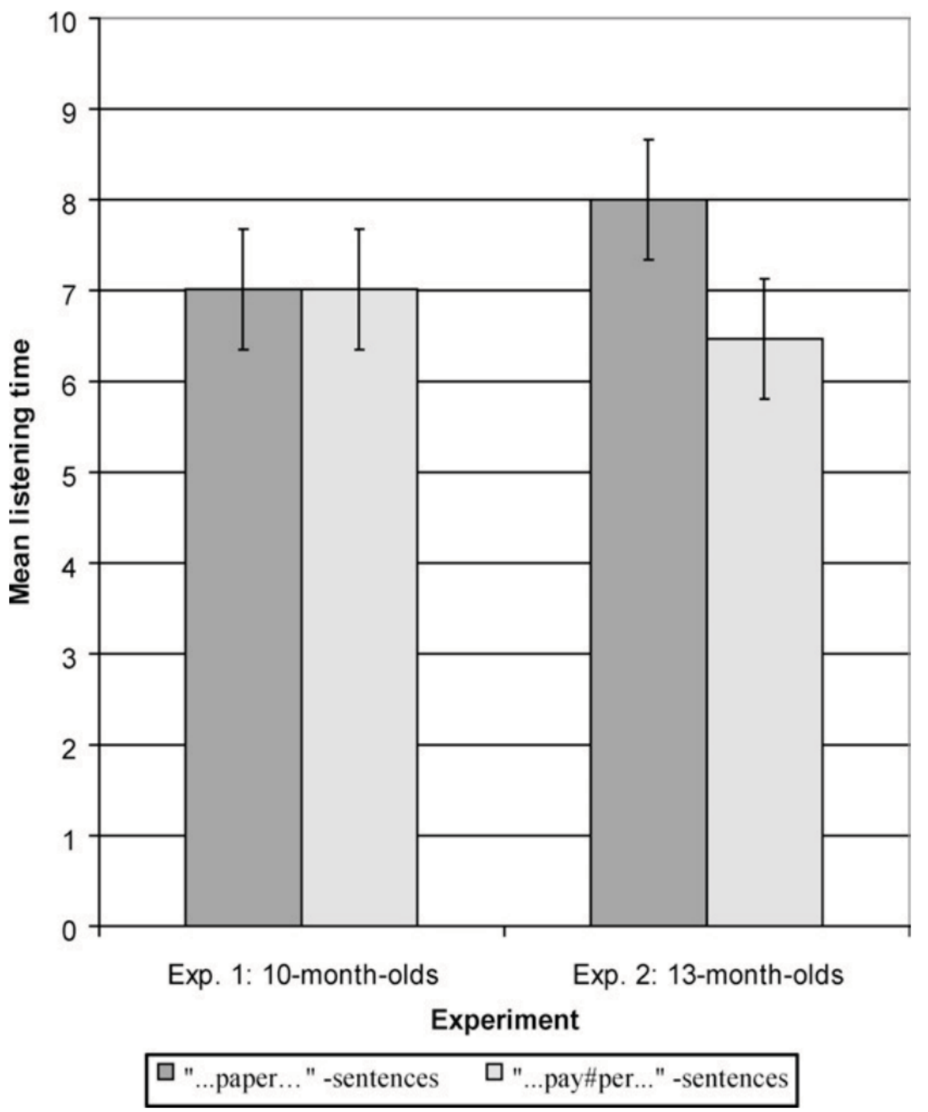


Fig. 1. Results from Experiments 1 and 2: mean listening time (seconds) for sentences containing the bisyllabic word (*paper*-sentences, dark-gray bars) and sentences in which it straddled a phonological phrase boundary (*pay#per*-sentences, light-gray bars). Error bars represent the standard error of the difference.

