The Role of Experimental Investigation in Understanding Sound Change

Alan C. L. Yu

The Oxford Handbook of Historical Phonology
Edited by Patrick Honeybone and Joseph Salmons

Print Publication Date: Nov 2015
Subject: Linguistics, Historical Linguistics, Phonetics and Phonology
Online Publication Date: May 2015 DOI: 10.1093/oxfordhb/9780199232819.013.010

Abstract and Keywords

Experimental methodologies and laboratory techniques have been employed to understand the problem of actuation, that is, the inception of phonological change. This chapter reviews major findings in experimental historical phonology and explore models of the speech production and perception that have been used to account for these findings.

Keywords: sound change, experiment, exemplar-based models, acoustics, speech perception, perceptual compensation

24.1 Introduction

> When we speak of the systematic effect of sound laws we can only mean that given the same sound change within the same dialect every individual case in which the same phonetic conditions are present will be handled the same. Therefore either wherever earlier the same sound stood, also in the later stages the same sound is found or, where a split into different sounds has taken place, then a specific cause—a cause of a purely phonetic nature like the effects of surrounding sounds, accent, syllabic position, etc.—should be provided to account for why in the one case this sound, in the other that one has come into being.
>
> (Paul 1880: 86; page numbers are from the 2nd edition of 1886)

The neogrammarian position on sound change, as summarized in Hermann Paul's quote above, consists of two assertions: sound change is *regular* and *purely phonetically conditioned*. While there has been widespread recognition that diachronic (and synchronic) phonological phenomena can be attributed to phonetic factors (Whitney 1867, Verner 1877, Paul 1880, Sweet 1888, Jespersen 1922, Hill 1936, Jakobson 1941, Martinet 1952, 1955, Baudouin de Courtenay 1895[1972]), the experimentalist's project on sound change, nonetheless, did not get underway in earnest until relatively recently. As Bloom-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

field remarked, '[a]lthough many sound-changes shorten linguistic forms, simplify the phonetic system, or in some other way lessen the labor of utterance, yet no student has succeeded in establishing a correlation between sound-change and any antecedent phenomenon: the (p. 411) causes of sound-change are unknown' (Bloomfield 1933: 385). As techniques and methodology in phonetic research become more sophisticated, systematic investigations into variabilities and biases in speech production and, more recently, speech perception offer more resources to investigate the relationship between diachronic sound changes and synchronic variation in speech, even though the impetus to apply experimental methodologies to the investigation of historical events is not always immediately obvious (and the interpretation of the effects of phonetic biases controversial). To be sure, the need to check the validity of any hypothesis concerning the cause of a particular sound change by testing the hypothesis in the laboratory is clear. As Ohala explains, '[i]f particular sound changes are posited to have a phonetic basis then one should be able duplicate the conditions under which they occurred historically and find experimental subjects producing 'mini' sound changes that parallel them. It is because of the posited phonetic character of sound change that a laboratory study is possible: were the initiation caused by grammatical and cultural factors, this would be more difficult or perhaps impossible' (Ohala 1993: 261). The question is what linking theory is needed to connect experimental findings to historical changes that took place in the past and where information regarding the precise conditions that gave rise to the specific changes are no longer available.

One important principle that is assumed, implicitly or explicitly, by all laboratory investigators of sound change is the principle of uniformitarianism (see Murray, this volume). As it applies in linguistics, uniformitarianism asserts that the same processes that operate in language now must have always operated in language in the past. That is, the same laws apply today as in the past. This assumption is crucial for any attempt to link experimental findings to causes of sound change in the past since there is no necessary *a priori* reason to think that observations made in the laboratory would have any bearing on our understanding of the past. While the uniformitarian principle supplies the larger conceptual framework, specific hypotheses about how phonetic conditions shape the emergence of new sounds and sound patterns are also needed. Hypotheses are generated from close examination of explicit theories. To this end, this chapter begins with an overview of theories of the phonetic origins of sound change in Section 24.2, wherein some controversy lies. While theorists differ in terms of their emphasis on which party of the communicative dyads (the speaker *vs* the listener) contributes most to the initiation of change, ultimately, approaches to sound change typically recognize that a major source of variation comes from constraints imposed by the human speech production and perceptual system. As Ohala (1993) argues, common sound changes attested independently in substantially the same form in unrelated languages are generally more likely to be the results of language universal factors such as the physics and physiology of the vocal tract and the nature of the human perceptual system. Thus, Section 24.3 reviews systematic sources of phonetic variation (phonetic biases) that have been associated with common sound changes and the experimental methodologies that uncover them. Section 24.4 discusses

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

recalcitrant problems, including how new variants are registered in the phonetic memory of individuals and how such variants propagate across the speech community, and suggests possible directions for future investigations. Some concluding remarks appear in Section 24.5. The approach taken here fits well with (p. 412) that of several other chapters in this volume, such as Jones (this volume) and Blevins (this volume); other chapters take different approaches, such as Dresher (this volume) and Holt (this volume).

24.2 Theories of the Phonetic Origins of Sound Change

One of the central questions in sound change is the so-called actuation problem (Weinreich et al. 1968). That is, why does a change occur in a particular manner in a particular time and space? This chapter focuses primarily on how variants come about, although some discussion regarding the diffusion of linguistic variants appears in Section 24.4.4. This section reviews two approaches to the phonetic origins of sound change.

24.2.1 'Ease of Articulation' and Sound Change

Most early theories of sound change appealed to the so-called 'ease of articulation' hypothesis to explain the source of new sounds and sound patterns (see Bybee, this volume). Phonetic precursors to sound change under this view consist of the interplay between the articulatory dynamics of speech production and the functional motivation of 'economy' (e.g. Paul 1880, Sievers 1901). To be sure, a simplistic appeal to 'ease of articulation' offers little insight into the complexity of language variation and change where stability is just as much a reality as change. A more nuanced articulatory view of sound change was most developed in the H & H theory of phonetic variation (Lindblom 1990). From the perspective of this theory, speakers adaptively tune their performance along the H(yper)-H(ypo) continuum according to their estimates of the needs of the listener in that particular situation. These needs include preferences to maximize the distinctiveness of contrasts and to minimize articulatory effort. Speakers hyper-articulate when listeners require maximum acoustic information; they reduce articulatory efforts, hence hypo-articulate, when listeners can supplement the acoustic input with information from other sources. Hyper forms tend to involve increased duration and amplitude of articulatory gestures as well as reduction in gestural overlap while hypo-forms show the opposite characteristics. Speakers and listeners dynamically adjust their production and perception to the communicative demands of the situation. Spectral variation observed in vowel reduction, for example, can be predicted given information about the vowel durations and adjacent consonant loci (Lindblom 1963). Vowel durations in turn are predicated in terms of phonological length, stress, and position within the utterance (Lindblom et al. 1981) and these factors are in turn determined by the morphological and syntactic specification of the utterance. From this perspective, sound change occurs when intelligibility demands are redundantly met or when (p. 413) the listeners focus their attention on the 'how' (signal-dependent) mode rather than the 'what' (signal-independent) mode of listen-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

ing (Lindblom et al. 1995). New phonetic variants are accumulated during the 'how' mode of listening.

