13 The Emergence of Sociophonetic Structure

PAUL FOULKES AND JENNIFER B. HAY

1. Introduction

Sociophonetics involves the integration of the techniques, principles, and theoretical frameworks of phonetics with those of sociolinguistics. Sociophonetic work also increasingly involves interaction with other fields including psycholinguistics, psychology, first and second language acquisition, theoretical phonology, dialectology, conversation analysis, and computational linguistics. The common focus of this diverse field is understanding and modeling the production and perception of speech, especially in natural contexts, as well as the cognitive storage and processing mechanisms that underpin speaking and listening (Hay and Drager, 2007; Foulkes, Scobbie, and Watt, 2010; Docherty and Mendoza-Denton, 2012). In this chapter we review evidence for the emergence of associations between linguistic elements and social entities, and how these associations in turn shape linguistic knowledge and processing. We consider issues in both phonetics and phonology, adopting a broad definition of "knowledge about sounds" (see further Docherty and Foulkes, 2000; Docherty, Foulkes, Tillotson, and Watt, 2006; Foulkes, 2010). That is, our focus is on how knowledge about sound units and processes is shaped by a person's socialization and experience of speaking and listening in social contexts.

1.1 Language as a social and cognitive phenomenon

Theoretical models of language have overwhelmingly focused on the categorical contrastive elements of language, i.e. the abstract system putatively shared by all speakers of a language. However, knowing a language involves not only cognitive representation of the elements and processes of that language (words, phonemes, rules) but also the ability to use the resources of the language to achieve communicative goals (Saussure, 1974 [1916]: 77). Natural speech is invariably situated in a social context – conversation or other forms of discourse – where communication of information is a prime concern. Whenever a speaker produces an utterance, two types of information are conveyed.

The Handbook of Language Emergence, First Edition. Edited by Brian MacWhinney and William O'Grady.

The first is what is usually considered the *linguistic* – the transmission of the referential meaning conveyed by words in grammatical context. In parallel with the linguistic meaning, and intertwined with it, is a rich layer of social or *indexical* meaning indicating (among other things) aspects of the speaker's demographic background, emotional state, attitude, and identity (Abercrombie, 1967; Foulkes and Docherty, 2006). A competent speaker knows how to make appropriate adjustments to their speech and language (including non-verbal behavior) in order to deliver *both* the linguistic and the indexical information. Competent listeners in turn know how to interpret the acoustic (and visual) signal and derive from it both strands of information. It is, we conjecture, impossible to do one without the other, although the relative importance of the linguistic versus the indexical may differ markedly depending on the context.

Children acquiring their first language(s) face the task not only of learning the elements and processes, but also of how to use and interpret the language(s) appropriately in different circumstances. Such knowledge is amassed through repeated experience: that is, engaging in conversations on different topics with different people and for different goals. As we show below, the act of conversing permits or encourages patterns of behavior to emerge between the participants to facilitate the interaction and exchange of information – both linguistic and indexical. As individuals' experiences grow, they build associations between linguistic forms and social characteristics of other people and situations. In time these associations come to form a very rich and complex body of indexical knowledge, permitting the individual to adjust their own speech appropriately, and to interpret layers of indexical meaning in the speech of others. For the mature language user, engaging in conversation on a daily basis, indexical experiences continue to accumulate. Most theoretical models overlook the effects of learning on the cognitive system after a hypothesized critical period. However, it is plain that knowledge about language continues to be updated at some level throughout life as experience feeds into the cognitive repository. For instance, we continue to meet new interlocutors and learn to recognize their individual voices and other indexical features. At times abrupt changes in circumstance might lead to abrupt changes in experience of life and therefore language, for example when people move to a new region where a different language or dialect is spoken, or where the norms of interaction differ markedly. To some extent these experiences are bound to affect linguistic knowledge, most clearly where prolonged exposure to a new dialect community leads to a shift in the speaker's own accent. A further implication of an experience-based model of language learning and use is thus that every individual has a unique cognitive representation of language at some level (Johnstone, 2000).

1.2 Emergence of sociophonetic associations

We take the above summary of language knowledge and use as largely uncontroversial. However, following the now long-standing theoretical division between *langue* and *parole*, or *competence* and *performance*, theoretical linguists have for the most part focused on investigating and modeling the linguistic dimension, with the indexical dimension reserved as the domain of sociolinguists. In adopting a broader stance on the issue of what it means to *know* a language, as we do here, we can identify a number of issues where understanding remains limited and where emergent processes appear important. Among these issues are the following:

- *How do sociophonetic associations emerge at the individual level?* What are the processes through which knowledge about indexical meaning is established? How does experience shape what is learned and represented cognitively? How do individuals generalize across their experiences, exploiting non-random associations by abstracting them into cognitive categories? To what extent does this process begin in childhood and continue in adulthood?
- *How do sociophonetic associations emerge at the group level?* How do groups of people converge on shared associations between phonetic forms and indexical meanings? How do sociophonetic associations evolve over time? What processes permit the propagation of associations from generation to generation?
- *What effects do indexical properties exert on linguistic elements and processes*? How is indexical information stored relative to "purely linguistic" information?

We now turn to these issues. We highlight in particular those areas we consider ripe for further investigation, and which offer provocative implications for modeling of linguistic knowledge. We also present a critical outline of aspects of exemplar theory as the framework currently best placed to model the interaction between linguistic and indexical knowledge.

2. Emergence of Sociophonetic Associations at the Individual Level

Learning about indexical properties of speech precedes the development of knowledge about linguistic structures. In fact, there is evidence it begins even before birth. The developing baby can hear sound emanating from outside the womb. The voice of its mother is probably the most frequent external sound it encounters, and patterns characteristic of this voice are learned *in utero*. Hepper, Scott, and Shahidullah (1993) played recordings of mothers' voices through loudspeakers to 36-week-old fetuses, and found evidence of discrimination between the voice speaking at a normal level and an amplified version. Babies process the maternal voice faster than other voices (Purhonen, Kilpeläinen-Lees, Valkonen-Korhonen, Karhu, and Lehtonen, 2005), while new-borns can distinguish the maternal voice from other voices at 2–4 days (DeCasper and Fifer, 1980). After only a few months there is a robust ability to identify changes in talker (Miller, 1983), and to discriminate individual voices (Houston and Jusczyk, 2000, 2003; Barker and Newman, 2004). E. Johnson, Westrek, Nazzi, and Cutler (2011), however, demonstrate that this infant ability to discriminate between different talkers is mediated by language experience. Infants can do it, but only in their native language. This result demonstrates a very early and robust entanglement between linguistic and indexical knowledge.

These studies all demonstrate that babies have encoded indexical information about particular talkers, voices, or vocal styles very early – well before they have mastered knowledge about the abstract properties particular to a language. (Of course, early knowledge of individual voices is not at the level shown by older children or adults: experiments show that the ability to identify talkers improves up to at least age 10 years; Bartholomeus, 1973; Mann, Diamond, and Carey, 1979.)

While individual talker differences may be the first indexical dimension encountered by the child, other dimensions soon follow. The initial awareness of individual talker differences may provide the grounding for subsequent emergence of knowledge about broader groups – children and adults, men and women – differentiated by stark acoustic differences which are transparently underpinned by biological differences in vocal tract properties (e.g. f0, heavily influenced by vocal fold size; Foulkes, 2010). Evidence in support of this position is provided by perceptual experiments with infants. Miller (1983) exposed 2- and 6-month-olds to male and female voices in a head-turn experiment. The 2-month-olds responded to changes in talker irrespective of the talker's sex, while the older infants only showed a difference in response when the sex of the talker changed. This result suggests the younger children were sensitive to individual talker differences while the older infants had generalized to a more abstract categorization of talker, presumably based on general cues such as f0.

Further information about indexical variation is available to the child from the outset as adults (in most societies, at least) modify their speech when addressing children (e.g. using wider f0 contours; Snow, 1995). These adjustments are argued to assist the child in locating relevant linguistic units in the stream of speech. However, the gross differences characterizing child-directed speech (CDS) also serve to introduce the child to a form of stylistic variation: patterns of speech behavior that are tailored to particular interlocutors or subgroups within the larger community. That is, children can be spoken to in ways that differ from those used to adults, showing that linguistic choices are oriented to the interlocutor or group to which they belong. The characteristic patterns of CDS also help the child recognize the value of speech as a social activity in which participants exchange information in conversation (Snow, 1995).

