Phonetic detail in connected speech

Katharine Barden

Paper 9: Foundations of Speech Communication
Week 7, Michaelmas 2009
Aims

• To consider how the phonetic detail of speech reflects linguistic structure and function

• To explore the hypothesis that phonetic detail is central to speech understanding

• To raise questions about whether the identification of particular phonological units is necessary for speech perception
Standard assumption (production):

- Words are stored as discrete strings of phonemes
- BUT, they don’t emerge as discrete invariant strings of phonemes in articulation and acoustics
  - physical constraints (instantaneous movement between phonemic targets is impossible)
    → coarticulation
  - phonological processes e.g. assimilation
  - allophonic variation
  - style, rate, emotion
  - probabilistic factors – frequency, neighbourhood density
  - prosodic influences (e.g. phrase boundaries, focus)
  - discourse context (e.g. turn-taking)
  - speaker differences (accent, age, health etc.)
Standard assumption (perception):

- During perception, the signal has to be converted back into discrete phonemes.
- Once this is done, the speech signal isn’t needed any further.
Variability: Noise or information?

• There is lots of variability in the signal that is not understood and that might be random.
• But…. much variation is **systematic** and provides linguistic (and non-linguistic) information that might be useful in speech perception.

• Contributes to speech intelligibility
• Can enhance meaning
What is (fine) phonetic detail?

• Introduced about 25 years ago by John Local and colleagues
• Since then, it has been applied to any systematic phonetic variation that is not considered a major, usually local, perceptual cue for phonemic contrasts in the citation forms of lexical items
• Some PD is indeed ‘fine’ and subtle, but other types are perfectly audible; they have just not been factored into the prevailing theory that perceptual processing of phonetic information is entirely aimed at identifying strings of features or phonemes that allow words to be distinguished
Phonetic detail indicates:

- coarticulation
- position in syllable; word boundaries; morpheme boundaries
- long-domain cues to phonemes/features
- grammatical status
- indexical information
- places you can join in a conversation
- prosodic structure
- rate and style of speech
- probabilistic characteristics, e.g. frequency
Long-domain coarticulation

- Coarticulation is not restricted to adjacent segments
- Some coarticulatory effects cross syllable and word boundaries
  - lip rounding
  - nasalisation
  - resonance effects of /ɹ/
    - can be perceptually useful
    - in a forced-choice test (e.g. belly vs berry), listeners could correctly identify a missing /ɹ/ when surrounding vowels and consonants were replaced by noise (West, 2000)
The segmentation problem: discrete words from a ‘continuous’ signal

- but Pat sawed them / but Pat’s awed them
- we dread the reviews / we’d read the reviews
Allophonic variation indicates word boundaries

Assuming that phonetic detail is discarded after the identification of phonemes...

Sounds like: \[k\dot{a}t\dot{e}n\alpha I?\] \[k\alpha ?\epsilon n\alpha I?\]

Mental lexicon?: \[/k\dot{a}t\epsilon n\alpha I/\]

Meaning is ambiguous: car tonight or Carter Knight?

→ Further word segmentation processes are required for disambiguation

A more parsimonious model of speech perception would make use of the systematic phonetic information in the signal to assign probable word boundaries
Allophonic variation indicates word boundaries

/katsaɪz/

*cat’s eyes or cat size*

- /s/ longer when in onset of second syllable
- quality and degree of diphthongisation in second syllable may differ

Hawkins and Smith, 2001
Evidence that infants use allophonic information

8.5-month-old English-learning infants were familiarised with *dice*.

In a preference task:
- passages containing *dice* > unrelated word: preferred
- passages containing *d # ice* = unrelated word (e.g. *cold ice*): not preferred

Mattys & Jusczyk, 2001
Grammatical status

• Phonetic detail differs between the ‘same’ words when they serve a different grammatical function e.g.
  – Function vs content words: ‘for’ much more reduced than ‘four’ (Lavoie, 2002)
  – Different grammatical roles: infinitival to tends to be shorter and more reduced than prepositional to
    complementizer that tends to be shorter and more reduced than pronominal that (Jurafsky, 2002)

• Listeners are sensitive to these differences between function words and content words (Baker, 2008)
Phonetic detail indicates grammatical structure

- Many ‘function’ words begin with /ð/: the, this, that, those, then, thus, therefore ...

- No content words begin with /ð/—and it occurs within only a few content words: mother, other, bother, father, smithereens, soothe, smoothie …
Connected speech processes differ for /ð/ in content words and function words

<table>
<thead>
<tr>
<th>content</th>
<th>function</th>
</tr>
</thead>
<tbody>
<tr>
<td>*ban thatch [banθatʃ]</td>
<td>*ban that [banæt]</td>
</tr>
<tr>
<td>[bʌˈnθætʃ]</td>
<td></td>
</tr>
<tr>
<td>*[banə : atʃ]</td>
<td></td>
</tr>
<tr>
<td>*ban zips [banzɪps]</td>
<td>*ban this [banə : ɪs]</td>
</tr>
<tr>
<td>[bʌˈnзɪps]</td>
<td></td>
</tr>
<tr>
<td>[bʌzɪ : ps]</td>
<td></td>
</tr>
<tr>
<td>*[bʌn : ɪps]</td>
<td></td>
</tr>
</tbody>
</table>