The listener-turned-speaker may then be free to select new forms from the pool of variants. The selection process is assumed to be governed by a host of different evaluative metrics. A new variant may be evaluated in terms of the economy of gesture (i.e. how much effort is required to produce it?) or it may be judged from the perspective of the listeners (i.e. how confusable is a variant?). The speech community may also evaluate a new variant in terms of its social value.

24.2.2 Listener-oriented Approaches

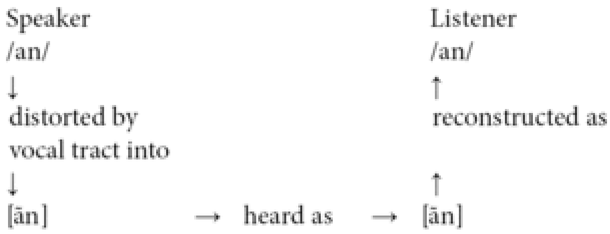

Fig. 24.1 Distortion of the speech signal and reconstruction of the intended signal.

Unlike the articulatory view of sound change, which affords a large agentive role to the speaker, many have emphasized the important role listeners play as well (Blevins 2004a, Beddor 2009), chief among them the misperception model advocated by John Ohala (1981, 1983, 1989, 1992, 1993, 1995). The basic premise of all listener-oriented views of sound change is the ambiguous nature of the speech signal. Such ambiguities, or 'noise', stem from articulatory, acoustic, auditory, and perceptual constraints inherent to the vocal tract and the auditory and perceptual apparatus. Consider, for example, the case of the development of contrastive nasal vowels from vowel+nasal sequences (VN > Ṽ), a sound change frequently attested in the world's languages. Nasalization is often observed on the preceding vowel due to low-level contextual anticipatory nasalization. When such 'unwanted' coarticulatory nasalization is factored out by the listener[1] via the relevant phonetic 'reconstructive rule' or the general mechanism of perceptual compensation, and an underlying VN sequence is reconstructed. Figure 24.1 provides a schematic representation of the chain of events that is characteristic of canonical transmission of coarticulated speech.

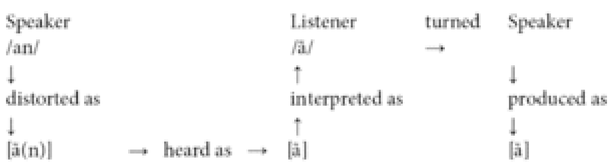

Fig. 24.2 Distortion of the speech signal and absence of reconstruction of the intended signal.

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

Under the scenario just laid out, listeners are assumed to be able to normalize for whatever contextual influences that distort the speech signal (indicated here in []) and to reconstruct accurately the intended message (forms in //). However, when listeners are unable to factor out 'unwanted' distortions in the speech signal, a mini sound (p. 414) change obtains. Figure 24.2 illustrates what might happen when listeners misperceive. If the listener misattributes the nasalization on the vowel as intended rather than as an artifact of coarticulated speech, a new perceptual norm may be established. Assuming a perception-production feedback loop (Pierrehumbert 2001, Oudeyer 2006, and Wedel, this volume), a mini-sound change is accomplished if the listener-turn-speaker begins to produce that lexical item with the new perceptual and production target.

24.2.3 Points of Convergence and Divergence

The models of sound change reviewed above are not mutually exclusive, especially given that both approaches emphasize the role of the individual in the inception of change (see Jones, this volume).[2] Lindblom's H & H approach to sound change, for example, does not rule out the possibility of misperceptions as a contributor to sound change; misperception is presumably only partial as successful access to the lexical level (i.e. the 'what' mode) is still possible even when the speaker has committed a phonetic 'error'. The main difference between these approaches resides in their treatments of phonetic variability. Variability, from the perspective of the H & H theory, is functional in nature and serves a communicative purpose. For Ohala (1993), on the other hand, variability is viewed as noise to be factored out. Sound change is taken to be phonetically abrupt but lexically gradual since misperception is assumed to lead to an immediate shift from one pronunciation to another; the seeming gradualness of sound change is assumed to be coming from its spread from speaker to speaker, speaking style to speaking style, and word to word. Lindblom, on the other hand, argued that the initiation and spread of sound change rest in the accumulation of new variants and speakers' decision to utilize such variants. Thus, like Ohala, sound change in Lindblom's model also spreads between speakers, speaking styles, and words. However, unlike Ohala, sound change is neither abrupt nor gradual. The addition of new phonetic memories (i.e. the introduction of a new phonetic variant) per se does not lead to change since the new variant is only one of many possible variants of a linguistic form. Linguistic innovations arise through deliberate selection of variants for perceptual, articulatory, or social reasons.

(p. 415) Evolutionary Phonology (see Blevins, this volume) can be seen as an attempt to synthesize the strengths of these two perspectives on sound change. Blevins proposes three mechanisms for sound change: CHOICE, CHANCE, and CHANGE. CHOICE refers to innovations grounded in articulatory variation along the H(ypo)-H(yper) continuum, along the same line as that proposed in Lindblom et al. (1995). Variation in the pronunciation of words like *memory* ([mɛməɹi]~[mɛmɹi]) and *camera* ([kæməɹə]~[kæmɹə]) is a case in point. Speakers may choose to hyperarticulate, thus realizing the medial schwa, or they can hypoarticulate, in which case, the schwa is absent. CHANCE refers to innovations based on intrinsic phonological ambiguity. Metathesis (i.e. the re-ordering of sounds; *-rd- > -*dr*-) is a prime example.[3] When the input signal is consistent with two or more phono-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

logical parses, listeners might choose a parse that is incongruent with the parse intended by the speaker. Finally, CHANGE refers to innovations that originate from perceptual bias-induced misperception. Nasal place assimilation in VNTV has been argued to be an instance of CHANGE. Because perceptually-speaking CV transition cues are stronger than VC ones (Fujimura et al. 1978, Repp 1978, Ohala 1990), listeners are biased toward confusing a weak VC transition cue for the place of articulation of a nasal (N) for the strong CV transition cue of the following obstruent (T). Blevins, however, stresses that sound change often involves more than one mechanism. For example, because speakers might show coarticulatory place assimilation in hypo-articulated speech, CHOICE might play a role in the emergence of nasal place assimilation in language.

