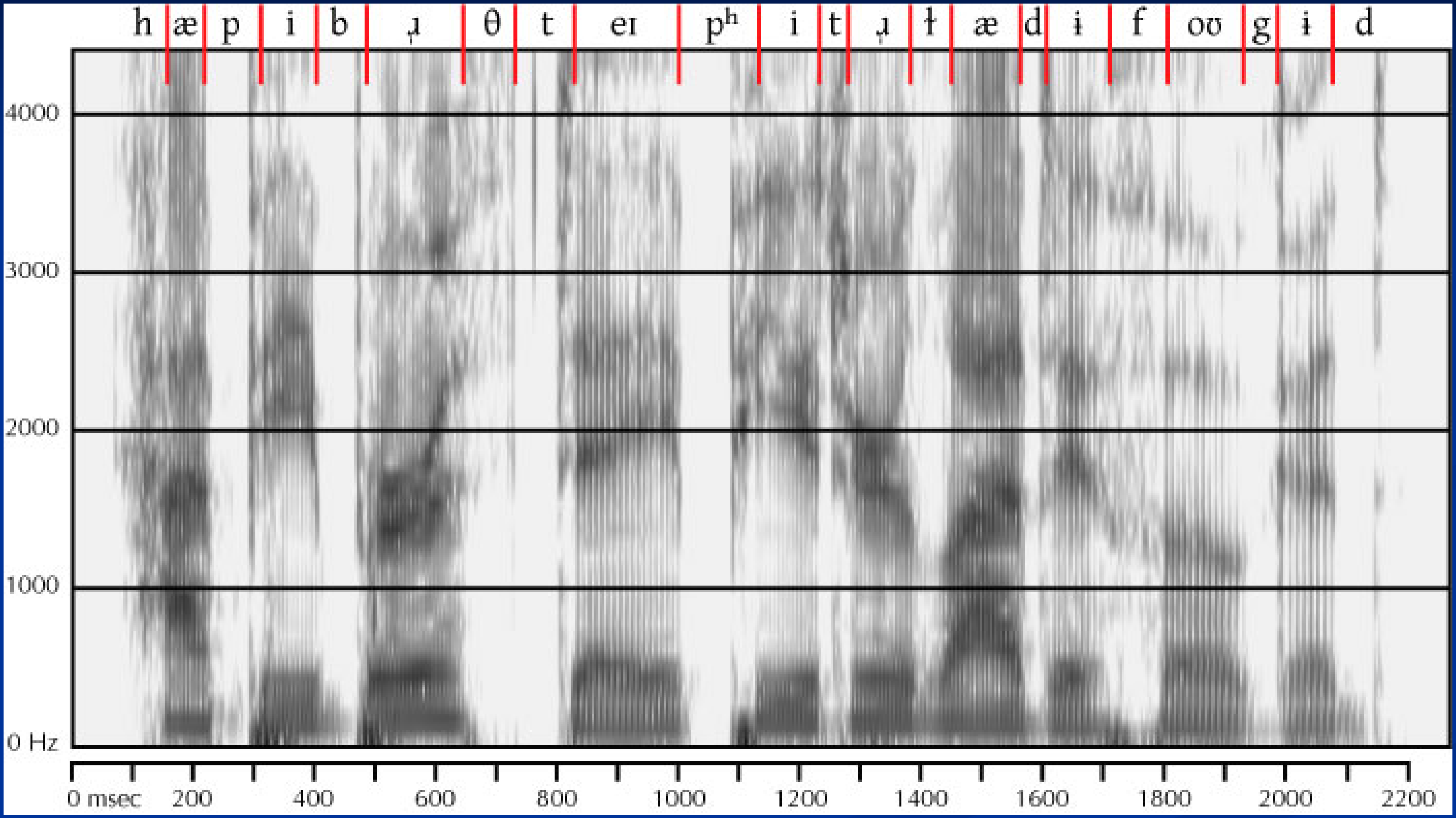
Even more on Speech Perception: It’s not just phonemes

