Prefinal version to appear in *Handbook of Laboratory Phonology*.

Chapter 13b. The relationship between synchronic variation and diachronic change

Jonathan Harrington, Institute of Phonetics and Speech Processing, University of Munich, Germany.

1. Introduction

The dramatic effect of sound change on phonology can suddenly be brought into sharp focus by place names in which the relationship between spelling and pronunciation can be especially opaque. Consider, then, the possibly amused reaction of a local inhabitant to an unwary visitor who asks for directions to the English villages of Cholmondeston, Happisburgh, or Wrotham, that is to /ˈtʃʌmstn̩/, /heɪzbrə/, or /ˈru:təm/. The extent of synchronic variation can be no less dramatic, as exemplified by the numerous phonetic forms associated with different meanings of *I do not know* that can reduce to little more than a sequence of three nasalized schwas differing minimally in phonetic height (Hawkins, 2003).

Diachronic and synchronic variation are evidently linked and especially at points in the utterance at which synchronic variation is high: moreover, as will be discussed at a later point in this chapter, diachronic change may be propagated by imitation. For example, the considerable variation synchronically in syllable-final compared with syllable-initial stops (e.g., Byrd, 1996) has a reflex in the greater tendency for diachronic place assimilation (Latin: *scriptu* > Italian *scritto*) and vowel-consonant blending (e.g, the development of nasal vowels in French) to occur in domain-final position (Hock, 1992; Ohala, 1990; Ohala & Kawasaki, 1984). Similarly, Beckman et al. (1992) show how many prosodically-induced diachronic changes such as intervocalic stop lenition and the deletion of weak vowels and syllables can be explained using the same mechanisms of gestural overlap, hiding, and truncation that form part of modeling speech production synchronically in the task dynamic model (see also Browman & Goldstein, 1992).

One of the aims in developing a phonetic model of sound change is not only to find evidence for such diachronic-synchronic relationships but also to explain the way in which fine-grained, continuous variability in speech communication can give rise over a much longer time-scale to the change from one abstract symbolic category into another. A question closely related to this is the following: if sound change is not planned i.e. is not teleological at the level of conversational interaction (Ohala, 1993; Lindblom et al., 1995), then how does it come to be that, far from being random, similar sets of patterns of phonological change such as vowel chain shifting can be observed in languages and their varieties? As a first step towards answering these difficult issues that repeatedly cause us to be confronted with one of the main issues in laboratory phonology of how phonetics and phonology are related, a brief overview will be given of the types of synchronic variation that are likely to be most relevant for understanding sound change.

2 The nature of synchronic variation

Synchronic variation is ubiquitous and occurs for a number of different reasons. First, the fact that we can speak intelligibly when performing different activities such as giving a lecture, talking while riding a bicycle, running, or taking part in immediate compensation experiments (Riordan, 1977) in which one or more articulators are artificially constrained, shows the plasticity of the speech production system in adapting to different environments (Lindblom, 1990). Secondly, variability is predicted from the non-linear relationship between articulation and acoustics: for example, a back vowel like [u] can be produced with a variable constriction location without the variability having very much effect on the resonances which are critical for the vowel's perceptual identification (Stevens 1989) - see **Iskarous, Chapter 20c**). Thirdly, the evidence for cue trading in perception experiments (Repp, 1981) suggests a certain degree of variability tolerated in the production of speech: thus the intervocalic stop voicing distinction can often be achieved by co-varying the extent of voicing in the closure, the duration of aspiration/frication and the direction of the first formant trajectory (Lisker, 1986).

A fourth type of synchronic variability, which has been central to Ohala's (1993) model of sound change comes about because of the biological and physical constraints on the speech production and perception mechanisms and the interaction between the two. Thus, laryngeal tension at the onset of voiceless stops tends to carry over synchronically into the following vowel causing the fundamental frequency to be raised: such variation has been shown diachronically to be related to the phonological development of tone in many Asian languages (Hombert et al., 1979). It is often possible to relate this type of synchronic variability to cross-linguistic patterns in the distribution of sounds. Consider that both the infrequent occurrence of high, compared with low, nasalized vowels and the tendency for high vowels to lower if they are nasalised diachronically (e.g., the development of Latin *una*/*unus* into feminine /yn/ but masculine /œ̃/, rather than /ỹ/, in French) can both be related to the same perceptual constraint introduced by nasalization: nasalized vowels produced with a high tongue position are nevertheless perceived to be phonetically lower because of the introduction of a nasal formant intermediate in frequency between the first two oral formants (Beddor et al., 1986; Wright, 1986).

