ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Journal of Phonetics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Journal of Phonetics

www.elsevier.com/locate/phonetics

Editorial

Modelling sociophonetic variation

This special issue of the Journal of Phonetics is dedicated to issues relating to the modelling of sociophonetic variation.

Phoneticians have made great progress in understanding the nature of variability in speech. Much effort has been spent on understanding and modelling factors such as the effects of co-articulation, rate of speech, phonemic context and vocal tract physiology on the acoustic properties of speech sounds. Socially driven phonetic variation, however, is also pervasive but has received much less attention. Most current models of speech perception and speech production can therefore not readily account for how we produce, perceive and interpret socially conditioned variability. The papers in this special issue demonstrate the tight interconnectivity of social and phonetic variation, and take on the challenge of modelling such effects.

Sociophonetic variation is phonetic variation in speech which is socially conditioned. Foulkes and Docherty (this volume) provide a wide range of examples of how phonetic variants can be associated with particular speaker groups, or particular speaker styles. The notion that variation in speech can be correlated with social structure is a core notion of sociolinguistics, and this variation is usually investigated within the Labovian framework of *Language Variation and Change*. Here, a primary goal of the analyst is to document the ways in which variation in the speech stream indexes social meaning. The data analysed within this framework is often, but not always, drawn from interviews with individuals from a particular speech community. Both language internal and external factors are taken into account to explain the variability found across speakers and contexts. While the data often shows gradient properties, the statistical methodology traditionally employed by variationists (Varbrul—see Tagliomonte, 2006) only allows for the analysis of categorical data. Therefore, even when the data is gradient, it tends to be coded in terms of binary choices (either the variable rule has applied or it has not applied). This methodology provides a ranking of the relative importance of the language internal factors (such as preceding or following segmental environment) and language external factors (such as social class, age, ethnicity or gender). Much of the analysis of phonetic variation, which is conducted within the rubric of sociolinguistics, reinforces the core insight regarding the importance of social structure in conditioning phonetic variability.

From a phonetician's perspective, much current variationist work could be viewed as lacking in methodological rigour both in terms of the analysis of fine phonetic detail, and the sophisticated statistical modelling of that detail.

On the other hand, from a variationist's perspective most phonetic work could be criticized for the focus on non-naturally occurring speech. Phoneticians traditionally work with experimental data, in which they tightly control for potential effects which could introduce unwanted variability (noise) into the speech of the participants. The lab-based focus of many phonetic studies strips away much of the socially conditioned variability which is present in spontaneous speech. In the literature on speech perception, social variation has often been treated as a source of noise—something which has the potential to interfere with the perception mechanism. While most phoneticians would acknowledge that socially driven phonetic variation exists, many would not regard it as something which is of any theoretical interest. We disagree. A growing body of results is suggesting that social and linguistic details are inextricably entwined in our lexical representations. Understanding how speakers simultaneously transmit social and linguistic messages, and how listeners

0095-4470/$ - see front matter
doi:10.1016/j.wocn.2006.08.001

Please cite this article as: Stefanie Jannedy, Jennifer Hay, Modelling sociophonetic variation, Journal of Phonetics (2006), doi:10.1016/j.wocn.2006.08.001.

accommodate socially driven variability in the speech stream has the potential to greatly enhance our models of speech perception and speech production.

The term 'sociophonetics' has been rapidly increasing in currency, and tends to be used by researchers who feel they straddle the divide between sociolinguistics and the phonetics. This could be because they are using the traditional methodologies and/or data sets from one field to inform questions traditionally associated with the other. Or it could be because the questions they are working on have important implications for both fields. Increasingly sophisticated acoustic analysis and experimental techniques are certainly making their way into sociolinguistics at an impressive rate. And phoneticians are beginning to investigate how phonetic detail indexes social categories and how such structured social indexing of phonetic detail might inform phonetic models.

It was in this context that we organized a panel at the 2004 Linguistic Society of America conference. The purpose of the panel was to take stock of what is known about socially conditioned variation in the speech stream, and to discuss how this knowledge might inform our models of speech perception and speech production. We brought together researchers working at the interface of the "social" and the "phonetic", to reflect upon issues relating to the modelling of sociophonetic variation. The presentations clearly demonstrated the potential benefits for both fields. With the theme 'modelling sociophonetic variation' we intended to ask how sociophonetic variation was represented by the individual. How is it that individuals can skillfully style-shift between phonetic variants—using speech in a way that expresses both linguistic and social meaning? How can we even begin to understand the linguistic messages that are being portrayed when they are so layered with large amounts of phonetic variability which indexes social meaning? How can sociophonetic variation inform cognitive models? The discussion was lively and productive, and resulted in the decision to organize a special issue of the Journal of Phonetics, which would tackle some of the questions raised.

