Devil or angel in the details?

Perceiving phonetic variation as information about phonological structure

Catherine T. Best
MARCS Institute and School Humanities & Communication Arts
University of Western Sydney, Australia
Haskins Laboratories, CT, U.S.A.

Labex EFL International Chair 2014
ISCA Distinguished Lecturer 2014-2015
Collaborators on the studies discussed:

**Perceptual Assimilation in Speech Perception:**
- Pierre Hallé *CNRS, Paris, France & Haskins Laboratories, USA*
- Ocke-Schwen Bohn *Aarhus University, Denmark*
- Andrea Levitt *Wellesley College & Haskins Laboratories, USA*
(numerous others, to be listed with next week’s talk)

**Adult Cross-Accent Word Recognition:**
- Jason Shaw *Linguistics and MARCS Institute, UWS*
- Jennifer Hay *Linguistics and NZILBB Institute, U Canterbury NZ*
- Gerry Docherty *Newcastle-upon-Tyne U, UK*
- Paul Foulkes *York U, UK*
- Bronwen Evans *U College London, UK*
- Jia Ying *Linguistics and MARCS Institute, UWS*
- Elizabeth Clancey *Psychology and MARCS Institute, UWS*

**Phonological Constancy in Toddler Word Recognition:**
- Christine Kitamura *MARCS Institute, UWS*
- Michael Tyler *MARCS Institute, UWS*
- Karen Mulak *MARCS Institute, UWS*
- Rikke Bundgaard-Nielsen *MARCS Institute, UWS*
- Julia Irwin *Southern Connecticut State U & Haskins Laboratories*
“The devil is in the details”

- it is the small details of a task that often impede achievement of a larger goal, by hindering or clouding what would otherwise be a straightforward process (Bartlett, 1992)

or ... could it be an *angel* in the details?

- Alternate version: “God is in the details”
  - attributed to: Flaubert; van der Rohe; Warburg
  - see Titelman (2000) for further discussion

- the universe (of words etc.) awaits discovery in the patterns of phonetic details in speech...
Main points of the talk

- **Givens**: are that native language (L1) experience …
  - shapes perception of speech segments/contrasts
  - guides acquisition and recognition of spoken words

- **Core Questions** re: these experience-based effects
  - *What do we learn from native speech experience?*
    - **Premise 1**: Language environment-specific relationships between surface **phonetic** patterns and more abstract **phonological** forms
  - *How do we learn those relationships?*
    - **Premise 2**: Naturally-occurring phonetic variation in spoken language input provides the material from which listeners discover native **phonetic-phonological relationships**
One core reason phonetic variation is central to perceptual learning:

1. **Necessity**: variability inherent to natural language
   - **Ubiquitous**
     - it pervades all levels and types of spoken utterances
   - **Multiple** sources of structured variation
     - *linguistic* (contrasts; phonetic contexts; phonological processes)
     - *indexical* (talkers, regions of origin; emotions, registers & styles)
   - **Systematic**
     - phonetic variation provides *information*
     - it is *not noise* that needs filtering or normalization

2. **Word identity**: the *nexus* of *linguistic* and *indexical* variation
   - constant, abstract phonological form for a given referent
   - recognizable over lexically-irrelevant phonetic variations
Evidence that indexical info is retained, not filtered/normalized

- **Multi-talker performance decrements**, e.g.,
  - words in noise ↓ accuracy ↑ RT (Mullennix et al, 1989)
  - ↓ speeded phoneme classif. (Mullennix & Pisoni, 1990)
  - ↓ rapid serial recall (Goldinger et al, 1991; Martin et al, 1989)
  - See e.g. Pisoni (1997); Pisoni & Levi (2006); Pisoni & Lively (1995)

- **Multi-talker performance improvements**, e.g.,
  - ↑ learning/recogn of novel words in noise (Nygaard et al, 1994)
  - ↑ list-initial word recall (rehearsal) (Goldinger et al., 1991)

- **Regional accent information affects perception:**
  - shifts in consonant/vowel category boundaries (Hay et al, 2006, 2010; Niedzielski, 1999; Strand & Johnson, 1996)
  - changes in accuracy & latency of lexical decisions (Hay et al, 2010; Scharinger & Lahiri, 2010; Sumner & Samuel, 2009).
Second reason phonetic variation is central to perceptual learning:

