

Commentary: Is «Sociophonetic» Knowledge Special?

Benjamin Munson

Associate Professor
Department of Speech-Language-Hearing Sciences
University of Minnesota, Minneapolis

Visiting Erskine Fellow
Department of Linguistics
University of Canterbury, Christchurch



Commentary: A Complex Enterprise

- Summary
- Critique
- Non-Grandiose Position Statement



Part 1

Commentary as Summary



Common Themes

- Emphasis that listeners access far more phonetic detail during perception than coarse SPE-style representations would suggest.



Common Themes

- The suggestion that the apparent sensitivity to social information supports exemplar models of perception.



Common Themes

- Suggestion that listeners' knowledge and beliefs about speaker attributes both (a) allows them to identify speaker attributes from audio-only signals, and (b) biases their perception of other types of linguistic information (syntactic, phonemic)



Common Themes

- A less-obvious but nonetheless equally important theme: it's a small world.
- So small that I have to be careful not to comment on work on which I am a coauthor, sufficient to say I'm engaging every neuron in my frontal lobes to suppress my urge to talk about Smith et al.!



Part 2

Commentary as Critique

Observations and Critiques, Mostly Methodological
Did the Papers Respond to the Session Questions?



The Observation

- There is no widely accepted taxonomy of types of pronunciation variation beyond the Ladefoged and Broadbent (1957) typology.
 - Linguistic
 - Sociolinguistic
 - ‘Personal’



Foulkes

- The information that falls under the label of sociophonetics is highly dependent on a taxonomy.
- A case in point: how distinct is speaker-specific information from sociolinguistic information?
 - Are talkers (or even just a subset of talkers) identifiable as an assembly of social features? [+gay], [+Northern Cities US], [-young], [-old] = Me
- Can we say that information is social if it is only exemplified in one or two speakers in the person's lexicon of remembered voices?



Foulkes

- A particularly gnarly case related to overall stature
- Seems to fall under the ‘personal’ branch of L&B’s tripartite division, but...
- Munson et al. (2006) showed that women manipulate formant frequencies in a way that changes their apparent stature. These tendencies are correlated with self-stated sexuality (See also Van Bezooijen, 1995).



Foulkes

- An even gnarlier case: vocal emotion
- This is off of Ladefoged and Broadbent’s chart, presumably because it was thought of as transient and related to autonomic functions—an assumption that is ripe for testing, and which comes up in this commentary later.



Foulkes

- 9:08 am: Paul mentions this in his talk
- Good on ya!



The Observation

- The phonetic characteristics of real-world speech are the product of many factors.
- Degree of glottalization and diphthongization can be affected by broader prosodic structure
 - Glottalization marks the left edge of phrases (Pierrehumbert and Talkin)
 - Diphthongization is stronger in accented syllables than in unaccented ones (various papers by Ken de Jong)



Drager

- Suggests that glottalization is associated with grammatical functions and social-group membership; however, these phonetic differences in *like* may be the consequence of prosodic factors.
 - Drager reports that the tokens of *like* don't differ in their duration, but prosodic structure is more than duration.
- If so, then a perception experiment would work best if the listener were given more information with which to parse the prosodic structure of the utterance.
 - The Palmer (2007) study that Johnson cited in his talk seems to exemplify a better methodology.



The Observation

- Many factors influence word-recognition accuracy. A very robust factor is phonological neighborhood density.
- This variable influences perception...
 - (Vitevitch & Luce, 1999; Luce, Pisoni & Goldinger, 1990, inter alia)
- ...and production
 - (Wright, Labphon 6; Vitevitch, 2002; Munson, Labphon 9; Scarborough, Labphon 10; Bease & Goldrick, Labphon 10)



Clopper et al.

- The words in the real-word and nonword competitor conditions were not equivalent in phonological neighborhood density.
- The real-word competitor condition may have simply have had harder-to-perceive words than the nonword competitor competition.



The Observation

- Overt mention of categories affects sensitivity to variation.
- Compare “is this /s/ or /ʃ/?” experiments with “how different are these tokens?” experiments.



Mack, Smith et al.

- By mentioning categories like “gay,” “straight,” “over 20,” “higher social class,” etc., listeners may have...
- 1. parsed the continuous signal into categories like “gay” and “straight” into which they might not have otherwise parsed the signal, and
- 2. The cues that they used might have been more closely related to stereotypes about these variables than they otherwise would have been.



The Observation

- Measures of statistical significance are not identical to measures of effect size.
- Simply saying that an effect is significant doesn't tell you how strongly it predicts a behavior.