Gout et al. 2004

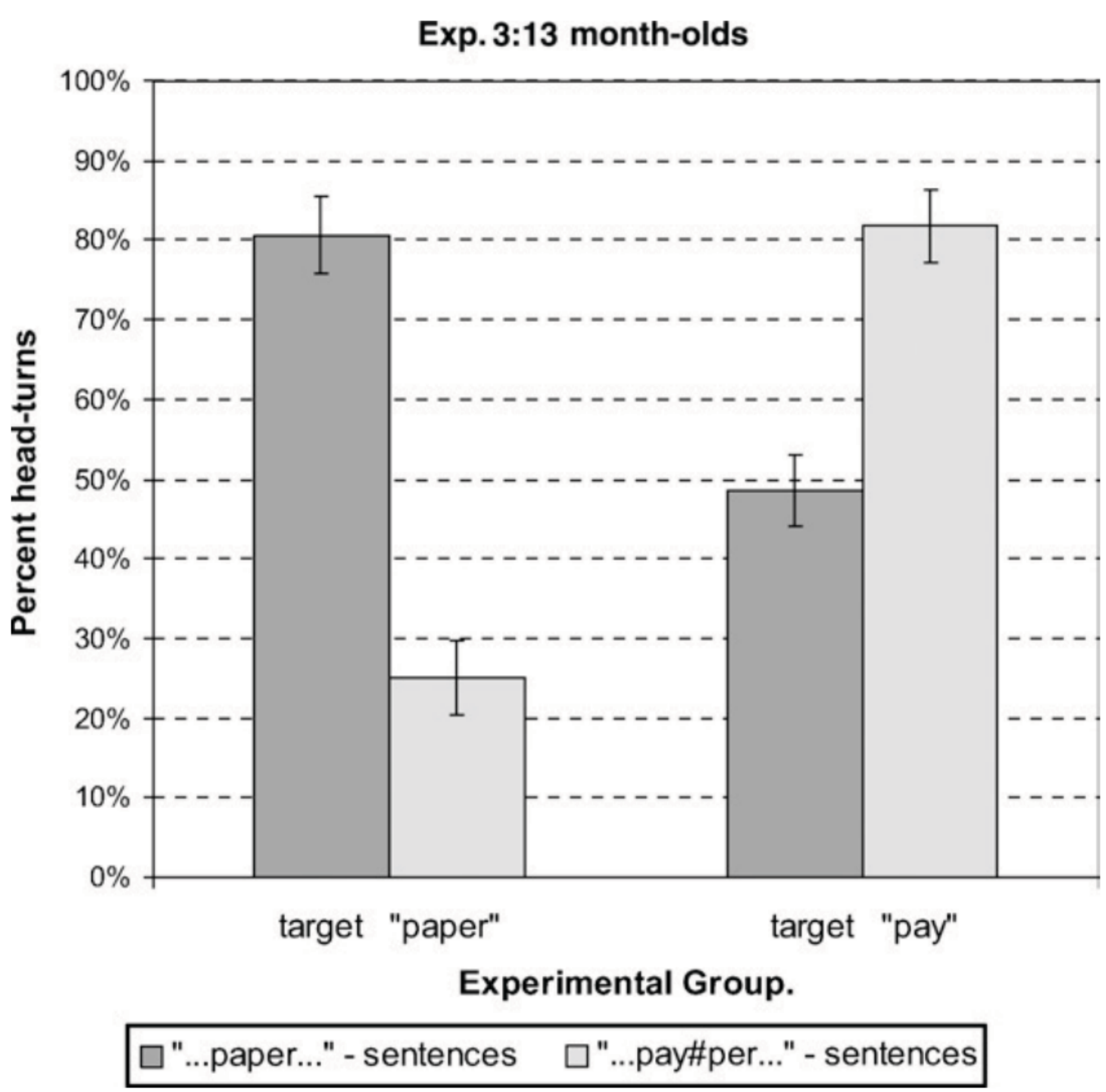
- Experiments 3 & 4
  - Paradigm: conditioned head-turn (CHT)
  - Infants are trained to turn when they hear a particular word
    - between groups: bisyllabic (paper, beacon) vs. monosyllabic (pay, bee)
  - If they turn at the appropriate time, it is "rewarded" with an interesting visual event at that location (conditioning)
  - Test:
    - Passages that did or did not contain targets
    - Bisyllabic words varyingly separated by ϕ-boundary

