PSYC 3640
Psychological Studies of Language

Speech Perception

September 18, 2007

1

PSYC 3640
Psychological Studies of Language

Today’s outline

• Administrative stuff
• Brief review of Lecture 1
• Altmann’s chapters 2 & 3
  – Techniques in testing infants
  – Physical and psychological properties of sound
  – Infant perception
  – Revisit: Is that a uniquely human behaviour?
• Vouloumanos & Werker (2007)

2

Brief review of Lecture 1

• Course outline, structure and related information
• Studying language from psychology (as opposed to linguistic, sociology or philosophy)
• History of “scientific” studies in Science
  – Early studies of language were not exactly scientific
  – Philosophical, linguistic
  – Is language uniquely a human behaviour?
  – Structures of (human) language

3

Speech Perception

• On a developmental trend, we know that speech perception precedes speech production.
• Speech perception starts not only before acquiring language, but even before birth!
• Is speech sound different from random noise?
• How do infants distinguish them?
• Methodologically, how do scientists study speech perception in infants?

4
Testing Infants

- What can babies do?
- Non-nutritive sucking:

Testing infants

- Habituation/dishabituation: Infant's sucking rate decreases after a stimulus is presented for some time. But the sucking rate increases again when a new stimulus is presented.
- Possible problems of this technique?

Hearing in utero

- Human auditory system starts to function at around 7 months from conception.
- But what's it like hearing sounds in utero?

Sound

- Vibration of air causes a vibration of a membrane in the inner ear
- Frequency: number of occurrence in a given duration.
- Amplitude: intensity of sound waves
- Hz = cycle per second
- Human (male and female combine) hearing frequency ranges from 20Hz to 20000Hz
- Human speech ranges from 100Hz to 4000Hz
Sound

- Psychological property:

![Sound Waveform](http://en.wikipedia.org/wiki/Frequency)

Human Ear

![Human Ear Diagram](http://www.seahi.org/images/the_ear.gif)

Hearing in utero

- Sounds are distorted in utero.
- Prosodic factors:
  - Intonation → melody of language
    - "We aim to please. You aim too, please" (Fromkin & Rodman, 1974)
  - Rhythm → depends on where the stress falls
    - a computer un ordinateur konpyu-ta
  - Stress → where the emphasis of a syllable falls
    - "chimpanzee"
- Prosodic variation: physical variation in sounds that triggers the psychological variation in intonation and rhythm.

Examples

Dear Mum and Dad: Hi! How are you? Well, here I am in the big city. Although the weather is nice at the moment, the forecast is for hail, but that should soon clear. I bought a new coat yesterday because they say it gets really cold. I have to stay at Aunty Deb's house for now, but I'm hoping to get a flat soon. The trip up was great, even though it took ten hours. Well, I must go. You know how rarely I write, but I will try to do better this year. Love Clare

[http://www.otago.ac.nz/anthropology/Linguistic/Accents.html](http://www.otago.ac.nz/anthropology/Linguistic/Accents.html)
Speech perception in infants

- (Mehler) Using the habituation/dishabituation method, it was shown that 4-days-old babies were able to distinguish two languages (French and Russian) based on familiarity before birth.
- (DeCasper) Let pregnant mothers read stories for the last 6 weeks of pregnancy. Can the babies distinguish the prosody of the stories? YES!! They preferred the familiar story.

Prosody

- Why is it so important?
- It tells us where does a word begin and end  \( \rightarrow \) word boundaries
- Syllables are the basic "sound boundaries" of a word.
  - Syllable by itself can be meaningful or meaningless
  - Given a few meaningful syllables, their combination may or may not mean the same thing by themselves.
  - Non-speech sounds do not have syllables  \( \rightarrow \) distinguishing speech from non-speech sounds

Syllable and Phoneme Perception

- Babies can distinguish /p/ and /t/  
- [pat] \( \neq \) [tap]
- [pst] = [tsp]
- Do you know of any word that has the syllables [pst] or [tsp]?
- Illegal syllables are not distinguished by babies.
- (Mehler) After adding a vowel that “legalize” one of the illegal syllables, [uptsu] vs. [utpsu], babies can differentiate the two syllables.

How do babies know?

- Phoneme or syllable gene?  \( \rightarrow \) Language gene?
- Well, sickness runs in families, but so do many other things, like recipes and wealth… (Pinker, 1994)
- Change in syllable \( \approx \) Change in prosody
- What’s in a syllable?
Infants vs. Adults

- Experience?
- Linguistic experience?
- Vocabulary?
- lexicon!
- But does speech perception require lexicon? Not really…
- Then, what's so special?