Common to Nearly All Studies...

- No paper in this session reported measures of effect size, nor did they compare the magnitude of social influences to the magnitude of other influences on perception.
- The contribution of social variables on perception may be quite small
 - Munson, Jefferson, & McDonald (2006): perception of fricatives appended to different talkers appears to be much more closely related to vocal-tract size than to social characteristics attributed to the speakers.



The Observation

- Equally appearing interval (EAI) scales are not the only way of eliciting gradient judgments.
- Direct Magnitude Estimation (DME) and Visual Analog Scale (VAS) provide an alternative to this.
- Moreover, comparing the same stimuli on an EAI and a DME scale allows you to assess whether listeners treat the variable of interest as varying in quantity or quality.



Common to Nearly All Studies...

- ...is the use of EAI scales to assess social categories, without an explicit test of whether EAI scales are appropriate.
- In particular, studies like Mack's might find not only that different constructs are predicted by different acoustic measures, but that listeners treat them qualitatively differently.



Summing up thus Far

- Sociophonetics is a relatively new area of inquiry.
- As such, it is important that everyone who works in this area be willing to do the sometimes-less-immediately-exciting methodological work before we make broad conclusions about our data.



The Questions Posed by Jen and Paul

- ...And what were we supposed to be making conclusions about? Three questions were meant to shape this session:
- *Is phonetic information in the lexicon accompanied by social information?*
- *How do social expectations about a speaker affect speech perception?*
- *Is speaker-specific detail stored in the lexicon?*



One Question:

- *How do social expectations about a speaker affect speech perception?*
- Are these based on experience or on stereotypes?
 - Drager's experiment 3 suggests that identification of CR and NCR girls is based on stereotypes.
 - Walker's poster (this conference) suggests that the relationship between phonetic form and grammaticality judgments is mediated by stereotypes about the way that people in different social groups speak.
 - Other work by Mack, Munson, and Kuntz (as seen at LSA 2008) also suggests that identification of men's sexuality in speech is highly dependent on stereotypes.



Two Questions are Harder To Answer

- *Is phonetic information in the lexicon accompanied by social information?*
- *Is speaker-specific detail stored in the lexicon?*
- A more test would be to examine whether social information can be accessed without accessing the lexicon.



Two Questions are Harder To Answer

- If the cue is suprasegmental, then one need not access the lexicon to identify the social characteristic.
 - Szakay, this conference, shows that one need not access the lexicon to make a judgment of how Pakeha- or Māori-sounding a person is.



What's in the Lexicon?

- One very course idea: can social categories be perceived through nonword stimuli?
 - Hard to test, as the task of producing nonwords might compromise how variants are produced and might increase the difficulty of the listening task, which mediates the perception of indexical features (McLennen & Luce, 2005)
- My strong suspicion: social categories will be readily perceptible through nonwords.
 - See also Pitt and McQueen, 1998



What's in the Lexicon?

- Another idea (currently being pursued by Munson & Kaiser, and Munson, Smith, & Schempp): Are phonotactic generalizations mediated by social indexing?
 - Similar work being pursued by Babel (dissertation in progress) and Graff, Nielson, & Zuraw (study in progress)



What's in the Lexicon?

- Phonotactic generalizations track lexical statistics (Albright, 2008; Coleman & Pierrehumbert, 2001; Edwards, Beckman, & Munson, 2004; Frisch, Large, & Pisoni, 2000).
- Do different generalizations emerge when phonetic variants co-vary with social variants?



What's in the Lexicon?

- Work on phonotactic learning by Onishi and colleagues suggest that the answer will be 'no'.
- Work by Docherty, Langstrof, & Foulkes (this conference) suggests that it is possible nonetheless to learn novel speaker attributes through phonetic variation.



Testing whether it's In the Lexicon

- More generally, we will have a much easier time answering these questions if we work from the models that have already been developed to test top-down and bottom-up processing.



Part 3

Commentary as Non-Grandiose Position Statement



The Issues I can't Get out of my Head

- 1. What is 'social information'?
- The proper phonetic parameterization of any indexical variable depends on the proper understanding of the universe of contrastive categories.
- Is there something wrong with the labeled categories that are part of traditional sociolinguistics?



The Issues I can't Get out of my Head

- An analogy: emotion
- The literature on vocal emotion is vast and has little agreement on the basic lexicon of emotions that individuals convey. These studies generally evaluate vocal emotion by examining correlates of different explicitly labeled emotions.
- The literature by personality psychologists on emotion suggests a two-factor (valence and affect) scheme for describing emotion.



