Acoustic-Perceptual Factors in the Actuation of Sound Change

James Kirby
Institute for Phonetics and Speech Processing, LMU Munich, Germany

Jonathan Harrington
Institute for Phonetics and Speech Processing, LMU Munich, Germany

The Wiley Blackwell Companion to Diachronic and Historical Linguistics. Edited by Adam Ledgeway,
Edith Aldridge, Anne Breitbarth, Katalin É. Kiss, Joseph Salmons, and Alexandra Simonenko.
Published 2026 by John Wiley & Sons, Inc.
DOI: 10.1002/9781119898023.wbcdl063

1 Introduction

Speech is characterized by ubiquitous variation. Although variation is often quite stable, both within speech communities and over long spans of historical time (see "Revisiting the Five Foundational Problems of Weinreich, Labov, and Herzog (1968)"; "Diachronic Linguistics: An Overview"), it is widely accepted that it is this systemic phonetic variation that provides the 'seeds' of sound change (Ohala 1989; Kiparsky 1995; Lindblom et al. 1995; Blevins 2004). But how does variation become change? The focus of this entry is on changes characterized primarily by the first stage of the life cycle model (see Bermúdez-Otero and Trousdale 2012; Bermúdez-Otero 2015; Ramsammy 2015; "Life Cycle of Phonological Processes"), that is, on the causes and conditions by which intrinsic phonetic variation becomes extrinsic or *phonologized*. In particular, we focus on four aspects of sound change actuation in which acoustic-perceptual considerations are paramount: acoustic-perceptual bias factors; perceptual parsing; enhancement; and perceptual learning.

Many of the classical examples of sound change involve *coarticulation*, that is, contextually conditioned variation in the acoustic realization of a speech sound (Ohala 1993a; Recasens 2018; Zellou 2022). Commonly cited examples include the emergence of umlaut due to vowel-to-vowel coarticulation (Ohala 1994; Grosvald and Corina 2012), the development of pitch-based (tonal) contrasts grounded in consonant-to-vowel coarticulation (Hombert, Ohala, and Ewan 1979; "Tonogenesis and the Evolution of Tone Systems"), or the emergence of contrastive vowel nasalization due to retiming of oral and velic gestures (Kawasaki 1986; Solé 1992; Cohn 1993; Beddor 2009; Carignan et al. 2021). Other changes, such as final obstruent devoicing (Lombardi 1991) or the rhotacism of voiced fricatives (Catford 2001), have plausible aerodynamic sources. Yet there are still other sound changes for which unambiguous articulatory or aerodynamic precursors are lacking. In Section 2, we discuss a number of these putatively purely acoustically based changes and consider different accounts of their actuation.

Because stable phonetic variation can persist over long periods of time, many researchers draw a distinction between variation and change. One way to conceive of the difference between synchronic acoustic-phonetic variation and sound change is that change involves the dissociation of synchronically co-varying acoustic cues which are both mechanically linked in production and *parsed* together in perception: "the sound change [a change in the pronunciation norm] is initiated when listeners cease to parse together phonetic cause and effect" (Ohala and Busà 1995, 3). For the purposes of this entry when we talk about 'sound change' we mean it in this sense: when a change in a production norm is accompanied by perceptual realignment. What 'parsed together' in perception means is a topic we will cover in more detail in Section 3 below.

It is frequently observed that sound change involves the exaggeration or *enhancement* of phonetic biases, either in terms of their magnitude or temporal extent. In Section 4, we discuss several definitions of enhancement and how they relate to sound change. Finally, in Section 5 we briefly review some of the main findings of the *perceptual learning* literature and consider the potential implications of this work for our understanding of sound change actuation.

2 Acoustic-perceptual bias factors in sound change

Historical linguists have often made a distinction between those changes to pronunciation having a clear phonetic conditioning *of some type*, and those that have no such obvious phonetic conditioning (Garrett and Johnson 2013; Garrett 2015). The set of changes with clear phonetic conditioning can be further subdivided by their putative phonetic origin, namely in terms of their articulatory, aerodynamic, or acoustic-perceptual properties. The development of contrastive vowel nasalization from historical VN sequences (Ohala 1975; Beddor 2009) is a classic example of a sound change with articulatory conditioning, while the spontaneous devoicing of voiced obstruents would be an instance of a change with a fundamentally aerodynamic source (the *aerodynamic voicing constraint*: Ohala 1983, 2011; Trouvain 2021).

