THE INDIVIDUAL AND THE ACTUATION OF SOUND CHANGE

Mary Stevens and Jonathan Harrington

Institute for Phonetics and Speech Processing
`mes;jmh@phonetik.uni-muenchen.de`

ABSTRACT

This paper reviews current experimental approaches to sound change. An ongoing challenge in sound change research is to link the initiation of sound change within individual cognitive grammars to the diffusion of novel variants through the community. The articulatory and perceptual phonetic forces that bring about the pre-conditions for sound change and that explain its directionality are always present in the transmission of spoken language, yet sound systems are remarkably stable over time. This paper describes how recent approaches to the actuation problem converge on the idea that variability between individuals may be the key to understanding how some synchronic variation can become sound change. It then reviews the evidence for individual differences based on four areas of phonetic research (speech production, speech perception/cognitive processing, the perception-production link, and linguistic experience and imitation). This evidence suggests that differences between individuals may help to explain why sound change is so rarely actuated even though the phonetic pre-conditions are constantly being generated in spoken language interactions.

Keywords: sound change, speech perception, speech production, individual differences, imitation.

RESUMEN

El presente artículo reseña actuales enfoques experimentales del cambio fonético. Un problema recurrente en la investigación del cambio fonético es vincular la iniciación del cambio fonético en la gramática cognitiva de un individuo con la cuestión de la difusión de nuevas variantes en la comunidad de habla. En la transmisión del lenguaje oral, las fuerzas articulatorias y perceptivas que hacen surgir las condiciones para el cambio fonético están siempre presentes. Al mismo tiempo, los sistemas fonológicos de las lenguas se caracterizan por ser considerablemente estables. En este artículo se presentan acercamientos actuales al problema de la iniciación del cambio fonético. Se mostrará que los diferentes enfoques acuerdan en la idea que la variabilidad interindividual puede ser la clave para comprender cómo de la variación sincrónica puede surgir cambio fonético. Se presentará una visión de conjunto de los diferentes factores que dan origen a la variación interindividual, basada en estudios fonéticos de cuatro áreas: 1) la producción del habla, 2) la percepción y el procesamiento cognitivo, 3) la relación entre la producción y la percepción y, 4) la experiencia lingüística y la imitación. Los resultados de estos estudios señalan que diferencias individuales pueden ser el motor del cambio fonético. Al mismo tiempo ayudan a explicar por qué el cambio fonético es iniciado sólo raras veces – a pesar de que las interacciones entre hablantes generan constantemente las condiciones para el cambio fonético.

Palabras clave: cambio fonético, percepción del habla, producción del habla, diferencias individuales, imitación.

1. INTRODUCTION

The question of how sounds change over time has fascinated speakers and linguists for centuries. Knowledge about sound change helps to reconstruct our linguistic past and it has always been a central part of historical linguistics. But sound change is an especially popular research topic at present with a biannual workshop series established in 2010 and at least five volumes dedicated to sound change appearing since that time (Recasens, Sánchez Miret & Wireback 2010, Solé & Recasens 2012, Sánchez Miret & Recasens

2013, Yu 2013, Harrington & Stevens 2014). The contributions in these volumes show that sound change research now incorporates aspects of many areas of linguistics and neighbouring disciplines including cognitive psychology, computational science, experimental phonetics, laboratory phonology, language acquisition, sociolinguistics, phonology and physics.

Sound change can be defined as change to the shared perception and production target for a speech sound within a speech community, a definition that encompasses changes that directly impact the number of categorical contrasts between sounds (e.g. neutralization) as well as changes that involve a shift in the pronunciation target for a speech segment without loss or introduction of a phonemic contrast (e.g. vowel chain shifts). The conditions that give rise to sound change are typically distinguished from those that have to do with its diffusion through a speech community (e.g. Ohala 1993, Joseph & Janda 2003). Broadly speaking, phonetic models tend to concentrate on identifying the perceptual and articulatory forces that provide the pre-conditions for sound change and that drive it in a particular direction (e.g. Ohala 1993). In order to identify these forces, experimental phonetic studies on sound change often factor out between-participant variability. On the other hand, individual speaker-listeners are crucial to the origins of sound change because, for sound change to occur, it is individual production and perception targets for speech sounds that must change. Or, as Milroy and Milroy (1985:347) noted "linguistic change must presumably originate in speakers rather than in languages". Recent phonetic studies have shown that people with similar linguistic backgrounds can differ in terms of their cognitive mapping between the auditory signal and perceptual categories (Beddor 2009, 2012) and in their production of speech sounds (Johnson 2006, Koenig et al. 2008). This article explores the idea that systematic differences between individual members of a speech community may play an important role in the early stages of sound change. Key importance has been given to heterogeneity between individuals in approaches to the diffusion of sound change. Labov's (1963) pioneering research with residents of Martha's Vineyard linked speaker attitude to participation in a sound change in progress and Milroy and Milroy's (1985) social network theory suggests certain individuals play a more crucial role in spreading sound change than others depending on their social position in the community. An ongoing challenge in sound change research is to link the initiation of sound change within individual cognitive grammars with the diffusion of novel variants through the community.

This paper first addresses the way that sound changes can originate in the everyday variability of spoken interactions in Section 2. Section 3 then focuses on the actuation of sound change and shows convergence between recent approaches on the idea that variability between individuals may be the key to understanding how some synchronic variation can become sound change. Expanding on this theme, Section 4 considers the evidence for individual differences based on results from four areas of phonetic research and the potential role of such individual differences in driving sound change.