Because of the differences in emphasis, functional and listener-oriented approaches often utilize different experimental evidence to substantiate their theories and to explain individual instances of sound change. For example, while all approaches to sound change must deal with the question of where phonetic variation comes from, the H & H theorists look for evidence of functional motivations (i.e. issues of articulatory effort and communicative efficiencies) for phonetic variation, while the listener-oriented approach generally seeks to uncover physical and perceptual biases in the speech production and perceptual system that could serve as phonetic precursors to sound change. The next section provides an overview of canonical sources of phonetic biases that give rise to precursors to sound change.

24.3 Phonetic Factors in Sound Change

24.3.1 Production Biases

Much variability in speech comes from constraints imposed by the speech production system. Two sources of production constraints are commonly observed: motor planning and coordination and speech aerodynamics.

(p. 416) 24.3.1.1 Motor Planning and Coordination

One of the most prominent contributing factors to phonetic variation comes from limitations imposed by the motoric aspects of the speech production system. The temporal coordination of the different articulators involved in speech offers a tremendous source of variability and, as a consequence, presents ample opportunities for misanalysis by the listener. Consider the production of nasal+fricative sequences. In order to produce a nasal consonant, the oral passage as controlled by the tongue or the lips must be closed and the nasal passage open. Likewise, for an oral consonant, the reverse condition is required. The production of apical fricatives in particular leaves open the apex as a possible vent. The production of a nasal+apical fricative sequence requires the simultaneous and opposite change of state of the nasal and oral passage. That is, the nasal passage must go from open to closed, while the oral passage must go from closed to open as the articulators transition from a nasal to an apical fricative. If both oral and nasal passages were closed during the transition between these segments, air flowing from the lungs would

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

accumulate in the oral cavity. When the oral constriction is finally released, a burst might become audible as a result of the pressure buildup inside the supraglottal cavity. If the listener misanalysed the audible burst as the presence of an intended obstruent between a nasal and an apical fricative, a new variant with an epenthetic stop will emerge (e.g. English *prince* ~ *prin[t]ce*; Ohala 1997b). A similar explanation applies to stop epenthesis in nasal+lateral sequences (e.g. Latin *tem**p**lus* < *tem-lo 'a section'; Ohala 1997b).

Another example of potential perceptual ambiguities caused by issues with gestural coordination comes from gestural overlap in connected speech processes. In their examinations of X-ray pellet trajectory data,[4] Browman and Goldstein (1990) demonstrated that perceived assimilation or deletion might arise as a result of subsequent consonantal gestures overlapping to such an extent as to either hide each other when they involve sufficiently independent articulatory gestures or blend their characteristics when they involve the same articulators. Thus in one of the utterances they examined, 'perfect memory', X-ray trajectories revealed, within a fluent phrase, the presence of the /t/ gesture, even though it overlapped with, and is masked by, the following /m/ gesture. The authors further pointed out that not all cases of assimilation include such residual articulations. 'Even deletion, however, can be seen as an extreme reduction, thus as an endpoint in a continuum of gestural reduction, leaving the underlying representation unchanged' (1990: 366).

The inter-dependence of individual components of the speech production apparatus also gives rise to phonetic biases that contribute production variation. An example comes from the production of voicing in stops. In addition to a steady transglottal flow, the production of voicing requires the vocal folds to be adducted to a suitable degree (p. 417) and the longitudinal tension of the fold must be adjusted within an appropriate range. Voicelessness, on the other hand, requires vocal-fold abduction and a decrease in pressure across the glottis. The coordination of articulatory gestures for these phonation differences often lead to articulatory by-products that may have perceptual consequences. Ewan & Krones (1974) found that the vertical position of the larynx differs for voiced and voiceless stops. Using data obtained by a special photo-electric device called the 'thyroumbrometer' invented by the authors, Ewan & Krones (1974) showed a higher position of the larynx for voiceless as opposed to voiced stops in French, English, and Thai, and f_0 was positively correlated with larynx position. The raising of the larynx is accomplished through the contraction of extrinsic muscles attached to the thyroid cartilage and hyoid bone. Such changes probably occur to increase the stiffness while decreasing the mass of the vocal folds, which would make it more difficult for the vocal folds to vibrate. Similarly, cricothyroid muscle activity has been shown to increase for voiceless consonants relative to voiced consonants. An increase in cricothyroid activity tenses the vocal folds longitudinally, which leads to an increase in the frequency of vibration (Löfqvist et al. 1989).

This pitch-perturbation effect has been linked to tonogenesis in many languages (see Ratliff and Hale et al., both this volume). That is, the development of tone splits (e.g. in Chinese, Vietnamese, Khmu) has been attributed to the exaggeration of physiologically-based consonantal voicing-induced pitch perturbations on the neighboring vowel

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

(Hombert 1978, Hombert et al. 1979) to such an extent that the pitch variation cannot be attributed entirely to the physiological properties of the preceding consonant's voicing (*pa > pá and *ba > bà where ´ indicates a high tone and ` indicates a low tone; Hyman 1976). Likewise, many languages show interaction between obstruent voicing and the height of neighbouring vowels.[5] For example, in Madurese, an Austronesian language of Indonesia, vowels are raised following a voiced obstruent (*m**ɛ**t̪ ɔŋ* 'AV.count' ~ *b**i**t̪ ɔŋ* 'count'; *ŋ**ɛ**pʰɨ* 'AV.carry' ~ *g**i**pʰɨ* 'carry'; AV = Agent Voice). No comparable change in vowel height is observed when the vowel is preceded by a sonorant or a voiceless obstruent (*n**ɔ**t̪ t̪ʰuʔ* 'AV.point' ~ *t̪ʰɔt̪ t̪ʰuʔ* 'point').[6] Such an interaction between consonantal voicing and vowel height has been argued to have its origin in the lowering of the larynx that accompanied the production of voicing (Bauer 2009, Moreton 2008), as one important acoustic effect of coarticulatory laryngeal lowering is the lowering of F1 (i.e. raising of vowel height).

