Plasticity in speech perception and production

Paper 9
Foundations of Speech Communication
Lent Term: Week 1

Katharine Barden
Plasticity

- Experience can alter how we respond to a stimulus
  ⇒ Categories are not fixed and immutable, they are dynamic

- Today’s lecture
  - Children (acquisition)
  - Adults (plasticity)
  - (Neural plasticity to be discussed in later lectures)
  - Modelling speech production and perception…
Acquisition…
Acquisition of categories

- Exposure to a language influences perceptual responses from as early as we can measure

- “Infants identify the key properties of the rhythmic organisation of their native language either prenatally or in the first few days of life” (p. 87, Jusczyk (1997))

- Mehler et al. (1988 and later):
  - French newborns sucked at higher rates when listening to French than to Russian (same talker)
  - Newborns from parents without French as a native language gave no evidence of distinguishing French and Russian
  - Utterances were low-pass filtered (removing segmental cues, but preserving prosodic characteristics)

- Preference can be ‘reset’
Acquisition of categories

- Evidence for the **emergence of native phonetic categories in the first year**

- Werker & Tees (1984):
  - Head-turn paradigm. Infants from English-speaking homes tested on:
    - English [ba]-[da]
    - Hindi [tə]-[ta]
    - Nthlakapmx [k’i]-[q’i]
  - 6-8 months: discrimination of all contrasts
  - 8-10 months: only some infants discriminate non-English contrasts
  - 10-12 months: hardly any evidence of discrimination of non-English contrasts
  - Hindi and Nthlakapmx infants had no difficulty discriminating their native language contrasts at 10-12 months

- Not all non-native contrasts decline in sensitivity e.g. Zulu clicks discriminated by English-speaking adults (Best et al., 1988 → Perceptual Assimilation Model)
Acquisition of categories

- Distribution of consonants that appear in babbling is influenced by native language
  - Comparison of babbling from French, English, Swedish and Japanese babies
  - By 10 months, proportion of stop consonants in both babbling and target language:
    Swedish > English > Japanese > French

Boysson-Bardies et al. (1992)
**Perceptual magnet effect** (reminder)

- **Kuhl**: By 6 months, babies have developed language-specific vowel categories with internal structure

<table>
<thead>
<tr>
<th>Discrimination by 6 month old babies</th>
<th>Exemplars of American /i/s: good bad</th>
<th>Exemplars of Swedish /i/s: good bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>American babies</td>
<td>poor good</td>
<td>no difference</td>
</tr>
<tr>
<td>Swedish babies</td>
<td>no difference</td>
<td>poor good</td>
</tr>
</tbody>
</table>

**Psychoacoustic space**
- With no phonetic category: no magnet effect
- With a phonetic category: perceptual magnet effect
Statistical phonetic learning

- Maye et al. (2002)
  - 6-month and 8-month old infants (from English-speaking homes)
  - Familiarised (2.3 mins) with a bimodal distribution or unimodal distribution of syllables on a VOT continuum (prevoiced to short-lag VOT)
  - Bimodal group $\Rightarrow$ increased discrimination of prevoiced vs short-lag VOT syllables
Statistical learning

- **Distributional information also** ⇒ **learning of larger categories** (reminder)
  
  - 8-month-olds are sensitive to transitional probabilities between sequences of syllables in ‘sentences’ from an artificial language

  **dibagudububibagubugibagu**

  *(Saffran, Aslin & Newport 1996)*
Summary

• Infants’ linguistic categories are extremely plastic
  – exposure to a language results in the development of native language categories

• Infants’ phonetic discrimination and ‘word segmentation’ is influenced by distributional information
Plasticity of speech production and perception in adults
Long-term changes in accent

• ‘Historical’ sound change not just due to successive generations, but also to changes in production within individuals over time

• Harrington et al. (2000)
  – Analyses of vowels in the Queen’s Christmas broadcasts in the ‘50s and the ‘80s showed a shift towards SSBE ‘80s accent
Plasticity of cues to phrase boundaries

Scott and Cutler (1984)

– Alveolar flapping is a cue to phrase boundaries in American English
  – For those of you who’d like to eat, early lunch will be served (no flap)
  – For those of you who’d like to eat early, lunch will be served (flap is optional)

– Does the presence of flapping facilitate comprehension of ambiguous sentences?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>American listeners</td>
<td>Yes</td>
</tr>
<tr>
<td>British listeners</td>
<td>No</td>
</tr>
<tr>
<td>British listeners long-term residents in US</td>
<td>A bit</td>
</tr>
</tbody>
</table>