In this section, we focus on the third type of changes: those which may come about due to perceptual similarity (symmetric or asymmetric) between speech sounds, due either to acoustic similarities between the intended and perceived utterances and/or biases inherent to the human perceptual system (both instances of CHANGE in the typology of Blevins 2004). While all sound change presumably requires perceptual reconfiguration at some point, whether due to error, cue reweighting, or gestural reparsing (see Section 3 below), here we focus on changes where there is no obvious articulatory or aerodynamic antecedent giving rise to acoustic variability. As the existence of this category has been called into question in a number of cases (Garrett and Johnson 2013; Recasens 2015), here we briefly review several of the most prominent examples of changes where 'purely' acoustic-perceptual factors have been implicated: velar palatalization (2.1), *th*-fronting (2.2), nasalization of aspirates (2.3), and the emergence of labiovelars (2.4). For further examples and discussion see: Blevins (2004, 2015, 2019); Lehnert-Lehouillier (2010, 2013); Garrett and Johnson (2013); Recasens (2015) and references therein.

2.1 Velar palatalization

Velar palatalization describes the process by which *k > tʃ (also *g > dʒ), typically before front vowels. Several examples are given in (1)–(3).

(1) Palatalization in Old English
OE *cinn* [kin] > MnE *chin*
OE *ciese* [kiːese] > MnE *cheese*

(2) Palatalization in Romance
Latin *centum* [ˈkɛntum] > Italian *cento* [ˈtʃɛnto]
Latin *cantare* [kanˈtaːre] > Old French *chanter* [tʃanˈter] (cf. Engl. *chant*)

(3) Slavic 'First Regressive Palatalization'
PIE _*kers-_ > OCS *černý*, Cz. *černý*, Ru. черный 'black' (cf. Skst *kr̥ṣṇa* 'Krishna')
PIE _*k^wetwer-_ > OCS *četyre*, Cz. *čtyři*, Ru. четыре 'four' (cf. Lat. *quadra-*)
PIE _*g^wḗn_ > OCS *žena*, Cz. *žena*, Sb. жена (cf. MnE *queen* < OE *cwēn*, Ger. *Königin*)
PIE _*g^wiH_3wo-_ > OCS *životŭ* 'life', *živŭ* 'alive', Cz. *život*, Ru. жизнь
(cf. MnE *quick* < OE *cwicu* 'alive')

Notably, this process is asymmetric: [ki] > [tʃi] is common, but the converse [tʃi] > [ki] is unattested (Guion 1998). As discussed by Guion as well as Garrett and Johnson (2013), gestural blending (Section 3.1) is implicated in velar palatalization inasmuch as the production of [k] will be affected by the front vowel target. Still, articulatorily fronted [k^{j}i] is both articulatorily and aerodynamically quite different from an alveopalatal affricate, and it is not clear how the incremental change from the use of tongue dorsum to tongue blade as the active articulator, together with the addition of a fricative release, are to be motivated in purely articulatory terms (Ohala 1992).

Evidence for this change having a fundamentally acoustic-perceptual basis comes primarily from speech perception studies in which a similar asymmetry is observed. Following up on suggestions by Ohala (1989, 1992), Guion (1998) showed that noise-degraded tokens of [ki] were often misheard as [tʃi], but the converse was not true. Guion's findings built on earlier work by Winitz, Scheib, and Reeds (1972) in which listeners identified plosive bursts extracted from [pi] [ti] [ki] sequences. They found that [k(i)] bursts were frequently heard as [t(i)], but not the other way round. Later work by Plauché, Delogu, and Ohala (1997) and Chang, Plauché, and Ohala (2001) showed that addition of a 3 kHz spectral peak (typical of /k/) to a /t/ spectrum did *not* lead to more /k/ percepts. They argued that the perceptual asymmetry arises because while listeners may fail to perceive the spectral peak characterizing /ki/, they are unlikely to 'introduce' it when confronted with an instance of /ti/. They further propose that the fact that /ki/ tends to become [tʃi] and not [ti] rests on a reanalysis of the aspiration noise as (af)frication (/ki/ > [kʰi] > [tʃi]; cf. Section 2.5).

Garrett and Johnson (2013) problematize the acoustic-perceptual account by (i) noting the absence of studies of ongoing changes that document a clear [kʲi] > [tʃi] shift without any intermediate variants, and (ii) suggesting that the affrication step could be the result of listener-oriented enhancements on the part of the speaker (see Section 5). Recasens (2015, sec. 1.2.3) also suggests that an articulatory account, whereby /k/ > /tʃ/ via an intermediate realization of [c], may be equally or more plausible.

2.2 *th*-fronting

Unconditioned change of [θ] > [f] (also [ð] > [v]) is a relatively common sound change attested in many English and Scots dialects (Wells 1982) as well as in Veneto Italian (Blevins 2006) and some Oceanic (Blust 2009), Semitic, and Athabaskan languages (Blevins 2019). Some examples are given in (4a-d).

