Infant Speech Perception: Learning Categories

The story so far:

Infants learn language-specific phonemic categories within the first year of life

A really important question

HOW?

Maye, Werker and Gerken 2002

" In the 17 years since these early language effects were first documented, there has been no demonstration of a definitive mechanism to account for this perceptual development"

Can they come from word learning?

*By 12 mo, L1 English infants can definitely discriminate [ba] vs [pa]

* But: "infants have not been shown to discriminate minimally-different meaningful words (e.g. bear vs pear) **before 17 mos**

See esp. Stager and Werker 97, Swingley & Aslin 2000

So what is directly in the speech signal?

1. Distributions of phones in the speech stream within and between words
2. Distribution of phones within phonemic categories

So what is directly in the speech signal?

-Stochastic (statistical) distributions vs. deterministic distributions

Distributions within and between words

- Word boundaries vs. word internal: frequency of C_1C_2
- Evidence they can be tracked by infants (and monkeys and starlings)
- More about this later – phonotactics and word segmentation

Distributions within Phonemic Categories

- Different tokens (exemplars) of sounds within each category
- Differences between sounds (phones) within a phonemic category smaller than differences between phones from different categories
- Sounds are comprised of multiple acoustic dimensions
- Some dimensions more reliable for between-category differences

Phonetic distributions (usually) reveal phonemic categories

"For any given phonetic category in a language, e.g. [b], actual tokens of the category vary considerably on several acoustic dimensions. However, despite this variation, along certain acoustic dimensions (e.g. VOT) most tokens of [b] … are more similar to each other than to tokens from some neighboring category, e.g. [p]."

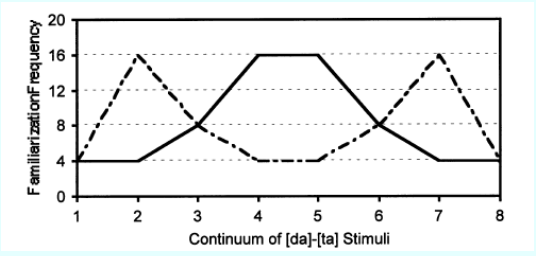
Maye, Werker & Gerken 2002

Purpose: experimentally test infants' abilities to use frequency to infer categories

Populations: 24 6 mos olds & 24 8 mos

Materials:
8 tokens on the da ~ ta VOT continuum, **unimodal** or **bimodal** distributions

Maye et al: Stimuli



Maye et al: Procedure

The method: *Preferential Looking*

Familiarization: 2.3 mins of steady stream of tokens, plus filler [la] [ma]

Testing:

Non-alternating vs. alternating trials

Alternating: half Tk3 ~ Tk 6
half Tk1 ~ Tk8

… *Same frequencies in familiarization!*

Maye et al Results

* No age effect
* Longer looking times for alternating trials in Bimodal infants

Table 1
Mean (SE) looking times for infants in each age group and familiarization condition on Alternating and Non-Alternating trials

	Alternating trials (s)	Non-Alternating trials (s)
6 months Unimodal	4.85 (0.47)	4.53 (0.51)
8 months Unimodal	4.98 (0.63)	5.20 (0.56)
6 months Bimodal	5.66 (0.44)	6.41 (0.32)
8 months Bimodal	5.45 (0.52)	6.15 (0.56)

Infant Speech Perception
Phonotactics and Word Stress

Learning Phonotactics

-What are phonotactics?
-List some English phonotactic patterns
-Phonotactics means that combinations between sounds are probabilistic – not all combinations equally possible or equally frequent
-When do infants pick up on this?