Evidence of adaptation to phrase boundary cues
Shorter-term phenomena

**Accommodation** – adaptation of speech in response to interlocuter’s speech

- **Style shifting**
  - Sydney teenagers in presence of female interviewer
  - Girls decreased use of non-standard syntactic forms
  - Boys increased use of non-standard syntactic forms (Eisikovits, 1987)

- **Convergence in speaking rate** during interviews (Street, 1984)

- **Phonetic convergence** during conversation (Pardo, 2006)
  - Map task
  - Perceptual similarity of speakers’ pronunciation assessed
  - More similar towards end of task
Map task

For eliciting sufficient instances of particular words/sounds in relatively natural speech

- 2 speakers sit opposite one another and each has a map which the other cannot see.
- Speakers are told
  - to reproduce the Instruction Giver's route on the Instruction Follower's map.
  - that the maps are not identical
Adaptation to long VOT

Nielsen (2005)
• Listeners heard /p/-initial words with extended VOT
• Pre-exposure and post-exposure productions of /p/-initial and /k/-initial words were compared
• Post-exposure VOT was longer for both /p/- and /k/-initial words
⇒ Adaptation of VOT production following auditory experience
⇒ Generalisation across PoA and to new words
Learning phonotactic constraints

Onishi et al. (2002)

- Listeners heard CVC syllables which had constraints on consonant position
e.g. \textit{bap} but not \textit{pab}
- Speeded repetition task containing ‘legal’ syllables (which followed the constraints) and ‘illegal’ syllables (which did not)
- Legal syllables repeated faster than illegal syllables

⇒ evidence for learning of phonotactics
Learning phonotactic contingencies

- Listeners also learned a phonotactic contingency
  - Where legal consonant position depended on the adjacent vowel
    - e.g. *bap* or *pib* but not *bip* or *pab*

- BUT they could not learn a contingency where the legal consonant position depends on the talker

⇒ Learning of some context-dependent constraints but not others
  ⇒ Limits on plasticity?
Perception-production links

• Speech perception and production are frequently modelled separately

• Influence of auditory experience on production even in adulthood

⇒ Perhaps an integrated model is more appropriate?
Perceptual learning

- Learning from perceptual experience...
- Evidence: a change in categorisation following exposure
- Knowledge of rapid changes in perception
- Has generated a lot of recent research
Perceptual learning
Norris, McQueen and Cutler (2003)

**Familiarisation:**
- Synthesised an ambiguous fricative between [s] and [f]
- Lexical decision task including /s/-final words and /f/-final words
- Between-subjects design:

<table>
<thead>
<tr>
<th>Group 1</th>
<th>/s/-final words</th>
<th>/f/-final words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>miss [mɪs]</td>
<td>Ambiguous</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>beef [bi?]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 2</th>
<th>/s/-final words</th>
<th>/f/-final words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambiguous</td>
<td>miss [mɪ?]</td>
<td>Normal</td>
</tr>
<tr>
<td>Normal</td>
<td>beef [bif]</td>
<td></td>
</tr>
</tbody>
</table>

**Testing:**
- Phoneme categorisation on an [f]-[s] continuum
- **Boundary shift:**
  - those who heard ambiguous fricative in /f/-final words were more likely to categorise ambiguous fricatives as /f/
  - those who heard ambiguous fricative in /s/-final words were more likely to categorise ambiguous fricatives as /s/
Perceptual learning

How specific is the learning? What kind of ‘category’ has adapted?

– Talker-specific?
– Phoneme-specific?
– Structurally-specific?
– Style-specific?

How long do perceptual learning effects last?
Talker-specific?

• Allen and Miller (2004):
  – Listeners learned to associate a relatively short VOT with one talker and a relatively long VOT with another talker
  – They were able to identify which of two variants of *town* spoken by a particular talker was consistent with their experience of that talker during training
  – (Restricted set of synthetic stimuli with extreme VOT values (170-180 ms for voiceless stimuli)

• Kraljic and Samuel (2007):
  – Mixed results
    • Shift on [s]-[ʃ] continuum = talker-specific
    • Shift on [d]-[t] continuum = not talker-specific
Talker-specific?

• May depend on what sound is investigated
  – Affected by range of variation of a sound?

• May depend on context
  – E.g. expectations influenced by knowledge

• Some voices are more similar than others
Phoneme-specific?

