Variation in First Language Development

Stoll & Lieven

Lecture 2:
Crosslinguistic studies of infant perception and production

Introduction

Language acquisition doesn’t start with syntax!

Issues:

- How does the child relate the acoustics of ambient language to articulatory realisation
- How does the child segment the ambient language into morphemes, words, phrases and clauses

Since each language differs in its phonetics, phonology, morphology, word stress patterns etc., this is where we should start to look for universals and particulars in acquisition

Theories

- An innate, universal set of distinctive features and realisation rules
  - Child uses these to develop the phoneme inventory of the ambient language, gradually losing distinctions and rules that are not relevant
- An acquired inventory of the relevant distinctions
  - based on mammalian perceptual features, distributional learning and human cognitive capacities

How do we work out what children know when?

Problem: Babies can't talk!

Methods

- **Conditioned head turn preference procedure (HTPP).**
  - Children trained that (for example) English audio is looped when they fix gaze on one loudspeaker; Russian audio another loudspeaker.
  - Can be used with very young children
  - Interpretation can be problematic. Listening time difference means children can *tell the difference* between two stimuli, but not necessarily anything more.
- **High amplitude sucking/kicking (HAS).**
  - Similar to HTPP but playback controlled by sucking or (pre-natally) kicking rate
  - Can be used with children only a few days old (or even unborn).
  - Can be used to investigate knowledge of language only at a coarse level (e.g., do infants prefer mother’s voice to that of another speaker).
- **Habituation.**
  - Used together with some looking/listening time measure.
  - Children become bored with repeated presentations of similar stimuli and looking/listening times decrease. If a stimulus from a different category causes recovery of looking times, this is evidence that children have formed the two categories.
  - Can also be used with very young children.
  - If some children show habituation (novelty preference) and others a familiarity preference, these may cancel each other out when group means are taken, potentially hiding effects.

Brain-Behaviour measures

- **Event-related potentials (ERPs).**
  - Recorded from scalp electrodes in a cap
  - Can be used with very young infants.
  - Good at detecting relationship between stimulus and response
  - Bad at detecting localisation.
  - Problems with relating infant measures to better specified adult measures
- **Functional magnetic resonance imaging (fMRI).**
  - Measures changes in blood flow relating to neural activity in the brain
  - Good at detecting localization
  - Difficult to use with infants; Hence usually carried out when infant asleep or sedated, but can be done on waking infants with careful protection procedures.
  - Poor temporal resolution, so hard to distinguish responses that occur within a short time window.

Early sensitivity to speech

- Newborns can discriminate their mother's voice from other female voices at birth
  (De Casper & Fifer, 1980)
- Newborns listen longer to their 'native' language
  (Spanish and English: Moon, Cooper & Fifer, 1993
- Newborns can discriminate two languages not their own
  (French children with Italian and English, Mehler et al., 1988)
- but 2-month-olds cannot
  (American infants with Russian-French, Mehler & Christophe, 1995)

Infants are responding to the 'prosodic package' of the language as a whole

Discriminating the metrical stress patterns of languages

- Stress-timed languages:
  - Stressed syllables with strong vowels alternate with unstressed syllables with reduced vowels (English, Dutch)
- Syllable-timed languages:
  - Each syllable receives equal stress (Italian, Spanish)
- Mora-timed
  - Evenly timed but the unit of timing is the mora, counted in terms of whether syllables are long or short (Japanese, Luganda)

- Newborns can only discriminate languages that do not come from the same rhythmic class
(Nazzi et al.)
- Also when one of the languages is their own
(Christophe & Morton, 1998)
- 4-5-month-olds can discriminate within rhythmic class provided one of the languages is their own
(Bosch & Sebastian-Galles, 1997; Nazzi, Jusczyk & Johnson, 2000)

Increasing refinement of discrimination as a function of experience with ambient language

Consonants

- Infants can detect differences in voice onset time, contrasts in manner and place of articulation and many other phonetic contrasts that occur in the world's languages
- These contrasts appear to be perceived categorically i.e. in terms of two clear categories although the underlying acoustical characteristics of the stimulus are continuous

→ Categorical perception (Eimas et al., 1971)

Categorical perception

(Eimas et al., 1971)

English:
If VOT less than 25msec, adults report hearing /ba/
If VOT more than 25msec, adults report hearing /pa/

Using the HAS procedure:

- Group 1: VOT: +20msec → /ba/
  +40msec → /pa/
- Group 2a: VOT: -20msec → /ba/
  0 → /ba/
- Group 2b: VOT: +60msec → /pa/
  +80msec → /pa/
- Group 3: Control group

Discriminating non-native contrasts

Infants can discriminate phonetic contrast that are not phonemic in their native language

e.g. Prevoiced-voiced contrast in stop consonants in Thai

(Aslin et al., 1981 using the HTPP procedure)