2. SYNCHRONIC VARIATION AND THE ORIGINS OF SOUND CHANGE

Variability is an inherent part of the transmission of language between speakers and listeners, and can occur due to a range of linguistic and non-linguistic factors. To illustrate with just one example, local speakers of an Australian English variety typically pronounce *Melbourne* as something like [ˈmæᵊbm̩], with a vocalized or completely elided /l/ alongside reduction of the second unstressed syllable (see e.g. Cox and Palethorpe 2007 on Australian English). Although deviating from the spelling and causing some comment among visitors, this pronunciation is not unusual in terms of typical patterns of synchronic variation. /l/ is more prone to vocalization in syllable-coda position than elsewhere (e.g. Recasens 2012) and segments in unstressed syllables are more

likely to be reduced due to gestural overlap and blending (de Jong et al. 1993). Frequency of use also influences pronunciation (e.g. Bybee 2002), so that one might expect reduction processes to affect the local pronunciation of *Melbourne* but not e.g. *(peach) melba*, a phonologically similar but comparatively rare item for this particular speech community. The link with historical sound change is that synchronic tendencies in the way speech is produced and perceived can, over time, cause permanent categorical change (e.g. Beckman et al 1992, Hura et al. 1992, Hansson 2008, Harrington 2012). This can be seen by comparing the pronunciation of *Melbourne* with English words like *salmon* and *talk*, for example, in which /l/ is no longer pronounced, or by comparing standard Italian words (e.g. *caldo* 'hot') with cognates from other varieties also descended from Latin in which /l/ has been modified or lost e.g. *caudo, cardo, cada* (examples from Rohlfs 1966:342). The similar way that pre-consonantal /l/ is subject to change in unrelated languages may have its origin in phonetic bias factors that are common to spoken languages. These biases mean that the phonetic variation that provides the input to historical sound change is both non-random and directional (Garrett and Johnson 2013). In a recent detailed overview, Garrett & Johnson (2013) group these bias factors into four main areas: motor planning, aerodynamic constraints, gestural mechanics, and perceptual parsing.

Not all instances of synchronic phonetic variation lead to permanent sound change, of course. Instead, sound change can be seen as a two-step process of variation and selection (Ohala 1981, 1993; Lindblom et al. 1995) or, similarly, of channel and analytic biases (e.g. Moreton 2008). For Ohala, novel variants arise constantly in production, and their eventual selection for sound change depends entirely on listener perception. Lindblom et al. (1995) instead argue that language users - as listeners and speakers - evaluate novel forms according to articulatory as well as perceptual and other criteria and that selection happens when listeners pay close attention to *how* something

is being said and actively choose to reproduce that variant in their own speech.

Building on a long tradition of linking sound change to listener perception (e.g. Paul 1888 [1886], Baudouin de Courtenay 1972 [1895]), Ohala's model was the first to provide a framework to test the evident parallels between synchronic variation and diachronic change in the laboratory. In speech perception, listeners normally compensate for the contextual effects on segments due to coarticulation (e.g. Fowler 2005). For example listeners typically take account of context effects such as the more fronted tongue position in *loot* than in *loop* in mapping the auditory signal to their cognitive representation for /u/. Ohala (1981) suggests that on some rare occasions the listener may be unable for whatever reason to attribute the contextual variation (in this case /u/-fronting) to its source (the /t/ in *loot*). Here the listener would interpret the high F2 as an inherent part of the vowel /u/, would update their cognitive model with this new 'hypo-corrected' variant accordingly (which may then also be introduced into the listener's speech production). Ohala argues that mini-sound changes of this kind within an individual's grammar happen randomly and frequently (2012:23 note 2), and that only a sub-set of these may eventually proceed to become a sound change at the community level. Sound change is rare; this is because, Ohala argues, listeners are normally very good at adjusting for contextual variation.[1]

Sound changes like /u/-fronting that involve the uncoupling of coarticulatory variation from its source are most apparent in the case where the source of the coarticulatory effect is eventually lost. A familiar example is historical vowel nasalization in French (e.g. Italian *pane* v. French *pain*), a sound change that has been extensively studied by Beddor (2009, 2012). Beddor's experiments suggest

---

[1] Ohala (1993) nominates children and L2 learners as being more likely to make perceptual errors and, by extension, to drive sound change. It is unlikely that novel variants arising in this way could lead to sound change at the community level because native listeners would identify them as errors and correct them accordingly.

that the phonologization (Hyman 1976, 2013) of vowel nasalization comes about via a stage where coarticulatory information for nasalization on the vowel gradually becomes a sufficient cue for listeners, while the information provided by the source (the nasal consonant) becomes less important and may eventually be lost. Alternatively, Kirby (2013) describes how the waning of one cue can drive the enhancement and phonologization of another. Kirby models this relationship with an agent-based simulation of a sound change underway in Seoul Korean involving the phonologization of F0 and dephonologization of voicing.

Key to Ohala's model is the idea that many sound changes have a perceptual origin without involving a change to speech production. However as Beddor (2012:51) points out "perceptual grammars only contribute to sound change […] if they are publicly manifested". Browman & Goldstein (1991) suggested that interactions between perception and production over time at the level of the individual language user cause the shared target for a speech sound to move in a particular direction. The idea that sound change, at its origins, involves articulatory factors in combination with perceptual considerations is supported by the results of a recent physiological study on /l/-vocalization in American English. Lin et al. (2014) find that a small degree of articulatory reduction in the apical gesture for /l/ in words like *milk* and *help* (measured as tongue tip aperture on ultrasound displays) can have major acoustic consequences, causing the first two formants to merge. Lin et al. point out that this small articulatory shift in tongue aperture could make a quantal difference in perception and raise the idea that this might be a driving factor in causing /l/-vocalization. In other words some articulatory changes may only be slight on the part of the speaker, but nonetheless enough to cause the listener to reinterpret the sound category a la Ohala's model.

Ohala (2012: note 2) explicitly distinguishes initiation from actuation and chooses to account only for the former within his model. Indeed, bridging the gap between models of grammatical change at the level of the individual listener/speaker and change at the community level is difficult. [1] The articulatory and perceptual phonetic forces that bring about the pre-conditions for sound change and that explain its directionality are always present in the transmission of spoken language, yet sound systems are remarkably stable over time. How is it that sound change does not always happen whenever the necessary pre-conditions are met? What causes sound change actuation?