The Issues I can't Get out of my Head

- Perhaps this endeavor can start by examining why indexical knowledge emerges.
- Why do we need to identify social-indexical categories?



The Issues I can't Get out of my Head

- One of the functions of phonological categories is because of the immense pressure of having to construct (roughly) as many wordforms as there are concepts to be conveyed, and
- Abstract phonological categories emerge in development in response to the pressures exerted by the developing lexicon (at least in part)
 - Ferguson & Farwell, 1975; Edwards, Beckman, & Munson, 2004; Munson, Swenson, & Manthei, 2004; Walley, 1993



The Issues I can't Get out of my Head

- Perhaps indexical categories serve a similar function of mediating complexity.
- Over what universe do we generalize the socioindexical categories? A 'lexicon' of voices? If so, then 'social information' might be the categories that emerge to structure the lexicon of voices.
- There might be substantial individual differences in the size of the lexicon of voices.



The Issues I can't Get out of my Head

- 2. What do listeners bring to the task?
 - Knowledge (=expectations)
 - Needs
- Knowledge of how different social categories are represented and how this should influence perception.
- The need that the variation be interpretable and not compromise other aspects of the signal.



The Issues I can't Get out of my Head

- What is the consequence of failing to learn sociophonetic variation?
- There is evidence that deficits in encoding the parametric phonetic space is a causal factor in developmental pronunciation problems (Munson, Baylis, Kruase, & Yim, Labphon 10; Munson, Edwards, & Beckman, 2005) and that deficits in inferring categories from the input leads to broader morphosyntactic problems (Beckman, Munson, & Edwards, Labphon 9; Munson, Windsor, & Kurtz, 2005)



The Issues I can't Get out of my Head

- Perhaps the greatest impact of sociophonetic variation might be on the kinds of subjective judgments that are made of people's speech
 - I.e., if children don't show the expected social variation, their peers, caregivers, and educators might react differently to them than if they do.
 - This notion is currently being explored by Munson, Bauer, and Zittnan (2008) relating to gender and perceived narrative quality.



The Issues I can't Get out of my Head

- 3. How does sociophonetic variation influence acquisition?
- Variation facilitates phonological acquisition (i.e., Bradlow et al. on /r/-/l/ acquisition by L2 speakers of English).
- Variation might facilitate acquisition only when the listeners can yoke the variation in social categories.
- Memor to Foulkes and others: I would caution against interpreting CDS as a sociophonetic phenomenon.



Surprise!

Part 4

Grandiose Position Statement



¡Viva La Variación!, o,
¿Porque la «Sociofonética»?

- Our knowledge of the sound structure of language exists at multiple levels of abstraction and specificity (Pierrehumbert, 2003; Beckman, Munson, & Edwards, 2007)
- A comprehensive understanding of the sound structure of language must include an understanding of both abstract and specific aspects of speech-sound production and knowledge.



¡Viva La Variación!, o,
¿Porque la «Sociofonética»?

- The questions posed by this session seem to suggest that...
- ... (a) the ability to perceive social categories through speech provides a privileged measure of the ability to perceive fine phonetic detail in speech, and
- ... (b) there is something special about the ability to convey and perceive social categories, compared to variation arising from other sources.



¡Viva La Variación!, o,
¿Porque la «Sociofonética»?

- Is this really the case?
- Comparison cases: emotion and pathology.
- These variables behave similarly to social variation: perception is biased by expectation.



¡Viva La Variación!, o,
¿Porque la «Sociofonética»?

- Why, then, have we focused on social knowledge?
- It feel comfortably within the ‘scope of practice’ of licensed, carry-carrying linguists, because of the long history of sociolinguistics.
- Perhaps for this reason, sociophonetics has opened up new areas of inquiry for laboratory phonologists. This is illustrated nicely by all of the talks in this session.



¡Viva La Variación!

- But if our broader goal is to model specificity and abstraction in phonological knowledge, we need to consider all of the sources of variation in each of these domains. Social variation is but one.
- Moreover, if our goal is to examine exemplar models, then we should be using more direct tests of whether exemplars are represented in memory.



Acknowledgments

- Many of the ideas in this talk arose from conversations with my colleagues and collaborators, most especially Mary Beckman and Jan Edwards, neither of whom is responsible for the content of this commentary, and both of whom probably disagree with substantial portions of it.
- This work was supported by the McKnight Foundation through a University of Minnesota Presidential Fellowship to Benjamin Munson, and an Erskine Fellowship to Benjamin Munson from the University of Canterbury.