Word Segmentation

PL Chapter 3

LDER Chapter 4


h
æ
p
i
b
ɹ̩
θ
t
eɪ
pʰ
i
t
ɹ̩
ɫ
æ
d
ɨ
f
oʊ
g
ɨ
d
4000
3000
2000
1000
0 Hz
0 msec
200
400
600
800
1000
1200
1400
1600
1800
2000
2200

How do I find words?

iamforcedintospeechbecausemenofsciencehaverefusedtofoll
owmyadvicewithoutknowingwhyitisaltogetheragainstmywillth
atitellmyreasonsforopposingthiscontemplatedinvasionofthea
ntarcticwithitsvastfossilhuntanditswholesaleboringandmelting
oftheancienticecapsandiamthemorereluctantbecausemywarn
ingmaybeinvaindoubtoftherealfactsasimustrevealthemisinevit
able;yetifisuppressedwhatwillseemextravagantandincrediblet
herewouldbenothinglefthehithertowithheldphotographsbotho
rdinaryandaerialwillcountinmyfavorfortheyaredamnablyvivida
ndgraphicstilltheywillbedoubtedbecauseofthegreatlengthstow
hichcleverfakerycanbecarriedtheinkdrawingsofcoursewillbeje
eredatasobviousimposturesnotwithstandingastrangenessofte
chniquewhichartexpertsoughttoremarkandpuzzleover

Detecting Word Boundaries

- Infants don’t know any words
- How can they be found (and then learned)?
- What information in the speech stream can a baby use to find words?
  1. Frequent sounds
  2. Frequently co-occurring sounds
  3. Phonotactics (combinations of legal sounds in words)
  4. Prosodic Patterns

Frequent Sounds

I'll give you a pot of sugar
Put some tea in the pot
His pot is full of water
Pot of tea or coffee?
Put the red pot on the table

- Infants were habituated on sentences like these
- Tested with "pot" and "car"
- Showed preference for "pot"
- Infants extracted sound pattern – despite not knowing meaning!

Frequently co-occurring sounds

- Sounds that co-occur may form words
  - Transitional probabilities for syllables within words higher than for between words
  - (given a syllable X, what is probability that next syllable will be Y)

- *pretty baby*
  - Probability of –by following ba- is higher than probability of –ba following –ty
  - (compare pretty doggie, pretty mommie, pretty flower, etc.)

Saffran, Aslin and Newport

- Created 3 syllable nonsense words
  - bidaku padoti golabu tupiro
- Strung them together in two minute block
  - bidakupadotigolabubidakutupiropadoti…
  - Words were arranged in random order
    - (transition probability between words is lower than within words)
  - Tested for infants' listening preference for words (tupiro) vs. non-words (dapiku)
  - Infants (8 months) preferred words to non-words

Phonotactics

- Certain sequences of phonemes are not legal within a word, or at the beginning or end of a word
  - “tb” is not legal at the start of an English word
    - Possible word boundary between /t/ and /b/?
  - “sp” is legal at start of English word (*special*) and middle (*especial*) and end (*lisp*)
  - “sp” is not legal at start of Spanish word
- Phonotactic cues are language specific

Prosodic Patterns

- Certain patterns are legal both within and across words
  - -rimen-
    - expe**<u>rimen</u>**tal vs. ve**<u>ry</u>** **<u>men</u>**acing
  - Subtle differences in prosody between the two
  - Infants can detect the differences between
    - -rimen-
    - -ry_men-
- Infants are also sensitive to stress patterns
  - 6-9 month old (English-learning) infants prefer the typical strong-weak stress pattern (TAble, CARpet)
  - 10-11 month old infants can identify weak-strong pattern found less typically (girAFFE)

Words in sequential and abstract (structural) patterns

LDER Chapter 4

Learning the form of language

- Using Artificial Languages
  - Real language is really complex!
  - We don’t fully understand it
    - We don’t know exactly what input the infant was exposed to
    - Meaning can’t easily be separated from grammar
- With Artificial languages
  - We control the input
  - We know exactly what the infant was exposed to
  - We can examine grammar separately from meaning
- Can infants start to learn grammar based just on form (before they learn what words mean)?