3. THE ACTUATION PROBLEM

The link between innovation within an individual's grammar, on the one hand, and widespread change for most members of a speech community on the other, is sound change actuation. There is no clear consensus in the sound change literature on the precise meaning of the term *actuation* and how it should be distinguished from *initiation* and *spread*. Ohala draws a distinction between the initiation of sound change versus its actuation/spread, whereas in Yu's account (2013:201) sound change actuation is seen as comprising variation and selection (i.e. as comprising initiation). The terms initiation and actuation are used synonymously in some sources, for example Hansson (2008:8) who refers to the "initiation (actuation) phase of sound change", whereas the terms actuation and spread are used synonymously in others. Baker et al. (2011:348), for example, suggest that a sound change is actuated when an individual (listener) reproduces a perceived novel variant in their own speech; they claim at this point "the sound change has begun to propagate around the community". Hansson and Baker et al. agree, then, that a sound

[1] Bybee (2012:221) suggests "a change does not have to start with […] just one person". Bybee argues that articulation, not perception, drives sound change and that all members of a speech community are biased towards the same kinds of articulatory reductions and casual speech processes. At this general level one could also argue that all listeners would make the same perceptual 'error'. In any case, articulatory ease cannot account for all kinds of sound changes (cf. e.g. Hansson 2008 on this point).

change is not only initiated but also actuated in the mind of the individual listener. Elsewhere in the literature this stage of listener-turned-speaker production is not considered to be sound change but rather to be an innovative "act of an individual speaker, regardless of whether or not it later catches on in a speech community" (Joseph and Janda 2003:17-18). Joseph and Janda argue that sound change should be "strictly defined as an innovation that has been widely adopted by members of […] a community" (cf. also Milroy and Milroy 1985).

Weinreich et al. (1968) famously identified the actuation problem by asking "what factors can account for the actuation of changes? Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?" For these authors and others since (e.g. Campbell 2013:193), actuation is key to explaining linguistic change because it relates to all other sociolinguistic and structural factors. Ohala (2002: note 2) is instead of the view that attempting to explain actuation, i.e. whether, where or when change may actually happen can be likened to "asking why a coin flip results in 'heads' and not 'tails'".

At least three recent studies have nevertheless addressed the actuation explicitly, namely Baker et al. (2011), Garrett and Johnson (2013) and Yu (2013). Baker et al. (2011) propose that actuation comes about via interactions between individuals whose targets for a speech sound fall at opposite ends of a production continuum, taking as a case study /s/-retraction in American English. Baker et al. classified speakers as "retractors" and "non-retractors" according to the proportion of their tokens that were judged (by the authors) to be retracted or not (62.50% and 35.48%, respectively). Baker et al. show that even speakers classified as "non-retractors" (i.e. those who pronounce e.g. *stream* more like [stɹ] than [ʃtɹ]) show a lower centroid frequency for /s/ in clusters, and especially so for clusters containing /r/. As such they argue that the phonetic motivation for /s/-retraction (assimilation to retroflex /r/ across the intervening /t/) can be considered to be generally present amongst English speakers. Given that /s/-retraction has been phonologized in some (e.g. New Zealand English) but not all varieties, the difficulty, as Baker et al. (2011:350) point out "lies in creating a theory in which sound change can plausibly occur, but without making sound change inevitable". Baker et al. (2011) propose that variability becomes sound change, i.e. a sound change is actuated, because individuals differ in the degree to which they coarticulate (retract). Speakers who do not tend to retract /s/ will interpret [ʃ]-like productions as a distinct target, and might imitate them accordingly. Crucial to this model is the idea that coarticulatory effects must be sufficiently perceptible if they are to be imitated by other members of the speech community (i.e. ambiguous tokens would be unlikely to be imitated because "if every speaker has essentially the same amount of coarticulation, then there is nothing to imitate in the first place" (2011:350)). This model of Baker et al. also generally predicts that coarticulatory effects that vary across speakers would be more likely to undergo sound change than those common to all speakers within a speech community. Baker et al. present only production data but it would be possible to test these predictions about the role of inter-speaker differences in sound change actuation with standard shadowing experiments. For example, one could recruit a group of speakers from a variety of English without /s/-retraction. Following Baker et al. these speakers could be classified as "retractors"/"non-retractors" based on the proportion of *str-* productions judged to be [ʃ] or [s]. After exposure to words containing extreme [ʃ]-like tokens, only those speakers whose own productions typically fall at the [s]-end of the continuum should show a change in the direction of [ʃ], whereas individuals with ambiguous productions between [ʃ] and [s] should not shift their production target.

Garrett and Johnson (2013) propose that some individuals are more likely to attach social meaning to linguistic differences than

others and that this is a driving force in the actuation of sound change. They draw on experimental work by Dimov (2010; also Dimov et al. 2012) that found a link between social and personality traits and the extent to which participants compensated for altered auditory feedback (for /u/). Dimov and colleagues' experimental work showed that the more powerful subjects judged themselves to be, based on responses to survey questionnaires, the less they compensated in the experimental task (i.e. they were either less finely attuned to phonetic variation or less willing to modify their own production in order to compensate for it). Garrett and Johnson (2013) build on this experimental research together with Giles et al.'s (1991) work on accommodation to suggest that individuals who wish to identify with a group may be more likely to interpret intrinsic phonetic variability (primarily due to coarticulation) as indexing group membership. As a result they would attach a social significance to phonetic properties that previously had none. An exemplar-based model predicts that if listeners attach social meaning to coarticulatory information then they will store both components i.e. the coarticulatory information together with the social meaning that it is assumed to convey. On the other hand, listeners who compensate for coarticulation would discard this contextual information before storing the exemplar in their cognitive grammar. This model therefore predicts that individuals who attach social significance to coarticulatory information are more likely to participate in sound change than those individuals who do not, and whose exemplar clouds (and pronunciations) should remain stable over time. Garrett and Johnson (2013) simulate their model with two groups of autonomous agents who were exposed to phonetic tokens for /z/ whose variants included a small number that were affected by an articulatory bias so that they sounded more like an approximant /r/. One group was modelled to compensate for the articulatory bias, essentially discarding the novel approximant variants, whereas the other group did not, and was thus intended to represent individuals with more sensitivity to social differences. After more than fifty iterations the two groups' productions diverged according to how they responded to the novel variants; the group that did not compensate for the articulatory bias came to produce /r/-like tokens in their own output. In this way Garrett and Johnson model the update of novel variants according to whether or not they are cognitively stored in individuals.