An important characteristic of the above fourth type of variability is that it arises involuntarily due to factors like biomechanical inertia and limitations on the perceptual system (and this is one of the main reasons why Ohala, 1993 argues that so much of sound change is not cognitive and therefore not teleological). But there is a fifth type of synchronic variation that evidently does not fit into this non-cognitive category and that is more directly associated with a range of meanings that are conveyed by the speaker. Speakers evidently vary the clarity of their speech in relation to how predictable the speech signal is for the listener (Lindlom, 1990). This is different from the fourth type of variability, both because a speaker has control over the extent of reduction of a phrase or word (as the earlier example of Hawkins', 2003 *I don't know* demonstrates) and because it can provide listeners with cues about the 'newness' of the information (Fowler & Housum, 1987). Included within this group might also be phonetic variation due to different kinds of prosodic structure, such as the syllable-dependent 'clarity' of /l/ in many English varieties, the degree of strengthening of consonants at different phrase boundaries (Keating et al., 2003), as well as the numerous cues arising from conversational interaction (Local, 2003), such as the phonetic markers to indicate whether a speaker has reached the end of a speaking turn. It is this fifth type of variation that is primarily implicated in Lindblom et al.'s (1995) model of sound change discussed in further detail below.

The preceding type of variation is primarily linguistic. But there is also a sixth type of non-linguistic variation that provides information about the speaker, including the emotional state and attitude of the speaker, as well as regional and social information which have been the primary data in many sociolinguistic investigations of sound change (Labov, 1994, 2001). But this type of variation has not until recently found its way into phonetic models of sound change primarily for the reasons amplified in Docherty (**chapter 10, this volume**) that sociophonetic variability has been marginalized in developing cognitive models of speech production and perception.

3 Phonetic models of sound change

One of the most influential accounts of the relationship between synchronic variability and diachronic change is due to Ohala (1981, 2003) in which, as discussed in further details by Chitoran (**Chapter 13a, this volume**), many sound changes are attributed to the listener's misparsing of coarticulation. This idea is also central to the analysis of sound change in Browman & Goldstein (1991) based on temporally overlapping gestures in the framework of articulatory phonology (Browman & Goldstein, 1992), also discussed by Chitoran.

Like Browman & Goldstein (1991), Lindblom et al. (1995) give greater emphasis to the role of the speaker in sound change as well as to the idea of volition. A central aspect of their model is that sound change arises along the continuum from hypo- to hyperarticulation. In normal conversational interaction, listeners typically attend to *what* is being said (the 'what' mode) whereas *how* something is said is not usually the focus of attention. It is when the how mode is especially active that a listener may sample a new pronunciation variant and add it to the lexicon. A prediction of their model is that a listener might add a new pronunciation variant at points of hypoarticulation when the variability in the speech signal is high: this is because, given that these also tend to be points of low information content, the what mode is to a large extent disengaged as a result of which the how mode is active. Thus, a very interesting aspect of the model of Lindblom et al. (1995) is that it makes quite explicit how information redundancy, high production variability, perception, and sound change might be inter-connected. In Lindblom et al. (1995), the lexicon is assumed to include multiple variant pronunciations sampled from those that are perceived from language use in everyday conversation and it is this aspect of their model that also foreshadows the similar idea in exemplar theory (Pierrehumbert, 2003, 2006) that the lexicon stores considerable amounts of non-redundant information and fine phonetic detail.

Since words and phrases of high frequency are more likely to be produced in a semantically redundant and therefore hypoarticulation context, and since this is also one of the main contexts in which new pronunciation forms are presumed to be absorbed into the lexicon, then, according to Lindblom et al (1995), high frequency words should undergo sound change earlier than low frequency words (see Hooper, 1976 and Philips, 1984, 1994 for compatible evidence). The idea that the progress of sound change is linked to lexical frequency is also central to Bybee's (2008) usage-based model which is founded on the idea that linguistic structure is created as language is used. In this model, sound change takes place in words and phrases as a result of the reduction of gestures through repetition. The change to lexical items is modeled in an exemplar framework in which a cluster of new phonetic variants for a word are updated. Bybee's (2008) model and indeed any conceivable model of sound change in exemplar theory is founded on the '*fact* [emphasis added] that articulatorily motivated sound change takes place earlier in high-frequency words than in low-frequency words' (Bybee, 2008, p. 115). However, the empirical evidence showing that sound change applies earlier in high-frequency words is still somewhat sparse. For example, the analysis by Dinkin (2008) of an extensive amount of formant data from the Telsur survey of American English (Labov et al., 2006) found little evidence to suggest that the Northern cities vowel shift is affected by lexical frequency and it is also questionable whether these effects, if they exist at all, really are lexical (Cohn, 2005).