(Local, 2003)
ban these [baŋ : iz]

all that [ɔ : l : atʰ]

Courtesy John Local
Manuel (1995, *J. Phonetics*)

*win those*

- /nð/ is not realised as [nð]
- but as [n : ]

- Listeners report hearing /ð/ in tokens like these
- They can generally distinguish *win noes* from *win those*, even when the fricative is completely nasalized
- Manuel suggests that low F2 at the release of the nasal is a cue to dental rather than alveolar place of articulation
Grammar and phonetic detail

*lime vs. I’m*

| content | pro + be |

- in ordinary lexical items, word-final *m* does not usually assimilate to the place of articulation of the following consonant in English, while *n* does

- *the line goes there* but not *the lime goes there*
Grammar and phonetic detail

*lime* vs. *I’m*

content pro + be

- However, in the grammatical chunk *I’m …* (= ‘I am’) assimilation regularly happens in everyday talk

\[
\text{aǐm} \quad \text{aǐn} \quad \text{aǐŋ} \quad \text{aǐw}
\]
Grammar and detail: *lime vs. I’m*

+ assimilation of nasal
  - I’m blowing [aɪm]
  - I’m throwing [aɪn]
  - I’m going [aɪŋ]
  - I’m watching [aɪw]

- assimilation of nasal
  - *lime throws*
  - *climb goes*
  - *crime wave*
  - etc

*I’m* contrasts with:
  - you’re, she’s, he’s,
  - we’re, they’re, it’s

*lime* contrasts with:
  - lamb, land, live, line, rime, right,
  - ride, lemon, ..... & lots more ......
  - always diphthong & [m]
True and pseudo prefix errors (Morphemic structure of words in natural sentences)

<table>
<thead>
<tr>
<th>Productive</th>
<th>True prefix (Pr)</th>
<th>mistimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unproductive</td>
<td>Pseudo prefix (Ps)</td>
<td>mistakes</td>
</tr>
</tbody>
</table>
True and pseudo prefixes
(Morphemic structure of words in natural sentences)

- Same phonemes: /mɪʃt/
  but
- PD systematically reflects morphological status:
  - properties of periodic part
    - duration
    - abruptness of [mɪ] boundary
    - F2 frequency (etc)
  - relative durations e.g.:
    - periodicity : aperiodicity (fric)
    - fricative : silence (closure)
    - VOT

<table>
<thead>
<tr>
<th>True prefix (Pr)</th>
<th>mistimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo prefix (Ps)</td>
<td>mistakes</td>
</tr>
</tbody>
</table>

Baker, Smith, Hawkins ICPhS 2007;
Baker 2008 PhD dissertation
True and pseudo prefixes
(Morphemic structure of words in natural sentences)

<table>
<thead>
<tr>
<th>True morpheme 1</th>
<th>mistimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo morpheme 2</td>
<td>mistakes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>match (control)</th>
<th>mismatch (exptl)</th>
</tr>
</thead>
</table>

- intelligibility in noise worse when mismatched
- details of sound patterns can signal morphemic status

Phonetic detail and prosodic structure

Many phonetic cues to prosodic structure...

– more next term for paper 3
  prosody (intonation, rhythm, stress)

– more next term for paper 9
  prosodic-segmental interactions
Phonetic detail and discourse structure

(See Lecture on Phonetics in Conversation)

• Turn-taking
• Turn co-completion

• New vs. given information
  – “Words are pronounced with greater duration and clarity when they introduce new information into a discourse than when they refer to given information” (e.g., Bolinger, 1963; Chafe, 1974; Hawkins & Warren, 1994)

Summary so far...

- Phonetic detail maps onto many different abstract units of linguistic structure
  - one sound chunk can signal many abstract units (with different probability levels)
- Phonetic detail can be perceptually useful
  - Models advocating early abstraction from the detail of the speech signal do not make optimal use of the available cues
  - Phonetic detail must be represented in memory and related to abstract linguistic structure
Summary so far...

For any given unit $U$
(e.g. segment, syllable, foot, word...)
• $U$ is functionally inseparable from its context
• If something in the context changes, there will probably be consequences for the phonetic realisation—and perception—of $U$

• Only when whole context is taken into account is the systematicity of phonetic detail evident
• Taking account of context can help to make sense of variation that would otherwise appear random
Context

• “Context” does not just refer to linguistic structure
• also probabilistic factors
  – word frequency
  – neighbourhood density
  – predictability
Context

• “Context” does not just refer to linguistic structure
• also probabilistic factors
  – word frequency
  – neighbourhood density
  – predictability
• also register and attitude
Very reduced/casual speech

I…do…not…know
I do not know
I don’t know
I dunno
dunno

speech style contributes to (non-linguistic) meaning
• Is the identification of sublexical units always a necessary stage in speech perception?

• Is the identification of words always a necessary stage in speech perception?
Summary

• Phonetic detail signals all sorts of linguistic categories, and linguistic and communicative functions

• Taking a broader context into account can help make sense of seemingly random variation – linguistic structure, prosodic structure, probabilistic factors, register, attitude, speaker, situation...
Summary

• Much phonetic detail is learnable and perceptually salient, and at least some of it makes speech easier to understand in at least some circumstances.

• If we take ‘understanding meaning’ to be the aim of listening to speech (rather than phoneme or word identification), it may ultimately be easier to make sense of the variation in the speech signal.