Yu (2013) addresses the role of individual cognitive processing style in the actuation of sound change. He compares the categorization of CV stimuli (where C is an /s…ʃ/ continuum and V = /i, u/) with personality and social traits as measured by a number of standard questionnaires. Yu reports that neurotypical listeners with fewer autistic traits (more specifically a lower Autism Quotient (AQ)) are less likely to link coarticulatory information during the fricative to its source in the following /u/ (i.e. more "ʃu" responses) than listeners with a high AQ who tend to compensate for context (and give more "su" responses accordingly). Yu argues that by failing to compensate for coarticulation, participants with a low AQ may be responsible for the creation of novel variants. Yu then relates AQ (and its sub-components of attention and social skills) to a number of personality and social traits. Ultimately Yu suggests that the same individuals who are likely to undergo mini-sound changes might also be more likely, due to their (more extroverted and agreeable) personality and social profiles, to spread such innovations within their social networks.

These three approaches together suggest that the selection of sound changes from a pool of synchronic variation depends on differences between individual members that make up a speech community. These differences are not due to chance but rather involve factors that are identifiable and generalizable to other groups of language users. For Baker et al. the systematic difference lies in production, whereas for Garrett and Johnson it involves sensitivity to social factors and for Yu it involves cognitive and social traits. Common to Garrett and Johnson (2013) and Yu (2013)

is the way that individual listener interpretation provides the catalyst for sound change and speaker productions do not change - at least initially. Differences across individual speaker productions are instead crucial to Baker et al.'s (2011) model, and can only subsequently be perceived and possibly imitated by listeners. This echoes Lindblom et al. (1995:16) who also suggested that there must be "significant change in the phonetic pattern" for a variant to be noticed by listeners and to eventually undergo sound change.

The role of phonetic similarity in driving sound change actuation deserves further experimental investigation. In contrast with Baker et al., Garrett and Johnson hypothesize that slight phonetic differences in production may be *more* likely to lead to sound change than larger differences, because they would not be detected by listeners and would therefore be included amongst stored exemplars. Garrett and Johnson hypothesize that more dramatic phonetic differences due to e.g. production errors (here one could also consider Baker et al.'s exaggerated coarticulations) would not be automatically stored in a listener's representation and would need to take on a socio-indexical meaning to participate in sound change. Evidence from the imitation literature (described in more detail below) favours the idea that phonetically similar variants would be more likely to undergo sound change. For example in spontaneous conversations, Kim et al. (2011) report more imitation between speaker pairs who both shared the same dialect background than between pairs with different linguistic backgrounds. Olmstead et al. (2013) report that native Spanish and English listeners' imitation of an 11-step [ba]–[pa] continuum showed less convergence outside the bounds of their native pronunciation range (i.e. English listeners converged less in the prevoiced region and Spanish listeners converged less for the long lag region).

4. SYSTEMATIC INDIVIDUAL DIFFERENCES AND SOUND CHANGE

Labov (2006:508) has argued that we should not seek to code linguistic variation between individuals for its own sake. He points out that "some further justification for the description of variation is required; otherwise there will be no stop to the enterprise and we will be plunged into an endless pursuit of detail". Labov goes on to point out that variation is crucial to language change, in particular, and indeed the further justification here is that we do not yet have a model of how sound change is actuated. That is, we need to link the initiation of sound change in an individual's cognitive grammar and widespread change at the group level. The approaches outlined above suggest that actuation is dependent on variation between individuals that make up speech communities. An ongoing challenge is to identify the factors responsible for variability within groups of individuals who interact on a daily basis.

4.1. Speech production differences

Speech production is idiosyncratic and differences between speakers can be attributed to learned behaviour as well as to physiology (Ladefoged and Broadbent 1957, Johnson et al. 1993, Koenig et al. 2008). There are reports in the phonetic literature of speaker-specific strategies for achieving articulatory goals for stable phonemic categories (Koenig et al. 2008 on fricatives and Beddor 2009 on nasals, both for American English). Synchronic lenition is also reported to be speaker-specific. For example lenition of /p t k/ in Florentine and other varieties of Italian spoken in Tuscany is most prominent for velar /k/, which is typically reduced to /h/ and can be elided altogether. Yet there is evidence (e.g. Dalcher 2008) that certain speakers resist /k/-reduction. Dalcher (2008) attributes this variability to external social factors, in this case to the extent to which individual speakers identified positively with being "Florentine" and therefore chose to lenite /k/ (echoing Labov's 1963 seminal study of speaker attitude and vowel centralization in Martha's vineyard). In contrast to the *Gorgia toscana* which is a stereotypical feature of Tuscan speech (e.g. Giannelli 1997, Bertinetto & Loporcaro 2005), lenition of /b d g/ in contemporary spoken Danish does not appear to carry any social meaning but does

nonetheless also show speaker-specific patterns in terms of place of articulation (Pharao 2011). Pharao analysed group-level patterns of reduction of /b d g/ using mixed effect models with individual speaker as a random factor, the results of which suggest reduction of /b/ and /g/ (but not /d/) belong to the same target undershoot process. However six of the 22 speaker participants did not conform to the group-level pattern, and instead showed divergent tendencies for reduction of /b/ and /d/. Both of these studies illustrate that lenition processes are not automatic, because otherwise individual speakers would show similar place-governed patterns.

