How infants begin to extract words from speech
Peter W. Jusczyk (1999)

Presented by: Mihan Mohagheghzade
23 May 2024

Introduction

When listening to speech in an unfamiliar language, many of us have experienced difficulty hearing where one word ends and another begins.

Why should it be so difficult to find the beginnings and endings of words for a foreign language?

Words are run together without reliable pauses between them.

What native listeners have learned to do is to use their knowledge of certain regularities in the sound structure of the language to predict the boundaries of words.

Because the nature of these regularities is tied to the particular sound structure of the language, knowledge of such regularities in one language is not always helpful in predicting word boundaries in another language.

Infants need to learn the particular features of the sound structure that are most predictive of word boundaries.

Cues to word boundaries in fluent speech

Fluent speech refers to the ability to produce speech smoothly and effortlessly. It involves multiple processes, including conceptual framing, lexical access, phonological encoding, and articulatory control.

There are several potential sources of information that a listener could use in segmenting words from fluent speech:

- Typical Prosodic Patterns of Words in the Language: some languages such as Polish have very regular accent patterns. Knowledge of these basic patterns could help listeners of these languages locate the likely beginnings and endings of words.
- Metrical Segmentation Strategy (MSS): Listeners identify word onsets with the occurrence of strong (stressed) syllables in fluent speech: adults listening to English naturally use stress patterns and rhythmic cues to help them determine where one word ends and the next word begins during speech comprehension.
- Phonotactic Cues: In English, a cluster, such as /mg/ is not permitted within a syllable. Nor does the sequence /kt/ appear at the beginnings of words in English, although this sequence occurs at the beginning of Russian and Polish word
- Statistical and Distributional Properties: For example, in the sequence ‘happy boy’, the co-occurrence relation between the two syllables hap and py is greater than the one between py and boy, because happy can be followed by many other words (e.g. ‘happy man’, ‘happy dog’). Thus, matching known lexical items to the input could help in isolating other words from fluent speech.

When do infants begin to segment words from speech?

Around 6 months old:	**By 9 months:**
Infants start to pay more attention to the sounds of their language. They begin to group sounds in a way that matches how adults in their language speak. They begin learning about how sounds typically form patterns in words in the language. They listen equally long to words with permissible or impermissible sequences for their native language.	Dutch and English learners favor words with the permissible sound sequences. English learning nine-month-old infants listen longer to words with the predominant (strong/weak) stress pattern than to ones with the less common (weak/strong) stress pattern.

Which Experiments have been done?

1. Jusczyk and Aslin. (1995)
2. Houston, D., Jusczyk, P.W. and Tager, J. (1998)
3. Jusczyk, P.W., Houston, D. and Newsome, M. (1999)
4. Saffran, J.R., Aslin, R.N. and Newport, E.L. (1996)
5. Mattys, S.L. et al. (1999)
6. Hohne, E.A. and Jusczyk, P.W. (1994)
7. Jusczyk, P.W., Hohne, E.A. and Bauman (1999)
8. Fernald et al (1998)
9. Stager, C.L. and Werker, J.F. (1997)
10. Swingley, D., Pinto, J. and Fernald, A. (1999)

1. Experiment - Jusczyk and Aslin

Participants	**Method**
24 American infants (11 male, 13 female) (7.5 months of age)	• Word Monitoring: listeners are presented with a word and upon hearing a series of sentences, must indicate whenever they detect the occurrence of a particular word. • Auditory Priming: A repeated auditory stimulus is followed by a test of recognition using degraded versions of the familiar and unfamiliar stimuli. • Headturn Preference Procedure (HPP): Present long speech passages to infants. • They familiarized infants with a particular word that was repeated several times in isolation and then tested different responses to sentences **containing the word** versus those without the familiar word. They used the monosyllabic words: feet, bike, cup, and dog as target items. But why? • Short passages (6 sentences long) were created for each of the target words. • They wanted to evaluate the ability of infants to detect words in fluent speech contexts in general and not at a specific locus in a sentence, so they **varied** the position of the target word within the sentences so that they occur twice at the beginning, middle, and end of two sentences in each passage. • They familiarized half of the infants with two of the words, and the other half with the other two words. • Then they were able to determine if familiarization was important or not.

1. Experiment - Jusczyk and Aslin

But why?

- ❏ They are content words containing stressed syllables
- ❏ Each has a well defined onset and offset.