24.3.1.2 Speech Aerodynamics

Aerodynamic principles govern the movement of air in the vocal tract which, in turn, generates the acoustic output we call sound. The production of speech sounds thus requires not only the movement of the vocal apparatus but also the coordination of the aerodynamic engine that powers sound production. Aerodynamically, the vocal tract (p. 418) has two relevant air cavities, the lung cavity and the supraglottal cavity. Many acoustic cues, such as voicing, frication, and trills, require a sustained and continuous supply of airflow. In addition to the muscular activities that govern the configuration of the glottis (i.e. the glottis must be adducted), voicing—vibration of the vocal cords—is possible only when a pressure differential exists between the lung and the supraglottal cavities and sufficient air flows past the glottis. Sustained production of voicing is difficult when there is a blockage along the vocal tract downstream from the glottis, as in the case of an oral stop, since the pressure differential across the glottis will eventually equalize at which point airflow will cease and vocal fold vibration will stop. This so-called 'Aerodynamic Voicing Constraint' (AVC; Ohala 1983) has been used to explain why cross-linguistically voiced stops are disfavoured compared to voiceless ones and why voiced stops (b, d, g) tend to give way to voiceless stops (p, t, k) or voiced spirants (β, ð, ɣ). That is, voicing during stop closure might cease as a consequence of the AVC or, if voicing is to be maintained, other aspects of stop production might have to be sacrificed (i.e. spirantization). Even in languages with voiced stops, the velar [g] is often missing (Maddieson 2008b), presumably due to the smaller oral volume for [g] which causes voicing to cease sooner relative to voiced stops of other place of articulation. To be sure, rather than succumbing to the AVC, speakers might employ other compensatory strategies such as larynx lowering and nasal venting to help sustain voicing. Implosives, which are produced in part with the lowering of the larynx, in Sindhi, an Indo-Aryan language, are said to have originated from former voiced geminates (e.g. Prakrit *pa**bb**a > *pa**ɓ**uɳi* 'lotus plant fruit', *bʰa**gg**a* > *bʰaː**ɠ**u* 'fate'; Ohala 1983). Likewise, voiced stops in many languages are reported to have (spontaneous) prenasalized variants (Iverson & Salmons 1996b, Solé 2009) or have developed into prenasalized stops (Ohala 1983), presumably a result of anticipatory nasal venting.

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

The production of fricatives and trills is also tightly constrained by aerodynamic factors. The production of fricatives, as well as the fricative release of affricates and the burst of stops, requires the generation of audible turbulence (i.e. noise) in the vocal tract. For the supraglottal fricatives, a delicate balance between subglottal and supralaryngeal pressures is needed to allow airflow through the glottis as well as across the narrow oral constriction downstream from the glottis. That is, subglottal pressure must be higher than supraglottal pressure, which in turn must be higher than atmospheric pressure. Sound change could occur if this delicate balance is disturbed. For example, if supraglottal pressure behind the oral constriction is vented through the nasal cavity, turbulence is endangered. Such a scenario might underlie the development of voiceless nasals in Burmese which has been argued to have come from historical *sN* sequences where *N* represents a nasal of any place of articulation (e.g. Written Tibetan *sna*/Written Burmese *hna* ~ Modern Burmese /n◦ a/ ‘nose’; WT *smin-po*/WB hmań’ ~ /m◦ ɛ̃’/ ‘ripe’; Ohala 1975: 295).

Audible frication might also be endangered as a result of variability in articulatory strength in different prosodic positions. Diachronically, syllable-final fricatives often despirantize, which could result in vocalization/gliding (Latin *no**s***, *vo**s***, *pō**s**t* > Italian, Romanian *no**i***, *vo**i***, *po**i*** ‘we’), aspiration (Latin *fe**s**ta, ve**s**pa, di**s**jejunare* > Lombardian (p. 419) ['fɛ**h**ta], ['vɛ**h**pa], [di**h**'na] ‘holidays’, ‘wasp’, ‘to dine’), assimilation (Gascon *es **l**ansos* > *e**l** **l**ansos* ‘the sheets’), rhotacism (PIE *au**s**- > English *ea**r***, German *Oh**r***), or elision (Latin *no**s***, *vo**s*** > French *nous* [nu], *vous* [vu] ‘we’, ‘you’). Solé (2010) argued that such weakening and loss of final fricative might be the result of audible frication being more difficult for speakers to produce and for listeners to detect on account of a decreased oral gesture in syllable-final position and/or lowered subglottal pressure in utterance-final position. Such modification of the aerodynamic conditions is supported by her findings that compared to onset fricatives, coda fricatives aerodynamically and acoustically exhibit a slower oral pressure build-up, a lower pressure peak, a delayed onset of audible frication as well as lower intensity of frication.

In addition to the patterning of voicing and frication, aerodynamic factors have also been argued to be important for understanding the sound pattern of trills (Solé 1999, 2002*a*,*b*). The production of apical trills requires tongue-tip vibration, initiated by muscle contraction of the tongue to be in the appropriate position, shape and elasticity, and a sufficient pressure difference across the lingual constriction. Once trilling is commenced, tongue-tip vibration is maintained as a self-sustaining vibratory system. In an effort to explain the tendency for lingual fricatives to assimilate to the following apical trill (e.g. Iberian Spanish /s (#) r/ *la**s** **r**ojas* [la'roxas] ‘the red ones’, *O**sr**am* ['oram]), Solé (2002b) showed that lingual trills exhibit highly constrained articulatory and aerodynamic requirements; small variations may result in a lack of tongue-tip vibration. Her investigation showed that, in /s, ʃ + r/ sequences, the generation of audible turbulence of the neighbouring fricative might be imperiled when speakers produce an early onset of the lingual movements for the trill, presumably to attain the tongue configuration and aerodynamic requirements for tongue-tip vibration.

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

24.3.2 Perceptual Biases

There are three commonly recognized perceptually-motivated sources of phonetic variation: perceptual confusion, hypocorrection, and hypercorrection.