Jusczyk et al 1993: materials

* read by Dutch/English bilingual

* look at the kinds of phonotactic differences

* 10 lists of each

English	Dutch
vacate	structuur
avoid	waardig
lengthen	geslacht
brutal	oprecht
jostle	nerveus
trustworthy	efferent
admission	revolutie
thistle	hersteld
exotic	uitsteeksel
lavish	woestijn
abundant	obstructie
jury	eggen
fluctuate	anderzins
usage	verwant
impact	lading

Jusczyk et al 1993: Exp 1

Population
L1 E 9 month olds

Method
Preferential looking

Results
Longer looking times
Eng > Dutch
(8.94 vs 5.83 secs)

English	Dutch
vacate	structuur
avoid	waardig
lengthen	geslacht
brutal	oprecht
jostle	nerveus
trustworthy	efferent
admission	revolutie
thistle	hersteld
exotic	uitsteeksel
lavish	woestijn
abundant	obstructie
jury	eggen
fluctuate	anderzins
usage	verwant
impact	lading

Why the diff in looking times?

Exp 2: ***Prosody differences****?*

Lists low-passed filtered @ 400Hz
NO diff. in looking times (8.63 vs. 8.67 s)

Exp 3: ***L1-independent preferences?***

Exp 1 lists played to 6 mos olds
NO diff. in looking times (9.39 vs. 9.19 s)

Jusczyk et al revise lists- Exp4

-Lists differ only in illegal **sequences**, not ***sounds**

-More X-ling violations (80%)

SAMPLE WORD LISTS FOR ENGLISH/DUTCH PHONOTACTIC STUDY

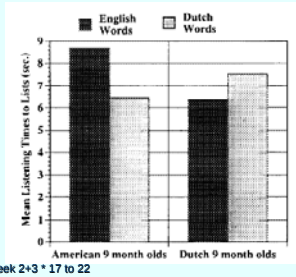
English	Dutch
kudos	zweten
aglow	vlakte
cubeb	kneizen
butane	zalmpjes
skutcheon	knoest
futile	bekomst
aboard	zwendel
cusec	toekomst
dudgeon	knotten
stewed	zwetsen
abound	knabbelen
newfangled	vangst

Comparing English & Dutch

Results:

Sig. interaction btwn list and L1

* L1 E: clearly look longer at E lists
* L1 D: trend to look longer at D lists (p = .24)



Why are Dutch kids less sure?

... Dutch children hear a lot more English than English children hear Dutch!

Checking other explanations…

Exp 5: ***Prosody differences****?*

Exp4 lists low-passed filtered @ 400Hz
NO diff. in looking times (6.48 vs 6.02s)

Exp 6: ***L1-independent preferences?***

Exp4 lists played to 6 mos olds
NO diff. in looking times (8.32 vs. 7.98 s)

Learning word stress

What is there to learn?

What stress is
Stress and syllables
Its phonological effects (V-reduction)

Basic discrimination:
Jusczyk & Thompson (78):
bAda vs baDA @ 2 months

Infants learning word stress

English stress tendencies:

Initial syllables likely to be stressed

* 2-syllable trochees most common
* Cutler & Carter (87): 4% of spoken BNC content words had initial weak syllables
  - and that 70% of strong sylls wd-initial
* Kelly and Martin (1994) on motherese:
  - claims that 95% strong sylls are wd-initial

Jusczyk, Cutler and Redanz 93

Population
L1 E @ 9 months

Method
Preferential looking

Materials
8 wS lists, 8 Sw
same stressed V

SAMPLE LISTS

W/S	S/W
comply	pliant
befall	falter
condone	donor
comport	comet
pomade	neighbor
abut	butter
define	final
restore	stalwart
resent	gentle
assign	sinus
caprice	rhesus

JCR 1993: Exp 1

Results:
9 month olds looked significantly longer at the Sw words ($p < 0.001$)

7.45 secs for Sw
5.43 for wS

What about 6 months?

JCR 1993: Exp 2

Results:
6 month olds showed NO significant difference in looking times

7.68 secs for Sw
7.69 for wS

Q: could this be just *phonotactics*?

JCR 1993: Exp 3

Same lists as Exp1, low-pass filtered

(Note: amplitude was maintained constant from Exp1)

Do you think 9 month olds will still prefer the Sw words?