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2020)

CONSONANTS (PULMONIC) ©①◎ 2020 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

- ❏ These words have different vowel quantities (sounds and lengths):
  - ❖ "Feet" has a long vowel sound,
  - ❖ "bike" has a diphthong (a complex vowel sound),
  - ❖ "cup" has a short vowel sound, and
  - ❖ "dog" has a short or slightly rounded vowel sound depending on the accent.
  - ❖
- ❏ These words are easy and common enough that they can be used in straightforward sentences that one might use when speaking to a child

1. Experiment - Jusczyk and Aslin

TABLE 1
The Four Six-Sentence Passages Used in All Four Experiments

Cup:	The cup was bright and shiny.
	A clown drank from the red cup.
	The other one picked up the big cup.
	His cup was filled with milk.
	Meg put her cup back on the table.
	Some milk from your cup spilled on the rug.
Dog:	The dog ran around the yard.
	The mailman called to the big dog.
	He patted his dog on the head.
	The happy red dog was very friendly.
	Her dog barked only at squirrels.
	The neighborhood kids played with your dog.
Feet:	The feet were all different sizes.
	This girl has very big feet.
	Even the toes on her feet are large.
	The shoes gave the man red feet.
	His feet get sore from standing all day.
	The doctor wants your feet to be clean.
Bike:	His bike had big black wheels.
	The girl rode her big bike.
	Her bike could go very fast.
	The bell on the bike was really loud.
	The boy had a new red bike.
	Your bike always stays in the garage.

1. Experiment - Jusczyk and Aslin

Stimuli	Design	Hypothesis
A female talker who was a native speaker of American English recorded 4 different passages. There were also 8 other filler passages.	Half of the infants heard the words: cup and dog during the familiarization phase. The other half heard the words feet and bike. The caregiver was seated in a chair in the center of the test booth. Each trial by blinking the green light on the center panel until the infant had oriented in that direction. When the infant made a head turn at least 30 in the direction of the loudspeaker, the stimulus for that trial began to play and continued until its completion. The test phase began immediately after the familiarization criterion was attained.	If infants at (7.5 month) have developed some capacity to recognize words in fluent speech contexts, they would listen longer to passages containing the familiar target words.

1. Experiment - Jusczyk and Aslin

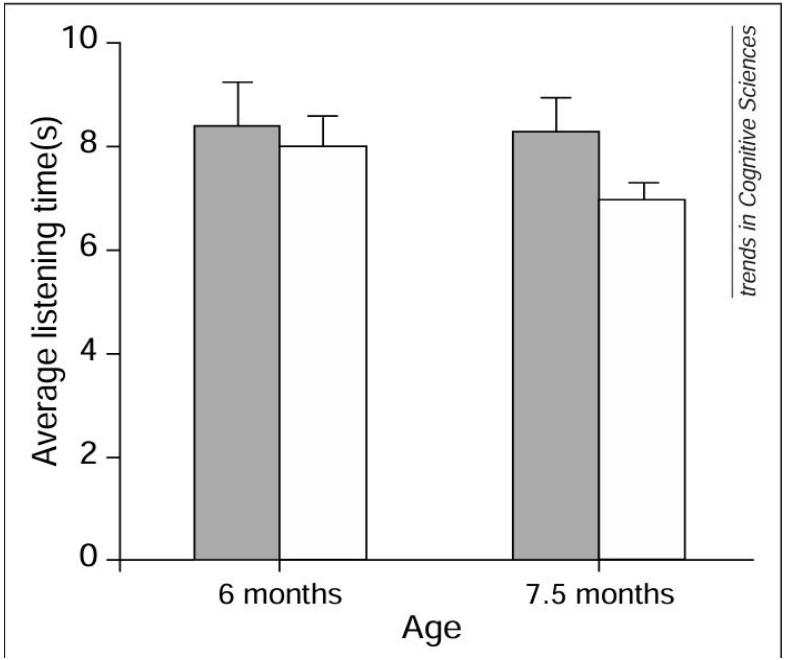
Mean listening times to the 4 passages were calculated:

- ★ If an infant hears a word like "dog" said by itself (in isolation), they will later be able to pick out the word "dog" when it is embedded in a full sentence or a longer sequence of words. So infants can segment and recognize familiar words within the continuous speech they hear.