What is grammar?

- A system for ***<u>generating</u>*** an infinite number of phrases and sentences from a finite set of words
- The grammar of a language enables you to describe which combinations of words belong to the language (are 'grammatical') and which don't.
- A good theory of grammar enables a simple, elegant description of how such phrases and sentences can be produced

Precedence

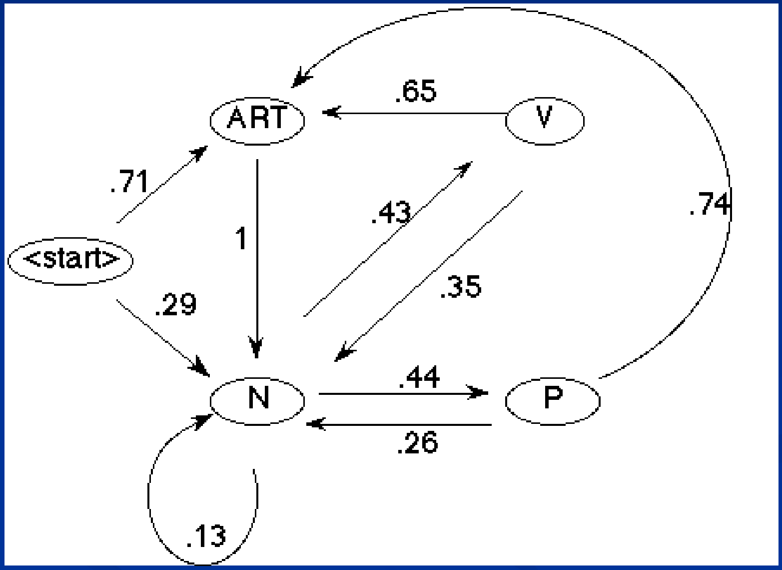
- What order do words go in?
  - big red apple
  - *red big apple
- What order do constituents go in?
  - SVO in English
  - SOV in Japanese; SVO and SOV in German

Dominance

- Language structures are hierarchical
  - High vs. low attachment
    - I saw the robber with binoculars
    - Who has binoculars? (me or the robber?)
- And recursive
    - I saw the robber who saw the burglar with binoculars
    - Who has binoculars? (me, robber, burglar?)

Simple Grammars

- Finite state grammars
  - Simplest grammar that can produce (infinite) recursive sequences of linguistic elements
  - Sequential transition probabilities between successive nodes ("states")
  - Not adequate for real languages!



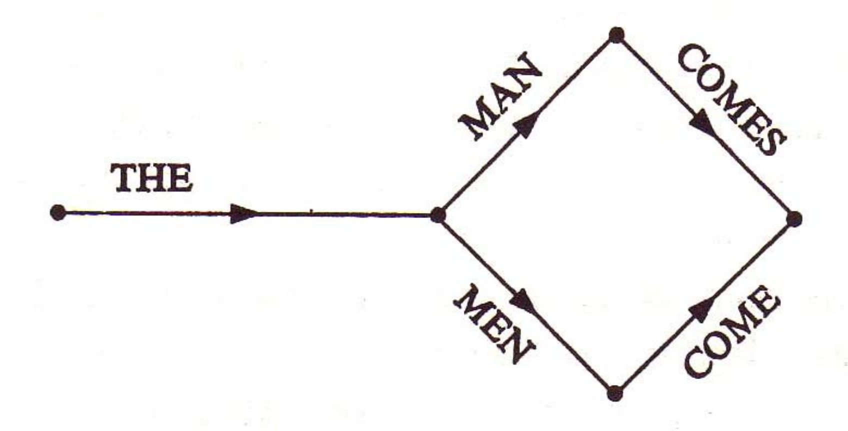
- ART N N V ART N P N
- N P N V N P ART N
- etc.