Stylistic variation is further developed at the level of sociolinguistic variables. From their studies of mother−child pairs in Buckie, Scotland, for example, J. Smith, Durham, and Richards (2013) show that highly localized dialectal variants are used more frequently in situations of intimacy and play, while variants typical of standard Scottish English have higher incidence in situations of instruction or discipline. The children begin to mirror the patterns shown by their mothers after the age of about three years. Stylistic variation, linked to the communicative goal of interaction, reveals to the child that alternative phonetic forms exist and are associated (statistically if not in absolute terms) with particular settings, contexts, topics, and/or speakers. Labov (2001) argues that stylistic variation of this sort is the principal route through which children acquire knowledge of all sociolinguistic patterning, gradually extending their awareness from situational style to associations with class, gender, and age, for example. However, in a study of /t/ variants in Newcastle upon Tyne, Foulkes, Docherty, and Watt (2005) show that sociolinguistic variation correlating with gender is also inherent to the input received by each child: mothers tailor their speech in line with the emerging gender of their child. Boys hear more local variants than girls, who by contrast are exposed to more standard forms. Boys and girls begin to diverge in their production patterns, in line with gendered differences in their community, from around age 3;6 (Roberts, 2002; Foulkes and Docherty, 2006). At around the same age children demonstrate ability to use stylistic variation in role play or to signal a particular discourse style (Andersen, 1990; Gillen and Hall, 2001). Acquisition of these and other sociolinguistic dimensions is reviewed by Nardy, Chevrot, and Barbu (2013).

Indexical knowledge about group properties may then continue as children gradually learn the associations between more arbitrary pronunciation variants (e.g. vowel or consonant variables) and broad demographic groups. Subtle and wholly arbitrary associations take considerably longer to learn and may never be mastered. These include associations between phonological forms and regional groups or subcommunities, such as the regional or class patterning of rhoticity in English, which differs markedly across dialects (Foulkes, 2010). This brief review of sociophonetic variation in the early acquisition process shows that systematic variation is present in the raw material from which the child learns its language(s). Evidence for associations between forms and people or communicative contexts is there from the very beginning.

How, then, do children learn these associations? There is growing agreement in many fields of cognitive science that patterns of association apparent in stimulus material can be identified by children, and thus learned via stochastic or probabilistic learning (Saffran, Aslin, and Newport, 1996; Maye, Werker, and Gerken, 2002; Maye, Weiss, and Aslin, 2008). Probabilistic learning has been shown for the acquisition of linguistic units such as phonemes, and also for social structures (Smith and Zárate, 1992). There is no reason to doubt that the same process permits learning of associations *between* the linguistic and social.

Some patterns of association may be much more apparent to a child than others by virtue of frequency of exposure (e.g. to the maternal voice), or through phonetic consistency and transparency (e.g. the voices of adult males, females, and children differentiated by large f0 differences, with relatively discrete partitioning between groups relative to narrow variation within groups). Exposure to some types of variation may in turn allow the child to learn other types of association (extrapolating from individual voices to group patterns, or from stylistic variation to patterns correlating with other demographic differences). Learning of indexical information is apparent from the very earliest stages, and in some ways precedes the acquisition of strictly linguistic information. Indeed, early indexical learning might provide a foundation for the learning of linguistic structures as well as continued indexical learning. By virtue of the bias to the maternal voice in prenatal experience, the child enters the world with an elementary framework for categorization of sound: the child has an ability to distinguish speech from other sounds, and the maternal voice from other forms of speech (Foulkes, 2010; cf. conclusions drawn with respect to word and phoneme learning by Studdert-Kennedy, 1983; DeCasper and Spence, 1986; Houston and Jusczyk, 2000, 2003).

As the child develops, we assume that learning about indexical properties emerges largely through the same basic processes that are essential for the learning of linguistic information: exposure to spoken language in communicative settings. Just as exposure to a stimulus is vital for a child to identify the sound units and processes of its language, so does this stimulus provide the necessary information to recognize linguistic alternatives and the norms of their usage. As the growing learner engages in an ever-widening range of speaking tasks with more and more interlocutors, more information becomes available about the statistical associations between linguistic forms and particular speakers, groups or contexts of use.

There is, of course, much that is still unknown about the precise processes underpinning the emergence of sociophonetic associations. Which types of indexical associations are learned most easily? To what degree is learning affected by factors such as phonetic transparency or weight of statistical evidence? And how does the learning of indexical associations proceed alongside the learning of linguistic structures? For example, might the processes of indexical learning evident in infants atrophy as the child's mastery of linguistic structures improves, perhaps especially at periods characterized by rapid development of linguistic knowledge? If so, does indexical learning resurface later once the child begins to use language for a wider range of communicative purposes?

There are still more questions than answers about the trajectory of acquisition for socio-indexical knowledge. However, some associations between forms and people or communicative contexts are learned early, before the learning of linguistic structures, and may in fact support the latter. Emergentist models do not assume this learning process ends after a "critical period" in acquisition. Instead, it is assumed to be an ongoing process throughout life, although the effects of ongoing learning may well have a considerably reduced impact on the cognitive representation of linguistic units once these are well established. Some evidence from real-time studies does show, however, that production patterns may change throughout adulthood. A notable study in this regard is that of Queen Elizabeth II, showing that her production of some vowels has shifted over a 50-year period towards innovative norms in the broader RP speech community (Harrington, Palethorpe, and Watson, 2005). Experimental studies also show that adults are able to identify novel non-random associations between phonetic forms and social groupings. Docherty, Langstrof, and Foulkes (2013), for example, played listeners word stimuli in which consonant and vowel variants were associated with one of two "tribes" – novel social groups that were not distinguished by any of the usual demographic divisions such as sex or age. The variants were either categorical (e.g. Tribe 1 stimuli had 100% [ʔ] for /t/, Tribe 2 had 100% [t]), or statistically biased (80%–20% for Tribe 1, 20%–80% for Tribe 2). These results demonstrate that individuals can learn structure in the input, with different degrees of success depending on the statistical distribution and phonetic salience of the variants. There was also considerable variation across respondents, an issue to which we return in section 6. below.

For adults in everyday life, a key driver in the ongoing learning process, as in the earliest stages of linguistic development, is conversation. As speakers talk to each other they engage in a partnership through which patterns of use are agreed, and associations emerge of how phonetic variation can be used to signal both indexical and linguistic meaning. Research in conversation analysis (CA) reveals remarkably sophisticated on-line processing through which speakers plan their utterances and understand the utterances produced by the interlocutor, monitoring both their own performance and that of their partner in order to construct a dialogue in which meaning is conveyed successfully (e.g. Couper-Kuhlen and Selting, 1996). The mutual negotiation of talk in interaction shows that participants monitor linguistic, pragmatic, and indexical properties simultaneously. In many areas of sociophonetics indexical meanings are often assumed to be attached to linguistic units (phonetic variants or lemmas). A different perspective is held by many researchers in CA: variable phonetic forms are seen as resources that can be deployed to achieve communicative goals, signaling stance and attitude, the structure of the conversation, and thus the nuanced "meaning" of the interaction conveyed as a nexus of linguistic and pragmatic components (Local, 2007; Ogden, 2012). Conversation may ultimately yield information about consistent associations between social entities and linguistic forms, such that at some level speakers might "know" that, for example, glottalization of /t/ signals casual speech or a speaker of lower social status. However, as Ogden (2012) argues, CA shows that such associations, and the entities they are linked to, are the product of social action. That is, social meanings are not fixed, and their encoding and decoding are not automatic (cf. Eckert, 2008). Rather, to some extent participants in conversation generate and decode indexical and pragmatic information on the fly. Just as associations may emerge between phonetic forms and certain types of people, style or setting, so might associations emerge if a variant regularly occurs at a particular point in a conversational sequence where certain kinds of actions are possible (e.g. (dis)agreement, repair, turn-taking; Ogden, 2012). The cognitive processes by which such on-line work is achieved remain largely ignored in most linguistic models, including those that characterize the relationship between sociolinguistics and phonetics (Local, 2003).