24.3.2.1 Perceptual Confusion

Errors in perceptual parsing could be precipitated by the intrinsic perceptual similarity between segments. An intriguing property of similarity-based confusion is that it is not necessarily symmetric. That is, while *a priori*, when segments X and Y are similar to each other, we would expect X to be equally likely to be confused as Y and vice versa. This is, however, not always the case. A prime example comes from asymmetry in consonantal changes. Diachronically, velar palatalization before front vowels is found in many languages, including in the histories of Slavic, Bantu, Indo-Iranian, Mayan, Salish, and Chinese (Guion 1996, 1998). Sound change in the opposite direction, i.e. from an alveopalatal to a velar before front vowels, are rare, if attested at all. The unidirectionality of this change has been argued to be the result of confusability between velars and alveolars in the context of front vowels (though see Recasens 2011 for an (p. 420) articulatory explanation). Velars before front vowels are similar acoustically to palatal alveolar affricates in the same environment and the acoustic similarity is greater in faster speech than in citation (Guion 1996, 1998). Crucially, velars before front vowels are perceptually more easily confused with alveolars and palatals but not the other way around (Guion 1996, 1998, Chang et al. 2001). The asymmetry in perceptual confusability helps explain the unidirectionality of neutralizing sound changes (see also Kawasaki 1982, Foulkes 1997). Chang et al. (2001) attributes this perceptual asymmetry to the degradation of the compact mid-frequency spectral peak as a cue to differentiate [ki] from [ti]. The otherwise robust F2 formant transition cue for distinguishing consonant place of articulation is neutralized because of the raising F2 of the following high vowel. Thus if the remaining non-robust mid-frequency spectral peak, which indexes the front cavity resonance, was degraded to the point of losing its contrastiveness, listeners are likely to confuse [ki] as [ti]. The confusion of [ti] for [ki] is unlikely since it would entail listeners erroneously inserting a non-existent cue into the speech signal for [ti].

24.3.2.2 Hypocorrection

Hypocorrection is characterized by a scenario where fortuitous results of the speech production process are misanalysed as part of the pronunciation norm. The development of Buchan Scots vowel height harmony has been argued to be a case in point. The modern Buchan dialect of north-east Scotland has two high vowels (i, u) and seven non-high vowels (e, ɛ, ɜ, a, ʌ, ɔ, o). Suffixes and clitics that contain an unstressed vowel /i/ or /ɜ/ exhibit height harmony with the vowel of the root (*hair-y* [here], *lassie* [lase], *rocky* [rɔke], but *beanie* [bini], *dearie* [diri] and *housie* [husi]; Paster 2004). Paster (2004) argued that vowel height harmony in Buchan Scots is a consequence of the phonologization of coarticulation in tongue height. Vowel-to-vowel coarticulation, including vowel height, frontness, and rounding, is rampant in natural speech and has been the object of many studies (e.g. Manuel 1990, Beddor & Yavuz 1995, Majors 1998, Beddor et al. 2002). Paster hypothe-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

sized that the lowered tongue position used to produce non-high vowels coloured the pronunciation of the following high vowels so much so that they were produced with the tongue body slightly lower than in high vowels in other contexts. When listeners hypocorrected for this coarticulatory effect and misanalysed the height variation as intentional, a phonologized harmony process might develop as a result where high vowels were lowered to non-high after a non-high vowel. Additional evidence in support of this perceptually-driven account comes from blocking in this language. Vowel height harmony in Buchan Scots is blocked when the root-final consonants that intervene between the triggering and target vowels are voiced obstruents (*Eddie* [ɛdi], *lovey* [lʌvi]) or combinations of voiced obstruents with each other and with other sonorants (*bendy* [bɛndi], *hardly* [hardli]). Paster (2004) accounted for the blocking effects in terms of the larynx lowering effect of voicing, which in turn lowers F1. Given that the magnitude of larynx lowering is greater at or near the end of the stops, the F1 lowering effect of voicing should have a more significant effect on the vowel following rather than preceding the consonant in question. (p. 421) Thus at the stage when suffix vowels are phonetically lowered (i.e. raised F1) when preceded by non-high root vowels, the phonetic effect of vowel lowering (F1 raising) might have been negated when the voiced obstruent intervened because the maintenance of strong, unattenuated voicing in obstruents requires significant laryngeal lowering, which lowers F1. Blocking can thus be explained as the lack of phonologization of vowel height harmony when voiced obstruents intervened between the trigger and target vowels.[7]

Other sound changes that have been argued to be the results of hypocorrection include umlaut (Ohala 1994), vocalic nasalization (Hajek 1997), and /u/-fronting (Harrington et al. 2008). In such hypocorrective changes, listeners are assumed to have failed to perceive or attend to the conditioning, or triggering, environment, thus hypocorrective changes often result in the simultaneous emergence of new segments and loss of the conditioning environment. The emergence of contrastive nasal vowels, for example, often takes place with the simultaneous loss of the conditioning nasal (VN > Ṽ). Of course, cases where the conditioning environment remains are not uncommon as well. Beddor (2009) argued, with experimental support, that the coarticulatory gesture might come to be reinterpreted as distinctive from its source segment as a result of the articulatory covariation between the duration of the coarticulatory source (N in the case of the VN > Ṽ change) and the temporal extent of its effects (Ṽ), and the perceived equivalence between the source and the effects. The loss of the conditioning environment might be attributed to the temporal extent of vowel nasalization overwhelming the duration of the triggering source to such an extent that listeners can no longer recover the presence of the triggering nasal.

24.3.2.3 Hypercorrection

A hypercorrective change takes place when listeners misattribute intended cues as superfluous and thus incorrectly factor them out. Dissimilation, which refers to restrictions between similar segments co-occuring adjacent to each other or at a distance, has been argued to be an example of hypercorrective change. For example, in the history of Cantonese, when two labial segments occur in the same syllable, the second consonant lost

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

its labiality (e.g. Ancient Chinese *pjam > Cantonese *pin*). Similarly, in the history of Greek and Sanskrit, aspiration in the first consonant disappeared when two aspirated segments occur within the same form (Grassmann's Law; Proto-Indo-European ***b̤**en**d̤**> Sanskrit ***b****an****d̤*** 'blind', cf. Cser, this volume). Dissimilation like Grassmann's law takes place on this approach when the listener incorrectly parses the incoming signal and misattributes aspiration on the first consonant as coarticulatory aspiration from another aspirated consonant occurring later in the word. Since dissimilatory changes tend to be long distance, phonetic features that dissimilate tend to have long stretching realization, such as labiality, retroflexion, palatality, laryngealization, (p. 422) and pharyngealization (Ohala 1993), presumably because these are precisely the type of phonetic cues that are difficult to localize. Another type of sound change that has been analysed as the result of hypercorrection is metathesis, although recent studies have suggested that there might be multiple contributing factors that give rise to metathesis (Blevins & Garrett 1998, 2004, Hume 2004). Like dissimilation, the so-called perceptual metathesis (Blevins & Garrett 2004) involves features of intrinsically long duration, such as retroflexion. When such features are spread out over a sequence of multisegmental strings, they might be reinterpreted in non-historical positions. For example, in Classical Armenian, the linear order of stop (or affricate) + *r* clusters was regularly inverted in initial (Armenian *a**rt**asu* 'tear(s)' < *brewr < Indo-European **b^{h}rē**wr) as well as medial positions (*k*^{h}*i**rt**n* 'sweat' < *k^{h}itrn < *swidros; Blevins & Garrett 2004: 129).