JCR 1993: Exp 3

Results:
9 month olds STILL showed longer looking times for Sw words
(though it was a bit weaker: $p < .05$)

8.25 secs for Sw
7.37 for wS

Summary of Infant speech perception thus far

- Infants can discriminate phoneme contrasts in (all?) languages when 6-8 mos; only in their language at 10-12 mos
- Infants are sensitive to distribution frequency patterns of phones and can use this to form phonemic categories at 6-8 mos

Summary of Infant speech perception thus far

- Infants are sensitive to distributional frequency and co-occurrence patterns of phones *within words* at 9 months
  - in their native language
- Infants are sensitive to predominant prosodic patterns of words in their native language at 9 months

Infant Speech Perception
More on Perceiving Phonotactics

What kind of learning is phonotactic learning?

What have L1 E 9 month olds learned?
* contingent probabilities,
seg by seg? (P (n|k) = 0.000...)

* ... or featural generalizations?
(No Stop-Nasal onsets)

Saffran and Thiessen 2003 Exp1

Method: *Preferential looking, with twists*

Training: 2 mins. passive listening to nonwords
CVCV (group 1) CVCCVC (group 2)

Segmenting: Passive listening to continuous stream of 4 novel words:
2 CVCV, 2 CVCCVC (26 X each)

Testing: Preferential looking, two speakers, repeating1 of the 4 seg. phase words

Saffran and Thiessen 2003 Exp2

In Exp1, **9** month olds listened longer to *familiar words* -- the ones that matched the training phase CVCV or CVCCVC template ($p < 0.05$)

Now that we know the methodology works:
can 9 month olds learn
new phonotactics this way?

Saffran and Thiessen 2003 Exp2

Training Materials

CVCCVC words:

* voiced onsets, voiceless codas OR
* voiceless onsets, voiced codas
(i.e. *todkad* or *dakdot*)

Saffran and Thiessen 2003 Exp3

Results

9 month olds looked significantly longer to *novel* words, that did NOT match the voicing template

- regardless of WHICH template!
(7.06 sec vs 6.05; p = 0.01)

Saffran and Thiessen 2003 Exp3

Materials -- the crucial study

CVCCVC words:

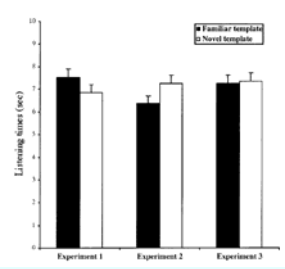
* onsets = btg & codas = pdk OR
* onsets = pdk & codas = btg
... these are NOT natural classes

Saffran and Thiessen 2003 Exp3

Results

9 month olds showed no significant difference in looking times to familiar vs. novel words, with either training phase (p = .74)

Saffran and Thiessen 2003 Sum



Cristia and Seidl 2008 Exp 1

Previous research used groups of sounds with prominent acoustic properties
Here: **stops** & **nasals** (natural) vs.
fricatives & **nasals** (arbitrary)

Familiarization Pseudowords:
C_1VC_2 words: C1 = /m, n, 2 stop/ (natural)
C2 = /m, n, 2 fric/ (arb.)
Stops: b,t,g,k Frics: v,f,z,S

Cristia and Seidl 2008 Exp 1

Population: L1 Eng 7 months

Procedure: Modified preferential HT

Familiarization: recording of pseudowords played continuously...

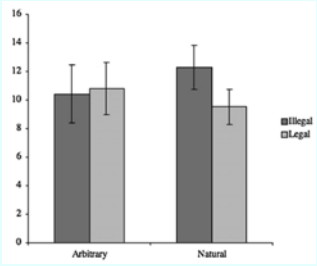
Testing: **Novel** lists activated by a HT:
* novel legal (e.g. tVC, gVC) for natural
* novel illegal (e.g. fVC, zVC) for natural

Cristia and Seidl 2008 Exp 1

Results:

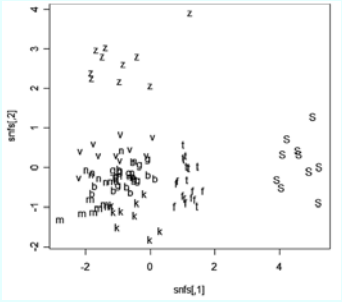
Sig. longer looking: illegal > legal

for Natural condition only



Alternative explanations?