If language were finite

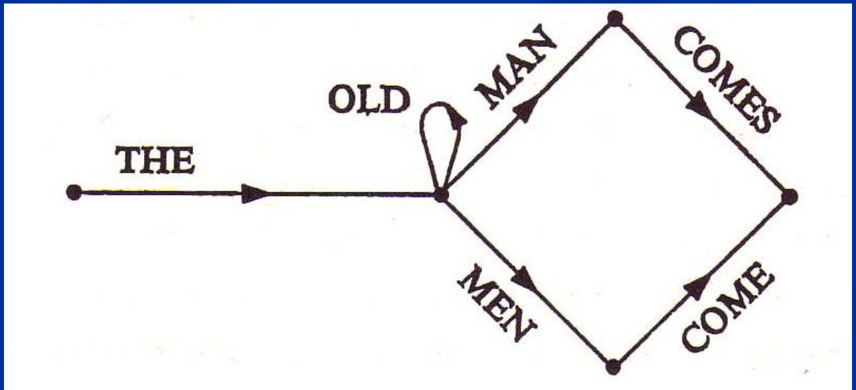
- Finite state grammars would be fine
- For a finite language, it would be possible to simply list the sentences of the language
- A finite state grammar could easily capture that list - but it would be uselessly complex

Assume language is infinite

- Language description gets simpler (generalize across sentences)
  - The man comes
  - The men come



- Add loops
  - The old man comes
  - The old old man comes
  - The old men come
  - The old old old men come
  - etc.



Imagine this without the loop!

Even loops fail…

- Finite state grammars (FSG) fail on two counts:
  - Cannot produce all and only the grammatical sentences of a recursively structured language

1) If an FSG produces all grammatical sentences in a language, it will also produce many ungrammatical ones (“over-generation”)

2) If an FSG is restricted so that it doesn’t over-generate, it will fail to produce many grammatical sentences (“under-generation”)

What can't FSGs handle?

- The distinctions between
  - If (sentence), then (sentence).
  - *If (sentence), or (sentence).

  - Either (sentence), or (sentence)
  - *Either (sentence), then (sentence)

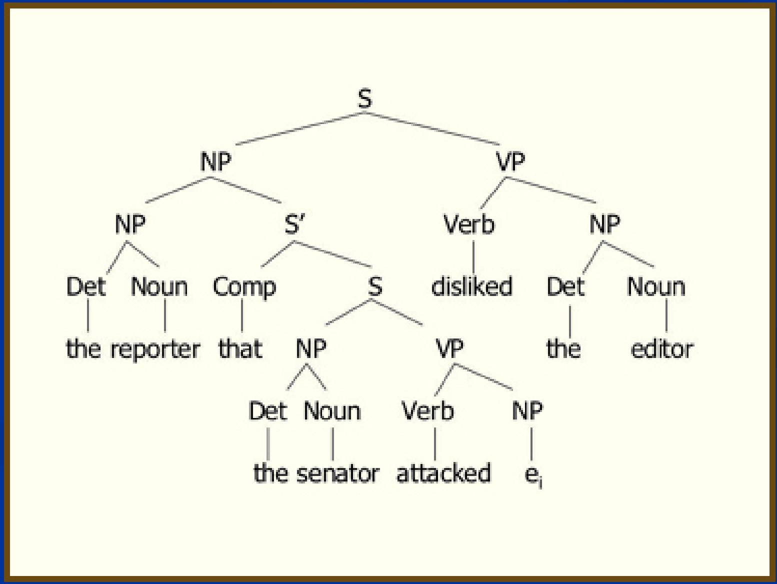
  - If, either (sentence), or (sentence), then (sentence)
  - *If, either (sentence), then (sentence), or (sentence)
- Long distance (and nested) dependencies

More complex Grammars

- Need a grammar for
  - Recursion
  - infinitely long sentences
  - long-distance (and nested) dependencies

S	→	NP VP
NP	→	ART NOUN
NP	→	NP PP
PP	→	P NP
VP	→	VERB NP
VP	→	VERB NP PP
ART	→	the
ART	→	a
NOUN	→	telescope
NOUN	→	man
NOUN	→	spider
VERB	→	saw
VERB	→	complimented
P	→	with
P	→	in