24.3.3 Perceptual Compensation for Coarticulation

The twin notions of hypo- and hyper-correction presuppose that there exists a normative process of perceptual compensation or 'correction' for coarticulation. Much research that appeals to hypo- and hyper-correction as explanatory mechanisms for sound change has focused on establishing a normative compensatory pattern. A classic demonstration of perceptual compensation is Mann & Repp's 1980 study of vocalic effects on sibilant perception (see also Nittrouer & Studdert-Kennedy 1987, Mitterer 2006, Yu 2010a). In this study, listeners were asked to classify sibilants in a series of CV continua where C is a synthesized continuum of /s/ to /ʃ/ and V is either /a/ or /u/. They found that listeners are more likely to report hearing /s/ before /u/ than before /a/, presumably because listeners take into account the lowered noise frequencies of /s/ in a rounded vowel context (see Strand & Johnson 1996, Strand 1999, for a potential role of social awareness in this type of perceptual normalization). Similar methodology has been applied to study perceptual compensation for other coarticulatory processes, such as vowel-to-vowel dependencies (Ohala 1994, Beddor & Yavuz 1995), intrinsic vowel pitches (Hombert 1977), intrinsic vowel duration (Gussenhoven 2004, Yu 2010b), and liquid/rhotic coarticulation (Abrego-Collier 2013). Beddor & Krakow (1999) found evidence of perceptual compensation using a metalinguistic rating task; listeners were accurate in rating acoustically identical nasal vowels as 'equally nasal' when both were in non-nasal contexts (e.g. [CṼC]-[Ṽ]) but were less accurate when only one of the two was in a nasal context ([NṼN]-[Ṽ]). Beddor & Krakow (1999) also found that, while listeners' judgements were least accurate when one vowel was in a nasal context and the other in a non-nasal one, the discrimination of vow-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

els in such pairs as [NṼN]-[V] was consistently above chance, suggesting that listeners attribute some but not all of the context-dependent variation to a coarticulatory source (i.e. the flanking nasal). Partial perceptual compensation of this sort, though not a case of outright misperception, might nonetheless provide seeds for future sound changes.

(p. 423) 24.4 Potential Limitations and Future Direction

Experimental investigations on sound change can offer tremendous insights into the physical, physiological, and perceptual mechanisms that potentiate sound change. These advances have also prompted new questions. In this final section, we discuss some of these new puzzles and, in some cases, offer preliminary answers.

24.4.1 Distinguishing between Phonetic Precursors and Gradient Phonologized Processes

An important development in phonetic research is the discovery of language-specific phonetic realizations of phonetic categories. Consider the case of voicing contrast in language. [+voice] consonants are characterized by the “presence of low-frequency spectral energy or periodicity over a time interval of 20–30 msec in the vicinity of the acoustic discontinuity that precedes or follows the consonantal constriction interval” (Stevens & Blumstein 1981: 29). This low-frequency property, as Kingston and Diehl (1994) call it, has multiple supporting subproperties such as voicing during the consonant constriction interval, a low F1 near the constriction interval, and a low f_0 in the same region, as well as enhancing properties such as the duration ratio between a consonant and its preceding vowel. These properties do not all surface in all positions. The contrastive feature [+voice] in English, for example, shows great variability in its phonetic realization.[8] In word-initial position, [+voice] stops are often realized as voiceless unaspirated, even when the preceding word ends in a vowel (Caisse 1982, Docherty 1989, 1992). Kingston & Diehl (1994) interpret such data as showing that speakers choose between two active articulations in producing initial [+voice] stops in English: delay glottal closure until the stop release, or close the glottis but expand the oral cavity to overcome the difficulty of initiating voicing. Such controlled variation is made possible by the fact that there are typically multiple, auditorily independent correlates that serve as distinct bases for a minimal phonological distinction.

According to Kingston & Diehl’s (1994) conception of the phonetics-phonology interface then, elasto-inertial, biomechanical, aerodynamic, psychoacoustic, and perceptual constraints delimit what a speaker (or listener) *can* do, but not what they *must* do. A phonemic contrast is thus taken to be ‘any difference in the feature content or arrangement of an utterance’s phonological representation which may convey a difference in semantic interpretation’ and allophones are ‘any phonetic variant of a distinctive feature (p. 424) specification or arrangement of such specification that occurs in a particular context’ (p. 420 fn.2). If the phonetic realization of phonological distinctions already admits this wide

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

a range of possibilities, it is only reasonable to assume that the range of possibilities between languages is considerably much bigger. Cho & Ladefoged (1999), for example, found the degree of between-language variation in voice onset time (VOT) realization is tremendous. Crucially, VOT values cannot be predicted from knowledge of the phonological contrasts within a language (Ladefoged & Cho 2001). A language lacking a contrast between /k/ and /k^h/ does not necessarily have the simplest possible VOT, nor would a language with a /k/ ~ /k^h/ contrast show the largest VOT for /k^h/. The mapping between measurable phonetic parameters and phonology is thus largely arbitrary; each language must choose a model VOT value for each voicing category (e.g. [voiced], [voiceless unaspirated], and [voiceless aspirated]) that must be specified in the phonology (Cho & Ladefoged 1999, Ladefoged & Cho 2001).

These differences raise methodological problems on two fronts for experimental investigations on sound change. Recall that an important assumption behind all attempts to employ synchronic evidence to explain sound change is the principle of uniformitarianism. It is the belief that whatever phonetic or linguistic principles that operate today must also have operated in the past. The fact that language-specific phonetic implementation is so pervasive raises concerns about the relevance of findings from laboratory investigations today to changes that have taken place at a different point of human history. Here, it is useful to point out that historical sciences, experimental studies of sound change included, necessarily offer only what Andersen (1989) calls a 'rational explication', rather than to provide actual 'causal explanation'. The problem of actuation can at best be answered probabilistically. All one could hope to do is to provide a rational account of how a particular sound change *could* have happened. While the actual phonetic parameters a language utilizes at a particular instance along the time-space continuum might differ, the principles that govern the physical, physiological, aerodynamic, perceptual, and psychological factors involved in language are presumably unbending, as humans are not likely to have evolved so dramatically within the reconstructible timeframe of linguistic history as to require new principles of speech production and perception.