Solé (2014) makes an explicit link between fine-grained inter-speaker differences in production and the origins of sound change. Using oral and glottal airflow measurements, Solé shows that some Spanish speakers show nasal airflow leakage in the production of voiced stops /b d g/, which serves to enhance the voiced status of /b d g/ by reducing supraglottal pressure and facilitating vocal fold vibration. Solé describes nasal airflow leakage in this context as an implementational feature and argues that such features can be distinguished from other kinds of phonetic variation because they are planned (rather than mechanical) on the part of the speaker who intends to produce a specific acoustic effect and because only some speakers use them in some contexts (i.e. they are not fully predictable). Notably, Solé's perceptual analysis shows that listeners have difficulty parsing nasal airflow leakage with the source, and can interpret it as a separate nasal segment. Solé proposes that low-level implementational features are more likely to undergo sound change not only because they are difficult for listeners to parse but precisely because they vary across speakers.

4.2. Perception and cognitive processing style

Functional and anatomical differences between human listeners' peripheral auditory systems are not normally considered to affect the long-term stability of shared sound systems. Johnson (2004:26), for example, points out "there is no reason to expect psychophysical thresholds for simple or complex stimuli to vary from language to language". However, since there is evidence that people with similar linguistic backgrounds can differ in the way that they hear and process auditory signals (including relatively simple tones as well as speech sounds) it is conceivable that such differences could play a role in sound change. Beddor (2012), for example, suggests that sound change can arise out of the idiosyncratic way in which coarticulation is perceived. Along the lines of Milroy and Milroy's (1985) notion of an innovative speaker, Beddor (2012:51) describes the innovative listener as one who comes to map the auditory input to abstract categories in a novel way that, if this is matched in production, could drive sound change.

There is an auditory illusion whereby human listeners normally hear a tone that is interrupted by a noise-filled gap as continuous, despite the intervening noise. This auditory illusion is due to perceptual restoration, an adaptive skill developed in late childhood (Warren & Warren 1971) and crucial to processing auditory information - including speech - in noisy environments. Yet recent evidence shows that individuals vary in the extent to which they experience this illusion: Vinnik et al. (2011) found that nearly one quarter of their 46 participants reported hearing an interrupted tone signal as discontinuous. This result shows that even for relatively simple auditory tasks listeners differ in the extent to which they weight the auditory signal against information from top-down perceptual restoration processes.

Speech perception is affected by a listener's native phonological system, which influences their ability to detect speech sounds (Mielke 2003) and to categorize them (e.g. Bohn et al. 2011, Davidson 2011). However, even within groups of people with similar language backgrounds, speech perception can differ from person to person. Beddor's (2009, 2012) well-known research reported idiosyncratic behaviour in the perception of nasalization in VNC sequences in American English, whereby listeners differed in their sensitivity to and

weighting of fine phonetic detail. Beddor suggests that listeners who compensate less for coarticulation could be more likely to initiate sound change (we saw earlier that Yu relates compensation for coarticulation to a person's Autistic Quotient).

Some recent research in our lab has also looked at individual differences in perception and the extent to which such differences might initiate sound change. Following reports that geminate /p: t: k:/ can be optionally produced with pre-aspiration in contemporary spoken Italian (e.g. Stevens 2012) Stevens & Reubold (submitted) investigated the impact of pre-aspiration on native listener perception of phonemic consonant length. Two continua were synthesized, one from short *fato* 'fate' to long *fatto* 'done' and the other in which a portion of the closure duration for the dental stop was replaced by pre-aspiration. The results of a forced-choice perception experiment (*n* participants =16) showed significantly more *fato* responses for the pre-aspirated continuum. Most listeners conformed to this group-level pattern but two showed no difference between the pre-aspirated and plain continua. In other words, these two participants parsed pre-aspiration with the consonant (= /t:/), whereas all other listeners parsed it with the vowel (= /t/). Since there are two different parsing strategies for pre-aspiration, then the type of resulting listener-driven sound change cannot be predicted (whereas perceptual confusion is typically asymmetrical, e.g. Garrett & Johnson 2013). Notably however, the two different parsing strategies correspond to dialect differences for pre-aspiration in Swedish (Wretling et al. 2003), suggesting that sound changes involving pre-aspiration might be directly influenced by individual perceptual patterns. Ohala's listener-driven model of sound change has been criticised on the assumption that all listeners must eventually make the same perceptual error (Baker et al. 2011, Bybee 2012). However, perhaps it is not necessary to assume that all listeners make the same error. Rather, sound change could be driven by interactions between listeners with different parsing strategies that would serve to weaken phonological category boundaries over time.

4.3. The perception-production link

Ohala's model of sound change initiation (and that of e.g. Baker et al. 2011) is implicitly founded on a direct relationship between perception and production in the sense that listeners turned speakers would reproduce novel perceptual targets in their own subsequent productions. However the experimental evidence of a direct link between perception and production at the level of the individual language user is mixed. Beddor (2009) compared results for one participant across perception and production tasks and found that they were aligned: the presence/absence of a nasal consonant was poorly discriminated in perception (in e.g. *bed* v. *bent*), and this participant also showed relatively more variability in production. In a larger-scale comparison involving nineteen participants, Perkell et al. (2004) report that an individual's ability to discriminate vowel phoneme pairs (e.g. *who'd* v. *hood*) in perception could predict the acoustic separation of these same vowels in production. There is also evidence of a close parallel between perception and production for sub-groups of listeners who differ in their linguistic experience due to age (Harrington et al. 2008) and socio-economic background (Hay et al. 2006). On the other hand Kataoka (2011) reports that for /u/-fronting in American English the extent to which a participant compensated for coarticulation in perception was not correlated with the degree of coarticulation in that same subject's productions. Stevens and Reubold (submitted) also compared the perception of pre-aspiration (described above) with the production of geminate /t:/ within each participant. Six subjects parsed pre-aspiration with the preceding vowel in perception and in production, but seven subjects showed a mismatch across the two experimental tasks (e.g. assigning pre-aspiration to the preceding vowel /at/ in production but to the consonant /at:/ in perception). This shows that sound change could originate not only form