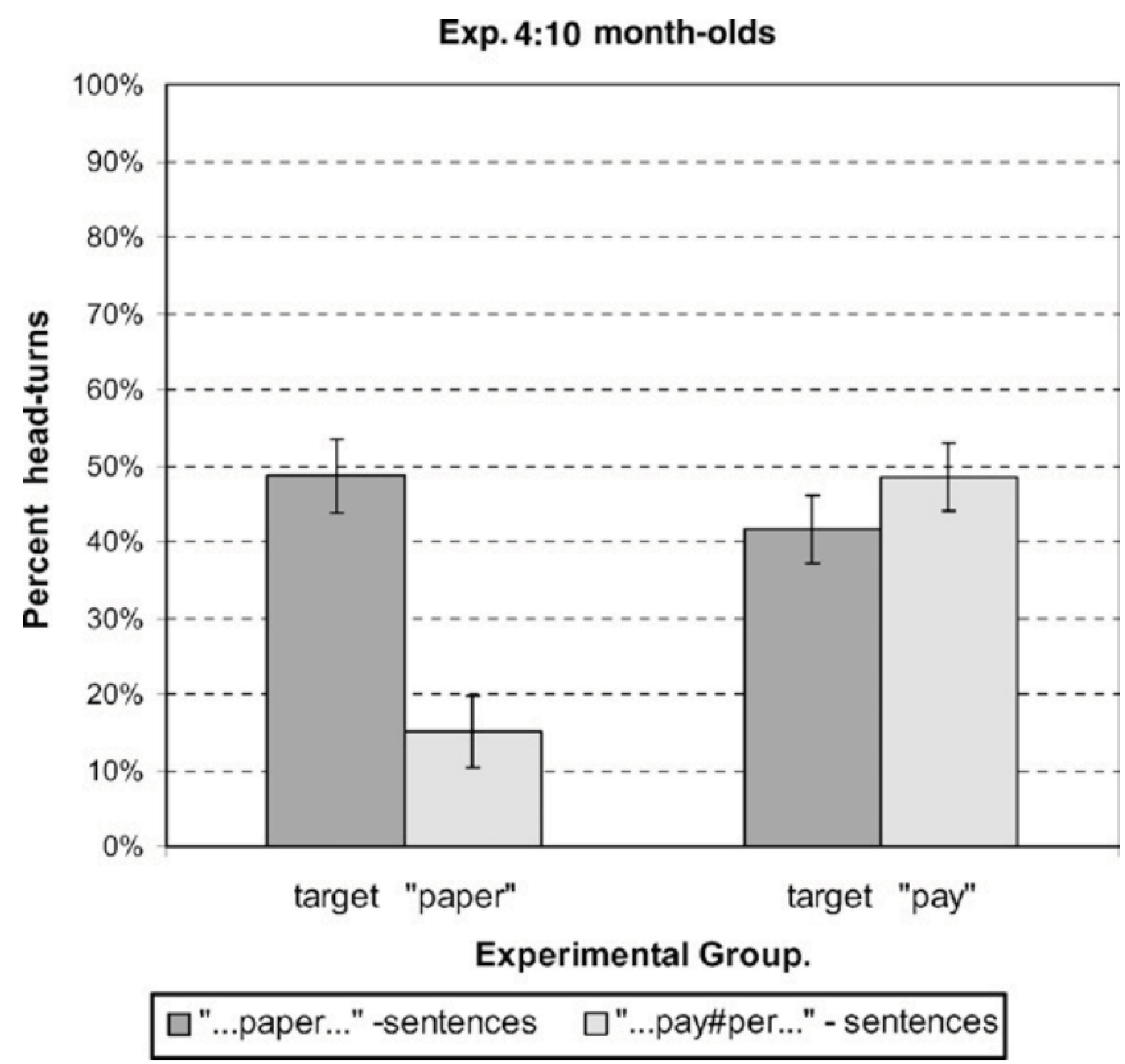
Gout et al. 2004

- Expectations
  - infants trained on bisyllabic targets should turn their head more often when the targets align w/ phonological phrase boundary than when they straddle one.
  - Infants from the monosyllabic group were tested on the same sentences; should turn more often when targets straddle phrase boundary (??)

Gout et al. 2004

- Experiments 3 & 4





Gout et al. 2004

- 13- and 10-mos use phonological phrase boundary cues to constrain lexical access
  - words do not straddle ϕ-boundary
- Note: weaker biases in the other direction for both groups
  - a single ϕ-domain may contain multiple words ([pay per] or [paper])

Going further…

- Shukla et al. (2011): prosody constrains TP-based word ***extraction***

Shukla et al. 2011

- Can infants extract a statistically defined, novel auditory word form from running speech and simultaneously map it onto a visual referent?
- How, if at all, does this process interact with structural properties such as prosodic constituency?