2. Logic:

- two complementary principles identify the relationships between *phonetic variants* and *phonological forms*:
  - **Phonological Distinctiveness**
    - contrastive functions:
      phonetic differences that *do* alter phoneme/word identity
  - **Phonological Constancy**
    - categorical functions:
      phonetic differences that *do not* alter phoneme/word identity
Relation of phonetic variability to categorical information in speech

- Speech encodes *discrete* (categorical) information about two complementary sources of variation …
  - *Phonological*: linguistic, talker-independent
  - *Indexical*: social, talker-dependent

- … along *continuous* (gradient) phonetic dimensions

- **Premise 3**: Phonological organization is reflected in dynamic invariance relations among multiple gradient phonetic dimensions (Shaw et al, 2009, 2010, 2011)

☞ Under this assumption, it follows that: *variability* along multiple parameters is *needed* to specify, and to learn, phonological organization
Clues from speech production data: Syllable structure and timing

Languages differ in how they parse a series of word-onset consonants as syllable constituents:

**Complex onset parse**, e.g. #(C)(C)CV

(English, French) /la/ → /pla/ → /spla/

**Simplex onset parse**, e.g. #(C)(C).CV

(Berber, Moroccan Arabic) /bu/ → /s.bu/ → /ks.bu/

Q: Stable phonetic index of these 2 syllable parse types?
Previous H0: Syllable timing invariants

Complex onset syllables

- Center to anchor more stable (lower % variability) than right/left edge to anchor.

<table>
<thead>
<tr>
<th></th>
<th>left edge to anchor</th>
<th>center to anchor</th>
<th>right edge to anchor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10%</td>
<td>3%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Simplex onset syllables

- Right edge to anchor more stable than center/left edge to anchor

<table>
<thead>
<tr>
<th></th>
<th>left edge to anchor</th>
<th>center to anchor</th>
<th>right edge to anchor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12%</td>
<td>7%</td>
<td>4%</td>
</tr>
</tbody>
</table>

but .. not perfect, & exceptions for some words & talkers
Shaw et al’s ALTERNATIVE H0: Onset parsing type affects phonetic variation along multiple dimensions

In complex onset lang’s, adding a C to a word onset ...
  e.g., row → grow, lit → flit,
  ... attracts two well-known prosodic modulations:

(1) consonant shortening
  — same consonant is shorter in CCV than in CV

(2) syllable compression (aka word length effect)
  — same syllable is shorter in CCV than as CV

$H_{\text{alternative}}$

Underlying syllabic organization will affect phonetic variation along 2+ onset-vowel timing dimensions
Dynamic invariance

Center → Anchor stability (y) re: syllabic compression (x)

Simulated data

- Dynamic invariance: reliable relationships between dimensions of natural variation expose phonological organization … not singular stable invariants

- Different onset parses can be distinguished in these data because they structure variability in different ways

Shaw et al, 2009, 2011; Shaw & Gafos, 2010
Perceiving dynamic invariance in speech: 
Phonological Constancy
Phonological Distinctiveness

- **Role of phonetic variation in speech perception:**
  - Perceiving dynamic invariance in phonemes and words
  - Acquiring phonological constancy and distinctiveness

- **Perceptual assimilation of non-native contrasts**
  - Categorical perception of non-native phonetic variation
  - Perceived phonological form of non-native syllable onsets
    - Multiple languages: Late L2 versus early L2 bilinguals

- **Cross-accent assimilation in spoken word recognition**
  - Adult word recognition across L1 vs. L2 regional accents
  - Early development of L1 word recognition across accents
Regional accent differences and native spoken word recognition

- Key source of systematic phonetic variation in word and phoneme pronunciation
  - adults are highly sensitive to accent variation for sociolinguistic purposes
  - yet adults also recognize phonological constancy of words even across striking accent differences

Complementary principles of phonetic variation:

[naɪs] [næs] = “nice”

phonological constancy

[naɛs] [nɪs] ≠ “nice”

phonological distinctiveness
Perceiving dynamic invariance in speech: 
*Phonological Constancy*

*Phonological Distinctiveness*

- **Role of phonetic variation in speech perception:**
  - Perceiving dynamic invariance in phonemes and words
  - Acquiring phonological constancy and distinctiveness