The collaborative work observable in conversation is linked to a more general process of accommodation, or imitation, by which people engaged in conversation tend to converge in aspects of their language usage. Imitation effects are very robust, and are found not only for many different features of language (Branigan et al., 2007; Balcetis and Dale, 2005; Babel, 2010, 2012; Giles, Coupland, and Coupland, 1991; Bosshardt, Sappok, Knipschild, and Hölscher, 1997; Pardo, Gibbons, Suppes, and Krauss, 2012), but also for many other aspects of behavior (Dijksterhuis and Bargh, 2001; De Fornel, 1992; Shockley, Santana, and Fowler, 2003; Sato and Yoshikawa, 2007). Imitative behavior can be both socially mediated and socially influential. Research in the gestural domain, for example, shows that individuals report experiencing interactions more positively when there is a high rate of gestural mimicry involved (Chartrand and Bargh, 1999). In the speech domain, it has been shown that speakers accommodate most to talkers whom they regard positively and who they rate as socially desirable (Natale, 1975a, 1975b). Documented cases of the opposite process, divergence, are much less numerous (but probably less sought out). However, divergence has been reported when there is social distance between two speakers and motivation to increase or mark that social distance (Bourhis and Giles, 1977; Abrego-Collier, Grove, Sonderegger, and Yu, 2011). Khattab (2009, 2013) documents in detail convergence and divergence patterns of bilingual children growing up in Yorkshire with L1 Arabic-speaking parents. The bilinguals' English is in most circumstances indistinguishable from monolingual children in the community, and not marked by L2 influence. To that extent the bilinguals differ from their parents, and diverge from them in preferring English to Arabic, and native features to L2 features. However, their experience of interacting with non-native speakers also enables them to use non-native features to converge. For example, when searching for an Arabic word they often code-switch to English, but use L2 features such as tapped /r/ and rhoticity, especially in situations where their parents are trying to persuade them to speak Arabic. This evidence demonstrates that experience of linguistic forms used by a minority of interlocutors in the community is not wholly ignored or filtered out of memory (contra the claim by Chambers, 2002). Instead it leaves its mark on memory, and social associations are attached to word forms or pronunciations, enabling those forms to be retrieved and used for communicative purposes.

Some aspects of imitative behavior appear to be automatic. For example, people shift their pronunciation to resemble a disembodied recorded voice that they have just been exposed to, a context where there can be little social motivation for imitation (e.g. Delvaux and Soquet, 2007). However, this is clearly mediated by a complicated raft of social factors, and is not guaranteed to occur (Abrego-Collier et al., 2011; Babel, 2012). There are also individual differences in the degree to which speakers accommodate, both within interactions, and over the lifespan. Very little work has probed this speaker-based variation, though Peek and Watson (2013) show that the tendency to imitate is related to other dimensions of non-linguistic perceptual sensitivity.

Through repeated interactions, speakers also learn associations between particular speaking styles and topics. This can lead to a general tendency to produce particular phonetic or linguistic forms when talking about particular topics (Gordon, Campbell, Hay et al., 2004; Mendoza-Denton, Hay, and Jannedy, 1999; Foulkes and Hay, 2013; Love and Walker, 2013). Bell (1984) directly ties such topic-based effects to associations between particular topics and particular interlocutors. If we tend to talk about a particular topic with a particular type of interlocutor, then convergence patterns associated with that interlocutor will come to be associated with the topic more broadly.

Bourdieu (1991) argues that individual variation in production is linked to the social "capital" that accrues from linguistic choices in a linguistic "marketplace." More recent work on variation within a speaker across contexts and topics has concentrated on the identity work that phonetic variation can play. Choices of phonetic variants signal social "stances" and attitudes adopted by a speaker, in context (Mendoza-Denton, Hay, and Jannedy, 1999; Eckert, 2000; Johnstone, 2009; Llamas and Watt, 2010). Through their largely subconscious linguistic choices speakers negotiate social positions, stances, and alignments in the course of their interactions with others.

In all of this work, it is clear that phonetic production is strongly influenced by conversations, and the social contexts in which they occur. Individuals make socially meaningful phonetic choices throughout the course of any interaction. Such effects are also mediated at the level of the individual's own production. It is well known that words are produced and perceived differently according to their frequency and predictability in context (Bybee, 2001; Bell, Brenier, Gregory, Girand, and Jurafsky, 2009). Repeated production of predictable information permits behavior to become relatively automatic, leading typically to reduction of spoken forms. However, speakers engaged in conversation monitor their production to ensure communication is successful. Degrees of reduction are not wholly automatic, but reflect the perceived needs of the interlocutor (Lindblom, 1990).

Across the lifespan repeated participation in conversations comes to shape the overall sociophonetic repertoire available to an individual. If speakers tend to converge in conversation, then the inventory of conversations that an individual speaker participates in will result in longer-term effects, affecting the overall distribution of that speaker's phonetic productions. This can be most clearly seen when there are abrupt life changes such as moving country or dialect area. The micro-level adjustments that are made in individual conversations accumulate, so that an individual who has moved to a new dialect area eventually begins to acquire aspects of the local phonetic norms (Munro, Derwing, and Flege, 1999; Evans and Iverson, 2007). However it is also clear that such lifetime adjustment to local context occurs even in the absence of abrupt shifts. Because sound change is constantly underway, local norms are in a constant state of flux. Thus even repeated conversations within a local context involve a constant updating of experience, leading to small adjustments throughout the lifespan.

3. Emergence of Sociophonetic Associations at the Group Level

As outlined above, an individual's phonetic repertoire is shaped and formed through a series of interactions. These interactions shape the individual's overall repertoire through the cumulative effect of repeated imitative behavior. In turn, associations are learned between phonetic implementation and social meaning. Geeraerts, Kristiansen, and Peirsman (2010: 5) summarize neatly the relationship between the individual and group:

> the same mechanism that allows the existing collective regularities to enter the individual minds is also the one that allows regularities to emerge to begin with, viz. mutual influence in social interaction. People influence each other's behavior, basically by co-operative imitation and adaptation, and in some cases by opposition and a desire for distinctiveness. Paying attention to what others do, however subconsciously, thus creates a mental representation of the collective tendencies in the behavior of the community; adapting one's own behavior to those tendencies, reaffirms and recreates the tendencies.

The question of how associations between phonetic variants and social meanings emerge is intricately entwined with questions relating to the dynamics of social networks. Individuals have very many interactions, with very many interlocutors. However, the distribution of typical interlocutors for a particular individual is a non-random selection of people. We tend to interact with particular types of people more than others. because of where we live, what work we do, how we spend our free time, and what types of people we like. Similarly, individuals tend to talk about a non-random distribution of topics. These different experiences shape language use in different ways. Varying experiences demand different linguistic performance, and shape experience differently as a reflection of the social differences in the broader society. At the level of a society or language community, then, we expect to find complex variation in social experience, and thus also in language use and the resultant cognitive representations of language.

While complex, such variation is not random but structured, and highly sensitive to social network patterns. Groups or communities may be small entities such as families (Hazen, 2002) or communities of practice (Eckert, 2000), or much larger entities including demographic divisions, speech communities or regional dialect groups. Membership of these groups may be self-selected (communities of practice) or given (family), temporary or permanent, and an individual moves within many groups at any time in life. Lifestyle patterns and opportunities tied to these groups yield different types of linguistic experience. For example, Milroy (1987) investigated variation and change in three neighborhoods of Belfast, Northern Ireland, differentiated by ethnicity (based largely on religion) and social network structure: open networks with relatively free flow of individuals in wide collectives, versus closed networks with relatively fixed membership and limited physical mobility. Traditional working-class communities are generally characterized by the latter: people live, work, and socialize in a compact area, largely with the same people. Degrees of social and physical mobility affect the range of contacts available to individuals, and in turn their experience of diverse linguistic forms and social associations (Kerswill, 2003; Britain, 2011). Routine behaviors (Giddens, 1984), encouraged by a lack of mobility, serve to reinforce linguistic patterns through repeated actions (Britain, 2013). Mobility enables speakers to broaden their experiences of language use. In turn this may reveal different or wholly new patterns of association between forms and social groups or interactional tasks, and thus different or novel social evaluation of those associations (Johnstone, Andrus, and Danielson, 2006). Stability and closed networks by contrast act to reinforce norms of usage, encouraging transmission of forms and norms intact (Milroy and Milroy, 1992).