Shukla et al. 2011

- Familiarization with two utterance-types, as the target object moved along the table, of the form **xAByz**, where AB is the target nonce word, whereas the syllables x, y, and z vary.

ʒə-muː-raː]-lei-sə…ʒə-muː-raː]-ni-sə

vs.

ʒə-muː-]raː-lei-sə…ʒə-muː]-raː-ni-sə

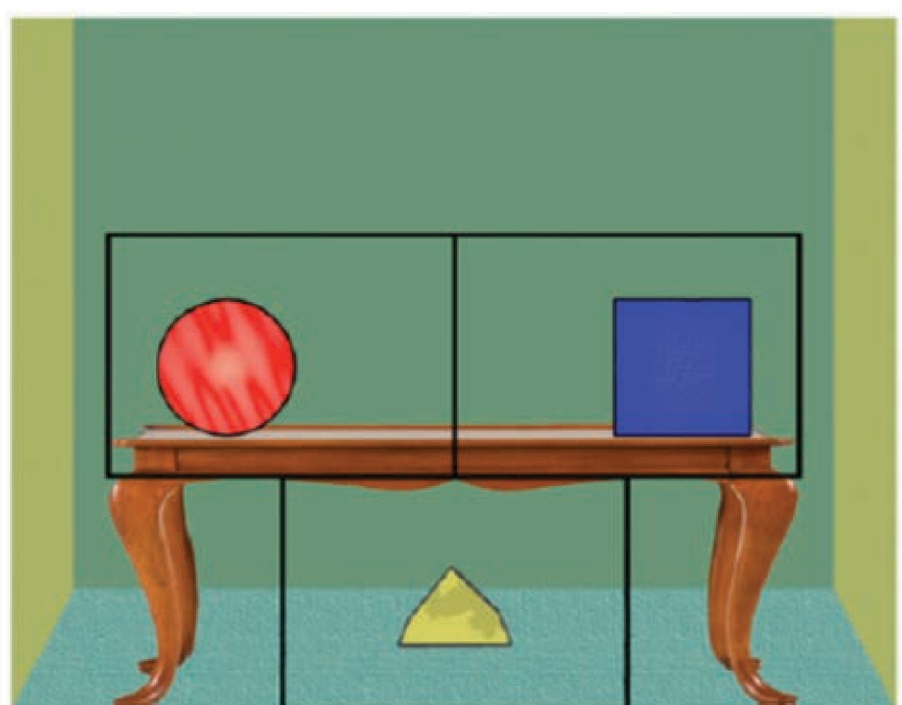
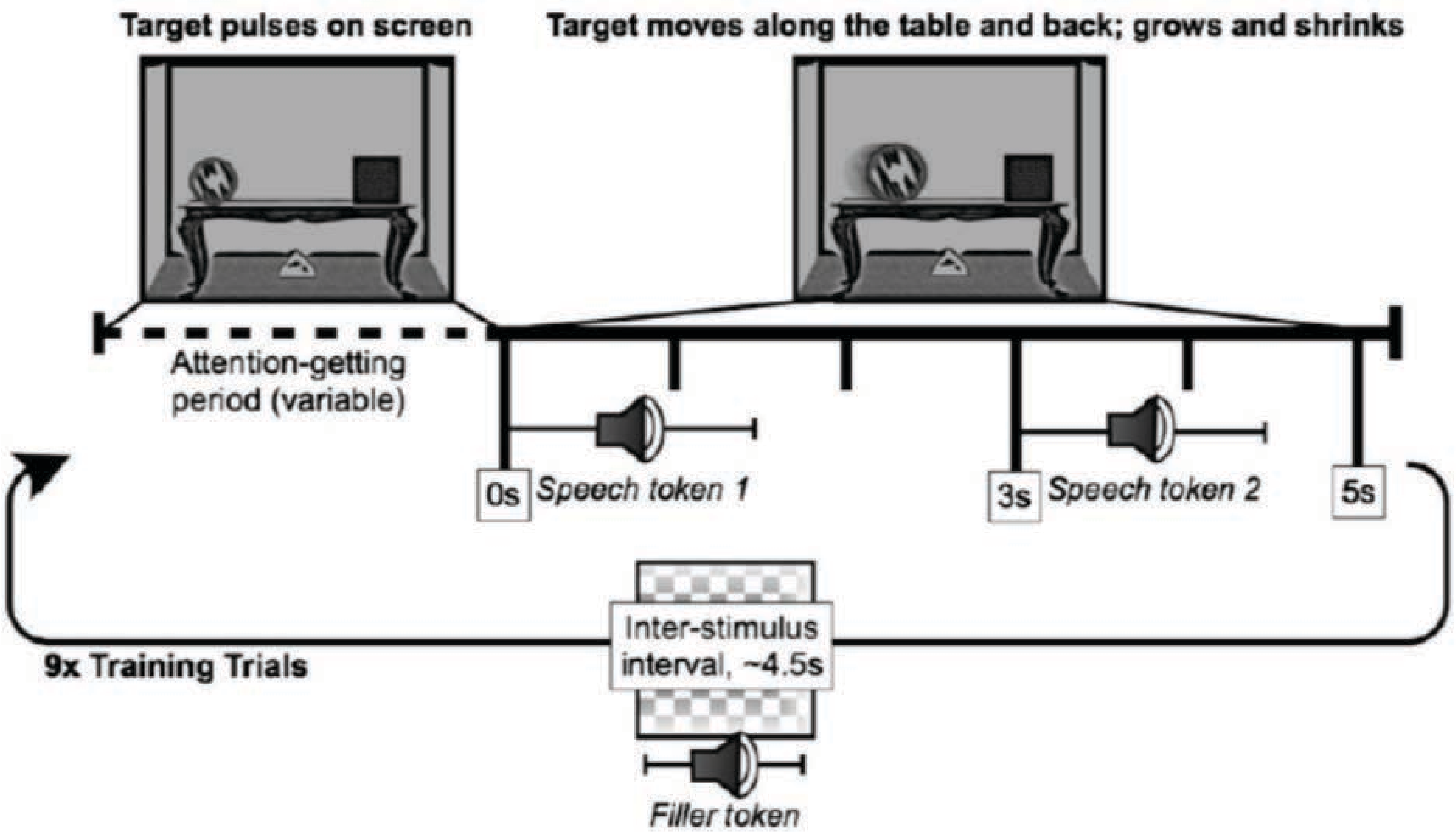