- ★ 7.5-month-old, but not six-month-old infants did listen longer to the test passages with the familiarized targets, so the older infants recognized the occurrence of the targets in the fluent speech contexts.

- ★ Infants familiarized with an item (e.g. tup) phonetically similar to a target word in the passage (e.g. cup) did not listen longer to these passages in comparison with control passages. So infants recognize and respond to words based on more complex and detailed characteristics of the words rather than just simple acoustic features like vowel sounds.

  These segmentation abilities contribute to the development of a lexicon.



2- Experiment- Houston et al

7.5-month-old infants, familiarized with target words on one day and tested 24 hours later, listen longer to passages with the targets than to ones without them.

Infants appear to encode information into memory about the sound patterns of words that occur frequently in speech directed to them.

How do infants segment words?

As mentioned before, there are several different sources of information in the speech signal that could be helpful to infants in segmenting words.

Several recent investigations have focused on the ability of English-learning infants between **7 and 11 months** of age to use one or more of these sources of information in word segmentation.

3- Experiment- Jusczyk, P.W., Houston, D. and Newsome, M.

Participants	**Method: Stress-Based Strategy**	**Result**
7.5-month-old infants	They first display sensitivity to the **predominant stress pattern of English words** (strong/weak as in fallen) at some point between **6 and 9** months of age. Q: How do English learners segment words with and without the predominant stress pattern?	• They correctly segmented bisyllabic words with the predominant stress pattern, but not words with a less frequent stress pattern (i.e. weak/strong). • Infants familiarized with words such as kingdom and hamlet listened longer to passages containing these words than to control passages. • Infants familiarized with words with weak/strong stress patterns, such as devices and guitars, did not give evidence of detecting these words in passages. • They mis-segment the weak/strong words at the strong syllable boundary. Hence, when familiarized with tar and vice, they listened longer to passages containing guitar and device than they did to control passages • By 10.5 months, English learners detect familiarized weak/strong words in fluent speech contexts. So, they do not rely just on stress cues to segment words from fluent speech.

4- Experiment - Saffran et al

Participants	**Method:** **Statistical Learning Strategy**	**Test Design**
24 infants (8-month-old) from an American-English language environment.	**Familiarization-preference** procedure developed by Jusczyk and Aslin. Infants were **familiarized** with a **two-minute** string of continuous synthetic speech consisting of **4** different **three-syllable** sequences (nonsense words). Word order is random, but a word is never followed by itself. Question: Are 8-month-old infants capable of exploiting **statistical regularities** in the input as **word segmentation cues**?	Infants heard isolated versions of **two of the words** in the sequence, together with two '**part-words**' composed of the **last syllable of one word** plus **the first two syllables of another word from the familiarization step**. The part-words were three-syllable sequences that the infants had heard during exposure period. **Test stimuli**: pabiku, tibudo, tudaro, pigola During the test phase, infants heard isolated versions of two of the words in the sequence (e.g. tibudo and pabiku), together with two 'part words' composed of the last syllable of one word plus the first two syllables of another word from the familiarization sequence (e.g. tudaro and pigola).

4- Experiment - Saffran et al

Condition A **Test stimuli**: pabiku, tibudo, tudaro, pigola	**Condition B** **Test stimuli**: pabiku, tibudo, tudaro, pigola
Words: pabiku, tibudo, golatu, daropi The first two strings (pabiku, tibudo) are words and the second two strings (golatu, daropi) are part-words. The part-word **pigola** spanned the word boundary between daro**pi**#**gola**tu and thus was heard during exposure.	**Words**: tudaro, pigola, bikuti, budopa. The first two strings (tudaro, pigola) are part-words and the second two strings are words(bikuti, budopa). gola**tu**#**daro**pi ⇨ **tudaro**

4- Experiment - Saffran et al- Result

- The listening preferences indicated that the eight-month-old infants did distinguish the words from the part-words.
- They treated the part-words as novel items.
- When such statistical regularities are present in the input, infants can use this information to segment possible words from a stream of speech.