While it is common for linguists to focus their attention on stable individuals in stable communities, modern urban societies are in fact characterized by complex patterns of mobility. In the UK, for instance, national statistics from the census and the National Health Service (NHS) show that as much as 10% of the population move region each year (Britain, 2011; Rhodes, 2012). The regions used to classify the statistics are similar to broad regional dialect divisions. These statistics indicate an extremely high level of mobility and dialect contact in an era characterized by dialect leveling, the weakening influence of Received Pronunciation, and the emergence of supralocal standards (Kerswill, 2003).

An important and growing branch of research in computational modeling explores the dynamic processes by which individual agents interact with one another, and the effect on simulated linguistic experience of different types of interaction between different types of individuals. This work demonstrates that repeated small-scale effects occurring in individual interactions can amplify over time and lead to complex emergent phenomena at the group level (see for example Epstein and Axtell, 1996; Fagyal, Swarup, Escobar, Gasser, and Lakkaraju, 2010). In computerized agent-based models, the community is modeled as a social network, in which links represent meaningful connections between people (Wasserman and Faust, 1994). Repeated interactions amongst connected people create feedback loops, which reinforce and amplify some patterns at the expense of others. Castellano, Fortunato, and Loreto (2009) give a good overview of work which clearly demonstrates that, in models incorporating imitative behavior, the structure of the social network plays a very powerful role in influencing resultant patterns. The emergent patterns are affected both by the overall structure of the network, and by the type of interactions taking place (see also Watts, 2002).

Beckner, Blythe, Bybee et al. (2009) point out that many models assume that transmission across generations occurs from speakers in one generation who have a fixed system (Nowak, Komarova, and Niyogi, 2002; K. Smith, Kirby, and Brighton, 2003), and argue (as we would) that it is important for agent-based models to allow for nuanced personal histories in terms of exposure patterns. Lu, Korniss, and Szymanski (2009) exemplify this approach. They present mathematical simulations showing how subgroups in a social network can maintain distinct vocabularies, even if they remain in contact with other groups. Distinct phonetic patternings would emerge and be maintained in a directly analogous fashion. Work which particularly models social effects on phonetic forms includes Wedel and Van Volkinburg (n.d.), who model the interaction between two different groups. They show that divergence and convergence may occur simultaneously across different linguistic features. Kenny and Stanford (2012) likewise demonstrate that endowing agents with different sources of linguistic and economic "power" permits commonly observed gender differences to emerge.

To date, agent-based modeling has for the most part deployed relatively socially naive agents in the models, and adding more sophistication in this area is an important issue for future research. A handful of studies have attempted to track the emergence of new sociophonetic associations in groups of humans. Witten (2013), for example, presents a unique study of sociophonetic emergence in a real-life community: an on-line social networking community using mainly written communication. Her study focuses on the pronunciation of the "netologisms" *MeFi* and *MeFite*, respectively the abbreviation of the Internet community's weblog, MetaFilter.com, and the name for its users. Members vary between several different pronunciations, with four possibilities for the first syllable alone ([mi- mɛ- mei- mai-], the last illustrating vowel harmony in [maifai], cf. *wifi*). Through observations based on blog discussion and questionnaire investigation, Witten establishes that some forms are preferred in particular regions, and reflect ideology or identity.

Conceptualizing the structure of both society and language in this way enables us in turn to understand and explain how and why language changes. Synchronic linguistic variation, structured in line with numerous non-linguistic factors, pervades every speech community. This synchronic variation is a necessary precursor to the emergence of longer-term linguistic change (Weinreich, Labov, and Herzog, 1968). A society is composed of a mosaic of individuals with overlapping but varying experiences and knowledge. Change in linguistic forms, or indexical meaning of those forms, may ensue where there are changes in the composition of the society, or the relationships that hold between individuals in that society.

A number of recent studies have attempted to model the emergence of change computationally. For instance, Fagyal et al. (2010) simulate a model of interlocutors, and demonstrate that language undergoes change (with the characteristic S-shaped curve; Weinreich, Labov, and Herzog, 1968) only when the network of interlocutors is relatively heterogeneous in terms of interaction patterns. Change emerges when the community contains both well-connected and relatively isolated individuals, and also when all individuals are biased to pay closer attention to the well-connected individuals (see also Blythe and Croft, 2012). Fagyal et al. argue that innovation is driven by isolated individuals, whereas community norms are established by the well-connected "leaders." Both elements are needed for change.

Stanford and Kenny (2013) use agent-based modeling to explore transmission, diffusion, and incrementation in sound change, arguing that such processes can be modeled by incorporating simple exemplar-based learning into a model involving variable density of interactions. However, imitation and social networks alone cannot account for the speed and vigor with which pronunciation can change across generations. Computational simulations of sound change show that something beyond simple imitation is needed to account for observed sound changes. For example, Baxter, Blythe, Croft, and McKane (2009) implement a model of language change in early New Zealand English, demonstrating that several changes occurred much faster than would be expected on the basis of social networks and propagation of variants in relation to their frequencies alone. Some social dynamics must also be involved.

Labov (2010: 344) argues that each generation actively adopts and exaggerates ongoing change:

> In the incrementation of change, children learn to talk differently from their parents and in the same direction in each successive generation. This can happen only if children align the variants heard in the community with the vector of age: that is, they grasp the relationship: the younger the speaker, the more advanced the change.

Just as other socio-indexical relationships emerge, so can the relationship between particular variants and age. Each generation has an interest in positioning itself as "young," and thus may exaggerate the pronunciation in the direction of the desired position on the "vector of age." Note, though, that an observable statistical relationship between age and pronunciation need not be linear, or even monotonic. Indeed, there is good evidence that changes in progress display an "adolescent peak," with adolescents adopting particularly advanced phonetic variants (Labov, 2001; Tagliamonte and D'Arcy, 2009). The existence of the adolescent peak reinforces the idea that indexical associations may fluctuate in salience across the lifespan. Indexical associations are emergent, are constantly in flux, and may vary and change both within individuals and across communities.

The studies reviewed above collectively show that emergent structure comes from repeated interactions amongst individuals who are embedded in social networks. Individuals can learn this structure, and exploit it in speech production and perception. This capacity sets up a feedback loop between the macro-level emergent structure at the group level, and socially sensitive agents at the individual level. Language change may emerge as agents or groups themselves change (see further Keller, 1994).

The cumulative effects of individuals interacting and imitating in social networks, together with frequency effects and lenition effects, can lead phonetic distributions to start to drift. Such changes are emergent from the mechanics of the production–perception loop, operating in a social context. Once a phonetic distribution has shifted enough for an indexical association to emerge between certain variants and certain social characteristics or stances, this association can be exploited by individuals. This can accelerate a change-in-progress beyond what would be deduced from frequency and imitation alone.

4. Linguistic Consequences of Sociophonetic Associations

In the preceding sections we have considered the mechanisms by which associations emerge between linguistic forms and social groups or contexts: what we have termed sociophonetic associations. It is clear that individuals observe, store, and exploit these associations in the course of everyday interactions. Speech production, speech perception, and conversational interaction are all mediated by indexical knowledge. In this section we consider the effects of sociophonetic associations on linguistic knowledge: storage of information about sounds and words and accessing that information in speaking and listening.

The repeated exposure to words in interactions comes to shape the remembered distributions of words. One very clear way in which word-level storage is affected by the cumulative effect of patterns of interaction is in the effect of word frequency. The frequency with which individuals encounter words clearly affects storage, and is influential in both speech production and perception. In perception, more frequent words are more quickly and more readily understood (Savin, 1963; Forster and Chambers, 1973). In production, words that are encountered frequently tend to take on more reduced phonetic forms (e.g. Bell et al., 2009). Gahl (2008) demonstrates that the effect of word frequency even extends to apparent homophones such as "thyme" and "time." The latter, which is more frequent, is more prone to reduction than the former.

