

RUNNING HEAD: Socially Meaningful Variation

Producing Socially Meaningful Linguistic Variation

Molly Babel

Department of Linguistics, University of British Columbia, Vancouver

Benjamin Munson

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

1. Introduction

It is axiomatic that spoken language is highly variable, and that variability can be observed at nearly every level of linguistic structure, from the acoustic instantiation of speech sounds to the information structure of long stretches of spoken discourse. Many of the chapters in this volume discuss empirical and theoretical studies of the production of linguistic structures. A challenge to these has been to model the generation of variable linguistic forms. This chapter focuses on one type of variation, that which is related to social categories and social functions. We intentionally define 'social' broadly in this chapter. Consequently, this chapter reviews numerous bodies of literature, including studies of linguistic differences between groups that differ in macrosociological categories like gender, age, social class, and regional origin, linguistic differences that reflect stances, attitudes, and ideologies, and linguistic differences that are elicited in experimental tasks that manipulate different social variables.

The purpose of this chapter is to consider empirical and theoretical studies of language variation, and to suggest how they might be linked to the psycholinguistic models of production that are the focus of this volume. The specific focus of this chapter is on variability in speech sounds. The reason for this focus is partly practical and partly theoretical. Practically, this is because of the expertise of the authors: both of us are laboratory phonologists who study social variation in speech-sound production, processing, representation, and acquisition. Another practical reason relates to the content of the existing literature on variation: there is simply a larger body of research on speech-sound variation than on syntactic variation. The other motivation is theoretical, and relates to the imbalance in existing literature. Early variationist studies of language were based on linguistic models that posited a strong disconnect between 'competence' and 'performance'. Sociolinguistic variation was seen as variation in performance

in a group of speakers whose competence was identical. In this view, phonological variation was seen as cases of variation in the performance of phonological categories whose abstract competence-level representations were equivalent across dialects. In contrast, syntactic variation was less amenable to this kind of modeling, as syntactic structures were seen as the product of a finite-state grammar that operated at the level of linguistic competence (e.g., Lavendara, 1978; Romain, 1984). The strict competence-performance distinction in phonology has been critiqued extensively in recent work (e.g., see Munson, Edwards, and Beckman, *in press*, and Pierrehumbert, 2003). Though some work has extended this critique to syntax (e.g., Manning, 2003; Troutman, Clark, & Goldrick, 2008), research on socially meaningful syntactic variation remains very much in its infancy, and will not be discussed in this paper.

By way of introduction, imagine the task of modeling the production of variants of /æ/ vowel, as in the word *cat*. Regional vernacular pronunciations of this vowel vary substantially. In the Inland North (using the geographic regions defined by Labov, Ash, & Boberg, 2006), the pronunciation is highly diphthongal, gliding from a high-front vowel to a mid- to low-front vowel, notated by Labov and colleagues as /iə/. In North California varieties, the vowel is typically lower and more central than in other dialects. In some Northeastern dialects, this vowel is pronounced differently in different phonetic contexts: prior to heterosyllabic non-dorsal nasals, it is pronounced as in the Inland North, while in other phonetic contexts it is similar to its pronunciation in the pan-dialectal 'standard' used in news media. Within dialects, pronunciation is also highly variable. In one small-scale study of the North dialect region, the use of different /æ/ variants was found to differ as a function of self and self-stated sexual orientation (Munson, McDonald, DeBoe, & White, 2006).

At first glance, two factors appear to govern this variation: regional dialect and social identity. If this were true, then modeling the production of dialectal variants would be easy to accomplish. Dialectal variants could be modeled by assuming that individuals were simply emulating the parametric phonetic detail that they were exposed to during acquisition. Modeling the effects of social identities on production would be similarly trivial if it were presumed that the acquisition of these variants was due to selective attention to selected speakers during childhood and consistent emulation of the phonetic characteristics of those speakers.

Unfortunately, this simple solution to modeling variation is likely to fail very quickly. A closer look at individuals' pronunciation patterns shows even more complex patterns of variation. Consider, for example, the pronunciation patterns of widely-known US television host Oprah Winfrey. Hay, Jannedy, and Mendoza-Denton (1999) showed that the degree of monophthongization in /aɪ/ in Ms. Winfrey's speech varied as a function both of word frequency and of the identity of the person with whom she was speaking, with monophthongal variants being favoured in high-frequency words and words spoken to African-American talk-show guests. Ms. Winfrey grew up in geographic regions where /aɪ/ monophthongization is common (i.e., rural Mississippi), and, presumably, in speaker communities where it is prevalent (i.e., African-American English-speaking communities, which we infer from the racial divisions in the Southern US during the time Ms. Winfrey was growing up). As an adult, she lives in a community in which monophthongization is not prevalent (i.e., among speakers in Upper Class Chicago and in the media community). Hence, her variable pronunciation of /aɪ/ is consistent with the variable models that she has encountered in different speaking contexts across her lifespan. However, the *systematic* nature of this variation suggests that a much more complex

process is at play than simply matching the probabilities of the different forms that have been encountered across a lifetime of language use. Specifically, it suggests that both cognitive and social processes work in concert to select different forms for production depending on the linguistic content, the meaning of the message to be conveyed, and the social context. A second telling example concerns the production of the word *Iraq*. Hall-Lew, Coppock, and Starr (2010) showed that differences in the production of the second vowel in the word *Iraq* were correlated with political attitudes, with /æ/ occurring more often in the speech of U.S. Congressmen and women who support the Iraq war and /a/ in those who opposed it. The relationship between the pronunciation of *Iraq* and political attitudes was most pronounced in speakers from dialect regions that don't have a global preference for /æ/ or /a/ in the second syllable of similar words. Again, this variability suggests that pronunciation variability is the product of a complex set of interactions among a variety of forces.

The structure of the remainder of this chapter is as follows. Section 2 reviews studies of linguistic variation, chronicling a brief history of sociolinguistic and stylistic variation in studies that use quasi population-based samples, and in ethnographic studies of variation within individuals and communities. Section 3 reviews laboratory studies of variation, including both studies of the perception of variation and studies of its production. Section 4 provides suggestion for future research on this topic. The chapter concludes in Section 5 with a brief summary.

2. Variation.

2.1 Linguistic and physiological variation

A necessary prerequisite to building models of the production of variable forms is to understand some of the natural sources of variation in the signal. That is, even without variation based on social and cultural categories like class or race, given the nature of how speech is

produced and how sounds are strung together into hierarchical sequences, significant variation in the signal is inevitable. In a classic article, Ladefoged and Broadbent (1957) presented a taxonomy that argued that variation can be roughly classified as originating in one of three sources: social, physiological, and linguistic. There is considerable overlap between these three sources of variation, such that one speech-sound variant cannot necessarily be classified as a member of just one of these categories. Since social variation is the theme of the chapter, the majority of the discussion in this section will be on that type of variation. Before moving on, however, we will briefly describe the physiological and linguistic sources of variation in turn.

The first in Ladefoged and Broadbent's taxonomy is called linguistic variation, for lack of a better term. This class of variation should be very familiar to readers of this handbook, and includes any variation that is secondary to linguistic organization. In the domain of speech sounds, this includes such phenomena as final strengthening and coarticulation. These processes are theoretically common to an entire speech community, regardless of *who* says them. Final strengthening, for example, is the process by which segments at word, phrase, or even syllable boundary become longer as a result of their prosodic position (Klatt, 1976; Fougeron & Keating, 1997; Cho & Keating, 2001; Keating et al. 2003; Byrd et al., 2000, 2005). Coarticulation is the way in which linguistic segments do not exist as actual individual segments, but inevitably involve the acoustic qualities of surrounding segments. In other terms, each phonetic segment sounds different depending on the segmental context in which it is embedded. Some theories treat coarticulation as a mechanical consequence based on inertia of the articulators (Lindblom, 1990; Browman & Goldstein, 1990), but research has found that coarticulation is largely planned (Nolan, Holst, & Kühnert, 1996; Whalen, 1990) and its degree varies across languages (Hombert, Ohala, & Ewan, 1979; Manuel, 1990).

When one understands how speech is produced, it becomes apparent that physiological variation will inevitably be reflected in sounds' spectral characteristics. The acoustics of the speech signal are determined by two main factors: the noise source and the filter through which that noise passes. In the production of voiced sounds with a relatively open vocal tract, the vibration of the vocal folds is the noise source and the supraglottal cavity is the filter. Differences in the size and shape of the vocal folds contribute to inter-speaker differences in pitch and voice quality. A talker with bigger vocal folds will have a voice that is lower in pitch than a talker with smaller vocal folds. Voice quality refers to whether a talker's voice is breathy or creaky, and it is determined largely by the closure duration of the vocal folds in the course of vibration, and by whether full vocal-fold closure is achieved. The size and morphology of the supraglottal cavity determines the resonances of the filter. Again, the larger this oral cavity, the lower the resonant frequencies produced by the vocal tract; this property determines the position of a speaker's vowels within the acoustic-phonetic space of possible human vocalizations. Thus, individual anatomical and physiological differences contribute significantly to the variation within the signal, without necessarily acting as an ideological or social identity marker.