In both Ohala (1993) and Lindblom et al (1995), sound change at the level of the individual is phonetically *abrupt* because the change from the old to the new pronunciation takes place in one discrete jump, rather than in gradual increments between the two. On the other hand, regular sound change for Labov (1981, 1994) is the result of the *gradual* transformation of a single phonetic property in a phonetic space. Mowrey & Pagliuca (1995) also favor an interpretation of sound change as gradual and consider that claims for abruptness are an artifact of orthographically based, categorical approaches to sound change. In Blevins (2004), sound change that originates from ‘choice’ or ‘chance’ in her model can occur ‘without noticeable changes in pronunciation or with gradual changes in pronunciation.’ Mowrey & Pagliuca (1995) present arguments based on neuromuscular activation to show how even metathesis may be gradual. Recent experimental evidence from performance errors is relevant in this regard in showing how many perceived categorical errors such as /k/ for /t/ substitutions are gradient in which the tongue-dorsum gesture for /k/ intrudes incrementally upon the tongue-tip raising gesture for /t/ (Goldstein et al., 2007; Pouplier, 2008).

Fig. 1 about here

The acoustic analysis of the Christmas broadcasts of Queen Elizabeth II in Harrington et al. (2000) and Harrington (2006, 2007) showed not an abrupt, but incremental phonetic change within the same individual over a fifty year period. Moreover, these changes in the Queen’s vowels were gradual and over a long time period (in some cases of over 30 years) and quite possibly at such a slow rate (e.g., an estimated 60 Hz per decade for /u/) that they may well be imperceptible within a short time period. Mielke (2007) quite rightly comments that these averages per decade might obscure different changes within words or within individual years, and proposes instead an analysis by year or by word. It is of course very difficult to do this, because the Queen did not always produce the same content words in abundance from one year to the next. However, some formant data for the most frequently occurring word with an /æ/ nucleus per annum in the Christmas broadcasts, *family*, is shown in Fig. 1 over a 20 year period between 1952 and 1972. These data are suggestive of a gradual change of around 250 Hz between 1952 and 1964 in /æ/ within the same individual and producing the same word.

The change is not likely to be due to the physiological effects of aging (in which F1 lowers rather than raises - see Harrington et al., 2007) and it is in the same direction as the community change in which the /æ/ was reported to be becoming more open in this period (e.g., Gimson, 1966).

This type of diachronic phonetic change found within the same individual seems to be quite reminiscent of the Neogrammarians' analysis of sound change as incremental and quite possibly not perceptible, at least not from year to year.

4 The relevance of imitation for modeling sound change

Both the above studies of the Christmas broadcasts as well as other longitudinal investigations (e.g. Sankoff & Blondeau, 2007) show that adults beyond the critical age of language acquisition are influenced by diachronic change taking place in the community. Labov (2006) comments that the extent of adaptation is greater in children than in adults and may even diminish in adulthood with increasing age. This is, as Docherty (**Chapter 10, this volume**) notes, an issue that needs further investigation. Another is how these kinds of phonetic adaptations to community changes in adults come about at all. Some results and conclusions from recent studies on speech imitation may begin to provide a solution to this issue. For example Silverman (2006) argued on the basis of acoustic and perceptual data that the sound change by which rounding spreads diachronically across velars, but not alveolars, in Trique could be accounted for by imperfect copying or imitation, but not by an intentional modification of phonetic variants.