However, our knowledge about our experience with words extends far beyond statistics regarding frequency of use. The non-random association between phonetic forms and social groupings, or social meanings, can also have long-term consequences for the storage of words. This is illustrated by Walker and Hay's (2011) experiment, which investigates words that have skewed distributions of usage across speakers of different ages. Some words tend to be used more by older speakers, and some by younger speakers. If older and younger speakers produce different phonetic forms, and if the representation of words is influenced by previous experience with those words, then "older" words will be represented by "older" phonetic forms. Walker and Hay show that "old" words are processed more quickly and accurately in "old" voices, and "young" words are processed more quickly and accurately in "young" voices. From socio-indexical phonetic meaning, then, long-term consequences for word-storage emerge. Hay and Walker (2013) replicate this finding for words which have usage patterns skewed by gender.

There is also evidence that listeners constantly monitor the social context of incoming speech. This monitoring can be seen in two ways. First, listeners can observe and understand the social meaning embodied by the phonetic material in an incoming speech stream. They can automatically and readily make social judgments based on fine phonetic detail (Campbell-Kibler, 2009, 2011; Drager, 2010). Second, listeners use socio-statistical information to help decode a linguistic message. They are sensitive to the speaker-specific characteristics of voices with which they are familiar (Nygaard, Sommers, and Pisoni, 1994; Smith and Hawkins, 2012). Listeners' inferences about a speaker's sex (Johnson, Strand, and D'Imperio, 1999), age (Drager, 2011), and socioeconomic status (Hay, Warren, and Drager, 2006) all influence processing which appears quite low-level, such as the positioning of perceived phoneme boundaries or perception of vowels undergoing merger. In other words, knowledge about how a speaker is "likely" to talk, based on social characteristics, influences how a listener processes their speech. Note that these findings present a considerable challenge to perceptual models that assume some level of acoustic invariance.

Performance in experimental tasks can also be affected by social primes which are not directly associated with the task. Hay and Drager (2010), for example, show that female New Zealanders appear to make different choices in a phonetic-matching paradigm when there are toy stuffed kiwis in the room (presumably cueing that the speaker they are listening to is also a New Zealander) than when there are toy kangaroos or koalas (associating the speaker instead with Australia). This would seem to indicate automatic priming of social cues which affects speech perception. Even subliminally presented pictures of kangaroos and kiwis appear to have an effect. Similarly, MacFarlane, Hay, and McAuliffe (submitted) show that social cues carried by a voice can affect listeners' performance in a subsequent non-linguistic reaction-time task. In their work, the age of a recorded voice affected reaction time in a subsequent shape/color classification task. Older listeners were speeded by older voices, and younger listeners were speeded by younger voices. Such results indicate that non-linguistic social cues affect linguistic tasks, and that linguistically carried social cues affect aspects of non-linguistic processing. However, we should tread with extreme caution in inferring that such priming effects have a widespread effect in everyday life. As yet, all that we know is that such results can sometimes be found in very constrained laboratory experiments involving rather artificial tasks. Finding a way to investigate how far such effects extend into real-world behavior is a significant challenge.

A further problem with most current lab-based experimentation is that the design of such tasks usually encourages the participant to be relatively disengaged. There is little motivation for a listener to attend to, or extract, social meaning when listening to a disembodied recorded voice reciting isolated words. Yet social motivation clearly matters. Hammersley and Read (1985) show that listeners are much better at recognizing the voice of a conversational partner than a voice they have listened to passively. It is thus likely that tasks that vary the degree of social engagement or motivation on the part of the listener will show different degrees of involvement of sociophonetic meaning. Processes of production and perception will be affected by the social context, and the communicative needs of the participants. This is an important area for future research.

Observing speech production behavior "in the wild" is, of course, a somewhat more straightforward task. A large number of studies have documented the extremely rich sociophonetic territory that individuals traverse with apparent ease. An individual's repertoire of associations between phonetic forms and social meaning repeatedly plays out as they navigate and build this space in the course of repeated interactions. The phonetic choices of an individual show very rich social knowledge, as they constantly adjust their implementation according to addressee (Bell, 1984), topic (Love and Walker, 2013), and formality (Labov, 2001), in order to signal their identity and stance in the local context (Eckert, 2000; Mendoza-Denton, 2008), to balance the interactional needs of the speaker and listener (Lindblom, 1990), and to perform complex conversational and interactional work (Local, 2003, 2007; Ogden, 2012).

5. Discussion

Acts of speaking involve highly complex interaction of many forces: we plan a message, encode it in linguistic units at multiple layers of structure, choosing as we do so among alternatives that signal issues of personal identity, attitude, emotion, and stance, and orientation to both the interlocutor and the conversation. The phonetic output is constrained by biological factors as well as social ones, in terms of the physical constraints imposed by the vocal tract and limitations of memory, and by cognitive factors such as attention. As speakers we learn intricate control to tailor speech appropriately. As listeners we must learn to decode the indexical layers of meaning just as we decode the abstract linguistic units and their referential meaning. Knowledge of language reflects the capacity to encode and decode indexical as well as referential meaning, and there is abundant evidence that cognitive storage and processing of abstract units are affected by emergent processes and indexical associations.

Recognition of these facts is now widespread, and there is a growing acceptance of some role for statistical learning and processing from many different corners of linguistics (e.g. Pierrehumbert, 2003a, 2003b; Chater and Manning, 2006; Cutler and Weber, 2007; Jackendoff, 2007; Coetzee and Pater, 2011). Of course, this is not a problem for theoretical models in which it is stipulated that "performance effects" are necessarily outside the grammar. However, there is now overwhelming evidence to demonstrate emergent properties and their effects on storage and processing. This being so, it is our view that a fully elaborated linguistic theory should embrace a wider representation of what speaker–listeners know and how they use their knowledge than is traditional in structuralist-generative linguistics. Despite this, we remain a long way from such a model. Various usage-based models are in development, as is clear from other contributions to this volume. Exemplar theory (ExT) appeals to us as the model best equipped to handle the emergence of sociophonetic knowledge (Johnson, 1997; Pierrehumbert, 2003a, 2003b, 2006).

At the core of ExT is the notion that linguistic representations are phonetically detailed, and are constantly updated with experience. The "word" is the unit which is most commonly operationalized as the primary level of representation and access. A word's representation is taken to be a rich distribution of episodic memories which is both phonetically and socially detailed. Generalizations across entries in this episodic word store yield abstractions both in the linguistic domain (e.g. phonemes, allophones) and the social domain (e.g. speaker groups, stereotypes, stances). Perception involves associating an incoming signal with the best-matching stored distribution; production involves identifying a contextually relevant subpart of the target distribution and generalizing over that phonetic space to create a production target (see Pierrehumbert, 2006 for a good overview). For researchers struggling to understand the rich complex interplay and co-variation between phonetic and social factors (as outlined in this chapter), ExT provides an attractive model. It provides an account of the intricate bundling of social and phonetic meaning in memory, and of how representations and meanings can be constantly evolving in the context of repeated and varied complex interactions.

ExT has been the subject of increasing attention in linguistics, albeit with fairly widespread misunderstanding of some of its key tenets and predictions, and a disheartening culture of bandwagonism in the place of empirical testing and theoretical refinement. Some of our own work has tested key predictions of the model and found supporting evidence, as noted above (Walker and Hay, 2011; Foulkes and Hay, 2013). However, there are a number of dimensions of ExT that remain underdeveloped (Docherty and Foulkes, 2000, 2013; Foulkes and Docherty, 2006; Foulkes, 2010).