On the surface, linguistic and anatomical variation appear to explain some of the speech-sound variation we see between different social groups. First, consider sex. Men's voices cluster together in having lower pitch and lower formant frequencies than women's voices. These differences are likely a consequence in part off the fact that men's vocal tracts and vocal folds are typically larger than women's. Next, consider age. Anatomical and physiological changes with age also naturally affect speech production. The extrinsic muscles that support the larynx become slack with age, and the mucosal tissue that covers the vocal folds loses its elasticity. This tends to change voice quality, and lower overall pitch and formant frequencies.

Thus, at least some age and gender differences in the speech-sound production plausibly arise from sex- and age-based anatomical and physiological variation.

Again, a closer look at a large set of data shows that variation in anatomy and physiology cannot account for all of the observed differences in groups that vary in gender or age. Classic studies describing the vowel patterns of males and females (Peterson & Barney, 1952) indeed find that women exhibit the predicted higher pitch and higher formant frequencies than men. However, a recent cross-linguistic study of male-female differences in vowel production found that the extent of these sex differences vary across languages, even when population height is controlled for statistically (Johnson, 2006). This suggests that some of the gender-based differences in speech production are the result of learned, socially and culturally specific norms. This conjecture is supported by the findings of Van Bezooijen (1995), who examined the perception of vocal pitch in Japanese- and Dutch-speaking listeners. Van Bezooijen argued that Dutch and Japanese cultures place different weights on producing a canonically masculine or feminine voice. Given this, we might predict greater pitch differentiation between men and women in Japan than in the Netherlands, a finding partially supported by Yamazawa and Hollien (1992).

Moreover, children acquire some sex-specific speech characteristics before puberty, which is generally assumed to be the point at which the anatomical differences between males and females emerge (Sachs, Lieberman, & Erickson, 1973; Perry, Ohde, & Ashmead, 2001, though see Vorperian, Wang, and Chung, 2009 for recent evidence of sex differences in vocal-tract anatomy prior to puberty). Hence, despite the pre-existence of sex-related differences in speech production, the magnitude of the differences between men and women across cultures,

and the age at which these differences are acquired indicates that many of the speech patterns partly are learned, socially and culturally specific behaviours.

The inadequacy of anatomic and physiological variation in explaining the full range of variation in speech is illustrated well by studies of sexual orientation and speech. Studies have found that self-identified members of the gay, lesbian, or bisexual (GLB) community use speech styles that are distinctive from those of the broader speech community (Gaudio, 1994; Crist, 1997; Linville, 1998; Smyth, Jacobs, & Rogers, 2003; Moonwomon-Baird, 1997; Pierrehumbert, Bent, Munson, Bradlow, & Bailey, 2004; Munson, McDonald, DeBoe, & White, 2006) and that listeners are indeed sensitive to their variation. That is, GLB individuals can be identified as such at above chance levels from audio-only content-neutral speech samples (Linville, 1998; Carahaly, 2000; Smyth, Jacobs, & Rogers, 2003; Gaudio, 1994; Munson, McDonald, DeBoe, & White, 2006, see Munson and Babel, 2007, for a much broader and more in-depth treatment of sexual orientation and speech.) The specific speech differences that have been observed in these studies cannot be reduced to simple anatomic or physiologic differences between GLB and heterosexual people, at least insomuch as we can infer from acoustic analyses of their speech. Hence, they must reflect in part learned, socially and culturally specific processes. We imagine that this situation is particularly complex when considering cultures in which sexualities are classified differently from how they are classified in English-speaking communities in North America (i.e., labels that go beyond the gender and sexuality labels used in Western cultures, like *man, woman, gay, straight, and bisexual*, such as the *hijras* of Indian society who have been shown to use distinctive linguistic forms (Hall & O'Donovan, 1996; Hall, 2002).

2.1 The Sociolinguistic Method

Having described what social variation is *not*, we now turn to what it *is*. The answer to this question is well illustrated through a brief summary of a seminal study on this topic, specifically Labov's pioneering (1963) work on the distribution of speech-sound variants in Martha's Vineyard. This study arguably set the stage for the examination of how social structures motivate linguistic variation. In that study, sociolinguistic variation in a community was defined as differences in the incidence of linguistic forms across major sociological categories like gender, age, and race. Labov (1963) chronicled the centralization of the diphthongs /ai/ and /au/—hallmarks of the local dialect—as a function three such variables: age, ethnic, and occupational groups. Labov's study occurred during a critical shift in the demographics in Martha's Vineyard, when the traditional fishing lifestyle was supplanted by a burgeoning tourism industry. Labov found that Vinelanders most threatened by tourism—men aged 31 to 45 years—had the highest rates of diphthong centralization. He interpreted this finding as evidence that centralization had a social meaning, which can be summarized broadly as "Vineyarde." That is, the fisherman's use of this variant was seen as marking their identity through their use of more-conservative pronunciation variants.

The general method in Labov (1963) fast became the norm in language variation research: identify a variable undergoing a sound change, examine the incidence of different variants of this linguistic variable in different major sociological categories in production tasks varying in complexity. The sociolinguistic stratification observed by Labov (1963) is regarded by many researchers as distinct from a second type of variation, stylistic variation. Stylistic variation is illustrated by Labov's (1966) in study of New York City English. In that study, spoken language was elicited from individuals in sociolinguistic interviews that had a variety of components intended to elicit categorically different productions: casual conversation, careful

speech, the reading of a short passage, word lists, and minimal pair lists. The speakers in Labov's study came from a range of social classes and the linguistic variable of interest was the presence or absence of post-vocalic /ɹ/. The principle finding of this study was that, while socioeconomic class groups differed *sociolinguistically* with respect to the use of post-vocalic /ɹ/ – with /ɹ/ being more frequent in the middle class groups than the working and lower class groups – all social groups exhibited the same *stylistic* pattern. For all groups the lowest rates of post-vocalic /ɹ/ were found in casual speech and the highest rates were found in readings of a minimal pair wordlist. This is taken as evidence that the minimal pair word-list elicits the most careful and formal speech style.

Giles (1973) argued that style shifting observed in sociolinguistic interviews like those in Labov's New York City study was an experimenter effect. That is, the subjects of the sociolinguistic interviews were accommodating their speech to the interviewer, who was modifying *his* speech style. Giles supports this argument with the results a study he conducted in Bristol, where Bristol-accented speakers were either interviewed by a native speaker of Received Pronunciation (RP, the prescriptive standard for British English used by, among others, members of the British Broadcasting Corporation and by high-ranking public officials) or a native speaker of the regional variety spoken in Bristol. An independent group of listeners judged the voices of those who had interacted with the native Bristol interviewer as having stronger Bristol accents than those who had been interviewed by the native RP speaker. This result was based on running speech; hence, the accentedness judgments could have been based on the use of phonetic, syntactic, or lexical variables. Even without knowledge of *what* cued listeners' accentedness judgments, Giles demonstrated that shifts in speech style may be prompted by the speech style of

the interlocutor. Trudgill (1981) revealed a different and more intricate pattern in his sociolinguistic interviews. In interviews with speakers of Norwich English, the sociolinguist interviewer was found to move toward the level of /t/-glottalization of his interviewees, but not to their degree of /a/-fronting, two linguistic features of Norwich English that vary as a function of speech style within speakers of that dialect. These two linguistic features, /t/-glottalization and /a/-fronting, differ with respect to their social salience in the speech community, with /a/-fronting being a linguistic variable that is ostensibly not consciously monitored by speakers or listeners. From his results, Trudgill concludes that accommodation, but not necessarily style shifting, occurs with only socially salient linguistic variables, and that the two processes are fundamentally different.

Variation in speech style is also illustrated in Coupland's (1980) case study of a single female travel agent in her workplace. Coupland tracked her use of five phonological variables while she interacted with clients, co-workers, family, and friends about a variety of topics, either in person or on the telephone. The frequency with which the travel agent used particular linguistic variables was correlated with the complex contextual dynamic – what she was speaking about, to whom, and through what communication medium.