In the imitation paradigm, subjects' speech is compared before and after they have performed a task such as shadowing or listening to another speaker whose speech may have been artificially manipulated in some way. There are now various experiments that have shown that the subjects' speech production is shifted subtly towards the speech that they have listened to, after they have engaged in the task. One of the first of these was due to Goldinger (1998) who demonstrated a shift, as judged from whole-word perception experiments, towards the speech being shadowed. This experiment was replicated by Shockley et al. (2004) who additionally showed that the imitation in the shadowing experiment was phonetic: subjects shifted their VOT in the direction of shadowed material in which VOT had been artificially lengthened. More recently Nielson (2007) extended these experiments by demonstrating just such imitation in a listening task in which subjects were recorded before and after they listened to a speaker whose VOT in /p/-initial stops had been lengthened. Nielson (2007) showed not only that subjects' VOTs were lengthened after listening to these stimuli, but also that the imitations generalized to /k/, even though /k/-initial words had not been part of the stimuli they had listened to. Interestingly, Nielsen (2007) did not find an imitation effect when subjects listened to stimuli with shortened VOTs possibly because any such imitations would encroach too much upon the acoustic-phonetic space of the corresponding voiced stop phoneme. In quite a different kind of experiment, Pardo (2006) demonstrated a phonetic convergence between interlocutors who took part in a conversation in the MAP task paradigm (Anderson et al., 1991). Finally, subjects in Delvaux & Soquet (2007) had to name an ideogram, X, in sentences like *il y a une X dans le pot* produced in a different variety of Belgian French. Their attention was therefore on X, but what was measured was the /o/ in *pot* which was repeated from trial to trial. Thus, imitation took place in a repetitive and semantically predictable context. Recall from the discussion earlier, that this is exactly the kind of context in which the pool of variants is likely to be sampled leading to a potential sound change in the model of Lindblom et al. (1995): therefore, this is a context in which the 'how' mode is strongly activated allowing novel pronunciations to be suggested to the listener, just as they would have been in the experiment by Delvaux & Soquet (2007), given the strong phonetic differences between the two varieties in the production of /o/.

The production-perception mechanisms that might be responsible both for these kinds of subtle unwitting shifts in imitation and regular sound change are interpreted by Goldinger (1998) in terms of an episodic/exemplar model of speech perception in which lexical items are built out of auditory traces of words accumulated in long-term memory. However, Pardo (2006) rejects an exemplar-based analysis of the greater convergence because her speakers' imitations were not tied to any specific lexical item. Moreover, since features and phonemes in exemplar theory are supposed to be emergent statistical generalizations across word-based stored exemplars, then it is not clear how the results in Nielsen (2007) or Shockley et al. (2004) in which phonemes or even features are imitated over a short time scale could be accounted for by a shift of stable feature or phoneme-based generalizations that have been built up over the speaker-hearer's lifetime, or at least not without invoking an adaptable phonological processor in speech perception and production that is independent of the lexicon (e.g. McQueen et al., 2006).

In a direct-realist model by contrast, imitation can be modeled as a natural consequence of the presumed 'common currency' of layered gestures that are invoked in both speech production and speech perception. As discussed in Fowler (2000), perceiving gestures might serve as a prime or goad for their imitation in production analogous to the spontaneous imitation of facial expression (see also Fowler et al., 2008 and Shockley et al., 2009 for a further discussion). Moreover, based on analyses of the remarkably slight discrepancy between choice and simple reaction times in speech shadowing tasks, Fowler et al. (2003) propose that speech perception can have an immediate influence on speech production without recourse to cognitive processing. For these reasons, a certain degree of imitation is predicted to occur in the direct realist framework as an automatic consequence of perceiving the same abstract speech gestures that control speech production. This direct link between perception, action and imitation is also central to Sancier & Fowler's (1997) explanation of the slight shift in the VOT of Portuguese and American English stops produced by a bilingual Portuguese-English speaker after the speaker had left the U.S.A. to spend several months in Brazil: they model this change as the result of a realignment of the speaker's laryngeal-supralaryngeal phase relationships induced by the perception of gestures in the ambient Portuguese language environment.

5 Imitation and sociophonetic constraints

At the same time, imitation cannot be entirely automatic. For example, Mitterer & Ernestus (2008) found that it was primarily phonologically relevant detail that was imitated, suggesting a somewhat looser coupling between production and perception than implied by Fowler et al. (2003). In addition, Babel (2009) has recently shown that the extent of subjects' imitation is conditioned by social factors: for example, they tend to imitate speakers more if they have a positive attitude towards them. As far as sound change within the individual is concerned, we also have to explain, as Labov (2006) comments, not only why younger speakers seem to adapt their speech to a greater extent in moving to a new community than older speakers, but also that some adults adapt very little or perhaps not all. Moreover there is so far no explanation for the incompleteness of sound change in the Christmas broadcasts: in Harrington et al. (2000), we found that there was shift towards, but not an attainment of, less aristocratic, mainstream RP vowels and certainly no evidence that the Queen adopted what were, for much of the 20th Century, stigmatized phonetic variants such as the London Cockney glottal stop in place of syllable-final /t/.