The HTTP procedure

1. Conditioned to turn head to a visual reinforcer in response to one audio stimulus but not another
2. Conditioning is established (e.g. 9 out of 10 ‘correct’ turns to the reinforced stimulus
3. Present a new, non-reinforced stimulus

If children turn their heads to the new stimulus
→ cannot tell the difference

Vowels

- Vowels show graded rather than categorical contrasts in perception
- 1m and 4m can distinguish vowels in isolation and in syllables, even when not in their ambient language (Trehub, 1973)
- Prototypical structure develops (Kuhl, 1979, 1983):

$\boxed{\text{perceptual magnet model}}$

Developmental effects of the native language

Increased sensitivity to native language:

- 6m tested on vowel prototype experiment showed the prototype effect but 4m did not (Polka & Werker, 1994)
- 9m prefer to listen to consonant clusters permitted in their native language but 6m did not (Best et al., 1988)

Decreased sensitivity to non-native contrasts:

- 6-8m could discriminate non-native consonant contrasts but 8-10m less so and 10-12m not
  - English /ba/,/da/
  - Thompson /k'i/, /q'i/ (Werker & Tees, 1984)

Interim conclusions

- Categorical perception is probably innate
  - but not specifically linguistic
- The development of prototypes may be a unique feature of human cognition
  - but may not be specifically linguistic
- Clear developmental progression:
  - prosody → metrical stress → native consonants and vowels → native phonotactics
  - showing an increasing effect of ambient language on these features

So is there evidence for full adult representations in young children?

- Early discrimination
- Children can detect differences they cannot produce (Smith 1973)
- Children can only detect some changes to familiar words
  - 7m familiarised on */dog/* did not detect change to */bawg/,* Juszcyk & Aslin, 1995)
  - onsets but not offsets: 11m Swingley, 2005, 16m, Vihman et al., 2004)
- Children familiarised on one sound to object mapping do not fully succeed in discriminating phonetically similar words until 17m (Stager & Werker, 1997; Werker et al., 2002)

Theory: A universal phoneme inventory?

- Innate distinctive features
  - Smith – generative phonology
  - Stampe and Donegan – natural phonology
  - Prince & Smolensky - optimality theory

- Bottom-up development of sound inventory
  - Pierrehumbert –exemplar-based theory
  - Mielke – emergent language-specific distinctive features
  - Blevins – evolutionary phonology

Segmenting words

- Many cues to word segmentation:
  - Isolated words
  - Stress patterns
  - Phonotactic and allophonic regularities
  - Transitional probabilities

Metrical segmentation strategy

Many languages have a predominant stress pattern. In English, it's strong/weak

Table, Football, Biscuit, Sofa

Metrical Segmentation Strategy

Positing a boundary after each strong/ weak syllable pair would often yield the correct segmentation

JohnnywatchesfootballeveryTuesday

Metrical Segmentation Strategy

Can infants (7.5 mnths) actually do this?
(Jusczyk, Houston & Newsome 2000):

Training: Words in isolation (chosen to be unfamiliar): Kingdom , Hamlet

Test: Measure listening time to passages that do and do not contain these words.