Not necessarily phoneme-specific

  - Familiarisation using ambiguous /d/ or /t/ phonemes
  - Shift in the categorisation boundary for [b]-[p] continuum showing generalisation across PoA

- Smith (2007) – talker-specific word-segmentation cues
  - Intelligibility in noise of sentences that contained hard-to-segment sequences (e.g. /patsɔːd/, which can correspond to Pat sawed or Pat’s awed)
  - Testing occurred before and after training with a voice
  - Subjects who heard the same voice during training as during testing showed more improvement than those who heard a different voice
Sensitive to linguistic structures?

• Smith – word boundaries
• Alveolar flapping study – prosodic/syntactic structures
• Not much work in this area…
• KB’s work:
  – Results not very clear cut, but suggest…
    • Listeners can adapt to talker/accent-specific phrase-final lengthening
    • Listeners are sensitive to talker/accent-specific stress-conditioned allophonic detail
Style-specific?

It can be…

• Nygaard and Pisoni (1998)
  – Familiarisation with talkers
  – Intelligibility of words/sentences in noise when spoken by familiar vs novel talkers
  – Familiar = more intelligible
  – Training on sentence-length utterances improved the intelligibility of words presented in sentences, but NOT words presented in isolation

⇒ the talker-specific cues which listeners adapt to may not always be consistent across different styles of speech
How long do perceptual learning effects last?

- Eisner and McQueen (2006)
  - [f]-[s] boundary shift remains robust after 12 hrs
- Goldinger (1996)
  - Talker-familiarity effect when identifying words masked by noise even after a delay of a week

• Does the perceptual system ever return to the state it was in before perceptual learning took place?
Summary

• **Experience can alter perceptual and production categories**
  – over the longer term (e.g. adaptation to a new accent, or sound change in a community)
  – rapid adaptation to an individual talker or experimental situation

• **Production-perception links**
  – Shown by effect of auditory experience on production
Questions

• What are the **limits** on plasticity? (presumably there must be limits...)

• How **specific** is perceptual learning?

• Does the **same process** govern plasticity in infancy and in adulthood?
  – acquisition of categories vs. adaptation of categories?
Infants to adults – different processes?

• Assumption of dynamic representations (children) vs. static representations (adults) may arise partly from separation of research areas:
  – acquisition
  – speech production and perception

• Theoretical motivation for different processes:
  – Critical period hypothesis
    • Feral children
    • Imperfect second language acquisition

• Alternative
  – Is it just that wider experience of a language leads to more stable categories in adults than children?

• Or both?
Implications

• Models of speech production and speech perception need to include (or at least be compatible with) mechanisms that can account for
  – development of categories during acquisition
  – continuing plasticity of categories in adulthood

⇒ Exemplar models (more next week)
Modelling speech perception and production...
Modelling speech perception and production...

Models should be plausible from the perspectives of:

- Linguistics
- Phonetics
- Biology/neuroscience
- Social aspects (language in use)
Linguistics

- linguistically relevant units (which theory?)
- basic processes compatible with all languages?
- yet accounts for language communities developing, including specific accents
- accounts for, or at least compatible with, properties of language in development and breakdown
Phonetics

• Plausible in terms of what we know about the complexity of articulations e.g. coarticulation, the same articulator contributing to different kinds of contrast

• Accounts for context-sensitive phonetic detail
  – Production: Mechanisms for higher-level structures to influence articulations
  – Perception: Mechanisms to account for the use of context-sensitive phonetic detail in accessing meaning
Context-sensitive categories

• Revision…
  – Phonetic context – local and long-domain
  – Lexical context
  – Sentential context
  – Syllabic context
  – Morphology
  – Grammar – function/content words
  – Prosodic context (stress, turn-taking – more on segmental affects of prosody later)
  – Visual context
  – Semantic and pragmatic (expectations/predictability)
  – Talker
Biology/neuroscience

• biologically plausible e.g. re types of processes, when they occur, what is innate and what is learned
• acknowledges and allows for memory, learning, attention and other basic cognitive processes
• integrates the biology/neurology of language impairments with their linguistic consequences
Social aspects (lang in use)

• Can account for the production and perception used in all social contexts e.g. very informal conversation vs reading in lab
• Function – understanding meaning rather than just word/phoneme recognition
• Allows for language change
• Acquisition and aging
• Multimodal
Next week...

• Models of speech perception
Reading

Acquisition:

Plasticity in adults:
• Follow up any specific references that interest you

Critical period:
References


References