- **Perceptual assimilation of non-native contrasts**
  - Categorical perception of non-native phonetic variation
  - Perceived phonological form of non-native syllable onsets
    - Multiple languages: Late L2 versus early L2 bilinguals

- **Cross-accent assimilation in spoken word recognition**
  - Adult word recognition across L1 vs. L2 regional accents
  - Early development of L1 word recognition across accents
Perceptual assimilation of phonetic variation in native vs non-native speech to native contrasts

**Assimilation of Non-native Contrasts**

**NATIVE PHONOLOGICAL SPACE**

- Two-Category Assimilation
- Single-Category Assimilation
- Category Goodness Difference
- Uncategorized-Assimilation
- Uncategorized²-Assimilation

Perceptual assimilation of phonetic variation in native vs non-native speech to native contrasts.
Perceiving *native* phonological contrasts with *non-native* phonetic distributions

- French listeners: categorization and discrimination of synthetic continua for American English /r/-/l/ /w/-/r/ /w/-/j/.
- French and English both employ /r l w j/ as phonemes and use all the associated *phonological contrasts*. HOWEVER …
- French and English [r] and [l] display different phonetic distributions.
- French and English [w] and [j] show ~the same phonetic distributions.

![Graphs comparing discrimination of stimulus pairs for English and French.](chart.png)
Why is English /w/-/j/ discriminated better by L2 than by L1 listeners?

- Hypotheses
  - Hallé et al. (1999): L1 **consonants** include +1 approximant (/w j/ + /ɥ/)
  - Bohn & Best (2012): L1 **vowels** include +1 phonetic feature (+rounded)

- Do L1 **consonant inventory** or L1 **vowel features** affect perception of L2 phonetic variation in the English approximants /r l w j/?
  - German: -1 approximants *(no /wl/)* but +/- rounded front V contrasts
  - Danish: +4 approximants **but not /wl/**, and +/- rounded front V contrasts
Native phonotactic constraints and perceiving non-native phonetic patterns

- Both English and French listeners perceive phonotactically *impermissible* non-native onset clusters /dla/-/tla/ as *permissible* native /gla/-/kla/ onsets (e.g., Hallé et al, 1998; Hallé & Best, 2007)
- i.e., the most similar native onset clusters in articulatory terms
- As expected, both groups have difficulty with /dl-gl/, /tl-kl/ contrasts
Perception of voicing contrasts re: gestural coupling in non-native syllable onsets

(Best & Hallé, 2010)

- **Fricatives**
  - *Zulu lateral fricatives*

- **Affricates**
  - *Tlingit lateral affricates*

- **Clusters: stop + /l/**
  - *Hebrew coronal stop-/l/ cluster*

- **Our 2 listener languages differ in stop voicing**
  - *English vs French: glottal gesture timing*
Perceiving dynamic invariance in syllable onsets?

Multiple phonetic parameters:

+/- glottal gesture re: gestural complexity; gestural coupling

Zulu /ɪa/-/ɪa/

Tlingit /dɪa/-/tɪa/

Hebrew /dlæ/-/tlæ/

(Best & Hallé, 2010)
Dynamic invariance in perceived onset voicing
Negative correlation with gestural coupling

Listener Studies:

- Hebrew: 97.1%
- Zulu: 94.5%
- Tlingit: 69.8%

Stimulus Language:
Discussion

• **Phonological Distinctiveness and Phonological Constancy:**
  ◦ complementary principles for perceiving dynamic invariance in natural, systematic phonetic variations
  ◦ “mapping” native phonetic patterns to phonological forms

• **compatible with numerous research findings:**
  ◦ native phonetic influences on perception of phonological structure in non-native consonants and syllable onsets
  ★ cross-accent word recognition in adults, and emergence of phonological constancy in toddlers

• **relevance to L2 learning:**
  ◦ L2 goals re: comprehension of *spoken language*: learners must be able to handle natural variation in the L2
  ◦ *benefits of* listening training, esp. w/ High-Variability input

Continuing on from here ...

Three further Labex talks

• **April 29 (17-19h)**: Cross-language speech perception: Naive, second-language and bilingual listeners

• **May 13 (16-18h)**: Spoken word recognition across regional accent variation: I. Native and second language adults

• **May 20 (16-18h)**: Spoken word recognition across regional accent variation: II. Development in young children