For example, a common error in discussions of ExT is to presume that every experienced word heard yields a detailed phonetic exemplar. This assumption in effect means that listeners memorize a spectrogram (or audiogram; Johnson, 1997) of each word they hear. In fact, exemplars are likely to be mediated by various factors, including degrees of attention, and the prior experience that enables listeners to attend to particular features of the incoming signal and extract information of various types from it (Goldinger, 2007). Moreover, ExT predicts that only *important* experiences are committed to memory (Pierrehumbert, 2006: 525). Thus, highly frequent items such as function words are unlikely to be memorized as often as less frequent but highly informative content words. It remains unclear, however, how we model this process. Further contributions from psychology would appear essential in this respect, to better understand the limitations imposed by attention and memory (Goldinger, 1998; see also Holland, Murray, and Kensinger, 2013 on the effects of emotion on episodic memory). Distinguishing different types of memory processes may also be important. For example, priming work from Sumner and Samuel (2009) suggests that socially marked variants may be understood effectively, and facilitate short-term priming, but may not necessarily influence long-term representations. Similarly, in the visual domain, Musen and Treisman (1990) show that a single exposure to a detailed exemplar can influence implicit memory, but is less influential on explicit memory.

Raw frequency of exposure certainly does not account for all observed patterns in speech production, as social advantage might emerge from variant forms that are in relatively small supply (e.g. as adolescents enter new social groups and conform to new norms of pronunciation). How are minority variants tagged to take higher order? Attention and saliency have been identified as important considerations in modeling this process. The former is difficult to model and the latter remains an elusive concept.

Probably the most important outstanding issue in exemplar theory work is the relationship between experienced-based memory and abstraction. Statistical learning in an exemplar model yields abstractions (Pierrehumbert, 2003a, 2003b), probably at many different levels (Hawkins, 2012). It remains unclear what form(s) abstract phonetic or phonological elements take, or what roles the abstract representations play in relation to detailed information embedded in exemplars. While experiments show that phonetic details may influence linguistic processes (as reviewed in Section 4.), it is likely that such influence is more marked in certain types of language use, or to achieve particular goals in communication. The relative importance of the abstract and the detailed may well shift rapidly and constantly. For instance, in receiving a telephone call we need initially to pay close attention to the phonetic detail of the voice in order to identify the talker. As the call continues with this information established, those details are less important. The content of the call requires closer attention to be paid to the linguistic units that convey the propositional message, while we must also monitor the phonetic details signaling conversational structure, stance, and attitude of the interlocutor in order to manage the conversation successfully.

6. Outlook

We anticipate that advances will be made as ExT predictions are tested both experimentally and with reference to natural speech. Another essential tool in addressing these issues in the coming years will be further computational modeling of interacting agents. Crucially, such modeling needs to take better account of the social relationships that hold among agents, and to better integrate insights from psychology on memory and its limitations. We also foresee a greater contribution from studies of individual speakers and listeners, and from analysis of differences in the abilities and behaviors of individuals (Docherty and Foulkes, 2014). The individual has been rather marginalized in variationist sociolinguistics, following an assumption that individuals merely reflect the social groups in which they are located (Johnstone, 2000). However, monitoring the behavior of individuals in specific communicative settings is a vital component in exploring emergent patterns. Research into emergence is well served by rapid advances in technology, enabling the development of diverse corpora. New insights will emerge as researchers explore corpora of lesser-known languages, and as they develop new tools to investigate speaker and listener behavior, including with reference to non-speech gestures. Key to the ongoing development of this field is the recognition that language is fundamentally a human activity, and to understand it fully requires us to drop anchor in the real world, observing people engaged in using language to achieve their social, linguistic, and communicative goals.

Acknowledgments

We thank Gerry Docherty, Andrew MacFarlane, Marilyn Vihman, and the editors for helpful comments on the material presented here. Both authors' work on this publication was supported through a Rutherford Discovery Fellowship awarded to the second author. Hay's contribution was also made possible through the support of a sub-award under a grant to Northwestern University from the John Templeton Foundation. The opinions expressed in this publication are those of the authors, and do not necessarily reflect the views of the John Templeton Foundation.

REFERENCES

Abercrombie, David. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Abrego-Collier, Carissa, Julian Grove, Morgan Sonderegger, and Alan C. L. Yu. 2011. Effects of speaker evaluation on phonetic convergence. In Wai-Sum Lee and Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, pp. 192–195. Hong Kong: University of Hong Kong.

Andersen, Elaine S. 1990. *Speaking with Style: The Sociolinguistic Skills of Children*. London: Routledge.

Babel, Molly. 2010. Dialect convergence and divergence in New Zealand English. *Language in Society* 39: 437–456.

Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40: 177–189.

Balcetis, Emily E. and Rick Dale. 2005. An exploration of social modulation of syntactic priming. In Bruno G. Bara, Lawrence Barsalou, and Monica Bucciarelli (eds.), *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, pp. 184–189. Mahwah, NJ: Lawrence Erlbaum.

Barker, Brittan A. and Rochelle S. Newman. 2004. Listen to your mother! The role of talker familiarity in infant streaming. *Cognition* 94: B45–B53.

Bartholomeus, Bonnie. 1973. Voice identification by nursery school children. *Canadian Journal of Psychology* 27: 464–472.

Baxter, Gareth J., Richard A. Blythe, William Croft, and Alan J. McKane. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change* 21: 275–296.

Beckner, Clay, Richard Blythe, Joan L. Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann. 2009. Language is a complex adaptive system. *Language Learning* 59, Supplement 1: 1–27.

Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand, and Daniel Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60: 92–111.

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13: 145–204.

Blythe, Richard A. and William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88: 269–304.

Bosshardt, H.-G., C. Sappok, M Knipschild, and C. Hölscher. 1997. Spontaneous imitation of fundamental frequency and speech rate by nonstutterers and stutterers. *Journal of Psycholinguistic Research* 26: 425–448.

Bourdieu, Pierre. 1991. *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.

Bourhis, Richard Y. and Howard Giles. 1977. The language of intergroup distinctiveness. In Howard Giles (ed.), *Language, Ethnicity and Intergroup Relations*, pp. 119–135. London: Academic Press.

Branigan, Holly P., Martin J. Pickering, Janet F. McLean, and Alexandra A. Cleland. 2007. Syntactic alignment and participant role in dialogue. *Cognition* 104: 163–197.

Britain, David. 2011. The heterogeneous homogenisation of dialects in England. *Taal en Tongval* 63: 43–60.

Britain, David. 2013. Space, diffusion and mobility. In Jack Chambers and Natalie Schilling-Estes (eds.), The *Handbook of Language Variation and Change*, 2nd ed., pp. 471–500. Oxford: Wiley-Blackwell.

Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perceptions. *Language Variation and Change* 21: 135–156.

Campbell-Kibler, Kathryn. 2011. The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change* 22: 423–441.

Castellano, Claudio, Santo Fortunato, and Vittorio Loreto. 2009. Statistical physics of social dynamics. *Reviews of Modern Physics* 81: 591–646.

Chambers, Jack. 2002. Dynamics of dialect convergence. *Journal of Sociolinguistics* 6: 117–130.

Chartrand, Tanja L. and John A. Bargh. 1999. The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology* 76: 893–910.

Chater, Nick and Christopher D. Manning. 2006. Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences* 10: 335–344.

Coetzee, Andries and Joe Pater. 2011. The place of variation in phonological theory. In John A. Goldsmith, Jason Riggle, and Alan C. L. Yu (eds.), *The Handbook of Phonological Theory*, 2nd ed., pp. 401–434. Oxford: Wiley-Blackwell.

Couper-Kuhlen, Elizabeth and Margaret Selting. 1996. *Prosody in Conversation*. Cambridge: Cambridge University Press.

Cutler, Anne and Andrea Weber. 2007. Listening experience and phonetic-to-lexical mapping in L2. In J. Trouvain and W. J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, pp. 43–48. Dudweiler: Pirrot.

DeCasper, Anthony J. and William P. Fifer. 1980. Of human bonding: Newborns prefer their mothers' voices. *Science* 208: 1174–1176.

DeCasper, Anthony J. and Melanie J. Spence. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development* 9: 133–150.

de Fornel, Michel. 1992. The return gesture: Some remarks on context, inference, and iconic gesture. In Peter Auer and Aldo Di Luzio (eds.), *The Contextualisation of Language*, pp. 159–176. Amsterdam: John Benjamins.