There might therefore be a sociophonetic regulatory system that prevents imitation and sound change from applying blindly. The model of Lindblom et al. (1995) incorporates various forms of feedback to evaluate the potential sound change for its articulatory, auditory, and sociophonetic cost. The first two of these could be conceived of in terms of the regulatory feedback or feedforward systems that have been proposed at the level of the speaker and hearer (e.g., Guenther & Perkell, 2004) but Lindblom et al. (1995) also invoke a more abstract community-level feedback which measures and regulates the sociophonetic consequences of the potential imitation and sound change (see also Pierrehumbert, 2003 for a similar idea). The sociophonetic regulator would presumably be like a filter allowing imitation to pass from hearer to speaker except for a 'black list' of allophones whose copying was prohibited. But this seems to accord the speaker a great deal of volition in deciding what to imitate which may be incompatible with the results from some of the studies discussed above showing that imitation takes place largely without the speaker's awareness.

Perhaps the way forward is to abandon the idea of a sociophonetic regulator and instead to recast the mechanism of imitation and the transmission of incremental sound change as a by-product of the way that speakers use language in conversation to interact with each other in solving cooperative tasks (Giles et al., 1991; Clark, 1996; Garrod & Doherty, 1994). For example, Garrod & Pickering (2009) discuss how when subjects interact with each other in some form of co-ordinated action or task, then imitation and entrainment are likely "at many different levels, from basic motor programs to high-level aspects of meaning". They also emphasize that the influences are largely automatic so that interactants are typically unaware of the alignment processes. In this model, the linguistic and motoric alignment of speakers facilitates co-operative action: that is, the success of solving a task collaboratively requires the development of a macroscopic structure in which the individual speakers' action plans are fused in a common goal, and it is this shared plan between interactants that either brings about, or is facilitated by, an alignment between them at various linguistic and motoric levels.

As Krauss & Pardo (2006) comment, while these types of models are informed principally by linguistic imitation, they also make a number of interesting predictions concerning phonetic convergence that have for the most part not been tested. One of them is that the more speakers are able to engage successfully in co-operative tasks, then the more likely it should be that they influence each other resulting in greater phonetic adaptation. Conversely (and compatibly with the results from Babel, 2009 discussed above), adaptation should be less likely when co-operation fails, or in tasks involving speakers who are unsympathetic to each other. Also, the influence on adaptation and sound change of a more passive medium such as television might be expected to be comparatively negligible (but see Stuart-Smith, 2006) precisely because there is no macroscopic action plan between the television and the recipient. These issues might go some way to explain why speakers differ in their extent of adaptation when exposed to a different variety for a longer period. Finally, the reason why the Queen has not embraced Cockney-style glottal stops would not be because of an internalized sociophonetic monitor banning the uptake of these allophones, but instead because the opportunities for Her Majesty to engage in conversational, cooperative task solving with members of the Cockney speaking community have probably been quite scarce in the last fifty years.

6 Modeling sound change in an interactional, self-organizational system

The previous section suggests, then, that imitation is a consequence of co-operative interaction between two individuals and that sound change may be a derivative of such imitation. But how exactly can we explain the way in which gradient synchronic variability ultimately produces phonologization i.e. a categorical shift? From the above considerations, it is immediately apparent that if we view imitation and sound change as a consequence of joint cooperative action between speakers, then any model based on versions of the speech chain, in which the speaker is compartmentalized from the listener, is not likely to yield many interesting solutions to this problem.

An alternative approach - and one that is more compatible with models of cooperative interaction discussed in the preceding section - is to situate both sound change and the relationship between phonology and phonetics in terms of a model of self-organization (see Lindblom et al., 1984 for an early application of self-organization to speech). As discussed in Oudeyer (2006), a common theme in systems of self-organization, which have also been used to predict many phenomena in nature such as the formation of ice-crystals and the cathedral-like formation of termite nests, is that a macroscopic structure emerges as a consequence of the interaction between sub-components of the system (Shockley, et al., 2009). Thus for speech, computational models consisting of agents with simplified vocal tracts and hearing systems are sometimes used to show how phonological structure emerges from the cumulative effect of many similar, imperfect imitations over time between speakers and hearers (e.g, de Boer, 2001). In this kind of model, a language's phonology is not determined by innate principles of universal grammar but is instead just one of the many possible solutions to the way in which convergence arises from speaker-hearer interactions (see Bybee, 2008, for a related interpretation that linguistic universals and specifically the principal of structure preservation in Kiparsky's 1985 theory of lexical phonology are not innate but arise through the interaction and change). As a specific example, the notion in universal grammar of phonologically unmarked vs. marked is recast in Kochetov's (2008) self-organizational model as stability vs. instability respectively in the relationship between the production and perception systems. Notice also that there is no sense in this kind of model in which phonology precedes, or is translated into, phonetics. Thus in Gafos (2006) and Gafos & Benus (2006) both the discrete and continuous aspects of complex systems are related using the same formal language of the mathematics of non-linear dynamics whose properties of differential equations are used to express the relationship between category stability and change.