Labov, Trudgill, and Giles' studies lend themselves to a definition of speech style as variation in speech patterns in accordance with social context and demands. However, there is an active debate in the community of sociolinguistics on what is meant by style. Eckert (2001) defines style as any number of linguistics variables that indicate or mark a social identity on either an individual or group level, similar to the kind variation found in Coupland's study. Others view style as the sociolinguistic variation exhibited by a single speaker (Labov 2001, Bell 2001). Tied up in this latter meaning of style is the assumption that a variable will not be

available for stylistic variation unless it is already in the speech community as a sociolinguistic variable. More recent analysis of speech style highlight how a single linguistic variable may be used in multiple contexts, and have a different social meaning in each of these contexts. For example, Podesva's (2007) study of a single gay man communicating in both professional and social contexts showed that the same variable, final stop-consonant release, can have the meaning "erudite" in one context and "prissy" in another

Are speakers aware of the variation they are producing? In a great deal of sociolinguistic research, describing a process as social is often tantamount to describing the process as intentional. For example, Labov (2001:85) described style shifting as a "controlled device" of language use. Similarly, Eckert (2001:124) sees language use on the whole as a "fairly low-level process" but the social use of language is subject to "conscious manipulation". The earliest accounts of style, such as that of Labov (1966), argued that shifts to more standard speech styles would occur as a function of the amount of attention paid to speech. The key evidence for this claim was the higher proportion of non-standard forms in spontaneous speech than in scripted speech or word list readings. Presumably spontaneous speech has higher cognitive demands than does reading. Recent research has shown that this is clearly not always the case and that multiple factors may go into the selection of a speech style. Sharma (2005) illustrates that when speakers of Indian English in California pay more attention to their speech, they become *less* standard in their usage. Bell's (1984) investigation of four radio newsreaders in New Zealand also supports the argument that attention is not the sole determiner of style. In this study, Bell found that the four radio newsreaders used different stylistic variants when reading the news on a community radio station as opposed to a national radio station. In both contexts, the newscasters were reading the news and, therefore, using a style of speech demanding relatively high amounts of

attention, but it was the audience to which the message was being delivered determining the speech style and not the act of reading.

These data, and others like it, provided the backbone for an alternative model of stylistic variation, Bell's theory of audience design. Within this model, a speaker is thought to consider not only the addressee, but also individuals who are in the roles of auditor, over-hearer, eavesdropper, and social group (Bell 1984, 2001) in selecting his or her speech style. Bell (1984:167) suggests three ways in which a speaker may select a stylistic variant: (1) A speaker evaluates an addressee's personal traits and a speech style is selected accordingly; (2) A speaker evaluates the general speech style of the addressee and a speech style is designed accordingly; or (3) A speaker listens for the use of particular linguistic variables and selects a speech style that reflects the speaker's uses. The underlying mechanisms or the level of automaticity with respect to *how* a stylistic variant is selected is not provided in Bell's model, although it is acknowledged that a full understanding of a speaker's social and psychological networks, biases, and values are necessary to predict a speaker's behaviour (Bell 1984:169). Bell's model is unique in sociolinguistic theory in that it attempts to predict when variation will arise in the signal, in lieu of simply describing its existence. In a similar vein, we will reconnect the sociolinguistic concept of style with current speech production research below in Section 3.2.

3. The Laboratory Method

3.1 Perception

The vast majority of studies of sociolinguistic variation have used the methods illustrated by Labov's (1963) study of Martha's Vineyard and his work on New York City English (Labov, 1966). This section reviews laboratory research on the perception of social categories through linguistic variation, as well as the influence of socially based expectations on linguistic

processing. In the context of a volume on production processes, we see a section on perception as somewhat of a necessary evil. It is *necessary* because a model of production can only exist if there is first a model of the target message that talkers are attempting. Perception studies are one way of understanding the nature of this message. It is an *evil*—at least figuratively speaking—because the results of perception studies give at best an incomplete picture of production processes, and at worst a misleading one, a point we elaborate on below. This section is a much briefer and slightly more current summary of perception studies given by Thomas (2002). We refer readers to that article for a fuller discussion of these studies and the issues they raise.

The first set of perception studies we review examines the attributes that listeners associate with different linguistic variants. Just as much of the research on sociolinguistics has focused on the influence of major sociological categories like age, gender, and social class on variation, so has much of the research examining the perception of these same categories. Some groups are perceived extremely robustly from phonetic variation. For example, listeners can identify a talker's sex from very brief signals that have had all linguistic content effectively removed (i.e., Bacharowski & Owren, 1999). Listeners can identify a talker's ethnicity, at least in tasks examining the perception of speech produced by African-American and Caucasian talkers. This is true when the categorical linguistic variation characteristic of African-American Vernacular English is removed (e.g., Purnell, Baugh, & Idsardi, 1999); however, this ability is much more strongly compromised when the signal is degraded than is the identification of sex (Lass, Almerino, Jordan, & Walsh, 1980). Listeners are able to identify regional dialects at greater-than-chance levels. In one study of American English (Clopper & Pisoni, 2004a), naïve listeners in one Midwestern US town were found to be able to discriminate among three broad dialect regions. A subsequent study with the same stimuli found that individuals who had lived

in multiple dialect regions were more accurate in classifying talkers than were those who had only lived in the Midwestern US (Clopper & Pisoni, 2004b). More recent research has examined social categories that were not examined in early sociolinguistic research in the variationist tradition. Munson, McDonald, DeBoe, and White (2006) found that listeners could identify talkers' sexual orientation at greater-than-chance levels from productions of single words. Not all social variables are equally perceptible, even to familiar listeners. Drager (2010) reports on an ethonographic study of phonetic variation among different social groups in a single girl's high school in Christchurch, New Zealand. She found that the degree of glottalization in the vowel in the word *like* varied as a function of the word's function and of the social group of the speakers who produced it. A subsequent perception study found that listeners from the same population (i.e., students at the same school) could not identify either the function of like or the social group of the speaker who produced it from the degree of glottalization. A general tactic in all of these studies is to analyze which acoustic parameters are associated with which social judgments. For example, Clopper and Pisoni (2004a) found that a weak or absent postvocalic /ɹ/ was associated with listener judgments of a speaker being from the Northeast or New England. Munson et al. (2006) found that talkers who were identified as gay-sounding produced tokens of /s/ with a high peak frequency and a compact spectrum. This was also found to differ between self-identified gay and heterosexual men.

A second set of perception studies have shown that listeners calibrate their regular linguistic perception based on knowledge of or expectations about socially meaningful linguistic variation. One particularly striking example of this is given by Jansen and Schulman (1983), who examined the perception of a synthetic /ɛ/-/æ/ continuum by speakers of Swedish. The /ɛ/-/æ/ contrast is preserved in some, but not all, dialects of Swedish. Jansen and Schulman found

that listeners perceived this contrast when they thought they were listening to a speaker from a dialect that preserved the contrast, but not when they thought they were listening to a speaker from one that did not. Similar results are presented by Drager (2011), Hay, Warren, and Drager (2006), Johnson, Strand, and D'Imperio (1999), Niedzielski (1999), Staum Casasanto (2008), and Strand and Johnson (1996). In each of those studies, listeners' identification of phonemes was affected by an experimental manipulation suggesting a particular attribute about a speaker.

Interestingly, this calibration does not extend to all perception tasks. Walker (2008) found that listeners did not judge nonstandard syntactic structures differently depending on whether they thought they were spoken by someone from a social group (here, New Zealanders with lower socioeconomic status) in which they are commonly used. Kraljic, Brennan, and Samuel (2008) and Kraljic, Samuel, and Brennan (2008) showed that listeners rapidly adapt to pronunciation characteristics of particular speakers and generalize them to new speakers who are presumed to share the same source of pronunciation variation (i.e., ones who they believe to speak the same dialect). Kraljic and colleagues also demonstrated that listeners are able to discern between sources of variation that are unique to speech communities (like dialects), to individuals (like those related to individual vocal-tract morphology), and to idiosyncratic conditions that led to a particular production (like having a pen in the mouth while speaking).