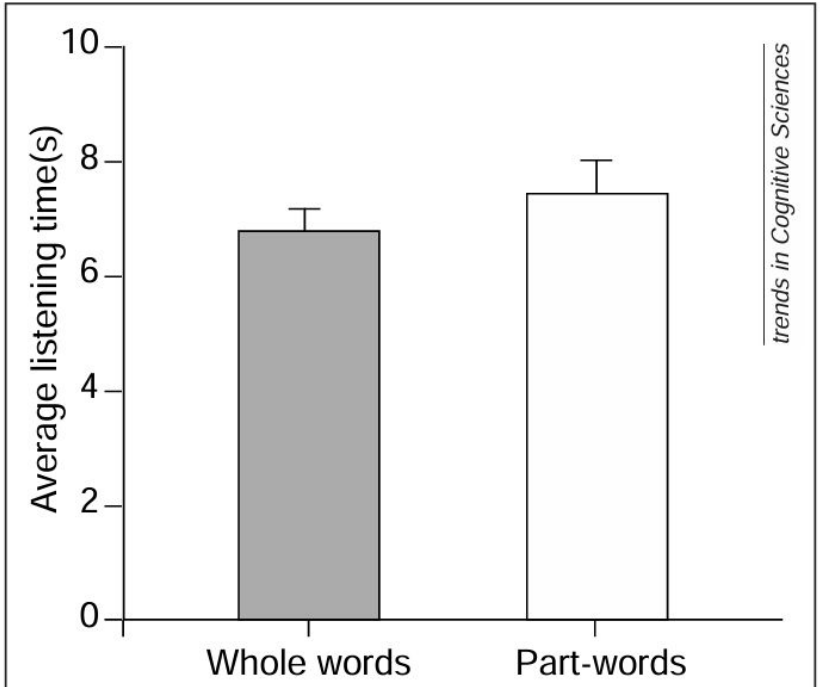


Fig. 3. Segmentation from statistical regularities in speech.

5- Experiment - Mattys et al.

Participants	**Method:** **phonotactic constraints Strategy**	**Hypothesis**
24 infants from monolingual American-English-speaking homes (13 males and 11 females), approximately 9 months of age.	Q: How do phonotactic constraints (rules governing permissible combinations of sounds in a language) help infants identify word boundaries in speech? Infants were presented with two types of lists containing nonsense CVC · CVC sequences whose second vowel, although unstressed, was fully realized. Both lists contain strong/weak stress patterns. They differ in the internal CC sequence. Researchers chose pairs of bisyllabic nonwords (e.g., "nongkuth" [ˈnɔŋ · köθ]). The focus was on the cluster of consonants (C · C) that appear between the two syllables.	If 9-month-olds perceive the CVC · CVC stimuli with strong/weak stress patterns as single word-like units, they should prefer the stimuli with the within-word clusters because these are phonotactically better formed than the stimuli with the between-word clusters.

5- Experiment - Mattys et al.

Task Design	**Result**
All chosen C · C clusters are common in English speech but differ in their likelihood of appearing within a word versus between words. Within-Word Condition: The C · C cluster has a high probability of occurring within a single word but a low probability of occurring between two words. Example: "nongkuth" [ˈnɔŋ · köθ] - The cluster [ŋ · k] is more likely to be found within words. Between-Word Condition: The C · C cluster has a high probability of occurring between two words but a low probability of occurring within a single word. Examples: "nongtuth" [ˈnɔŋ · töθ] or "nomkuth" [ˈnɔm · köθ] - The clusters [ŋ · t] or [m · k] are more likely to be found between words.	Sensitivity to the distribution of allophonic cues (auditory variants of the same phoneme) within words appears to develop more slowly in English learners. Sensitivity to how allophonic cues are distributed within words seems to develop in English learners between 9 and 10.5 months.

6 & 7- Experiment
Hohne, E.A. and Jusczyk, P.W. (1994)
Jusczyk, P.W., Hohne, E.A. and Bauman (1999)

Q: Can infants use **allophonic differences** in the words '**nitrates**' and '**night rates**' to detect these words in fluent speech contexts?

Design: They **familiarized** infants with **isolated versions of one of these words and another word** and then tested them on **passages that either included or did not include these targets.**

Results:

- Although an earlier investigation (Hohne 1994) infants could discriminate the allophonic differences between nitrates and night rates, **nine-month-old infants** did not indicate using this information to locate the familiarized target word in the passages.
- Nine-month-old infants familiarized with nitrates listened equally long to the test passage with night rates as they did to the one with nitrates.
- In contrast, 10.5-month-old infants did listen significantly longer to the test passage that contained the familiarized item.

Why multiple cues are necessary for word segmentation

Towards the end of the first year, English learners are sensitive to several different possible sources of information about word boundaries in fluent speech.