Blevins & Wedel (2009) present a model of sound change based on self-organization, albeit within the exemplar framework (see also Wedel, 2007). In their model, sound change is the outcome of the opposition between noise in the production-perception feedback loop which can cause category instability by introducing new phonetic variants and a so-called reversion to the mean which, through processes such as motor entrenchment (Zanone & Kelso 1997) and the perceptual magnet effect (Kuhl, 1991), maintains category stability. Blevins & Wedel (2009) demonstrate one of the main principles of self-organizational systems that sound change occurs when the phonological system is in an unstable state. Similarly, Kochetov (2008) used computer simulations to show that the combination of complex vowel inventories and secondary consonant articulations is unstable as far as production-perception relationships are concerned: as a result, the language self-organizes to a more stable system with either rounding contrasts in the vowels or secondary articulations in the consonants (or neither or these).

In a self-organizational model, change can, but need not, be driven by social forces. This is because, even within a socially stable system, there is nevertheless a randomness in the way that speakers and hearers interact with each other and this kind of noise can be shown mathematically to push a phonological system from one stable state to another (e.g. Gafos & Benus, 2006). Alternatively, a stable phonological system can be made unstable because of the changing speaker-hearer interactions that might result either from a realignment of the social structure in the community or from dialect contact. Notice that, in neither case is there any sense in which sound change is teleological or planned, precisely because the randomness of speaker-hearer interactions implies that there is an unpredictably in the phonological reorganization that they give rise to.

6 Concluding remarks

Sound change seems to be imperceptible and non-teleological at the level of speaker-hearer interactions but organized and apparently purposeful at a macroscopic level. The following components of phonetic models of speech production and perception have been especially useful for modeling this dichotomy. (1) Listeners make unintentional errors in processing the acoustic signal in which coarticulation is misparsed. (2) Phonetic variants may be especially salient at points of high information redundancy in the speech signal. (3) Speakers and listeners imitate each other unwittingly and this may be one of the mechanisms by which sound change is transmitted incrementally. (4) Imitation may be a consequence of shared macroscopic action plans between a speaker and listener in conversational interaction that may be affected by social forces. (5) A self-organizational model is most likely to be compatible with (4) in which phonological category stability and change are emergent properties of speaker-hearer interactions.

Finally, the gradient modeling of the shifts that result both from sound change (Browman & Goldstein, 1991) and the perception of the ambient linguistic environment (Sancier & Fowler, 1997) in a gestural model provides a way of thinking about sound change that goes beyond the categorical shift from one IPA allophone to another and that may turn out to be as fruitful as has been the recent recasting of categorical performance errors as gradient shifts of gestural intrusion and reduction (e.g., Goldstein et al.. 2007; Pouplier, 2008).

References

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S. and Weinert, R. (1991). The HCRC Map Task Corpus. *Language & Speech*, 34, 351-366.

Beckman, M., de Jong, K. Jun, S-A., and Lee, S-H. (1992) The interaction of coarticulation and prosody in sound change. *Language and Speech*, 35, 45-48.

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge University Press: Cambridge.

Blevins, J. and Garrett, A. (2004). The evolution of metathesis. In B. Hayes, R. Kirchner, and D. Steriade (Eds.) *Phonetically Based Phonology*. Cambridge University Press: Cambridge. (p. 117-156).

Blevins, J. & Wedel, A. (2009). Inhibited sound change: an evolutionary approach to lexical competition. *Diachronica*, 26.2.

Browman, C. and Goldstein, L. (1991). Gestural structures: distinctiveness, phonological processes, and historical change. In I. Mattingly and M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception.* Erlbaum: New Jersey. (p. 313-338).

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics,* 24, 209-244.

Cohn, A. (2005) Levels of abstractness in phonology and the lexicon: evidence from English homophones.paper presented at the 79 Meeting of the LSA, Oakland, CA, January 2005 and the 13th Manchester Phonology Meeting, May 2005. http://ling.cornell.edu/docs/CohnhomophonesHO.pdf

de Boer, B. (2001) *The Origins of Vowel Systems*. Oxford University Press: Oxford.