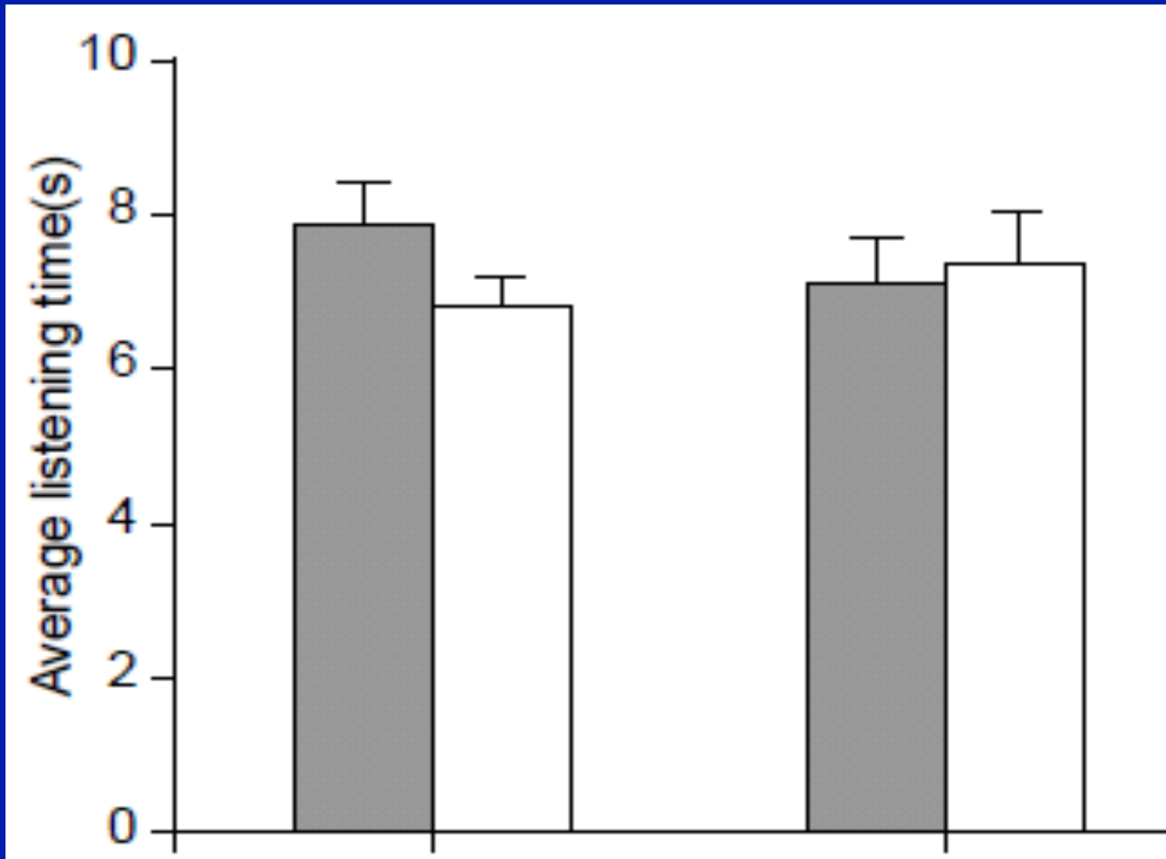
Result: “Prefer” those that do.

Metrical Segmentation Strategy

So, children could segment the test passages well enough to pick out the previously-presented words (e.g., kingdom , hamlet), but how do we know they were doing this on the basis of the strong/weak stress pattern (meter)?

When the training words had a weak/ strong pattern (e.g., device, guitar), the effect disappeared.

Metrical Segmentation Strategy



Test passage
Does
Doesn't
contain training words

strong/weak (e.g., kingdom/ hamlet)

weak/strong (e.g., device, guitar),

Metrical Segmentation Strategy

Problem 1: This works pretty well for English, as 90% of two-syllable words follow the dominant pattern (strong-weak)

But we don't know whether this is true for all languages. There may be some where the dominant pattern accounts for only slightly over 50% of cases or where there is no word stress pattern.

Metrical Segmentation Strategy

In fact, we know children must be using some other strategy as well, as by 10.5 months children can also "recognize" weak/strong words (e.g., guitar) in test passages (i.e, they can correctly segment passages containing words that violate the usual weak/strong pattern).

Another possible cue is….

Allophonic Cues

/t/ is aspirated at the beginning and internally in English words but unaspirated in codas

tap = aspirated   cat = unaspirated

Allophonic Cues

Nitrates or Night rates

Aspirated within/at start of word

Unaspirated t means it must start a new word

Allophonic Cues

So, can children do this?
(Jusczyk, Hohne and Bauman, 2000)

Children hear “nitrates” presented in isolation and then listen to

Passage a) - Contains nitrates
Passage b) - Contains night rates

Listen longer to (b) at 10.5 m, NOT at 9m

Statistical Cues

Transitional probability is the probability that, after a given syllable, a particular syllable will occur

E.g.,	pre→tty	0.8
	pre→dict	0.1
	pre→cise	0.1

If, after pre, the next syllable is tty 80% of the time, the TP of pre→tty is 0.8.

Statistical Cues

How does this help a learner?
If you hear a pair of syllables for which the transitional probability is low, there is probably a word boundary there.

E.g., ty→ti(me)	0.5
ty→a(nd)	0.5
ty→ba	almost zero

If the TP is high, there probably isn't (pre→tty)

Statistical Cues

If you hear a string such as prettybaby, you can use TPs to find the words

pre→tty→ba→by

TP of pre→tty is high, so don't posit a word boundary

TP of ty→ba is very low (almost zero) so DO posit a word boundary

Statistical Cues

So can children actually do this?
Saffran, Aslin & Newport (1996): 8m olds

bi-da-ku-pa-do-ti-go-la-bu-bi-da-ku

1.0 1.0 0.33 1.0 1.0 0.33 1.0 1.0 0.33 1.0 1.0

Only cue to word boundaries is the transitional probability information: High between syllables within a “word” (e.g., bi→da); low between syllables across a “word” (ku→pa). No gaps in stream!