How do these studies inform models of the production of social variation? Consider the case of sexual orientation and speech. Given the findings by Munson et al. (2006), we might conclude that the particular high-frequency, compact-spectrum /s/ used by gay men and associated with judgments of gay-soundingness means *gay*, much in the way that the sequence of articulations and their acoustic consequences which are denoted by the phonetic symbols /kæt/ means 'member of the class of animals *felis catus*'. Producing this /s/ could be modeled by

dictating that the articulatory spell-out of a lexical item would include a particular variant if that item was indexed to the social meaning 'gay'. One problem that this conjecture faces is that the same variant is potentially subject to many different interpretations. We can illustrate this with studies of two phenomena: the widely attested variation in the place of articulation of the nasal in the *-ing* morpheme by speakers US dialects of English and variation in the spectral detail of the vowel /æ/ by English speakers in the upper Midwestern US. Campbell-Kibler (2007) showed that the use of an alveolar variant in the *-ing* morpheme can increase the tendency for a speaker to be labelled as Southern, and decrease the tendency to be rated as gay, though these specific perceptions seem to be strongly tied to other attributes about the same talker. That is, simply saying an [n] in *-ing* does not guarantee that someone will be labelled as heterosexual and Southern, but instead interacts with the perception of other attributes about the speaker. Smith, Hall, and Munson (2010) examined the perception of two variants of /æ/: one with a high-front on-glide, characteristic of a set of pronunciations known collectively as the Northern Cities Chain Shift (NCCS), and low, retracted variant, characteristic of a different set of pronunciations, the California Chain Shift (CCS). Previous work by Munson et al. (2006) showed that gay men produced vowels that were broadly similar to those produced by CCS speakers, and heterosexual men produced vowels more like those of the NCCS. In a perception experiment using single words, listeners identified NCCS vowels as more heterosexual sounding and CCS vowels as more gay sounding. Smith et al. examined the perception of tokens of /æ/ produced by trained speakers that had intentionally exaggerated NCCS and CCS characteristics. While they failed to replicate Munson et al.'s original finding, they did show that different variants of /æ/ were associated with a variety of perception, including ones about the talker's age, height, and health habits (i.e., whether or not they smoked or drank alcohol). This finding was

quite unexpected—indeed, these items were included only as fillers in an experiment whose primary goals were to examine the perception of sexual orientation! Nonetheless, the finding is further valuable evidence that a single linguistic variant can be perceived as indexing multiple social meanings.

As discussed in Munson (2010), there are two logical interpretations of the findings reviewed in the previous paragraph. The first is that variants like NCCS /æ/ are essentially homophones, much like the homophones that exist in other form-meaning relationships. The other is that the interpretations that people make reflect a hierarchy of meanings associated with forms like NCCS /æ/ vs. CCS /æ/, or alveolar-in vs. velar-final -ing. This latter hypothesis has been discussed at length in recent works in sociolinguistics, including Silverstein (2003), Eckert (2008), and Campbell-Kibler (2009). Clearly, resolving the nature of the representation of socially meaningful variables is key to developing psycholinguistic models of its production.

3.2 Speech production

The focus of this section is on studies of phonetic accommodation by psychologists, sociologists, and laboratory-oriented linguists. However, as described above, sociolinguists have long been interested in accommodation (e.g., Giles, 1973; Trudgill, 1981) and have recently displayed renewed interest in its role in dialect contact (Trudgill, 2004, 2008). Hence, before discussing recent laboratory studies, it is worthwhile revisiting the foundations of sociolinguistics, as it relates to current work on variation in speech production. In perhaps its earliest inception, sociolinguistics was a field interested in modeling sound change. The original list of primary interests included the *actuation* problem and the *transition* problem (Weinreich, Labov, & Herzog, 1968). The actuation problem involves understanding why a particular sound change is introduced into a particular speech community at a particular time. That issue is

outside of the scope of this chapter. The transition problem, however, revolves around the path of the sound change, whether it behaves gradually or categorically. In addition to this original goal of the field, we can consider another type of path as well: how does the sound change transition through the community? In this section on speech production, we do not review the literature on macro-level sound changes, but rather we focus on laboratory- and corpus-based studies of phonetic convergence and imitation. This line of research has been fruitful in integrating methods and models of speech production that attempt to predict phonetic and phonological variation in the signal. This work is truly interdisciplinary, and represents an intersection of psychology, sociology, and several branches of linguistics. An understanding of an individual's social network also has bearing on the question of how a sound change spreads through a speech community (Milroy & Milroy, 1985), but from a different theoretical and methodological perspective.

From the beginning, studies of phonetic convergence have been concerned with the social and contextual effects of speech production. Phonetic convergence (sometimes referred to as *phonetic imitation* or *accommodation*) is the process by which interacting talkers come to be more similar, both acoustically and perceptually. Some of the earliest studies considered the psychological and social factors that affect the direction of talkers' convergent or divergent behaviour. For example, Natale (1975a) examined convergence of mean vocal intensity between conversational dyads. The intensity level of a confederate to the experiment was instrumentally manipulated three times through the course of the conversations; subjects generally converged toward towards the intensity level of the confederate. Natale's second task also used conversation dyads. Prior to participating, individuals completed the Marlowe-Crown Social Desirability test (Crown & Marlowe, 1964), which evaluates an individual's desire to be accepted

by society. Through the course of three conversations, participants converged and individuals' contribution to their dyad's amount of convergence was positively correlated with their Marlowe-Crown score. In another study, Natale (1975b) examined same-sex dyads' convergence of temporal patterns across two conversations. Again, using the Marlowe-Crown test, Natale demonstrated that participants' degree of convergence was predictable by their desire to be accepted socially. This result was only significant, however, in the second conversation. This suggests that increased familiarity and more than a passing level of social engagement with a conversational partner prompts convergent speech behaviour.

Gregory and colleagues (Gregory, Webster, & Huang, 1993; Gregory & Webster, 1996; Gregory, Dagan, & Webster., 1997; Gregory, Green, Carrothers, & Dagan, 2001) examined the role of vocal fundamental frequency (f_0 , changes in which are perceived as changes in vocal pitch) in phonetic convergence in conversations. Gregory et al. (1993) examined a corpus of twelve American English telephone conversations along with a corpus of eleven conversations from an Egyptian Arabic database. Spectral measures were averaged from a 62-192 Hz band-delimited region of the speech signal and with this long-term-average spectra (LTAS) measure, Gregory and colleagues found convergence between the dyads. American English listeners evaluated the conversations from the English-speaking corpus in terms of how smoothly they went. This independent listener group rated the interviews more favourably when accommodation had occurred, suggesting the convergence between interlocutors affects the perceived quality of the conversation.

The role of social status in accommodation was explored by Gregory and Webster (1996). Drawing upon Communication Accommodation Theory (Giles & Coupland, 1991, a model that emphasizes the role of language as a tool for asserting identity and social relations),

Gregory and colleagues interpreted the phenomenon of accommodation as a subconscious response to an interlocutor's higher status. They theorized that accommodation was an adaptation to another's communicative behavior in the service of achieving a social goal or being accepted. Gregory and Webster (1996) measured accommodation in interview excerpts from the Larry King Live television program (a popular television program where a host, Larry King, interviews a range of guests, including both high-profile celebrities and politicians, and non-celebrities). The social status of the guest was used to predict King's behavior, specifically whether King would accommodate toward high status guests while lower status guests would accommodate toward him. Gregory and Webster's analysis indeed found that King accommodated more toward high status guests than lower status guests. The results from the f0 data illustrate that Larry King modulated his pitch less during interactions with talkers of a lower social status.

In another study, Gregory et al. (1997) filtered the audio signal of one member of a conversational dyad. Dyads in the control condition and those who were low-pass filtered condition exhibited convergence, while dyads in a high-pass filtered condition where the f0 frequency region had been removed did not. Groups of listeners also rated the conversations for quality. Generally, judgments were more negative for conversations that had been filtered, although the low-pass group did receive slightly more favorable ratings than the high-pass group. Gregory et al. (1997) argue their result indicates that low frequency energy plays a significant role in convergence because of its role as a salient phonetic feature that transmits social information by conveying emotion and attitude. Using a similar design, Gregory et al. (2001) sought to determine how visual information influences accommodation. They replicated their finding that accommodation occurred in dyads in the control and low-pass filtered conditions.

Additionally, Gregory and colleagues found that along with not accommodating phonetically, dyads in the high-pass filtered condition did not look up and interact visually as much as the dyads in the other conditions. This finding suggests that accommodating an interlocutor is part of the social structure of communication and when it does not go as intended, the level of engagement declines.