This special issue brings together four papers which examine aspects of sociophonetic variation, and considers how these might be modelled. Two commentaries, one written by a sociolinguist, the other by a phonetician, provide different perspectives on the papers.

While the investigation of social conditioning in speech may be somewhat unusual territory for the Journal of Phonetics, the methodologies employed in the papers are fundamentally 'phonetic'. Attention to finely detailed sub-phonemic variation of consonants (Foulkes and Docherty), the use of reaction-time data, and the consideration of the relationship between speaker height and bark-scaled formant values (Johnson), the usage of the second coefficient of cosine transformations of vowel formants to detect differences in the curvature of formant trajectories (Harrington), and the use of Pillai Scores for measuring degree of overlap between almost merged vowel distributions (Hay et al.), all go beyond techniques that would tend to be used in traditional sociolinguistics research. One sense in which the papers in this volume represent 'sociophonetic' work, then, is that they bring phonetic methodologies to bear on sociolinguistic data and/or questions.

As the authors discuss the modelling of their data, a clear theme which emerges, is an investigation of the appropriateness of 'exemplar' models as one way of accounting for the representation, production and perception of sociophonetic variability. This theme is also picked up and discussed at length by both commentators. Thus, while the primary topic of this special issue is the modelling of sociophonetic variation, exemplar models of speech production and perception also emerge as a strong secondary theme.

Exemplar models of speech production and perception proceed from the assumption that lexical structure consists of detailed and acoustically rich remembered linguistic experiences. Because these memories are phonetically detailed, and because memories can be "indexed" with relevant social information, these models seem to provide a compelling account of the interconnectedness of the linguistic and social meaning which is embedded in phonetic detail.

Exploring the consequences for theoretical models in phonetics and phonology, Foulkes and Docherty provide an excellent overview of the body of work which is generally described as 'sociophonetic'. Sociophonetic variation, according to Foulkes and Docherty, refers to "variable aspects of phonetic or phonological structure in which alternative forms correlate with social factors". They carefully document different types of socially conditioned variation on the segmental, sub-phonemic and suprasegmental level, suggesting that variation is pervasive in "all levels of phonetic and phonological structure".

These authors' work on the acquisition of sociophonetic variability provides an intriguing and enlightening account of the inseparability of linguistic and non-linguistic meaning from the very earliest stages of language

Please cite this article as: Stefanie Jannedy, Jennifer Hay, Modelling sociophonetic variation, Journal of Phonetics (2006), doi:10.1016/j.wocn.2006.08.001.

acquisition. And their results on mothers' differential modelling of phonetic forms to boys and girls provide the clearest picture yet of how girls and boys come to acquire different phonetic variants from the very early stages of language acquisition. Their results also demonstrate that children older than three produce phonetic variants that suggest that they recognise and can actively exploit the social meanings carried by fine phonetic detail.

Foulkes and Docherty raise a valuable range of questions and future challenges for the exemplar model, and call for a shift of focus away from cross-subject designs to a focus on individuals, and how they "perform and interpret sociophonetic variability across a range of communicative settings".

The importance of a focus on individuals is certainly heeded by Harrington, who conducts a very careful study of changes over time in Queen Elizabeth II's use of the vowel at the end of words like *happy*. Work by Harrington and his collaborators, shows that the Queen's speech has gradually shifted in the same direction as modern RP. The paper in this volume goes beyond previous real-time studies of language change, by also seriously considering potential changes over time which may be due to effects of vocal tract maturation. Both sound change and physiological factors appear to be involved in changes in the Queen's vowels over time.

Harrington proposes that just as sound change progresses "gradually and imperceptibly", so too does it have a very gradual and imperceptible effect on individual speakers' production over a very long period of time. Harrington's work demonstrates that the Queen participates to different degrees in different shifts. Delineating the kinds of changes which adults are most prone to participate in will also be an interesting challenge for future work. Harrington speculates that social stigma and/or markedness may play a major role, and that adults are likely to resist innovations that carry strong social stigma.

The work by Hay et al. investigates a vowel merger-in-progress in New Zealand English, which is conditioned by both age and social class. Their experiments demonstrate that social factors such as the perceived identity of a speaker affect the accuracy with which listeners can discriminate between the vowels in question. Perceived identity is manipulated by associating the same voice with different photos across conditions. These effects exist alongside a complex set of interacting factors, which work together to influence participants' accuracy in the perception task.

Hay et al. argue that multiple levels of representation are required in order to explain their complicated set of results. Lexical exemplars seem to be the site of their frequency effects, and can explain why fully merged participants are actually quite accurate at the perception task despite the fact that they feel they are guessing. However the behaviour differences between merged and non-merged participants also seem to suggest a difference in the *phonemic* labeling of these lexical exemplars. The question of whether an exemplar model can or should accommodate the notions of 'phonemes' is a theme which is touched on by several authors, and explicitly examined by both commentators.