Fig. 1. Screen shot from the experiment, showing the setup of the objects. The three observation windows are overlaid as black outline rectangles.

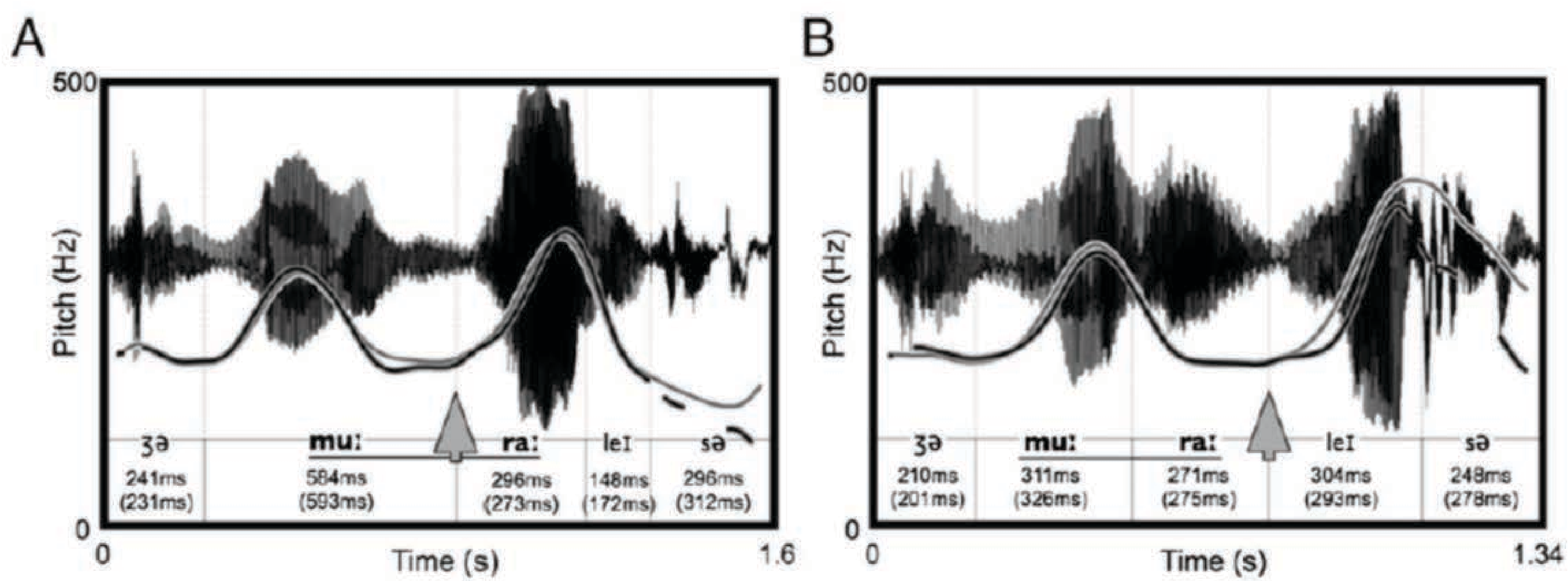
Shukla et al. 2011

- Familiarization with utterances of the form **xAByz** as the target object moved along the table, where **AB** is the target nonce word, whereas the syllables **x**, **y**, and **z** vary.



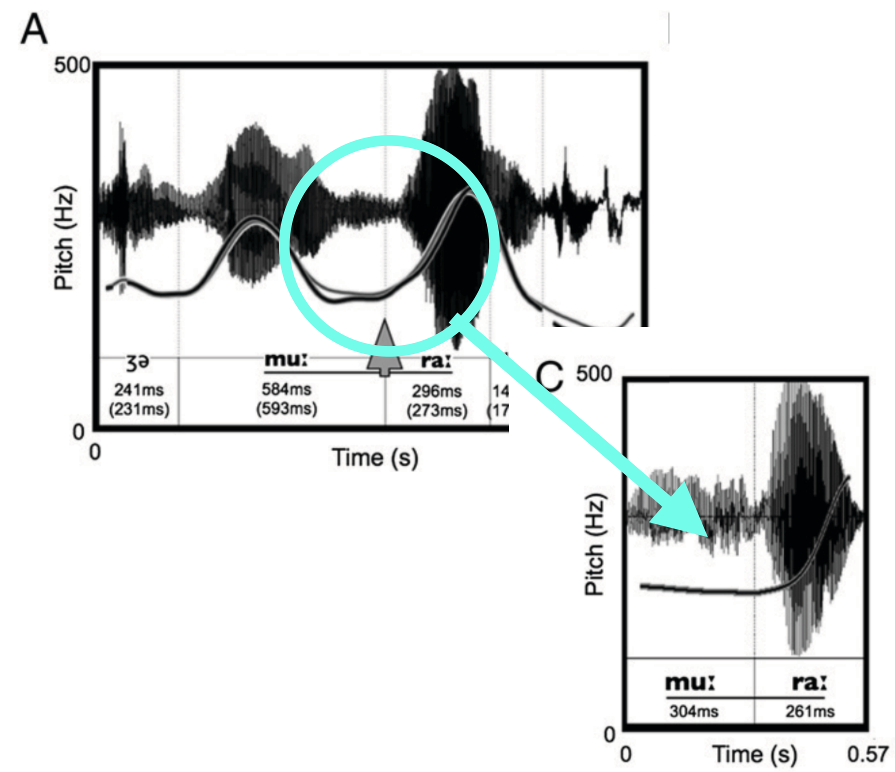
Shukla et al. 2011

- **2 between-subjects conditions:**
  - AB either was within a phonological phrase (one pitch accent; [ʒə-**muː-raː**]-[lei-sə]) or straddled two phonological phrases ([ʒə-**muː][-raː**-lei-sə])



Shukla et al. 2011

- Test
  - **AB** (mu:ra, the statistical "word") vs. **By** (ra:lei, the part-word)
    - NB: prosodically the test word had the intonation-patterns closer to the boundary-straddling sequence
  - Measure: looks to target (the object that moved in fam)