Statistical Cues

Children listened to this for 2 minutes. Then their listening time was measured for:

A) Repetitions of words: “bidaku, bidaku, bidaku” (1/2 sec gap between each)
B) Repetitions of non-word “kupado, kupado, kupado”

They preferred to listen to B.

Statistical Cues

Problems:

1) How do children know to compute transitional probabilities over syllables (as opposed to individual phonemes, for example)?

2) When run on actual child-directed speech, accuracy is low. 60% of “words” posited by a TP computer model are not in fact words, and 80% of words in the speech are not found (Gambell & Yang)

Statistical Cues

Why? It works well when the words are multisyllabic: Learner posits a boundary when there is a TP that is low relative to others in the stream (e.g., 0.33 vs 1)

In real CDS most words are monosyllabic. This means that the TPs are uniformly low, so the learner has no idea where to place the boundary:

Johnandbillsatonthefence

General Problems

- The cues are all probabilistic
- The child needs to know that what has been segmented IS a word
  - *Mum me* vs *Wozzat*

A chicken and egg situation

Children need to know the patterns followed by their language in order to identify words (e.g. whether their language follows the strong/weak or weak/strong pattern).

But they can't do this until they segment some words out of the speech stream and look to see whether they are weak/strong or strong/weak.

But they can't segment words out of the speech stream until they know…

Contrasting solutions

Nativist:

- The Unique Stress Constraint (Gambell & Yang, 2003, Yang, 2004)
  - Any utterance with just one stressed syllable is a word: *ChewBACca* vs *DARth VAder*
  - Stress/Phonotactic/Articulatory cues can begin to be learned as soon as a few words have been segmented by USC

- But many languages do not have word level stress (e.g. Japanese)

Contrasting solutions

Constructivist

- It does not matter if initial segmentations are not all words
  - It may be that much of the early lexicon is only roughly matched to that of the adult and consists in part of both mis-segmentations and unanalysed strings (Bannard & Matthews)
- Meaning is left out of the nativist account
  - When a child can pair a string with an entity they know well (e.g. *Mummy* or their own name), meaning fixes this as a word and this then allows them to make use of other cues (Tincoff & Jusczyk 1999)

Production

- Very early words similar across languages
  - Universal unfolding of contrasts between distinctive features (Jakobson, 1941)
  - Universal rules (Smith, 1973)
  - Universal initial ranking of constraints (Kager, Pater & Zonnefeld, 2004)

→ Rules or constraints operating on segments to produce harmony, metasthesis etc.

- The sounds produced in late babbling are the sounds produced in early words (Locke, 1983, Oller et al.1975, Boysson-Bardies & Vihman, 1991)
- There is an influence of ambient language on babbling and early words

Whole-word approach

(Vihman & Croft 2007)

- Children are trying to produce words with meanings
- They have articulatory constraints which is why early words sound similar across langauges
- They store underspecified whole words/strings with meanings
- They develop individual templates which have some language-specific characteristics but also show differences between children learning the same language
- They select words that fit the templates and adapt words that they want to produce to the templates

Two examples

- 18 months, P. learning English:

  A template with an initial ‘optional' stop, followed by a vowel followed by [ʃ].

  [byʃ] (*brush*), [dyʃ], (*dish*) [iʃ] (*fetch*) [iʃ] or [uʃ] (*fish*) and [uʃ] (*vest*) (Waterson, 1971).

- Madli, learning Estonian:

  A template that started with an ‘optional’ /p/ or /t/ followed by a vowel and ending with /s/.

  [is:] (for *isa, issi* ‘Daddy’), [as:] (for *kass* ‘kitty’), [pis:] (for *piss* ‘pee’), [us:] (for *suss* ‘slipper’), [tis:] (for *tiss* ‘teat’) and [us:] for (*uss* ‘snake’) (Korgvee, 2001)

Arguments for the whole word approach

- A child can produce sounds that are the same in the adult target differently in different words.
- It is difficult to account for the relationship between the adult target and the child production on a segmental basis
- A child’s words may be more similar to each other than is reflected in the differences between the adult targets

Conclusions

- Early perceptual sensitivity does not necessarily mean that infants either start with, or develop, a fully segmental phonology in advance of production.
  - But there are discrepancies between perception and production that need to be resolved through experimentation with the same children
  - Could representations for perception and production be separate?
- Despite widespread assumptions to the contrary, it is not yet clear how children segment linguistic units out of the speech stream
  - Modelling that varies probabilities and uses more than one method is required
- The point at which infants connect the sounds they are hearing to communicative function (roughly at 7-9 months) is crucial to the emergence of language-specific phonological contrasts.
- There is some rapprochement between some more recent versions of OT and templates approaches
- Urgent research is needed on the perceptual and productive development of children exposed to agglutinative and polysynthetic languages in which words are much longer and have meaningful internal structure