- Phrase structure grammars
  - A superset of finite state grammars



Comparing Grammars

- Finite State

$a \rightarrow b$

New elements are appended

Can be rewritten as:
$(ab)^n$

- Phrase Structure

$Z \rightarrow ab$

$Z \rightarrow aZb$

New elements are inserted

Can be rewritten as:
$(a)^n(b)^n$

Creating longer sentences

$(ab)^n$	n=	$(a)^n(b)^n$
ab	1	ab
abab	2	aabb
ababab	3	aaabbb
abababab	4	aaaabbbb

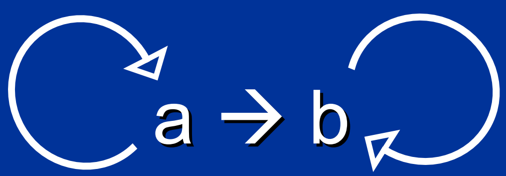
Can't finite state grammars do it?

- Phrase structure grammar can easily be written to produce ***just*** $(ab)^n$ sequences:

  Z → ab

  Z → abZ

- Finite state grammar cannot produce ***just*** the nested sequence (unless all nestings were listed - a very bad solution!)



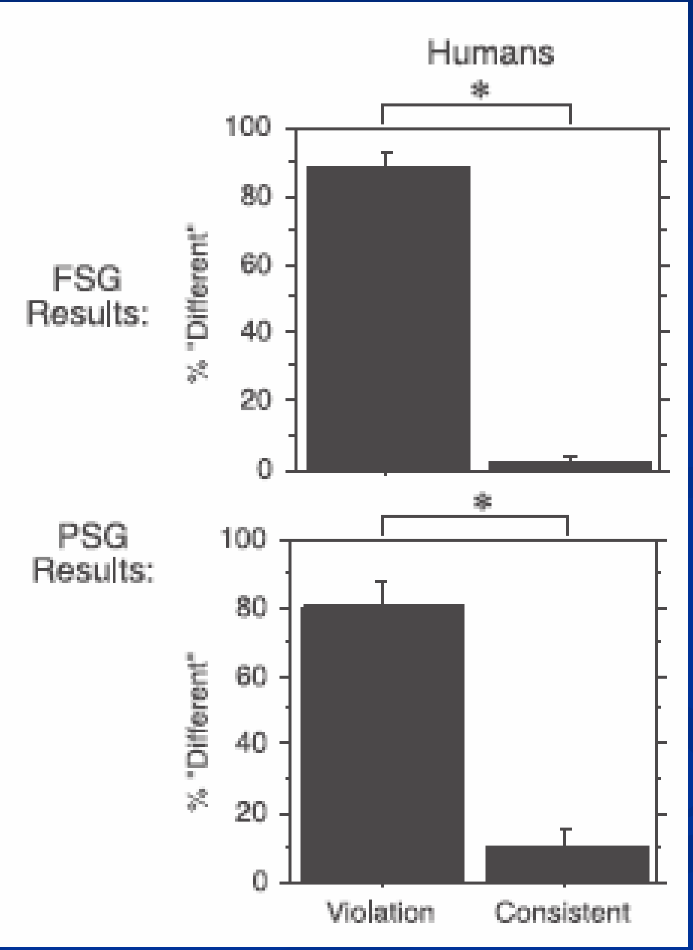
How does this grammar produce aaabbb but not aabbbb, aaaabb, or abbbbb?

Can college students learn grammars?

- Fitch and Hauser, 2004
- Created two grammars of CV syllables
  - Finite state grammar $(AB)^n$
  - Phrase structure grammar A^nB^n
- A = {ba, di, yo, tu, la, mi, no, wu}
- B = {pa, li, mo, nu, ka, bi, do , gu}
- ‘A’ syllables spoken by female voice
- ‘B’ syllables spoken by male voice
- n is restricted to be 2 or 3, to avoid processing limitations

How did the students do?