Phoneme

- Words/syllables with single different phonemes have different meanings:
  /mat/  /bat/
- /b/ and /p/ differs in subtle vibrating action of the vocal folds
- Voice onset time (VOT): The different timing when the vibrating action starts in the vocal folds. For voiced sounds, the vibration starts immediately. For voiceless sounds, it starts with a small delay.

Voice Onset Time (VOT)

[Graph showing voice onset time for "En pæ" and "En bli" with annotations p and b]

Phoneme perception illusion: The McGurk Effect

[Image of a person with a speech bubble saying "da" and a computer screen saying "ba"]

http://www1.ldc.lu.se/~logopedi/department/andy/Perturbations/VOT.html
Categorical Perception

- Vowel durations are generally longer than consonants.
- Unlike consonants, vowels are perceived continuously rather than categorically.
- (Studdert-Kennedy, 1975) Vowels carry stress, rhythm and prosody, which have an “echo” after production.

Phoneme Continuum

http://cfa-www.harvard.edu/~jbattat/a35/wavelength_color.html
Categorical Perception

• (Eimas) One-month-old babies can do it!
• Not only in their only “native” languages, but also in “foreign” languages!
• This ability is lost at about 10 mos.

Why categorical perception cannot be innate?

• Non-speech sounds such as musical tones can also be perceived categorically.

Æ categorical perception is not limited to speech sounds
Æ categorical perception only applies to consonants, not vowels
• Chinchillas do it too!
Æ not a uniquely human behaviour
Æ not speech-specific, but auditory-specific

Kuhl & Miller (1975)

Abstract: Four chinchillas were trained to respond differently to /t/ and /d/ consonant-vowel syllables produced by four talkers in three vowel contexts. This training generalized to novel instances, including synthetically produced /da/ and /ta/ (voice-on-set times of 0 and +80 milliseconds, respectively). In a second experiment, synthetic stimuli with voice-onset times between 0 and +80 milliseconds were presented for identification. The form of the labeling functions and the “phonetic boundaries” for chinchillas and English-speaking adults were similar.


Fixed Boundaries in Categorical Perception?

• Boundaries of the /b/ (< 20ms) and /p/ (> 40ms) are influenced by speech rate.
• Speech rate:
  – amount of time spent on articulating an utterance
  – number and length of pauses during utterance
• Rate ↑: vowel duration, VOT ↓
  → VOT ↓, the boundary between voiced and voiceless consonants shifted towards the shorter end, hence harder to differentiate
  /g/ → /k/ when rate ↑
Chapters 2 & 3
• Sensitivity to language starts before birth.
• Infants are sensitive to prosody in language(s) even before they are born.
• After birth, infants show sensitivity to the smallest unit of spoken language, phoneme.
• The ability to perceive phoneme categorically could be related to auditory system, not specially to speech.
• Boundary in phoneme categories are context-dependent and can be influenced by speech rate.

Vouloumanos & Werker (2007)
Listening to language at birth: Evidence for a bias for speech in neonates
• Developmental Science, 10, 159-171

Introduction
• Do babies show a bias to language, the communicative tool?
• Previous suggested neonates could differentiate
  – speech from non-speech sounds
  – Other linguistic properties of speech
• Brain
• Not surprising that neonates chose folk music to white noise.

Methods
• Use physically comparable speech and non-speech sounds as stimuli
• Non-speech sounds are sine waves modeled after natural speech
• Contingent sucking responses as preference for speech vs. non-speech sounds
• 22 neonates (1-4 days old)
• Tested 2 hours after feeding
• Baseline: sucking amplitude in 1min silence
• Stimulus presented when sucking amplitude is in the 80% of the baseline range
**Timeline**

- Baseline
- Experimental Block 1
- Experimental Block 2
- Silence
- Speech and non-speech stimuli alternate every minute

**Speech vs. Non-speech Stimuli**

**Results**

- First 4 mins
- Last 4 mins

**Conclusion**

- Human neonates have a listening preference for speech.
- Similar to other species’ adaptation to auditory signal from the same species.
- Children who were later diagnosed to have language difficulty do not show this bias.
- Question 1: prenatal or experiential?
- Question 2: what speech aspect was preferred?
Rosen & Iverson’s commentary

- Results crucially rely on the speech and non-speech stimuli.
- Revised conclusion: Neonates prefer to list to full-blown speech sounds compared to sine-wave analogues.
- Poor controls... → there was no voice melody (prosody??) in the non-speech stimuli.
- “Human neonates are biased to listen to sounds with a strong voice melody”
- Preference develops in utero

V&W’s response

- Voice melody (pitch) is a subjective perception. The component chosen in the stimuli was an appropriate formant to differentiate multiple natural speech.
- Prenatal ≠ innateness
- Using low-pass filtered (LPF) sounds stimuli, no preference was shown.
- Information for discrimination is from high frequencies, which are not available in utero.