* Fricatives are just harder to perceive as belonging to the same group?
* Acoustically, some but not all frics are weird: (graph)



Cristia and Seidl 2008 Exp 2

Materials

Group 1: stop onsets only (e.g. bVC, kVC)
Group 2: fric onsets only (e.g. fVC, SVC)

Procedure otherwise the same
testing: tVC, gVC (novel legal)
vVC, zVC (novel illegal)

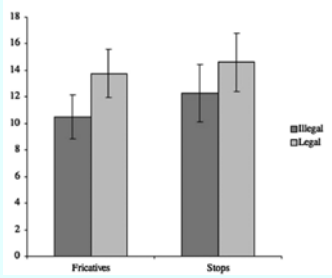
Cristia and Seidl 2008 Exp 2

Results

* NO interaction btwn group and legality

Main effect of legality

i.e.: both discriminated



Cristia and Seidl 2008 Sum

"In conclusion, the experiments presented here suggest that, since not all groups of sounds are equally learnable by young infants, it is unlikely that infants merely track the co-occurrence of sounds in phonological patterns in order to posit features that refer to those sounds."

Word Boundary Phonotactics

Two intertwined questions:

* How do you learn to pick words out of continuous speech?

* How do you learn the phonotactics that hold at *word boundaries*, which are often special? (e.g. 'sixth'…)

Word Boundary Phonotactics

Today:
When are infants sensitive to word-boundary phonotactics?

Next unit in the course:
When/how does word segmentation happen?

Friederici & Wessels 1993

When do infants become sensitive to word-boundary phonotactics?
(continuation of Jusczyk et al 93)

Then: when/how can they use those phonotactics to segment spech?

Friederici & Wessels 1993 Exp 1

When do infants become sensitive to word-boundary phonotactics?

Populations
L1 Dutch infants 4.5, 6 & 9 months

Method: Preferential HT

Friederici & Wessels 1993 Exp 1

Materials:
Monosyllabic nonsense Dutch words

Illegal clusters created by using opposite-edge cluster:

bref vs *febr
murt vs *rtum

Friederici & Wessels 1993 Exp 1

Results:

9 month olds looked significantly longer at legal lists (7.7 sec vs. 6.0)

4.5 & 6 month olds did not.

Okay, good.

Friederici & Wessels 1993 Exp 2

Using phonotactics to segment speech?
Method: Same preferential HT
Materials: exposure to 4 syllable strings
licit clusters: /migbrefpar/
mig bref par
Three licit "words" can be segmented

illicit clusters: /mifri**sgk**ur/
at least one word MUST be illegal when segmented

Friederici & Wessels 1993 Exp 2-5

9 months old *do* look longer at lists with embedded legal words... ($p < 0.05$)

But only:
with illegality restricted to codas
- with Infant direct speech
- with a short ISI (800 ms, not 1.25 s)

So... Clearly this skill is just emerging

Summary of Learning Phonotactics

- Infants are good statistical learners, but the do not seem to learn phonotactics strictly by transitional probabilities
- Infants' learning of phonotactics is phonologically relevant

References

Cristia and Seidl, 2008. 'Is Infants' Learning of Sound Patterns Constrained by Phonological Features?" *Language Learning and Development* 4(3): 203-227.

Friederici and Wessels, 1993. 'Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics* 54 (3): 287-295.

Maye, Werker and Gerken 2002. 'Infant sensitivity to distributional information can affect phonetic discrimination' *Cognition* 82: B101-111

Hillenbrand, 1984. Speech perception by infants: Categorization based on nasal consonant place of articulation. *Journal of the Acoustical Society of America*, *75*, 1613–1622.

Jusczyk, Cutler and Redanz 1993. 'Infants Preference for the Predominant Stress Patterns of English Words.' *Child Development* 64: 675-687.

Jusczyk, Friederici, Wessels, Svenkerud and A-M Jusczyk 1993. 'Infant sensitivity to the Sound Patterns of Native Language Words. *Journal of Memory and Language* 32: 402-420.

Saffran & Thiessen, 2003. 'Pattern Induction by Infant Language Learners'. *Developmental Psychology* vol 39 no. 3: 484-494.