- Listened to sequences of syllables that conformed to the grammar (implicit learning; 3 minutes)
- Tested with novel sequences that either conformed to the grammar or didn't

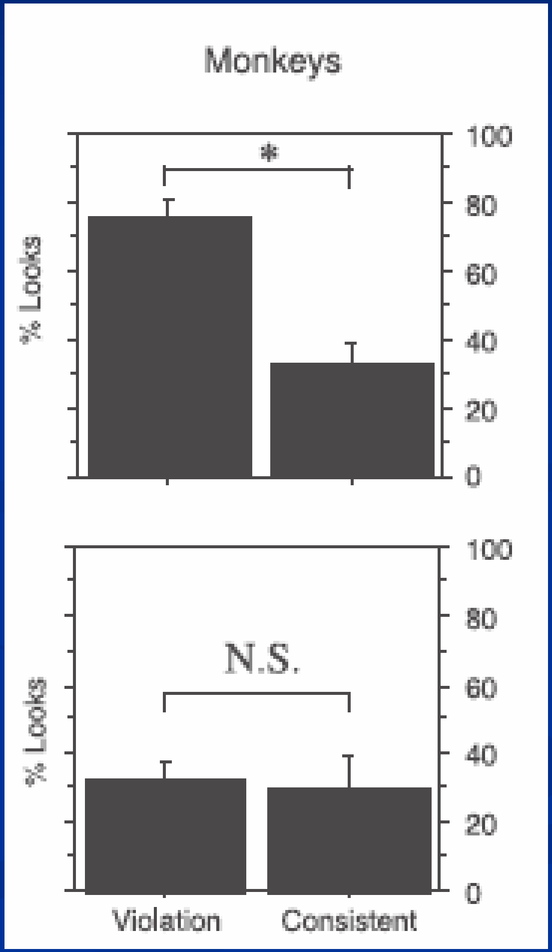


Monkey sequence, monkey do(n't)

- Fitch and Hauser, 2004
- Cotton top tamarin monkeys tested with same FSG, PSG used with college students
- Trained for 20 minutes
- Tested on novel stimuli - longer looks to violations indicate detection of violation

FSG

PSG



What about infants?

- Gary Marcus and colleagues
  - Infants (7 months old) trained on:
    - ABA patterns (wi-di-wi; de-li-de)
    - ABB patterns ((wi-di-di; de-li-li)
  - Infants were tested on (violation detection)
    - same patterns with different syllables (ba-po-ba; ba-po-po)
  - Infants were able to distinguish grammatical from ungrammatical strings, even though all test patterns were new to them
- Did infants learn an abstract grammatical rule?

A limitation?

- Are these syllables the same as language?
- ABA = Noun Verb Noun?
  - ba-po-ba (1st and 3rd elements perceptually identical)
  - Dogs eat pizza (1st and 3rd elements categorically identical)
  - John loves books (1st and 3rd elements categorically identical)
- Maybe the syllables are too simple?

Getting better at words

LDER Chapter 5

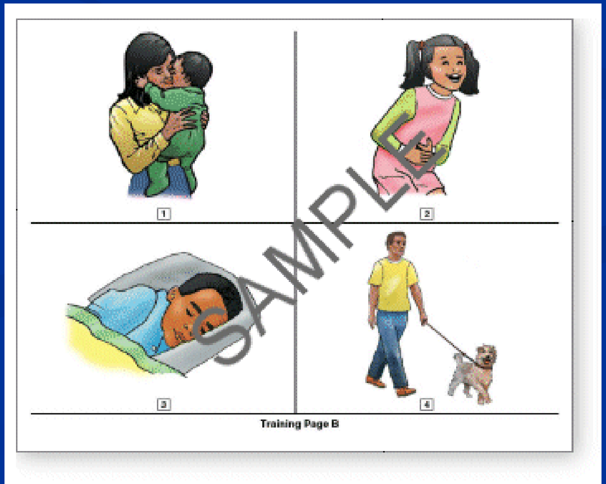
What about words?