One recent and influential study on phonetic convergence is reported by Pardo (2006). Pardo examined phonetic convergence in same-gender dyads involved in jointly completing a map task, where one member of the dyad was the giver of map directions and one was the receiver whose task it was to navigate the dictated path. Pardo used an AXB task to examine the extent to which listeners modified their productions to match those of their conversational partner. In this AXB task, triplets of words were played to naïve listeners. The middle "X" word was one person's production of a word elicited in a conversational dyad. The first and third "A" and "B" words were productions of the same word from the other conversational partner either in the conversation task itself (what Pardo called *task repetitions*), produced in a pre-task reading list, or produced in a post-task reading list. Listeners judged which of the A or B words sounded most like the X word. Convergence was assumed to occur when the task repetitions were rated to sound more like the target than the pre- and post-tasks readings were. Using data from this AXB task, dyads were perceived to have converged on 62% of the experiment trials. Female dyads were found to converge toward the speaker who was receiving instructions, whereas male dyads patterned oppositely; they converged toward the speech of the male talkers giving instructions. Pardo concludes that particular social factors dependent on the situational context of a conversation determine the direction of phonetic accommodation. Subsequent research by Pardo, Jay, and Krauss (2010) found that when one member of a map-task dyad was given the

instruction to imitate the other, social and interaction factors often overrode the instructions to imitate. While they replicated Pardo's original finding that male dyads converged more than female dyads, pairs were judged to have converged when the map-directions receivers were instructed to imitate, but when the givers of directions were instructed to imitate, most dyads were judged to have diverged.

Pardo (2010) reports on the acoustic analyses of the speech of the same-gender dyads along with presenting results from the mixed-gender dyads whose data are not discussed in Pardo (2006). Like same-gender dyads, mixed-gender pairs converged, albeit to a lesser extent; listeners perceived mixed-gender pairs to have converged on 53% of the trials. Pardo (2010) also analyzed F0 and duration data to determine the cues on which listeners based their judgments; the acoustic measures were taken as the difference of F0 and duration between each pair for each AXB trial. These values accounted for 41% of the variance for the female talkers, but only 7% of the variance for the male data. Pardo also compared formant frequencies of vowels produced by participants in pre- and post-task sessions. Talkers were found to diverge or converge with their conversational partner, depending on the talkers' role (receiver or giver) and vowel identity; the high vowels converged and low vowels diverged. Givers were found to centralize their vowel space more than receivers. These results are important because they indicate that convergent phonetic behaviour does not stop simultaneously with the conversational partner with whom a talker is converging. Moreover, the vocalic changes found by Pardo (2010) suggest that the talker's entire phonological system was affected by exposure to the map task partner.

Researchers have also explored phonetic convergence using an auditory-naming paradigm – this paradigm elicits productions from research participants by having them repeat

(i.e., name) words after hearing another talker say the words (i.e., the auditory object) over headphones or a loudspeaker. This paradigm uses single-word shadowing from a model talker as a way to elicit convergent or accommodating speech behavior. While this method removes most, if not all, of the social aspects of actual language interaction, it is a useful tool as it demonstrates how even in the absence of social motivation, phonetic convergence (often termed *spontaneous imitation* when using this method) still occurs. Goldinger (1998), for example, found that the degree of phonetic imitation interacts with word frequency and amount of exposure. Lower frequency words undergo more imitation than higher frequency words and the more often a model talker's production of a word is heard, the more a participant will spontaneously imitate its acoustic properties. This was later replicated in a study by Goldinger and Azuma (2004). Namy, Nygaard, and Sauerteig (2002) also replicated Goldinger's findings and expanded the result in revealing that female participants, in their study, imitated more than male participants. In using single-word production, researchers are able to probe in more detail what in the acoustic detail is imitated. With this in mind, Shockley, Sabadini, and Fowler (2004) has revealed that participants readily imitate lengthened voice onset time in aspirated American English stops. Nielsen (2011) extended this finding in demonstrating that participants imitate lengthened VOT and generalize lengthened VOT to all voiceless stops, not just those presented during an exposure phase. Following up on claims made by Gregory and colleagues, Babel and Bulatov (in press) used an auditory-naming paradigm to demonstrate that participants imitate less when presented with high-pass filtered single words, but that listener judgments of imitation and perceptual similarity do not simply correspond to the degree to which participants imitated f0.

We can gather from this research that imitation and convergent behaviour appear to be, perhaps, an inevitable phenomenon. This idea is supported not only by the fact that it occurs

across the language system (Pickering & Garrod, 2004), but by the fact that it is a pervasive behaviour across human behaviour in general (Chartrand & Bargh, 1996; Dijksterhuis & Bargh, 2001; Bargh & Williams, 2006). Therefore, recent research has used the auditory-naming paradigm to examine the role of social factors with respect to low-level imitative tendencies.

Babel (2010) presents evidence demonstrating the relationship between social affinity and phonetic imitation. In this task, native speakers of New Zealand English shadowed single-word productions from an Australian model talker. Before auditory exposure to the model talker, half of the New Zealand participants were presented with a short statement regarding the Australian model talker's pejorative feelings toward New Zealand. The other half of the participants were presented with a statement which stated the Australian thought positively of New Zealand. Contrary to predictions based on previous work (Bourhis & Giles, 1977), the condition to which participants were assigned had no bearing on the degree of accommodation. Participants in both conditions accommodated to the speech of the Australian model talker. Participants' levels of implicit bias toward New Zealand and Australia as measured through an Implicit Association Task (Greenwald, McGhee, & Schwartz, 1998) did, however, predict the amount of imitation. Participants with pro-Australia bias were more likely to accommodate to the vowels of the Australian model talker. This suggests that simple liking can bolster the perception-production link (cf. Dijksterhuis & Bargh, 2001). In this study, imitation was not equally distributed across vowels; the DRESS vowel – realized as [i] in New Zealand English and [ε] in Australian English – was the target of the most accommodation. This is worthy of mention as it is one of the vowels with the largest measurable psycho-acoustic differences between New Zealand and Australian Englishes, but is not perceptually one of the most salient differences to naïve listeners (Bayard, 2000; Hay, Nolan, & Drager, 2006). This finding

supports the claim that imitation and the social factors that mediate it are not the results of conscious choice, but are cognitive biases.

Using a participant population that was relatively homogenous with respect to dialect, Babel (under review) further explored the roles of cognitive and social bias in phonetic imitation. In this task, California-based participants were exposed with model talkers who were male native speakers of California English as well. Participants were presented with only one of the model talkers, one of whom was Black and the other was White. Two groups of participants were assigned to an auditory-naming task that included still digital images of the model talkers. In these tasks, listeners were therefore aware of the physical make-up of the model talkers. After the speech production task, participants in the visual condition completed an Implicit Association Task that measured Black and White racial bias. Both male and female participants were also asked to rate the model talkers' attractiveness on a scale of 1 (not at all attractive) to 10 (very attractive). All participants, male and female, succumbed to the same cognitive biases where the low vowels /a/ and /æ/ were imitated more than mid and high vowels /o i u/, but the genders differed with respect to the influence of the social measures on their degree of imitation. Female participants were more likely to imitate the White model talker when they were judged to be more attractive, whereas male participants were less likely to imitate when they judged the White talker as more attractive. The behaviour of female participants with respect to the White talker can be interpreted as following directly from the predictions of Communication Accommodation Theory (e.g., Giles & Coupland, 1991) where accommodation or imitation takes place to decrease the social distance between interlocutors. We can imagine that female participants who view the model talkers positively would desire a decrease in social distance. These findings and these views coincide nicely with more recent work by Bargh and colleagues who suggest that

liking is one of the simplest ways in which to intensify the perception-behavioural link (Chartrand & Bargh, 1996; Dijksterhuis & Bargh, 2001; Bargh & Williams, 2006).

Several key points fall out of the work on imitation. For one, the data suggest that participants default to imitating and converging, at least in laboratory settings. Indeed, recent modeling work demonstrates that convergence is inevitable unless there is in-group pressure to maintain linguistic features (Wedel & Van Volkinburg, in prep.). But, the fact that social factors mediate low-level behaviour such as imitation suggests that speech production is never without social influence. The second major point also relates to the social factors. Social factors such as the desire to fit in, talkers' social status, and liking influence the degree to which talkers imitate. If we consider the dynamic ways in which talkers are modifying their speech in the corpus studies and experiments we have reviewed above as a snapshot of style-shifting in real-world interactions, we can hypothesize that at least some of style-shifting is the result of behavioural imitation that is mediated by social factors.

Convergence, of course, does not always occur. In addition to there being clear instances of linguistic divergence (Bourhis & Giles, 1977), talkers are ineffective at accurately imitating themselves (Vallabha & Tuller, 2004). While there is likely a host of reasons for the lack of convergence, part of the explanation lies in the role of auditory feedback in speech planning. In studies where what listeners hear as their own voices has been resynthesized and manipulated, for example, by lowering the first formant frequency, listeners partially compensate by raising the first formant frequency. This compensation finding has been reported for vowels (Houde & Jordan, 2002; Purcell & Munhall, 2006) fundamental frequency (Jones & Munhall, 2005), and fricative centroids (Schiller, Sato, Gracco, & Baum, 2009). The amount of compensation seems to vary considerably across individuals and is argued to be due to individual differences in the

weighting of auditory and somatosensory feedback (Katseff, Houde, & Johnson, under review). Indeed, Larson, Altman, Lui, & Hain (2008) found that the amount of compensation in fundamental frequency manipulation increased when participants' vocal folds were anesthetized such that they were no longer privy to somatosensory feedback.