The other issue raised by the findings of language-specific phonetic implementation concerns the nature of 'change'. That is, how can we distinguish effects of constraints of the speech production and perception system (i.e. the seeds or phonetic precursors of sound change) from language-specific phonetic implementation? One way to discern genuine mechanistic causes of variation in speech from language-specific phonetic realization that entails a change in progress is to investigate the amount of control speakers have over the maintenance of subphonemic phonetic differences. Solé (2007), for example, found that English speakers actively maintain durational differences before voiced and voiceless stops regardless of speaking rates, while speakers of Catalan and Arabic do not. Her findings suggest that English has already partially phonologized the effect of consonant voicing on vowel duration, while Catalan and Arabic have not. Similarly, using aerodynamic evidence, Solé (1992, 1995) showed that Catalan speakers do not (p. 425) have as fine-grained a control as American English speakers have over the degree of vocalic nasaliza-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

tion from a tautosyllabic coda nasal, suggesting that vowel nasalization has been phonologized in American English but not in Catalan.

Evidence of phonologization can be established from a perceptual point of view as well. In an attempt to establish whether /u/-fronting, a sound change that has been in progress for the last 20–30 years in standard southern British (SSB), could be linked synchronically to the fronting effects of a preceding anterior consonant, Harrington et al. (2008) examined the production and perception of /u/ by speakers of two age groups, individuals between the ages of 18–20 and those over 50. They found that, for the younger speakers, of /u/ was phonetically more fronted and that the coarticulatory influence of consonants on /u/ was less. Crucially, they also found that younger listeners compensated perceptually less for the fronting effects of the flanking anterior consonants than older speakers. These findings suggest that younger speakers of SSB not only produce a fronter realization of /u/, they also have a fronter category boundary in perceiving this vowel. Such a change in speech production and perception is consistent with the hypocorrective view of sound change laid out above, assuming that young listeners gave up on compensating perceptually for coarticulation.

24.4.2 Structural Factors in Sound Change

While many experimental investigations have focused on identifying the phonetic precursors to sound change, some scholars have questioned whether non-phonetic factors might have a role in shaping phonetically-conditioned sound change (see, for example, Purnell & Raimy, this volume). One potentially important factor is the role structural constraints have in channelling the directionality of sound change. That is, when the listeners have to resolve ambiguities in the speech signal, the set of possible resolutions might not be equally available. Consider the case of phonotactic influence in speech perception. Listeners' perceptual responses are influenced by their knowledge of what are possible and impossible sound sequences in the language (Massaro & Cohen 1983, Pitt 1998, Hallé et al. 1998, Moreton 2002, Berent et al. 2007). For example, when listeners were asked to classify a synthetic /r/-/l/ continuum embedded in a C_i context where C = {t, p, v, s}, they were most likely to report the ambiguous liquid as [r] when C was either /t/ or /p/ and the least when C was /v/ or /s/ (Massaro & Cohen 1983), presumably because *tl*- and *vr*-/*sr*-sequences are phonotactically ill-formed in English. Thus, phonotactic knowledge might bias misperception-driven sound changes toward phonotactically licit outcomes.

Structural constraints need not be based on phonotactic information extracted from the lexicon, however. Some scholars have recently argued that there exist phonotactic restrictions that are a prioristic, rather than learned from experience (Berent et al. 2007, 2008, Berent 2009, Berent et al. 2009). In Berent et al. (2007), for example, subjects listened to CCVC and CəCVC words where the first two consonants either have a rising (*bn*), plateau (*bd*), or falling (*lb*) sonority profile and were asked to indicate whether the (p. 426) stimulus has one or two syllables. While all three clusters are illegal word-initially in English, subjects nonetheless showed a tendency for preferring sequences that follow the sonority sequencing principle (i.e. *bn* > *bd* > *lb* where X > Y should be interpreted as

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

X is more preferred than Y). The authors contended that their findings support an interpretation where knowledge of the sonority hierarchy comes from universal constraints on language learning; they argued that their findings cannot be reduced to statistical properties of the English lexicon and also rejected a purely phonetic explanation (cf. Peperkamp 2007).

24.4.3 Systematic Sources of Deviation in Modes of Speech Processing

Both the H & H and the listener-oriented models of sound change presuppose that new variants arise only when individuals deviate from their normative mode of speech perception. In the case of the misperception model, listeners must fail to take the coarticulatory context properly into account by either failing to correct for the coarticulatory influence or overanalysing the potential effects of articulation. In the H & H model, sound change is only possible when individuals focus on the 'how' mode of listening instead of the 'what' mode. A question that must be addressed is why such deviation occurs. Ohala (1993) argued that the listeners who fail to properly take coarticulatory contexts into account presumably do not have the necessary linguistic background. Thus, under his theory, new variants must necessarily be introduced by individuals who are acquiring the language for the first time, i.e. children acquiring their first language, or learners of a second language (compare Foulkes & Vihman and Eckman & Iverson, both this volume). Such an assumption, however, might be too restrictive as recent experimental evidence has suggested that the sound system of individuals might change throughout their life times (see Sancier & Fowler 1997, Harrington et al. 2000, Sankoff 2004, Harrington 2006, Evans & Iverson 2007, and also Bowie & Yaeger-Dror, this volume).

Recent studies, which show that perceptual responses to variation in speech may vary as a function of individual differences in cognitive processing style (Stewart & Ota 2008, Yu 2010a, 2013), offer potential answers to the question of systematic deviation in speech perception. Yu (2010a), for example, found that the magnitude of perceptual compensation for coarticulation, in this case, the effect of vocalic rounding on sibilant perception, is modulated by the level of 'autistic traits' the listener exhibits. Neurotypical individuals exhibiting few 'autistic traits' show the least amount of perceptual compensation for coarticulation. That is, an individual's overall Autism-Spectrum Quotient (AQ; Baron-Cohen et al. 2001) is positively associated with the way linguistic information is processed. Individuals with an imbalanced brain type, defined in terms of the difference in empathising and systemizing abilities (as measured by the normalized Empathy Quotient (EQ; Baron-Cohen & Wheelwright 2004) and the normalized Systemizing Quotient (SQ; Baron-Cohen et al. 2003)), have also been found to not perceptual compensate for coarticulation to the same degree as individuals with a balanced brain type (Yu 2013).