Delvaux, V. & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, 64, 145–173.

Dinkin, A. (2008) The real effect of word frequency on phonetic variation. *University of Pennsylvania Working Papers in Linguistics*, 14.1. Available from: http://repository.upenn.edu/pwpl/vol14/iss1/8/

Docherty, G. & Foulkes, P. (2006) The social life of phonetics and phonology. *Journal of Phonetics*, 34, 409-438.

Fowler, C. A. (2000). Imitation as a basis for phonetic learning after the critical period. *Paper presented at the Twenty-fifth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, California.

Fowler, C., Brown, J., Sabadini, L., and Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, 49, 396–413.

Fowler, C., Richardson, M., Marsh, K., and Shockley, K. (2008) Language use, coordination, and the emergence of cooperative action. In A. Fuchs & V. Jirsa (Eds). *Understanding Complex Systems*. Springer: Berlin. (p. 261-279).

Gimson, A.C. (1966). *An Introduction to the Pronunciation of English* (1st edition). Edward Arnold: London.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279.

Goldstein, L., Pouplier, M., Chen, L., Saltzman, E. & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition,* 103, *386-412.*

Guenther, F.H., and Perkell, J.S. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In: B. Maassen, R. Kent, H. Peters, P. Van Lieshout, and W. Hulstijn (eds.), *Speech Motor Control in Normal and Disordered Speech (pp. 29-49)*. Oxford: Oxford University Press.

Harrington, J., Palethorpe, S., and Watson, C. (2000). Does the Queen speak the Queen's English? *Nature*, 408, 927-928.

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 and J. Hualde (Eds.) *Laboratory Phonology 9*. Mouton de Gruyter:Berlin. (p. 125-143).

Harrington, J., Palethorpe, S. and Watson, C.I. (2007). Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers. *Proceedings of Interspeech* 2007, Antwerp.

Hock, H. (1992). Causation in language change. In W. Bright (Ed.) *Oxford International Encyclopedia of Linguistics*, Vol. 1.. London & New York: Oxford University Press. (p. 228-231).

Hooper, J. (1976). Word frequency in lexical diffusion and the source of morphophonological change. in Christie, W. (ed.) *Current Progress in Historical Linguistics*. Amsterdam: North Holland. (p. 96-105).

Hyman, L. (1973) The feature [grave] in phonological theory. *Journal of Phonetics*, 1, 329-337.

Keating, P., Cho, T., Fougeron, C., and Hsu, C. (2003): Domain-initial articulatory strengthening in four languages. In J. Local & R. Ogden (Eds.) *Papers in Laboratory Phonology 6*. Cambridge University Press: Cambridge. (p. 143-161).

Kuhl, P. (1991). Human adults and human infants show a 'perceptual magnet' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93–107.

Labov, W. (1981) Resolving the Neogrammarian hypothesis. *Language*, 57, 267-308.

Labov, W. (1994) *Principles of Linguistic Change. Volume 1 Internal Factors*. Blackwell: Malden, MA.

Labov, W. (2006) *The Social Stratification of English in New York City* (2nd Edition). Cambridge University Press: Cambridge.

Labov, W., Ash, S. and Boberg, C. (2006). *Atlas of North American English: Phonetics, Phonology and Sound Change*. Mouton de Gruyter.

Lahiri, A. & W.D. Marslen-Wilson (1991). The mental representation of lexical form:a phonological approach to the recognition lexicon. *Cognition*, 38, 245-294.

Lindblom, B. (1990) Explaining phonetic variation: a sketch of the H & H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modeling*. Kluwer: Dordrecht. (p. 403-439).

Lindblom, B., MacNeilage P., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie & O.Dahl (eds.) *Explanation for Language Universals*. Mouton: Berlin. (p. 181–203).

Lindblom, B., Guion, S., Hura, S., Moon, S-J., and Willerman, R. (1995). Is sound change adaptive? *Rivista di Linguistica*, 7, 5–36.

Martinet, A. (1955). *Economie des Changements Phonétiques*. Bern: Francke.

McMahon, A. (1994). *Understanding Language Change*. Cambridge University Press: Cambridge.

McQueen, J.M., Cutler, A. & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, 30, 1113-1126.

Mielke, J. (2007). Multiple mechanisms of change and influence: comments on Harrington, Gussenhoven, Gow and McMurray, and Munson. In J. Cole and J. Hualde (Eds.) *Laboratory Phonology 9*. Mouton de Gruyter:Berlin. (p. 229-240).