As an end to this section, let us visit a phenomenon that brings together social variation, accommodation, and self-monitoring of one's speech: new dialect acquisition in adults. When adults move to a new region, they generally accommodate their speech patterns to those of the new dialect, although they never sound completely native to the new dialect (Munro et al., 1999; Evans & Iverson, 2006). Without physical displacement, adults who are simply exposed to new social dialects also modify their speech patterns, a finding that has been illustrated in studies of changes in the pronunciation of HRM Queen Elizabeth II's pronunciation during her reign (Harrington, 2006, 2007; Harrington, Palethope, & Watson, 2000a,b). Moreover, pronunciation patterns can vacillate when speakers move back and forth between different language communities regularly, such as the adult that Sancier and Fowler (1997) studied, who live part-year in Brazil and part-year in the Northeastern US. We can hypothesize that acquisition of the new dialect features is not complete for adults for myriad reasons, but we would be remiss not to conclude that at least part of the issue revolves around the role of somatosensory and auditory feedback in self-monitoring. As a case in point, Howell, Barry, and Vinson (2006) show how adults who have acquired new dialects resort to their original native dialect when presented with shifted or delayed auditory feedback of their own voices. This finding was obtained with a dramatically different type of distorted feedback. Simply interfering with the normal self-listening process interfered with talkers' ability to monitor their speech style.

4. Areas of Future Research

We end this chapter by considering different emerging areas of research that are crucial for developing psycholinguistics models of the production of socially meaningful linguistic variation. The cornerstone of many models of production—including the classic Leveltian models of production that inspired this volume—is the notion that communication starts out with a representation of the intended message in the mind of the speaker and ends when the listener has accessed that intended meaning. Hence a critical endeavour for researchers modeling the production of socially meaningful variation is to delimit precisely the nature of the meanings associated with the variation that we have reviewed thus far in this chapter. While it is clear that much work needs to be done in this area, we can confidently make two generalizations from work that has been done previously. First, just because a listener *perceives* a linguistic variant as indexing a particular meaning does not mean that the talker *intended* it to index that meaning. Consider again the case of /æ/ variation, discussed in section 3.1. Presumably, speakers who use NCCS /æ/ do not do so because they intend to be perceived as overweight, heavy smokers, and over 60 years old, though that is what the listeners in Smith et al. (2010) perceived talkers who use NCCS to be. Similarly, just because someone is perceived as sounding gay does not mean that 'gay' was the intended meaning. This error of interpretation unfortunately pervades some work on the perception of socially meaningful variation, particularly work related to sexual orientation. For example, Rieger, Linsenmeier, Gygax, Garcia, and Bailey (2010) argue that gay-sounding speech is intentionally produced by gay people to signal sexual orientation to potential sexual partners. This ignores the fact that previous research predicts that the same voices would likely be identified as having many other attributes, such as being more-articulate and younger-sounding than heterosexual talkers, at least for men's voices.

The second conclusions that we can draw is that, if social categories are to be examined experimentally by psycholinguists, then the intended meanings that they index must be understood first. Understanding this will require rigorous investigations drawing on expertise from many different fields, including linguistic anthropology, psycholinguistics, sociolinguistics, and formal semantics, as argued by Smith, Hall, and Munson (2010) and Munson (2010), among others. The way to go about these investigations is by no means straightforward. There is at least one striking difference between the social meanings of linguistic variation and their 'regular' semantic meanings, namely, they are not reinforced by prescriptive instruction in the same way that regular meanings are enforced by language arts instruction in schools. Consequently, many of the investigation tools that require explicit knowledge of a meaning (i.e., providing definitions, giving judgments of similarity in meaning) are likely to be of limited utility in studying social meanings. In this sense, they parallel other types of meanings, such as the pragmatic 'meanings' that guide the interpretation of utterances relative to the larger discourse. Here, a particularly useful model for this is work on the meaning of intonational contours (e.g., Pierrehumbert & Hirschberg, 1990). Work by Pierrehumbert and Hirschberg developed a formalism to account for speakers' use different intonation contours to guide listeners' interpretation of speech. Such formalisms may prove fruitful for modeling the kind of meanings that are conveyed by social variation.

A second area of research that we regard as very important is understanding how socially meaningful variation interacts with and is affected by linguistic variation of the type discussed in section 2.1. There is a small body of research suggesting linguistic variation constrains socially meaningful variation. Consider first the general finding that words in highly semantically predictable contexts are spectrally and temporally reduced (e.g., Lieberman, 1963; Jurafsky,

Bell, Gregory, & Raymond, 2001). Clopper and Pierrehumbert (2008) find that speakers of the Northern Cities dialect of American English produced vowels that were more advanced with respect to the Northern Cities vowel shift in highly semantically predictable positions relative to the same words in unpredictable contexts. What is unknown from this study is whether the low semantically predictable production involves suppression of extreme vowel variants or whether the highly predictable position allows for more extreme-than-usual productions, perhaps due to lack of self-monitoring in such environments. Consider next the general finding that vowels in words that are in dense phonological neighborhoods (i.e., words that are only one-phoneme different from many other real words) are hyperarticulated relative to the same vowels in words in sparse neighborhoods (Munson & Solomon, 2004; Scarborough, 2010; Wright, 2004). This presumably reflects an attempt, either overt or tacit, to maintain the distinctiveness of words in dense neighborhoods: words from dense neighborhoods are generally harder to perceive than ones from sparse neighborhoods, and the hyperarticulated vowels partially counter this. Munson et al. (2006, see also Simpson, 2000 and Pierrehumbert, Bent, Munson, Bradlow, & Bailey, 2004) showed that larger-sized vowel spaces are associated with the speech of less-masculine sounding men and more-feminine sounding women. Munson (2007) showed an interaction between linguistic and social factors in vowel-space size: though more-masculine sounding men produced smaller-sized vowel spaces than did less-masculine sounding men, this difference was absent for high-density words.

Other research has shown that different patterns of socially meaningful variation can emerge in more-or-less difficult speaking tasks. Howell et al. (2006) showed that British English speakers who had relocated to other dialect regions in the UK and had reported losing their native accent were rated by listeners as sounding more like speakers of their original native

dialects when producing speech in two challenging conditions (producing speech with acoustically altered feedback). This finding and those by Clopper and Pierrehumbert and by Munson may stem from a single source, if it were shown that production conditions that favored the production of socially marked variants in those latter two studies were found to be more difficult than the conditions that did not. Clearly, more work in this area is needed. This work should also examine whether the ability to perceive and produce socially meaningful linguistic variation has a reciprocal, positive effect on other aspects of language and cognition. Work by Bialystok and colleagues (e.g., Bialystok, 1999, Bialystok & Martin, 2004) has shown that bilingual children have superior executive function abilities, particularly those related to suppressing task-irrelevant information. It may be that there is a similar non-linguistic cognitive advantage to perceiving and producing multiple, socially meaningful forms within a single language.

The third area of research we consider critical is to extend work on socially meaningful variation to various exceptional (i.e., clinical) populations. Speech production research has long examined production processes in exceptional populations as a means of testing different theories of production, as is reviewed by chapters by Thompson, Dell, and Buchwald in this volume. We consider two populations of great interest in understanding the cognitive bases of socially meaningful variation. The first of these is individuals with Autism Spectrum Disorder (ASD). One influential view of ASD is that it arises in part from deficits in a theory of mind—the ability to correctly infer what others' know about a topic being discussed. Given this, we might predict that individuals with ASD would not make the proper inferences of when to use different linguistic forms socially. Indeed, we might expect that they would be unable to learn the systematic correspondences between speaker attributes and linguistic forms. There is some

support for the latter hypothesis. Baron-Cohen and Staunton (1994) showed that children with ASD being raised by a non-native speaking parent were much more likely to emulate their mother's non-native accent than were children without ASD. The possibility that individuals with ASD might process socially significant phonetic variability differently from other populations is made all the more likely by recent findings by Yu (2010) and Stewart and Ota (2009). These investigators showed that performance on a self-report measure of tendencies associated with ASD by individuals in the normal population predicts sensitivity to acoustic-phonetic detail in speech-perception tasks. Individuals with more behaviors characteristic of ASD showed reduced tendencies to use contextual information in speech perception. The second population of clinical interest is adults with Acquired Brain Injury (ABI). In contrast to strokes, ABI typically results in diffuse brain lesions, often with significant damage to the part of the brain responsible for the executive functions related to cognitive control. Here we predict that individuals with ABI would be considerably poorer at typical individuals in switching among different socially meaningful linguistic variants.