Delvaux, Véronique and Alain Soquet. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64: 145–173.

Dijksterhuis, Ap and John A. Bargh. 2001. The perception–behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology* 33: 1–40.

Docherty, Gerard J. and Paul Foulkes. 2000. Speaker, speech, and knowledge of sounds. In Noël Burton-Roberts, Philip Carr, and Gerard J. Docherty (eds.), *Phonological Knowledge: Conceptual and Empirical Issues*, pp. 105–129. Oxford: Oxford University Press.

Docherty, Gerard J. and Paul Foulkes. 2014. An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142: 42–56.

Docherty, Gerard J., Paul Foulkes, Jennifer Tillotson, and Dominic J. L. Watt. 2006. On the scope of phonological learning: Issues arising from socially structured variation. In Louis Goldstein, Doug Whalen, and Catherine Best (eds.), *Papers in Laboratory Phonology 8*, pp. 393–421. Berlin: Walter de Gruyter.

Docherty, Gerard J., Christian Langstrof, and Paul Foulkes. 2013. Listener evaluation of sociophonetic variability: Probing constraints and capabilities. *Linguistics* 51: 355–380.

Docherty, Gerard J. and Norma Mendoza-Denton. 2012. Speaker-related variation – sociophonetic factors. In Abigail C. Cohn, Cécile Fougeron, and Marie K. Huffman (eds.), *The Oxford Handbook of Laboratory Phonology*, pp. 44–60. Oxford: Oxford University Press.

Drager, Katie. 2010. Sensitivity to grammatical and sociophonetic variability in perception. *Laboratory Phonology* 1: 93–120.

Drager, Katie. 2011. Speaker age and vowel perception. *Language and Speech* 54: 99–121.

Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12: 453–476.

Epstein, Joshua and Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge: Cambridge University Press.

Evans, Bronwen and Paul Iverson. 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.

Fagyal, Zsuzsanna, Samarth Swarup, Anna María Escobar, Les Gasser, and Kiran Lakkaraju. 2010. Centers and peripheries: Network roles in language change. *Lingua* 120: 2061–2079.

Forster, Kenneth I. and Susan M. Chambers. 1973. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior* 12: 627–635.

Foulkes, Paul. 2010. Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology* 1: 5–39.

Foulkes, Paul and Gerard J. Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34: 409–438.

Foulkes, Paul, Gerard J. Docherty, and Dominic J. L. Watt. 2005. Phonological variation in child-directed speech. *Language* 81: 177–206.

Foulkes, Paul and Jennifer Hay. 2013. The evolution of medial (-t-) in real and remembered time. Paper presented at UKLVC Conference, University of Sheffield.

Foulkes, Paul, James Scobbie, and Dominic J. L. Watt. 2010. Sociophonetics. In William J. Hardcastle and John Laver (ed.), *The Handbook of Phonetic Sciences*, 2nd ed., pp. 703–754. Oxford: Blackwell.

Gahl, Susanne. 2008. "Thyme" and "Time" are not homophones: Word durations in spontaneous speech. *Language* 84: 474–496.

Geeraerts, Dirk, Gitte Kristiansen, and Yves Peirsman. 2010. Introduction: Advances in cognitive sociolinguistics. In Dirk Geeraerts, Gitte Kristiansen, and Yves Peirsman (eds.), *Advances in Cognitive Sociolinguistics*, pp. 1–19. Berlin/New York: de Gruyter Mouton.

Giddens, Anthony. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley/Los Angeles: University of California Press.

Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. In Howard Giles, Nikolas Coupland, and Justine Coupland (eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics*, pp. 1–86. Cambridge: Cambridge University Press.

Gillen, Julia and Nigel Hall. 2001. "Hiya, mum!" An analysis of pretence telephone play in a nursery setting. *Early Years* 21: 15–24.

Goldinger, Stephen. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105: 251–279.

Goldinger, Stephen. 2007. A complementary-systems approach to abstract and episodic speech perception. In J. Trouvain and W. J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* pp. 49–54. Dudweiler: Pirrot.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury, and Peter Trudgill. 2004. *New Zealand English: Its Origins and Evolution*. Cambridge: Cambridge University Press.

Hammersley, Richard and J. Don Read. 1985. The effect of participation in a conversation on recognition and identification of the speakers' voices. *Law and Human Behavior* 9: 71–81.

Harrington, Jonathan, Sally Palethorpe, and Catherine Watson. 2005. Deepening or lessening the divide between diphthongs? An analysis of the Queen's annual Christmas broadcasts. In William J. Hardcastle and Janet M. Beck (eds.), *The Gift of Speech: A Festschrift for John Laver*, pp. 227–261. London: Lawrence Erlbaum.

Hawkins, Sarah. 2012. The lexicon: Not just elusive, but illusory? In Abigail C. Cohn, Cécile Fougeron, and Marie K. Huffman (eds.), *The Oxford Handbook of Laboratory Phonology*, pp. 162–173. Oxford: Oxford University Press.

Hay, Jennifer and Katie Drager. 2007. Sociophonetics. *Annual Review of Anthropology* 36: 89–103.

Hay, Jennifer and Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics* 48: 865–892.

Hay, Jennifer, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34: 458–484.

Hay, Jennifer and Abby Walker. 2013. Skewed experience with words affects lexical access patterns. Paper presented at Variation and Language Processing 2, University of Canterbury, Christchurch, New Zealand, January 16–18.

Hazen, Kirk. 2002. The family. In Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, pp. 500–525. Oxford: Blackwell.

Hepper, Peter G., D. Scott, and Sara Shahidullah. 1993. Newborn and fetal response to maternal voice. *Journal of Reproductive and Infant Psychology* 11: 147–153.

Holland, Brendan D., Alisha C. Murray, and Elizabeth A. Kensinger. 2013. Episodic memory and emotion. In Michael D. Robinson, Edward R. Watkins, and Eddie Harmon-Jones (eds.), *Handbook of Cognition and Emotion*, pp. 156–175. New York: Guilford Press.

Houston, Derek M. and Peter W. Jusczyk. 2000. The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance* 26: 1570–1582.

Houston, Derek M. and Peter W. Jusczyk. 2003. Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance* 29: 1143–1154.

Jackendoff, Ray. 2007. Linguistics in cognitive science: The state of the art. *Linguistic Review* 24: 347–401.

Johnson, Elizabeth K., Ellen Westrek, Thierry Nazzi, and Anne Cutler. 2011. Infant ability to tell voices apart rests on language experience. *Developmental Science* 14: 1002–1011.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson and John Mullennix (eds.), *Talker Variability in Speech Processing*, pp. 145–165. San Diego, CA: Academic Press.

Johnson, Keith, Elizabeth Strand, and Mariapaola D'Imperio. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27: 359–384.

Johnstone, Barbara. 2000. The individual voice in language. *Annual Review of Anthropology* 29: 405–424.

Johnstone, Barbara. 2009. Stance, style and the linguistic individual. In Alexandra Jaffe (ed.), *Stance: Sociolinguistic Perspectives*, pp. 29–52. New York: Oxford University Press.

Johnstone, Barbara, Jennifer Andrus, and Andrew E. Danielson. 2006. Mobility, indexicality, and the enregisterment of "Pittsburghese." *Journal of English Linguistics* 34: 77–104.

Keller, Rudi. 1994. *On Language Change: The Invisible Hand in Language*. London: Routledge.

Kenny, Laurence A. and James N. Stanford. 2012. Gender in a world without sex: An agent-based simulation of gender and language variation. Paper presented at New Ways of Analyzing Variation 41, Indiana University, October 25–28.

Kerswill, Paul. 2003. Dialect levelling and geographical diffusion in British English. In David Britain and Jenny Cheshire (eds.), *Social Dialectology: In Honour of Peter Trudgill*, pp. 223–243. Amsterdam: John Benjamins.

Khattab, Ghada. 2009. Phonetic accommodation in children's code-switching. In Barbara E. Bullock and Almeida J. Toribio (eds.), *The Cambridge Handbook of Linguistic Code-Switching*, pp. 142–160. Cambridge: Cambridge University Press.