idiosyncratic perceptual parsing strategies (in line with Beddor 2009, 2012) but also because not all subjects align their perception and production of coarticulation in the same way. Based on data from 28 participants, Grosvald & Corina (2012) also found that the perception and production of long-distance vowel coarticulation were not correlated at the level of the individual. Grosvald and Corina's data showed that subjects who were especially sensitive to long-distance coarticulatory effects on schwa in perception did not tend to produce more coarticulation in their own speech. This experimental result (together with those of Kataoka 2011 and Stevens and Reubold, submitted) appears to cast some doubt on models of sound change that assume that an individual listener would match novel perceptual categories in their own productions (e.g. Ohala 1993, Baker et al. 2011). Here Grosvald & Corina (2012:96) suggest a more nuanced interpretation whereby "it does not matter if the speech community at large exhibits a perception-production correlation or not" but rather that some small proportion of the community does. According to these authors, it is only this small proportion of the community that would be (a) especially sensitive to novel variants in perception and (b) likely to match or exaggerate these variants in production (Grosvald and Corina point out that for their experimental data, only one subject appears to fall into this category). The notion that the strength of the link between perception and production might vary between individuals and that very few individuals would meet both criterion (a) and criterion (b) is tantalizing because, taken together with the idea of sound change as a process of variation and selection, it implies that only a small portion of the speech community would be able to select sound changes from the pool of synchronic variation. This would help to explain why sound change is so rarely actuated even though the phonetic pre-conditions for sound change are constantly being generated in spoken language interactions.

4.4. Linguistic experience over a lifetime and imitation

Linguistic experience accumulated over an individual's lifetime is another potential source of sound change. A person's speech can change to reflect ongoing change taking place in the wider community, not just during the earliest phases of language acquisition but over the lifespan (Harrington 2007, Sankoff & Blondeau 2007). Sankoff and Blondeau (2007:584) examined /r/ production in Montreal French over a thirteen year period and showed that a minority of speakers changed their production. Some altered only the frequency of the two (apical or dorsal) variants but others replaced the apical variant "that they appeared to use spontaneously and unreflectingly with the innovative [… dorsal variant] characteristic of speakers younger than themselves". Such change over the lifespan is understood to come about because cognitive models are the result of statistical generalizations over linguistic experiences and are constantly being updated (e.g. Pierrehumbert 2003). Because no two people can take part in exactly the same conversations and have exactly the same linguistic experiences, everyone's cognitive model must be uniquely different. The effect of linguistic experience can be seen in speech processing, for example, which is affected by familiarity with different dialects (Clopper 2014). Moreover social knowledge and expectations about a speaker's social attributes affect the perceptual categorization of speech sounds (Hay, Warren and Drager 2006). Non-linguistic experiences such musical training can also affect the perception and production of linguistic contrasts (cf. e.g. Yu 2013:204 and references therein).

While individuals *can* update their sound categories, longitudinal studies (Sankoff & Blondeau 2007, Kammacher et al. 2011) show that not all individuals participate in the sound changes taking place around them. This brings us to imitation (or accommodation), which refers to the way that an individual's speech can come to resemble that of their interlocutor. Imitation has been documented in experimental tasks involving modified single

word tokens (Nielsen 2011) and between individuals who interacted over longer time periods (Pardo et al. 2012). Imitation is understood to be one of the factors by which a sound change can spread through a community; it is also thought to play a role in first language acquisition and dialect convergence (Trudgill 1986).

Empirical studies show that imitation is not automatic but rather that it is constrained by linguistic factors (e.g. Nielsen 2011) as well as social preferences (Babel 2012) including self-reported closeness to the interlocutor (Pardo et al. 2012) and novelty of the interlocutor's voice (Babel et al. 2014 for gender-atypical voices). Yu et al. (2013) report variability between subjects in the extent of imitation after exposure to a narrative containing extended VOT durations. Measured in terms of acoustic VOT, some individuals converged towards while others diverged from the narrator in ways that were found to depend on attitude and social/personality factors. Overall however, Yu et al. found no effect of imitation at the group level. Nielsen (2011) on the other hand, did report an overall effect of imitation after exposure to increased VOT in isolated words. Yu et al. (2013:11) note that "imitation might be more automatic […] in a context where the words are presented in isolation devoid of social significance" whereas their study allowed participants to make evaluative judgements about the speaker during a narrative. This observation is reminiscent of Lindblom et al.'s (1995) notion of bimodal listener perception, noted in §2 earlier, and the suggestion that the 'how' mode of listening (and not the 'what' mode) would feed sound change. The fact that imitation is more typical in tasks involving single word items than after exposure to narratives supports Lindblom et al.'s notion that imitation, and by extension sound change, might happen when listeners pay particular attention to *how* something is being said. Pardo et al.'s (2012) study involving college roommates shows that imitation also happens in more natural settings (and indeed it must, to play a role in sound change). All five participant-pairs in Pardo et al.'s study converged over an academic year,

as judged globally by naive listeners in an AXB task, but individuals varied in the degree to which they did so. Indeed Pardo et al. (2012:196) emphasize the complexity of imitation and raise the idea that "each individual talker might converge on a unique set of acoustic-phonetic attributes while diverging, varying randomly, or remaining neutral on others".