Mitterer, H. & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition*, 109, 168–173.

Mowrey, R. & Pagliuca, W. (1995). The reductive character of articulatory evolution. *Rivista di Linguistica*, 7, 37-124.

Nielsen, K. (2007). Implicit phonetic imitation is constrained by phonemic contrast. *Proceedings of the 16th International Congress of Phonetic Sciences*, 1961-1964. Saarbrücken, Germany.

Ohala, J.(1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, and M. F. Miller (Eds.) *Papers from the Parasession on Language and Behavior*. Chicago Linguistic Society, Chicago. (p. 178–203).

Ohala, J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston and M.E. Beckman (eds.). *Papers in Laboratory Phonology I*. p. 258-275. Cambridge University Press: Cambridge.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (Ed.) *Historical Linguistics: Problems and Perspective*s. Longman: London. (p. 237–278).

Ohala, J.J. and Kawasaki, H. (1984). Prosodic phonology and phonetics. *Phonology Yearbook*, 1, 113-127.

Oudeyer, P-Y. (2006). *Self-Organization in the Evolution of Speech*. Oxford University Press: Oxford.

Pardo, J. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119, 2382-2393.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46, 115–154.

Pierrehumbert, J. (2006). The next toolkit. J. *Phonetics*, 34, 516-530.

Pouplier, M. (2008). The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics*, 36, 114-140.

Sancier, M. & Fowler, C. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421-436.

Shockley, K., Sabadini, L., and Fowler, C. (2004). Imitation in shadowing words. *Perception and Psychophysics*, 66, 422-429.

Silverman, D. (2006) The diachrony of labiality in Trique, and the functional relevance of gradience and variation, in L. Goldstein, D. Whalen, and C. Best (Eds.), *Papers in Laboratory Phonology 8*, Mouton de Gruyter: New Haven. (p. 133-154).

Vennemann, T. (1993). Language change as language improvement, In Charles Jones (ed.), *Historical Linguistics: Problems and Perspectives*. London: Longman. (p. 319-344).

Wedel, A. (2007). Feedback and regularity in the lexion. *Phonology*, 24, 147-185.

Zanone, P. & Kelso, J. (1997). Coordination dynamics of learning and transfer: collective and component levels. *Journal of Experimental Psychology: Human Perception and Performance* 23, 1454–1480.

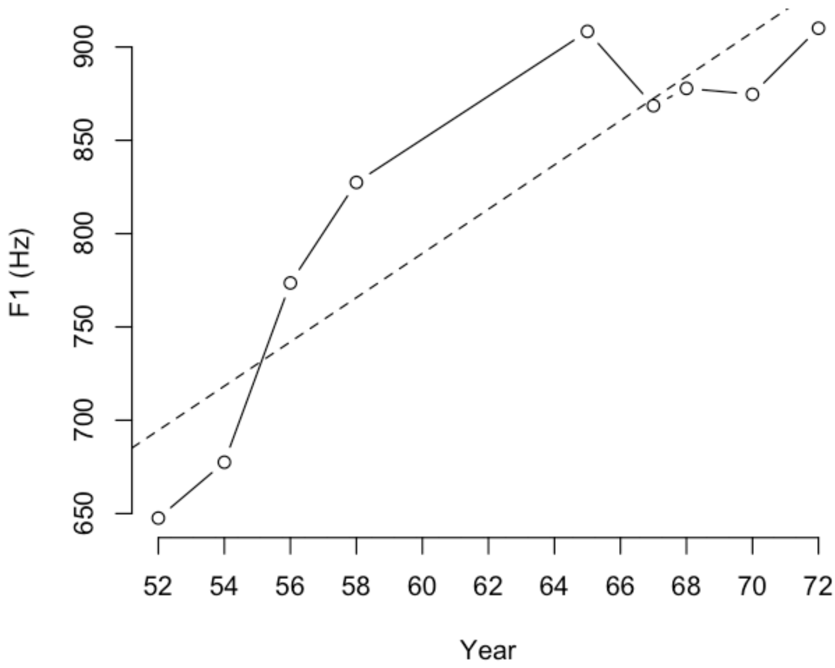

Fig. 1. F1 at the temporal midpoint of /æ/ averaged across all productions of *family* that occurred in any one year (the absence of a data point means that there were fewer than 3 tokens in any given year). The data are from Queen Elizabeth II producing the annual Christmas broadcasts (Harrington et al, 2000). All productions of *family* were prosodically accented. Data points are only shown when there were at least three tokens of *family* in any one year.