(p. 427) These findings have significant ramifications for theories of sound change, particularly in its potential for reconciling the tension between the H & H emphasis on the speaker and the misperception model's focus on the listener. Recall that the misperception model of sound change maintains that innovation in sound change resides in listen-

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

ers' failure to properly compensate for contextual variation in speech. The findings reviewed above suggest that variability in perceptual compensation need not be confined to accidental misperception at the level of the individual word or utterance, but might exist pervasively within a speech community as a function of inherent individual differences in cognitive makeup (as measured by individual-difference dimensions such as the AQ, EQ, and SQ). Since variation in cognitive processing style has been shown to covary with differences in the listener's response patterns during speech perception, particularly in the case of perceptual compensation for coarticulation, to the extent that such differences in perceptual response may ultimately lead to individual differences in perceptual and production norms, variability in cognitive processing style stands to be a major contributor to the creation of new linguistic variants in sound change. These findings also suggest an alternative source of new variants from the perspective of the H & H approach to sound change, which see sound change as the result of listeners prioritizing the 'how' mode of listening over the 'what'. If new variants are made possible as a result of individual differences in perceptual compensation, the listeners are presumably deemphasizing the influence of coarticulatory contexts in speech in their listening (i.e. the 'how' mode of listening') and may instead be focusing on the content of the signal (i.e. the 'what' mode). While further research is needed to ascertain the nature of the individual variability in perceptual compensation for coarticulation and in speech perception in general, studies on individual differences highlight the importance of understanding how speech perception and production operate at the level of the individual; much information might be obscured by averaging experimental results across individuals (see Jones, this volume).

24.4.4 Beyond Inception

A crucial aspect of the actuation problem concerns the question of how a speech community comes to adopt a new norm. In recent years, proponents of exemplar-based models of sound change (see Bybee, Phillips, and Wedel, all this volume) have argued that sound change may be modelled in terms of shifts in phonetic memory distributions, or exemplar 'clouds' (de Boer 2001, Pierrehumbert 2001, 2002, Wedel 2006, 2007, Yu 2007, Blevins & Wedel 2009), a view anticipated in Paul 1880 (see Murray, this volume). Such models assume that listeners retain fine phonetic details of particular instances of speech (see also Lindblom et al. 1995), new variants introduced by persistent bias factors would accumulate in such a fashion that eventually moves the distributions of exemplars in the direction of the biased variants, presumably as a consequence of convergence via imitation. That is, speakers' production targets are altered along some phonetic dimensions to become more similar to those of their fellow interlocutors (Goldinger 1998, (p. 428) Shockley et al. 2004, Pardo 2006, Babel 2009, Nielsen 2011, Yu et al. 2011). One source of evidence for phonetic convergence comes from studies on perceptual learning (Norris et al. 2003, Eisner & McQueen 2005, Kraljic & Samuel 2006). These experiments show that listeners can retune their perceptual categories when exposed to a series of oddly pronounced phonemes embedded somewhere in words in the language (e.g. *ob[?sf]ene* or *bro[?sf]ure* where [?sf] is half way between [s] and [f]). Perceptual learning has been shown to generalize across both speaker and test continua (Kraljic & Samuel 2006).

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

While the ability to imitate and retune perceptually is assumed to be innate (Dijksterhuis & Bargh 2001), imitation is not likely to be the lone driving force behind the systematic propagation of new variants throughout the speech community, since phonetic imitation is not an entirely automatic or unrestricted process. Kraljic et al. (2008), for example, showed that, while perceptual representations are flexible, such changes are 'pragmatic' in nature. Biological sex difference (Namy et al. 2002, Pardo 2006), speaker attitude toward the interlocutor (Yu et al. 2011, Abrego-Collier et al. 2011), and perceived sexual orientation (Yu et al. 2011) have also been associated with phonetic convergence and divergence, suggesting that social factors are important motivators for imitation (Dijksterhuis & Bargh 2001, Babel 2009). Rather than propagating aimlessly and blindly as implied by a simplistic conception of an exemplar-based model of sound change, these findings suggest that new variants are spread across the speech community when they come to be associated with social significance (Weinreich et al. 1968, Eckert 2000, Labov 2001). The continued engagement between phoneticians, laboratory phonologists and psycholinguists on the one hand, and sociolinguists and historical linguists on the other, should yield fruitful results concerning the nature of sound change propagation as part of the actuation process.

24.5 Conclusion

The marriage between historical phonology and experimental investigations in phonetics, laboratory phonology, and psycholinguistics has proven to be fruitful. Over the years, cross-pollination of ideas has pushed both enterprises to new advances and generated numerous useful insights. Students in historical phonology are no longer, and must not be, confined to the archives and libraries. The linguistic laboratory offers a wealth of tools, both conceptual and physical. As historical phonology matures as a science, the need for rigour in establishing sound changes must also rise with the time. Only through empirical confirmations can we be confident that conjectures regarding the nature of sound change stand on firm grounding.

Notes:

([1]) Scarborough (2004) argues coarticulation facilitates speech recognition and may therefore be intended by the speaker.

([2]) Neither approach is designed to address how members of a speech community come to share these new perceptual and production targets. For more discussion on the question of sound change propagation, see Sections 24.4.3 and 24.4.4 as well as D'Arcy, Foulkes, & Vihman, Eckman & Iverson, and Bowie & Yaeger-Dror (all this volume).

([3]) See Section 24.3.2.3 for more discussion.

([4]) The data came from the AT&T Bell Laboratories archive of X-ray microbeam data (Fujimura et al. 1973). The X-ray microbeam system tracks the position of (up to seven) small

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022

lead pellets placed on the lower lip, the jaw, the tongue blade, the tongue dorsum mid, the tongue dorsum rear, and/or the soft palate, in addition to two reference locations.

([5]) Blocking of height harmony in Buchan Scots to be discussed in Section 24.3.2.2 is also an example of this interaction.

([6]) Madurese examples are taken from Davies (1999 : 7). See also Stevens (1968), Cohn (1993).

([7]) The fact that -NT- and -lT- sequences also served as blockers was argued to be the result of analogical extension of the original voiced obstruent blocking pattern to -NT-and -lT- sequences where the T segment is phonetically voiced (Paster 2004).

([8]) Advocates of laryngeal realism (Honeybone 2005, Iverson & Salmons 2007) argue that the contrastive laryngeal feature in English is [spread glottis] rather than [voice].

Alan C. L. Yu

Alan C. L. Yu is Associate Professor of Linguistics and the College and the University of Chicago. He also directs the Phonology Laboratory and the Washo Documentation Project. His research focuses on phonological theory, phonetics, language typology, and language variation and change. He is the author of A Natural History of Infixation (2007, Oxford University Press) and the editor of Origins of Sound Change: Approaches to Phonologization (2013, Oxford University Press).

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2022. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: University of Georgia; date: 31 January 2022