Johnson explores a particular implementation of an exemplar model (an exemplar 'resonance' model) in the context of two sets of data. The first is a cross-linguistic survey of vowels produced by male and female speakers, demonstrating substantial gender differences across languages which cannot be ascribed to speaker height and associated vocal tract length. The second set of data demonstrates that listeners are quicker to process words when the speaker has a gender-stereotypical voice. His implementation of an exemplar-based model with reentrant pathways ('resonance') shows the very same sensitivity to voice stereotypicality.

All of the authors argue that both linguistic and socio-indexical information are encoded in the speech signal. In exemplar accounts, the acoustically rich signal is stored in the lexicon together with linguistic and social indexing. This results in exemplar 'clouds' (phonetic density distributions) which are active in speech perception and speech production. Such models can account for the apparent sensitivity of the listener to social information such as age and social class (Hay et al.), and to gender stereotypicality (Johnson). They also provide a clear account of the tight relationship between input (in speech perception), and output (in speech production)—a relationship that continues to exist throughout the life span. When infants acquire language, the phonetic detail of their speech patterns is heavily influenced by the input they receive from primary care takers (Foulkes and Docherty). As of age 3+, children are also able to generalize over the input they have received in order to exploit the social meaning of different phonetic variants. Through adulthood, the phonetic detail of speech production continues to be subtly influenced by the speech of those around us (Harrington). The relationship between speech perception and speech production, the degree to which we generalize over and weight input forms, and the degree to which this changes through the lifetime, all need to be captured by our models.

Please cite this article as: Stefanie Jannedy, Jennifer Hay, Modelling sociophonetic variation, Journal of Phonetics (2006), doi:10.1016/j.wocn.2006.08.001.

The two commentaries by Bill Labov and Janet Pierrehumbert provide valuable overviews. Both explicitly pick up on the theme of exemplar models, and provide their own perspective of their advantages and drawbacks. While the authors come from very different perspectives (a sociolinguist and a phonetician), there are also many threads in common. For example, the question of how phonemic information is to be represented in an exemplar approach plays a major role in both commentaries. Labov questions whether exemplar theory can really account for the relatively abstract, discrete categories which appear to participate in sound change. By focusing on the word as the unit of storage, argues Labov, it may be difficult for an exemplar approach to capture the "abstract and discrete representations that characterize linguistic structure and drive linguistic change in progress". Pierrehumbert also argues that we need to consider how our models will be able to account for discrete linguistic structure, and advocates a "hybrid model" which has "multiple levels of representation (like neo-generative models) while also having explicit mechanisms for statistical learning and situational indexing (like exemplar models)."

Labov and Pierrehumbert both discuss the central role of frequency in exemplar models. Labov discusses neogrammarian sound change, and the pervasive finding of regularity in sound change. He is sceptical that an exemplar approach, with its predictions relating to the centrality of frequency, can account for the overwhelming number of sound changes in which frequency appears to play no role. Pierrehumbert is more optimistic about the frequency predictions, and warns against any expectation that frequency effects should be straightforward: in "models with multiple levels of representation, [frequency] effects can play out in surprising ways". She goes on to discuss some of these ways in detail.

Both commentators caution against assuming that all social variation is necessarily stored in the cognitive system. Labov highlights the need to establish "what forms are socially evaluated and what social evaluations are retained in memory." Pierrehumbert points out that many social factors may affect speech production, but not emerge as cognitively salient categories because of the lack of sufficient sampling size. She therefore warns against assuming that all social variation is cognitively relevant, arguing "… social factors could be relevant to the phonological system without constituting social categories, because social factors could affect the propagation of phonetic characteristics through the population, without being observable by every speaker individually."

Between them, the papers and commentaries raise significant questions and challenges for the exemplar approach, and point in a variety of directions for what the best way forward might be. We believe that taking these challenges seriously is likely to benefit phonetic theory as well as sociolinguists who are interested in understanding the dynamics of language variation and change. The papers in this volume certainly illustrate the rewards of giving sociophonetic variation the attention it deserves, and carefully considering how it might be incorporated in our models.

In closing, we would like to thank all of the contributors to this special issue, the reviewers, and the Journal of Phonetics editorial team. We would also like to thank Nancy Niedzielski and Norma Mendoza-Denton, who took part in the original LSA panel from which this special issue was born.

Reference

Tagliomonte, S. (2006). *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Stefanie Jannedy, Jennifer Hay
Department of Linguistics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
E-mail address: jen.hay@canterbury.ac.nz (J. Hay)

Please cite this article as: Stefanie Jannedy, Jennifer Hay, Modelling sociophonetic variation, Journal of Phonetics (2006), doi:10.1016/j.wocn.2006.08.001.