5. FINAL COMMENTS

While phonetic research on the origins of sound change has tended to focus on group-level biases in speech production and perception, it appears that sound change should be seen as the result of interactions between individuals who have slightly different cognitive mappings, perceptual abilities, production strategies or sensitivities to particular social factors. Identifying these factors and the range of variability between individuals who make up a speech community is vital to understanding what causes certain sound categories, but not others, to become unstable over time.

6. ACKNOWLEDGMENTS

7. REFERENCES

Babel, M., McGuire, G., Walters, S., & Nicholls, A. (2014). Novelty and social preference in phonetic accommodation. Laboratory Phonology, 5(1), 123-150.

Baker, A., Archangeli, D., & Mielke, J. (2011). Variability in American English s-retraction suggests a solution to the actuation problem. Language variation and change, 23, 347-374.

Baudouin de Courtenay, J. (1972 [1895]). An attempt at a theory of phonetic alternations. In E. Stankiewicz (Ed.), A Baudouin de Courtenay anthology: the beginnings of

structural linguistics (pp. 144-212). Bloomington: Indiana University Press.

Beckman, M., De Jong, K., Jun, S.-A., & Lee, S.-H. (1992). The interaction of coarticulation and prosody in sound change. Language and Speech, 35, 45–58.

Beddor, P. S. (2009). A coarticulatory path to sound change. Language, 85(4), 407-428.

Beddor, P. S. (2012). Perception grammars and sound change. In M.-J. Solé & D. Recasens (Eds.), The initiation of sound change: perception, production, and social factors (pp. 37-55). Amsterdam/Philadelphia: John Benjamins.

Bertinetto, P. M., & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. Journal of the International Phonetic Association, 35(2), 131-151.

Bohn, O.-S., Best, C. T., Avesani, C., & Vayra, M. (2011). Perceiving through the lens of native phonetics: Italian and Danish listeners' perception of English consonant contrasts. Proc. ICPhS XVII. Hong Kong: 336-339.

Browman, C., & Goldstein, L. (1991). Gestural structures: distinctiveness, phonological processes, and historical change. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), Modularity and the motor theory of speech perception: Proceedings of a conference to honor Alvin M. Liberman (pp. 313-338). New Jersey: Lawrence Erlbaum Associates, Inc.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. Language variation and change, 14, 261–290.

Bybee, J. (2012). Patterns of lexical diffusion and articulatory motivation for sound change. In M.-J. Solé & D. Recasens (Eds.), The initiation of sound change. Perception, production and social factors (pp. 211-234). Amsterdam/Philadelphia: John Benjamins.

Campbell, L. (2013). Historical linguistics: an introduction (Third Edition). Edinburgh: Edinburgh University Press.

Clopper, C. G. (2014). Sound change in the individual: Effects of exposure on cross-dialect speech processing. Laboratory Phonology, 5(1), 69-90.

Cox, F., & Palethorpe, S. (2007). Australian English. Journal of the International Phonetic Association, 37(3), 341-350.

Dalcher, C. V. (2008). Consonant weakening in Florentine Italian: a cross-disciplinary approach to gradient and variable sound change. Language variation and change, 20, 275-316.

Davidson, L. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. Journal of Experimental Psychology: Human Perception and Performance 37(1), 270-282.

de Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. Language and Speech, 36(2-3), 197-212.

Dimov, S. (2010). Social and Personality Variables in Compensation for Altered Auditory Feedback. UC Berkeley Phonology Lab Annual Report, 259–282.

Dimov, S., Katseff, S. & Johnson, K. (2012). Social and Personality Variables in Compensation for Altered Auditory Feedback. In M.-J. Solé & D. Recasens (Eds.), The initiation of sound change: Perception, production, and social factors (pp. 185-210). Amsterdam & Philadelphia: John Benjamins.

Fowler, C. A. (2005). Parsing coarticulated speech in perception. Journal of Phonetics, 33, 199-213.

Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. In A. Yu (Ed.), Origins of sound change: Approaches to phonologization (pp. 51-97). Oxford: Oxford University Press.

Giannelli, L. (1997). Tuscany. In M. Maiden & M. Parry (Eds.), The Dialects of Italy (pp. 297-302). New York: Routledge.

Grosvald, M., & Corina, D. (2012). The production and perception of sub-phonemic vowel contrasts and the role of the listener in sound change. In M.-J. Solé & D. Recasens (Eds.), The Initiation of Sound Change: Perception, production, and social factors (pp. 77–100). Amsterdam/Philadelphia: John Benjamins.

Hansson, G. Ó. (2008). Diachronic explanations of sound patterns. Language & Linguistics Compass, 2(5), 859-893.

Harrington, J. (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. . In J. Cole & J. Hualde (Eds.), Laboratory Phonology 9 (pp. 125-143). Berlin: Mouton.

Harrington, J. (2012). The relationship between synchronic variation and diachronic change. In A. Cohn, C. Fougeron & M. Huffman (Eds.), Handbook of Laboratory Phonology (pp. 321 - 332). Oxford: Oxford University Press.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. Journal of the Acoustical Society of America, 123(5), 2825-2835.

Harrington, J., & Stevens, M. (2014). Editors' introduction: Cognitive processing as a bridge between phonetic and social models of sound change. Laboratory Phonology, 5(1), 1-8.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. Journal of Phonetics, 34, 458–484.

Hura, S. L., Lindblom, B., & Diehl, R. L. (1992). On the role of perception in shaping phonological assimilation rules. Language and Speech, 35(1,2), 59-72.

Hyman, L. M. (1976). Phonologization. In A. Juilland (Ed.), Linguistic studies presented to Joseph H. Greenberg (pp. 407-418). Saratoga: Anma Libri.

Hyman, L. M. (2013). Enlarging the scope of phonologization. In A. C. L. Yu (Ed.), Origins of Sound Change: Approaches to Phonologization (pp. 3-28). Oxford: Oxford University Press.

Janda, R. D., & Joseph, B. D. (2003). On Language, Change and Language Change. In R. D. Janda & B. D. Joseph (Eds.), The Handbook of Historical Linguistics (pp. 3-180): Wiley-Blackwell.