Khattab, Ghada. 2013. Phonetic convergence and divergence strategies in English-Arabic bilingual children. *Linguistics* 51: 439–472.

Labov, William. 2001. *Principles of Linguistic Change. Volume 2: Social Factors*. Oxford: Blackwell.

Labov, William. 2010. *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors*. Oxford: Blackwell.

Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H and H theory. In William J. Hardcastle and Alain Marchal (eds.), *Speech Production and Speech Modelling*, pp. 403–439. Amsterdam: Kluwer.

Llamas, Carmen and Dominic Watt (eds.). 2010. *Language and Identities*. Edinburgh: Edinburgh University Press.

Local, John. 2003. Variable domains and variable relevance: Interpreting phonetic exponents. *Journal of Phonetics* 31: 321–339.

Local, John. 2007. Phonetic detail and the organisation of talk-in-interaction. In J. Trouvain and W. J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, pp. 1–10. Dudweiler: Pirrot.

Love, Jessica and Abby Walker. 2013. Football versus football: Effect of topic on /r/ realization in American and English sports fans. *Language and Speech* 56(4): 443–460.

Lu, Qiming, Gyorgy Korniss, and Boleslaw K. Szymanski. 2009. The Naming Game in social networks: Community formation and consensus engineering. *Journal of Economic Interaction and Coordination* 4: 221– 235.

MacFarlane, Andrew (2014) Voice activated: Exploring the effects of voices on behaviours. PhD thesis, University of Canterbury.

Mann, V. A., R. Diamond, and S. Carey. 1979. Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology* 27: 153–165.

Maye, Jessica, Daniel J. Weiss, and Richard N. Aslin. 2008. Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science* 11: 122–134.

Maye, Jessica, Janet Werker, and LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.

Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Oxford: Wiley-Blackwell.

Mendoza-Denton, Norma, Jennifer Hay, and Stefanie Jannedy. 1999. Probabilistic sociolinguistics: Beyond variable rules. In Rens Bod, Jennifer Hay, and Stefanie Jannedy (eds.), *Probabilistic Linguistics*, pp. 97–138. Cambridge, MA: MIT Press.

Miller, Cynthia L. 1983. Developmental changes in male/female voice classification by infants. *Infant Behavior and Development* 6: 313–330.

Milroy, Lesley. 1987. *Language and Social Networks*. 2nd ed. Oxford: Blackwell.

Milroy, Lesley and James Milroy. 1992. Social network and social class: Toward an integrated sociolinguistic model. *Language in Society* 21: 1–26.

Munro, Murray J., Tracey M. Derwing, and James E. Flege. 1999. Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics* 27: 385–403.

Musen, Gail and Anne Treisman. 1990. Implicit and explicit memory for visual patterns. *Journal of Experimental Psychology: Learning, Memory and Cognition* 16: 127–137.

Nardy, Aurélie, Jean-Pierre Chevrot, and Stéphanie Barbu. 2013. The acquisition of sociolinguistic variation: Looking back and thinking ahead. *Linguistics* 51: 255–284.

Natale, Michael. 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, Michael. 1975b. Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills* 40: 827–839.

Nowak, Martin A., Natalia L. Komarova, and Partha Niyogi. 2002. Computational and evolutionary aspects of language. *Nature* 417: 611–617.

Nygaard, Lynne C., Mitchell S. Sommers, and David B. Pisoni. 1994. Speech perception as a talker contingent process. *Psychological Science* 5: 42–46.

Ogden, Richard. 2012. The phonetics of talk in interaction: Introduction to the special issue. *Language and Speech* 55: 3–11.

Pardo, Jennifer S., Rachel Gibbons, Alexandra Suppes, and Robert M. Krauss. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40: 190–197.

Peek, Mike and Kevin Watson. 2013. Personality and non-linguistic perceptual sensitivity predict phonetic convergence. Paper presented at Variation and Language Processing 2, University of Canterbury, Christchurch, New Zealand, January 16–18.

Pierrehumbert, Janet. 2003a. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115–154.

Pierrehumbert, Janet. 2003b. Probabilistic phonology: Discrimination and robustness. In Rens Bod, Jennifer Hay, and Stefanie Jannedy (eds.), *Probabilistic Linguistics*, pp. 177–228. Cambridge, MA: MIT Press

Pierrehumbert, Janet. 2006. The next toolkit. *Journal of Phonetics* 34: 516–530.

Pisoni, David. 1997. Some thoughts on "normalization" in speech perception. In Keith Johnson and John Mullennix (eds.), *Talker Variability in Speech Processing*, pp. 9–32. San Diego, CA: Academic Press.

Purhonen, Maija, Riitta Kilpeläinen-Lees, Minna Valkonen-Korhonen, Jari Karhu, and Johannes Lehtonen. 2005. Four-month-old infants process own mother's voice faster than unfamiliar voices: Electrical signs of sensitization in infant brain. *Cognitive Brain Research* 24: 627–633.

Rhodes, Richard. 2012. Assessing the strength of non-contemporaneous forensic speech evidence. PhD dissertation, University of York.

Roberts, Julie. 2002. Child language variation. In Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, pp. 333–348. Oxford: Blackwell.

Saffran, Jenny R., Richard N. Aslin, and Elissa L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274: 1926–1928.

Sato, Wataru and Sakiko Yoshikawa. 2007. Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition* 104: 1–18.

Saussure, Ferdinand de,. 1974 [1916]. *A Course in General Linguistics* (trans. Wade Baskin). London: Fontana.

Savin, Harris B. 1963. Word-frequency effect and errors in the perception of speech. *Journal of the Acoustical Society of America* 35: 200–206.

Shockley, Kevin, Marie-Vee Santana, and Carol A. Fowler. 2003. Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance* 29: 326–332.

Smith, Eliot and Michael Zárate. 1992. Exemplar-based model of social judgment. *Psychological Review* 99: 3–21.

Smith, Jennifer, Mercedes Durham, and Hazel M. Richards. 2013. The social and linguistic in the acquisition of sociolinguistic variation. *Linguistics* 51: 285–324.

Smith, Kenny, Simon Kirby, and Henry Brighton. 2003. Iterated learning: A framework for the emergence of language. *Artificial Life* 9: 371–386.

Smith, Rachel and Sarah Hawkins. 2012. Production and perception of speaker-specific phonetic detail at word boundaries. *Journal of Phonetics* 40: 213–233.

Snow, Catherine E. 1995. Issues in the study of input: Finetuning, universality, individual and developmental differences, and necessary causes. In Paul Fletcher and Brian MacWhinney (eds.), *The Handbook of Child Language*, pp. 180–193. Oxford: Blackwell.

Stanford, James N. and Laurence A. Kenny. 2013. Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities. *Language Variation and Change* 25(2): 119–153.

Studdert-Kennedy, Michael. 1983. On learning to speak. *Human Neurobiology* 2: 191–195.

Sumner, Meghan and Arthur G. Samuel. 2009. The role of experience in the processing of cross-dialectal variation. *Journal of Memory and Language* 60: 487–501.

Tagliamonte, Sali A. and Alex D'Arcy. 2009. Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85: 58–108.

Walker, Abby and Jennifer Hay. 2011. Congruence between "word age" and "voice age" facilitates lexical access. *Laboratory Phonology* 2: 219–237.

Wasserman, Stanley and Katherine Faust. 1994. *Social Network Analysis*. Cambridge: Cambridge University Press.

Watts, Duncan J. 2002. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* 99: 5766–5771.

Wedel, Andy and Heather Van Volkinburg. n.d. Modeling simultaneous convergence and divergence of linguistic features between differently-identifying groups in contact. University of Arizona. http://dingo.sbs.arizona.edu/%7Ewedel/publications/PDF/Wedel_VanVolkinburgSneetches.pdf. Accessed 14 April, 2013.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann and Yokov Makiel (eds.), *Directions for Historical Linguistics*, pp. 95–188. Austin: University of Texas Press.

Witten, Kim. 2013. Sociophonetic variation and enregisterment in an online community of practice: A case study of MetaFilter.com. PhD dissertation, University of York.