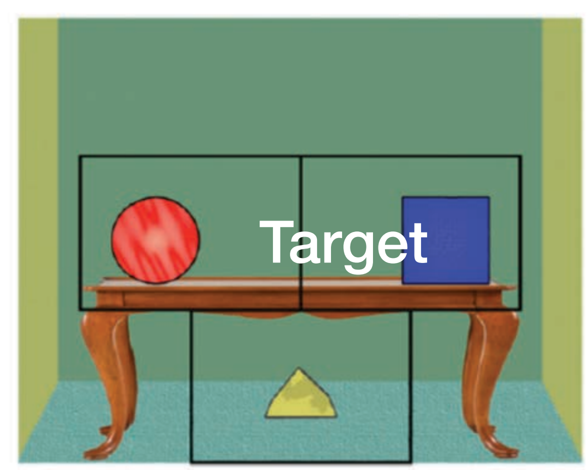


Fig. 1. Screen shot from the experiment, showing the setup of the objects. The three observation windows are overlaid as black outline rectangles.

Shukla et al. 2011

- Results:
  - more looks to target upon word vs. part-word…
  - …but *only* for the prosodically-aligned group!

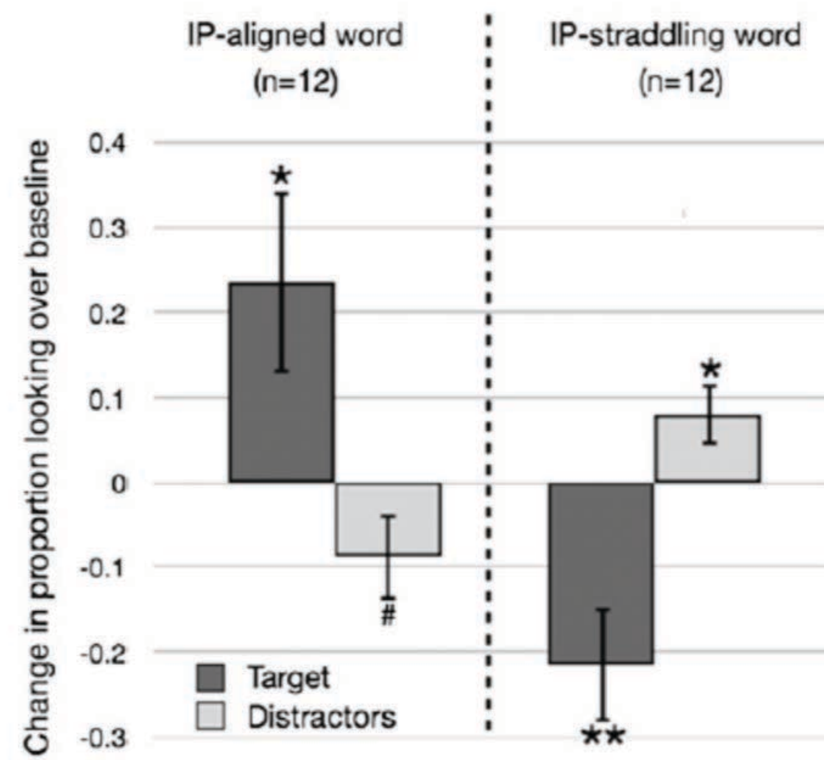


Fig. 5. Change in proportion of looking over baseline (±SEM) to the target and distracter objects upon hearing the nonce word in the two prosodic conditions (n = 12 each). Positive values indicate an increase in proportion of looking over baseline. $^{**}P < 0.01$; $^{*}P < 0.05$; $^{\#}P = 0.1$.

Shukla et al. 2011

- Infants in the prosodic-phrase-internal word group associated the high-TP test word to the target object, but infants in the prosodic-phrase-straddling group did not
  - even though the test items were more *perceptually* similar to the boundary-straddling sequence

Shukla et al. 2011

- "Our findings lend support to arguments that prosodic cues, which signal constituent edges, are critical for acquiring word forms and grammatical patterns in infants and adults."
- “Cognitive capacities of infants are appropriately constrained… language acquisition is most rapid when the structure of the linguistic input is well matched to these constraints."
- "Prosodically organized input may be an essential feature for optimal word learning."