5. Conclusions

This chapter reviewed social variation from a very broad perspective, surveying classic work in traditional sociolinguistics and phonetics, in addition to more recent research on social variation in speech. The trends of the current research on social variation in speech attempt to predict when particular variation will arise based on the speech style of the interlocutor, the semantic predictability of the utterance, and the feelings of the talker toward the interlocutor. We would like to suggest that future work continues along this line, in tandem with traditional sociolinguistic work describing complex patterns of variation in speech and attempts to predict its presence. This often requires using methodologies considered non-traditional in speech

research, but we hope the research reviewed in this chapter has convinced the reader that such a path is worthwhile.

References

- Babel, M. (under review). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. Under review at *Journal of Phonetics*. pp. 1-46.
- Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, 39, 457-456..
- Babel, M., & Bulatov, D. (to appear). The role of fundamental frequency in phonetic accommodation. To appear in *Language and Speech*, 1-36.
- Bachorowski, J., & Owren, M. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *Journal of the Acoustical Society of America*, 106, 1054-1063.
- Baron-Cohen, S., & Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language*, 42, 317-352.
- Bayard, D. (2000). The cultural cringe revisited: changes through time in Kiwi attitudes toward accents. In Allan Bell & Koenraad Kuiper (eds.) *New Zealand English*, 297-324. John Benjamins.
- Bell, A. (1984). Language style as audience design. *Language in Society* 13:145–204.
- Bell, A. (2001). Back in style: reworking audience design. In P. Eckert & J.R. Rickford (Eds.), *Style and Sociolinguistic Variation* (p. 139–169). Cambridge University Press.
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70. 636-644.
- Bialystok, E., & Martin, M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, 7, 325-339.

- Bourhis, R. Y. & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (ed.) *Language, ethnicity, and intergroup relations*, 119–136. Academic Press: London.
- Browman, C. P. & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M.E. Beckman (eds.) *Papers in Laboratory Phonology: Between the Grammar and Physics of Speech*. Cambridge University Press. 341-376.
- Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. Broe & J.B. Pierrehumbert (eds.) *Laboratory Phonology V: Acquisition and the Lexicon* (p. 70-89). Cambridge University Press.
- Byrd, D., Lee, S., Riggs, D., & Adams, J. (2005). Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America*, 118, 3860-3873.
- Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, 82, 32-64.
- Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change*, 21, 135-156.
- Carahaly, L. (2000). Listener accuracy in identifying the sexual orientation of male and female speakers, unpublished MA thesis. Columbus, OH: Department of Speech and Hearing Science, Ohio State University.
- Clopper, C. & Pisoni, D. (2004a). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111-140.

- Clopper, C. G., & Pisoni, D. B. (2004b). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change*, 16, 31-48.
- Crist, S. (1997). Duration of onset consonants in gay male stereotyped speech. *University of Pennsylvania Working Papers in Linguistics*, 4, 53–70.
- Crown, D. P., & Marlowe, D. (1964). *The approval motive - Studies in evaluative dependence*. New York: John Wiley & Sons, Inc.
- Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of social information processing goals: Nonconscious priming reproduces effects of explicit conscious instructions. *Journal of Personality and Social Psychology*, 71, 464-478.
- Coupland, Nikolaus. (1980). Style-shifting in a Cardiff work-setting. *Language in Society*, 9:1-12.
- Dijksterhuis, A. & Bargh, J. (2001). The perception-behavior expressway: Automatic effects of social perception on social behaviour. In M. Zanna (Ed.), *Advances in Experimental Social Psychology*, Vol. 33. 1–40. Academic Press: San Diego.
- Drager, K. (2010). Sensitivity to grammatical and sociophonetic variability in perception. *Journal of Laboratory Phonology*, 1, 93-120.
- Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54 (1), 99-121.
- Eckert, P. (2001). Style and social meaning. In P. Eckert, & J. R. Rickford (Eds.), *Style and Sociolinguistic Variation*, (pp. 119-126). Cambridge: Cambridge University Press.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12, 453-476.
- Eckert, P. & Rickford, J.. (2001). *Style and sociolinguistic variation*. Cambridge University Press.

- Evans, B. G. & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America*, 121:3814–3826.
- Fougeron, C. & Keating, P. (1997). Articulatory strengthening at the edges of prosodic domains. *Journal of the Acoustical Society of America*, 101. 3728-3740.
- Gaudio, R. (1994). Sounding gay: pitch properties in the speech of gay and straight men. *American Speech* 69.30–57.
- Giles, H. (1973). Accent mobility: a model and some data. *Anthropological Linguistics*, 15, 87–105.
- Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences*. Milton Keynes: Open University Press.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279.
- Goldinger, S.D. & Azuma,T. (2004). Episodic memory in printed word naming. *Psychonomic Bulletin & Review*, 11:716–722.
- Greenwald, A.G., McGhee, D.E., & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74:1464–1480.
- Gregory, S. W., Green, B.E., Carrothers, R.M., & Dagan, K.A. (2001). Verifying the primacy of voice fundamental frequency in social status accommodation. *Language Communication*, 21, 37–60.
- Gregory, S. W., Webster, S., & Huang, G. (1993). Voice pitch and amplitude convergence as a metric of quality in dyadic interviews. *Language Communication*, 13, 195–217.

- Gregory, S.W. & Hoyt, B.R. (1982). Conversation partner mutual adaptation as demonstrated by Fourier series analysis. *Journal of Psychological Research*, 11, 35–46.
- Gregory, S.W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, 70, 1231–1240.
- Gregory, S.W., Dagan, K., & Webster, S. (1997). Evaluating the relation between vocal accommodation in conversational partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior*, 21, 23–43.
- Hall, K. (2002). 'Unnatural' gender in Hindi. Gender across languages: the linguistic representation of women and men, ed. by M. Hellinger and H. Bussman, 133–62. Amsterdam, The Netherlands: John Benjamins.
- Hall, K., & O'Donovan, V. (1996). Shifting gender positions among Hindi-speaking hijras. Rethinking language and gender research: theory and practice, ed. by V. Bergvall, J. Bingand A. Freed, 228–66. London: Longman.
- Hall-Lew, L., Coppock, E., & Starr, R. (2010). Indexing political persuasion: variation in the Iraq vowels. *American Speech*, 85(1), 91–102.
- Harrington, J. (2006). An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics*, 34, 439–457.
- Harrington, J. (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. In J. Cole, & J. Hualde (Eds.), *Laboratory Phonology 9: Phonetics and Phonology* (pp. 125–144). Berlin: Mouton de Gruyter.

- Harrington, J., Palethrope, S., & Watson, C.I. (2000a). Does the Queen speak the Queen's English? *Nature*, 408, 927–928.
- Harrington, J., Palethrope, S., & Watson, C.I. (2000b). Monophthongal vowel changes in Received Pronunciation: an acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association*, 30, 63–78.
- Hay, J., Jannedy, S., & Mendoza-Denton, N. (1999). Oprah and /ay/: Lexical Frequency, Referee Design and Style. San Francisco, CA: Proceedings of the 14th International Congress of Phonetic Sciences.
- Hay, J., Nolan, A., & Drager, K. (2006). From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review*, 23, 351-379.
- Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34, 458-484.
- Hillenbrand, J., L. Getty, M. Clark, and K. Wheeler. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–111.
- Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55, 37-58.
- Howell, P., Barry, W., & Vinson, D. (2006). Strength of British English accents in altered listening conditions. *Attention, Perception, & Psychophysics*, 68, 139-153.
- Janson, T., & Schulman, R. (1983). Nondistinctive features and their use. *Journal of Linguistics*, 19, 252-265.
- Johnson, K. (2006). Resonance in an exemplar-based lexicon: the emergence of social identity and phonology. *Journal of Phonetics*, 43, 485–99.

- Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J.W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 145-165). San Diego: Academic Press.
- Johnson, K., Strand, E. & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, 27, 359-384.
- Katseff, S., Houde, J., & Johnson, K. (submitted). *Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback?* Manuscript under consideration.
- Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208-1221.
- Kraljic, T., & Samuel, A.G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51, 141-178.
- Kraljic, T., & Samuel, A.G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin Review*, 13, 262-268.
- Kraljic, T., & Samuel, A.G. (2007). Perceptual adjustments to multiple speakers. *Journal Of Memory and Language*, 56, 1-15.
- Kraljic, T., Brennan, S.E., & Samuel, A.G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*, 107, 54-81.
- Kraljic, T., Samuel, A.G., & Brennan, S.E. (2008). First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*, 19, 332-338.
- Labov, W. (2001). The anatomy of style-shifting. In P. Eckert, & J. R. Rickford (Eds.), *Style and Sociolinguistic Variation*, (pp. 85-108). Cambridge: Cambridge University Press.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19, 273-309.

- Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- Labov, W., Ash, S., & Boberg, C. (2006). *Phonological Atlas of North America*. New York: Mouton de Gruyter.
- Ladefoged, P. & Broadbent, D.E. (1957). Information Conveyed by Vowels. *Journal of the Acoustical Society of America*, 29 (1), 98-104.
- Lass, N.J., Almerino, C.A., Jordan, L.F., & Walsh, J.M. (1980). The effect of filtered speech on speaker race and sex identifications. *Journal of Phonetics*, 8, 101-112.
- Lavendara, B. (1978). Where does the sociolinguistic variable stop? *Language in Society*, 7, 171-182
- Lindblom, B. (1990) Explaining variation: A sketch of the H and H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling*. Kluwer.
- Manning, C. (2003). Probabilistic Syntax. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic Linguistics* (p. 289-342). Cambridge, MA: MIT Press.
- Manuel, S. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88, 1286–98.
- Milroy, J. & Milroy, L. (1985). Linguistic change, social network, and speaker innovation. *Journal of Linguistics*, 21, 339-84.
- Munro, M.J., Derwing, T.M., & Flege, J.E. (1999). Canadians in Alabama: a perceptual study of dialect acquisition in adults. *Journal of Phonetics*, 27, 385–403.
- Munson, B. (2007a). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech* 50.125–42.

- Munson, B. (2007b). Talker sex mediates the influence of neighborhood density on vowel articulation. Presented at Proceedings of the International Conference on Phonetic Sciences, Saarbrücken, German, August 6–10.
- Munson, B. (2010). Levels of phonological abstraction and knowledge of socially motivated speech-sound variation: a review, a proposal, and a commentary on the Papers by Clopper, Pierrehumbert, and Tamati; Drager; Foulkes; Mack; and Smith, Hall, and Munson. *Journal of Laboratory Phonology*, 1, 157-177.
- Munson, B., E. C. McDonald, N. L. DeBoe, and A. R. White. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34, 202–40.
- Munson, B., & Solomon, N.P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, 47, 1048-1058.
- Namy, L.L., Nygaard, L.C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: the role of perception. *Journal and Language and Social Psychology*, 21, 422–432.
- Natale, M. (1975a). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 32, 790–804.
- Natale, M. (1975b). Social desirability as related to convergence of temporal speech patterns. *Perceptual Motor Skills*, 40, 827–830.
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18, 62-85.
- Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39, 132-142.

- Nolan, F., Holst, T., & Kühnert, B. (1996) Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics* 24, 113-137.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119, 2382–2393.
- Pardo, J. S. (2009). Expressing Oneself in Conversational Interaction. In E. Morsella (Ed.), *Expressing Oneself/Expressing One's self: Communication, Cognition, Language, and Identify* (pp. 183-196). Mahwah, NJ: Lawrence Erlbaum Associates.
- Pardo, J.S., Jay, I.C., & Krauss, R. (2010). Conversational Role Influences Speech Imitation. *Attention, Perception, and Psychophysics*, 72 (8), 2254-2264
- Perry, T. L., R. Ohde, and D. Ashmead. (2001). The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America*, 109, 2988–98.
- Peterson, G., & Barney, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 175–84.
- Pickering, M.J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioural & Brain Sciences*, 27, 169–226.
- Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46, 115–154.
- Pierrehumbert, J. & Hirschberg, J. (1990). The Meaning of Intonation in the Interpretation of Discourse, in P. Cohen, J. Morgan, and M. Pollack, (Eds.) *Intentions in Communication* (pp. 271-311). Cambridge MA: MIT Press.
- Pierrehumbert, J.B., Bent, T., Munson, B., Bradlow, A.R., & Bailey, J.M. (2004). The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America*.

- Podesva, R.J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, 11, 478-504.
- Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experimentation on American English dialect identification. *Journal of Language and Social Psychology*, 18, 10-30.
- Rieger, G., Linsenmeier, J., Gygax, L., Garcia, S., & Bailey, J.M. (2010). Dissecting "gaydar": accuracy and the role of masculinity-femininity. *Archives of Sexual Behavior*, 39, 124-140.
- Romain, S. (1984). On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica*, 18, 409-39.
- Sachs, J., Lieberman, P., & Erickson, D. (1973). Anatomical and cultural determinants in male and female speech. Language attitudes, ed. by, R. W. Shuy and R. W. Fasold, 74–83. Washington, DC: Georgetown University Press.
- Sancier, M. & Fowler, C. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421-436
- Scarborough, R.A. (2010) Lexical and Contextual Predictability: Confluent Effects on the Production of Vowels. In press in *Laboratory Phonology X*. Mouton de Gruyter.
- Shepard, C. A., Giles, H., & LePoire, B.A. (2001). Communication accommodation theory. In W. P. Robinson, & H. Giles (Eds.), *The New Handbook of Lang. and Social Psychology*, (pp. 33–56). New York: John Wiley & Sons Ltd.
- Shiller D., Sato, M., Gracco, V., & Baum, S. (2009). Perceptual recalibration of speech sounds following speech motor learning. *Journal of the Acoustical Society of America*, 1103-1113.

- Shockley, K., Sabadini, L., & Fowler, C.A. (2004). Imitation in shadowing words. *Perception & Psychophysics*, 66, 422–429.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, 23, 193–229.
- Simpson, A.P. (2001). Dynamic consequences of differences in male and female vocal-tract dimensions. *Journal of the Acoustical Society of America*, 109, 2153-2164.
- Smith, E.A., Hall, K.C., & Munson, B. (2010). Bringing semantics to sociophonetics: social variables and secondary entailments. *Journal of Laboratory Phonology*, 1, 121-155.
- Staum Casasanto, L. (2008). *Experimental Investigations of Sociolinguistic Knowledge*. *Dissertation*, Stanford University Department of Linguistics.
- Stewart, M. & Ota, M. (2008). Lexical effects on speech perception in individuals with “autistic” traits. *Cognition* 109, 157–162.
- Strand, E., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference, Bielfeld, October 1996* (pp. 14-26). Berlin: Mouton de Gruyter.
- Thomas, E. (2002). Sociophonetic applications of speech perception experiments. *American Speech*, 77, 115-147.
- Troutman, C., Clark, B., & Goldrick, M. (2008). Social networks and intraspeaker variation during periods of language change. *Penn Working Papers in Linguistics*, 14.
- Trudgill, P. (2004). *New dialect formation: The Inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.

- Trudgill, P. (2008). Colonial dialect contact in the history of European Languages: On the irrelevance of identity in new-dialect formation. *Language in Society*, 37, 241–280.
- Trudgill, P. (1981). Linguistic accommodation: Sociolinguistic observations on a socio-psychological theory. In R. Hendrick, C. Masek, and M.F. Miller (eds.), *Papers from the Parasession on Language and Behavior*, 218–237. Chicago: Chicago Linguistics Society.
- Van Bezooijen, R. 1995. Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech* 38.253–65.
- Vorperian, H., Wang, S., & Chung, M. (2009). Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of the Acoustical Society of America*, 125, 1666–1678.
- Walker, A. (2008). *Phonetic Detail and Grammaticality Judgements*. Christchurch, New Zealand: MA Thesis, Department of Linguistics, University of Canterbury.
- Wedel, A. & Van Volkinburg, H. (in preparation). *Modeling simultaneous convergence and divergence of linguistic features between differently-identifying groups in contact*. Manuscript in preparation.
- Wells, J. C. (1982). *Accents of English*. Cambridge University Press, Cambridge.
- Weinreich, U., Labov, W., & Herzog, M. (1968) Empirical foundations for a theory of language change. In W. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press.
- Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, 18, 3–35.

- Wright, R. (2004). Factors of lexical competition in vowel articulation. In J. Local, J., R. Ogden, and R. Temple (eds.). *Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press.
- Yamazawa, H., & Hollien, H. (1992). Speaking fundamental frequency of Japanese women. *Phonetica*, 49, 128-140.
- Yu, A. (2010). Perceptual compensation is correlated with individuals' "Autistic" traits: Implications for models of sound change. *PLoS ONE*, 5 (8): e11970.doi:10.1371/journal.pone.0011950.