This is a fortunate development because none of these sources is sufficient for correctly segmenting all words from fluent speech.

Reliance on:

Prosodic cues: This would lead an English listener to miss the onset of words beginning with weak (unstressed) syllables.

Statistical regularities: could cause a listener who knows the word candle to make segmentation errors in contexts such as ‘can deliver’.

Phonotactic cues: although /zn/ occurs relatively infrequently within words, this sequence does occur in business.

Consequently, listeners must draw on some combination of these potential cues in segmenting words from English speech.

Segmenting words when extracting meanings

By the end of their first year, infants can detect familiar sound patterns within continuous speech.

However, when 7.5-month-old infants recognize the word "kingdom" in a passage, it doesn't mean they understand what it means. For infants to fully comprehend sentences, they need to learn the meanings of the words they identify.

Although it might seem logical that infants who are good at recognizing sound patterns would easily start understanding word meanings, this transition is not as straightforward as expected.

8- Experiment - Fernald et al

Participants	**Method** **Meaning Extraction Methodology**	**Result**
15-month-old infants learning English.	Infants were shown two objects on video monitors while hearing a sentence that included the name of one object. Q: whether the infants look at the named object more often?	When the target word (the name of the object) was **at the end** of the sentence, infants looked at the corresponding object **more frequently.** When the target word was in the **middle** of the sentence, infants **did not show** this preference. This finding is notable because previous studies with 7.5-month-old infants showed no difference in response based on the word's position in the sentence. 15-month-old infants find it hard to recognize words in the middle of sentences but by 18 months, they've developed better skills to handle this task. But Why do 15-month-old infants have difficulty with targets in non-final positions? The cognitive load of matching the word to the correct picture might overwhelm the word-segmentation abilities of 15-month-old infants. Another suggestion by Fernald et al. is that placing the word at the end of the sentence makes it more noticeable for infants, aiding in segmentation.

8- Experiment -Fernald et al

Meaning Extraction Strategy

They analyzed infants' **reaction times** to the **correct picture** after **hearing the target word** and documented that **24-month-old infants** were **316 ms** faster than **15-month-olds** and **148 ms** faster than **18-month-olds** in **shifting their gaze** from a distractor picture to the target picture.

Result: The older group was much faster at understanding the words that they heard.

9- Experiment - Swingley et al

Participants	**Hypothesis**	**Method**	**Test design**			**Result**
Infants (24-months old)	They are much better at recognizing words in sentences, similar to adults.	Q: Does the similarity of sounds at the beginning (onset) or end (offset) of words affect how quickly infants can recognize and locate the correct word in a sentence?		**Onset Overlap (distractor)**	**Offset Overlap (distractor)**	Onset Matters More: The study found that 24-month-old infants are more affected by similarity in the beginning sounds of words than the ending sounds. This suggests that the initial sounds of words play a more crucial role in early word recognition for infants.
			target word: (ball)	bat	call	
			Findings: Infants **.....** to recognize the target word when distractor words shared the same starting/ending sounds.	**took longer**	**did not take longer**	

Conclusion

- 7.5-month-old infants can segment words that conform to the predominant stress pattern of English words.
- The ability to segment words with other stress patterns requires using different sources of information about word boundaries.
- By 10.5 months, English learners display sensitivity to additional cues to word boundaries such as statistical regularities, allophonic cues, and phonotactic patterns.
- Infants' word segmentation abilities undergo further development during their second year when they begin to link sound patterns with particular meanings.
- By 24 months, the speed and accuracy with which infants recognize words in fluent speech is similar to that of native adult listeners.

References

1. Jusczyk. Peter. k, 1995, *Infants Detection of the Sound Patterns of Words in Fluent Speech*, Cognitive Psychology, Vol. 29, 1-23
2. Jusczyk. Peter, 1999, *How infants begin to extract words from speech*, Trends in Cognitive Science, Vol. 3, No. 9, 323- 328
3. Jusczyk. Peter, and Houston. Derek, 1999, *Cognitive Psychology,* Vol.39, 159-207
4. Thiessen. Erik D, and Saffran. Jenny R, 2007, *Learning to learn: Infants Acquisition of Stressed-Based Strategies for word Segmentation*, Vol. 3, No. 1, 73- 100

Any

Question

Thank You
For Your Attention