- At 12 months, babies just beginning to speak
- At ~18 months, vocabulary burst
- By 24 months, infants can produce 200-500 words

- But this focuses on what babies say, not what they understand!

How to measure?

- How can we measure what words a baby knows?
  - Ask parents
  - Ask child to choose a named object from several options
- Methods for measuring adult understanding of words much better – can we use them with infants?



Infant Eye Tracking

- Infants tend to look at a familiar object when it is named ("ball")
- Even when the name is embedded in a sentence context ("Over there there's a ball")

- With a very time-sensitive measure – we can ask:
  - How quickly does an infant recognize a word?

Results

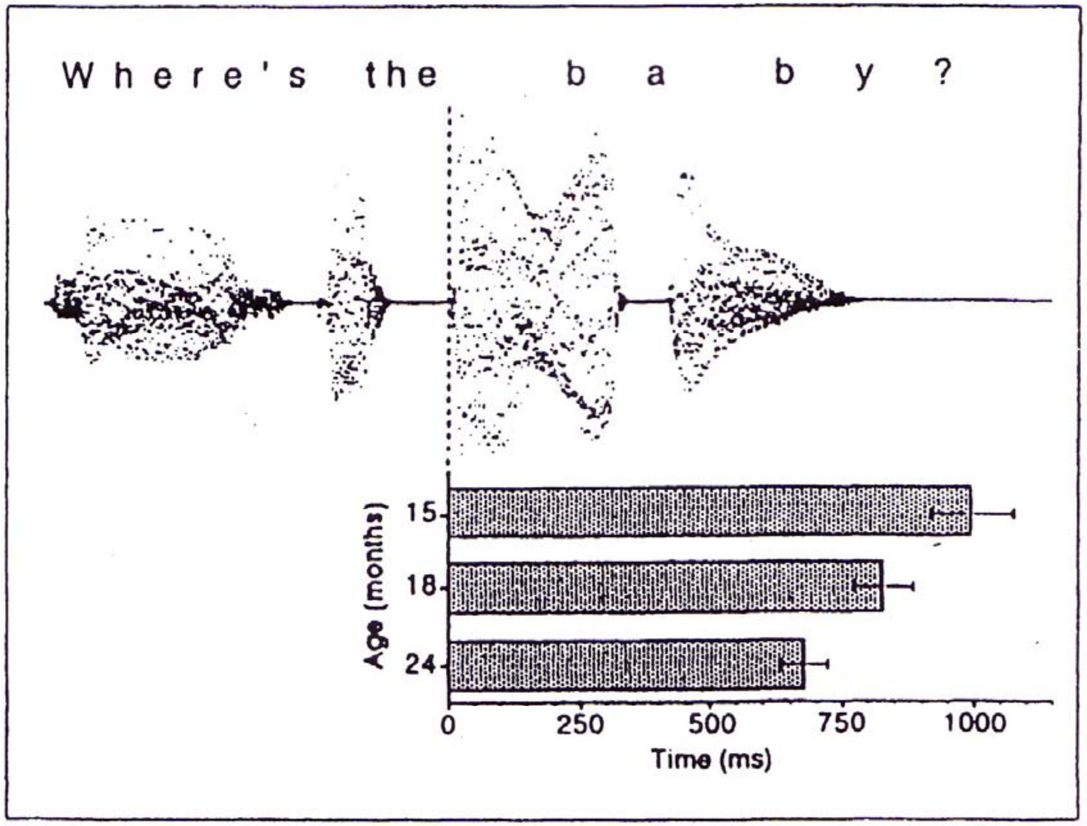


Figure 5.1 Mean latencies to initiate a shift in gaze from the distractor picture to the target picture, measured from the beginning of the spoken target word, for 15-, 18- and 24-month-old infants. This analysis included only those trials on which the infant was initially looking at the incorrect picture and then shifted to the correct picture when the target word was spoken. The graph is aligned with an amplitude waveform of one of the stimulus sentences.

Summary

- Infants gain productive vocabulary quickly towards end of second year
- Infants also get much faster at understanding words they hear!
- Next week we'll start to look at meaning…