Johnson, K. (2004). Cross-linguistic Perceptual Differences Emerge from the Lexicon. In A. Agwuele, W. Warren & S.-H. Park (Eds.), Proceedings of the 2003 Texas Linguistics Society Conference (pp. 26-41). Somerville, MA: Cascadilla Proceedings Project.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. Journal of Phonetics 34, 485–499.

Kammacher, L., Stæhr, A., & Jørgensen, J. N. (2011). Attitudinal and sociostructural factors and their role in dialect change: Testing a model of subjective factors. Language Variation and Change, 23, 87–104.

Kataoka, R. (2011). Phonetic and Cognitive Bases of Sound Change. PhD Dissertation. UC Berkeley.

Kim, M., Horton, W. S. & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. Laboratory Phonology 2(1). 125-156.

Kirby, J. (2013). The role of probabilistic enhancement in phonologization. In In A. C. L. Yu (Ed.), Origin of Sound Change: Approaches to Phonologization. Oxford: OUP.

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. Journal of the Acoustical Society of America, 124(5), 3158-3170.

Labov, W. (1963). The social motivation of a sound change. Word, 19, 273-309.

Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. Journal of Phonetics 34, 500–515.

Ladefoged, P., & Broadbent, D. (1957). Information conveyed by vowels. Journal of the Acoustical Society of America, 29(1), 98-104.

Lin, S., Beddor, P. S., & Coetzee, A. W. (2014). Gestural reduction, lexical frequency, and sound change: a study of post-vocalic /l/. Laboratory Phonology, 5(1), 9-36.

Lindblom, B., Guion, S., Hura, S, Moon, S.-J., & Willerman, R. (1995). Is sound change adaptive? Rivista di linguistica, 7(1), 5-37.

Mielke, J. (2003). The interplay of speech perception and phonology: experimental

evidence from Turkish. Phonetica, 60, 208-229.

Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. Journal of Linguistics, 21, 339-384.

Moreton, E. (2008). Analytic Bias as a Factor in Phonological Typology. In C. B. Chang & H. J. Haynie (Eds.), Proceedings of the 26th West Coast Conference on Formal Linguistics (pp. 393-401). Somerville, MA: Cascadilla Proceedings Project.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. Journal of Phonetics, 39(2), 132–142.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (Ed.), Historical Linguistics: Problems and Perspectives (pp. 237-278). London: Longman.

Ohala, J. (2012). The listener as a source of sound change. An Update. In M.-J. Solé & D. Recasens (Eds.), The initiation of sound change. Perception, production and social factors (pp. 21-36). Amsterdam/Philadelphia: John Benjamins.

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

Olmstead, A. J., Viswanathan, N., Aivar, M. P. & Manuel, S. (2013). Comparison of native and non-native phone imitation by English and Spanish speakers. *Frontiers in Psychology* 4, 1-7.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. Journal of Phonetics, 40, 190-197.

Paul, H. (1888 [1886]). Principles of the history of language. London: Swan Sonnenschein, Lowrey & Co.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., et al. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. J. Acoust. Soc. Am, 116, 2338.

Pharao, N. (2011). Plosive reduction at the group level and in the individual speaker. ICPhS XVI, Hong Kong: 1590-1593.

Pierrehumbert, J. B. (2003). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. Language and Speech, 46(2-3), 115-154.

Recasens, D. (2012). A phonetic interpretation of changes affecting dark /l/ in Romance. In M.-J. Solé & D. Recasens (Eds.), The Initiation of Sound Change: Perception, production, and social factors (pp. 57–76). Amsterdam/Philadelphia: John Benjamins.

Recasens, D., Sánchez Miret, F., & Wireback, K. (2010). Experimental Phonetics and Sound Change. München: Lincom Europa.

Rohlfs, G. (1966). Grammatica storica della lingua italiana e dei suoi dialetti: fonetica (Vol. 148). Pisa: Einaudi.

Sánchez Miret, F., & Recasens, D. (2013). Studies in Phonetics, Phonology and Sound Change in Romance. München: Lincom Europa.

Sankoff, G., & Blondeau, H. (2007). Language change across the lifespan: /r/ in Montreal French. Language, 83(3), 560-588.

Solé, M.-J. (2014). The perception of voice-initiating gestures. Laboratory Phonology, 5(1), 37-69.

Solé, M.-J., & Recasens, D. (2012). The initiation of sound change. Perception, production and social factors. Amsterdam/Philadelphia: John Benjamins.

Stevens, M. (2012). A phonetic investigation into "Raddoppiamento sintattico" in Sienese Italian Speech. Bern: Peter Lang.

Stevens, M., & Reubold, U. (submitted manuscript). The parsing of pre--aspiration in perception and production: implications for sound change.

Vinnik, E., Itskov, P. M., & Balaban, E. (2011). Individual Differences in Sound-in-Noise Perception Are Related to the Strength of Short-Latency Neural Responses to Noise. PLoS ONE, 6(2), 1-8.

Warren, R. M., & Warren, R. P. (1971). Some age differences in auditory perception. Bulletin of the New York Academy of Medicine 47(11), 1365–1377.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (Eds.), Directions for historical linguistics (pp. 95-198). Austin: University of Texas Press.

Wretling, P., Strangert, E., & Schaeffler, F. (2003). Preaspiration as a quantity feature. Proc.15th ICPhS, Barcelona: 2701-2704

Yu, A. C. L. (2013). Individual differences in socio-cognitive processing and the actuation of sound change. In A. C. L. Yu (Ed.), Origins of Sound Change: Approaches to Phonologization. Oxford: Oxford University Press. 201-227.

Yu, A. C. L., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic Imitation from an Individual-Difference Perspective: Subjective Attitude, Personality and "Autistic" Traits. PLoS ONE, 8(9), 1-13.