(4) Examples of *th*-fronting
  a. Gothic *þliuhan* > OE *flēon* 'flee' (Jones 2002; disputed)
  b. Cockney English *thin* [ˈfɪn], *brother* [ˈbrʌvə] (Wells 1982); also many British English varieties (Britain 2005) including Scottish English (Stuart-Smith and Timmins 2006; Schleef and Ramsammy 2013)
  c. South Slavey [θa] > Mountain Slavey [fa] 'sand', [θaj] ~ [fa] 'tent pole', [θɛ]- ~ [fɛ]- PERFECTIVE (Flynn and Fulop 2014)
  d. Veneto Italian [θémena] ~ [femena] (cf. Ital. *femmina* 'woman'); [θonc] ~ [fonc] (cf. Ital. *fungo* 'mushroom') (Mackay 1995, xvii)

Observers as early as Sweet (1874, 470) cast doubt on incremental articulatory explanations for these correspondences ("the not unfrequent change of *th* into *f* is no doubt purely imitative"). A change in active articulator (from tongue tip to lower lip) would appear to rule out incremental shift or gestural overlap, and the typically context-free nature of the change makes coarticulation an unlikely source. The fact that [θ] is asymmetrically misperceived as [f] (Miller and Nicely 1955; Harris 1958; McGuire and Babel 2012) points toward an acoustic-perceptual trigger. However, unlike velar palatalization, most phonetic studies have found relatively symmetric spectral similarities between the two sounds (Hughes and Halle 1956; Heinz and Stevens 1961; Tabain 1998; Jongman, Wayland, and Wong 2000), leaving the typological absence of *f > θ changes unexpected.

One possible explanation for this asymmetry is that visual cues may play an important role in [θ-f] confusions (Miller and Nicely 1955; Jones 2002; McGuire and Babel 2012). McGuire and Babel (2012) conducted a multimodal perception experiment in which participants identified tokens of /f/ and /θ/ in CV, VC, and VCV contexts in three vowel contexts. Participants were assigned to one of three blocks: one in which they received only audio stimuli, one in which they saw only video stimuli, and one in which they both heard and saw the stimuli. Participants were more sensitive to the contrast in the audio-visual condition than in both the audio-only and visual-only conditions, and responses for /θ/ were more variable than those for /f/, presumably due to the greater visual salience of the /f/ articulation. This supports the suggestion of Jones (2002) that listeners/learners may associate the visible lip movement of /f/ with weak frication, and summarily deploy it as an articulatory strategy to achieve the spectrally similar frication of /θ/.

Garrett and Johnson (2013, 71–72) again question the categorical absence of an articulatory precursor, suggesting that an intermediate [θ^w] stage may be involved. Although this is on the face of it similarly difficult to justify in purely incremental articulatory terms, Flynn and Fulop (2014) argue that the labialization enhances an acoustic cue ⟦grave⟧ to interdentals, so that this would be a case of actuation by auditory *enhancement* (Section 4.4), rather than the cause being fundamentally acoustic in nature. For further discussion see Garrett and Johnson (2013) and Blevins (2019).

2.3 Nasalization and aspiration

Perceptual origins have also been argued to underlie a number of sound changes involving 'spontaneous nasalization' in the vicinity of aspirated segments, where there is again no immediately obvious articulatory link. Matisoff (1975) describes a number of cases of allophonic nasalization of vowels following /h/ and glottal stop, which he terms 'rhinoglottophilia', including in Thai, Lisu, and English; other examples include Hayu (Michailovsky 1975) and Arakanese (Bradley 1985). In addition to /h/, voiceless fricatives and aspirated plosives are also implicated in spontaneous nasalization, for example, Hindi [sãp] from Sanskrit *sarpa* 'snake' or Hindustani *ā̃kh* < MIA *akkhi* 'eye' (Grierson 1922, 383). The reverse process, where a nasal is lost in the environment of a voiceless fricative, is also attested: compare German ∼ English pairs like *Gans* ∼ *goose* or *Mund* ∼ *mouth* (but *Hund* ∼ *hound*). There are also cases where NC sequences become aspirated Cs (e.g., Sprigg 1987; Hill 2007 on NC>C^h in Tibetan varieties).

The acoustic-perceptual basis for spontaneous nasalization can be traced to the fact that the lowered amplitude and increased F1 bandwidth of vowels near voiceless fricatives acoustically mimic nasalization, due to the sub-glottal cavity functioning like a branched resonator in a similar way that the nasal cavity does in the production of nasalized sounds. Similarly, nasal effacement (as in *goose* from *Gans*) may come about when listeners parse the nasalization as a contextually predictable coarticulatory effect and discount it entirely (see Section 3). Experimental evidence supporting this account comes from Ohala and Busà (1995), who showed that /-VNC/ sequences are judged as /-VC/ more often when C is a voiceless fricative than when it is a voiced fricative or voiceless plosive.

However, the similar acoustic effects of nasalization and aspiration also make it challenging to determine their articulatory source from acoustic data alone. Johnson (2019) studied spontaneous nasalization in Thai using a combination of electroglottography, oral airflow, and high-speed MRI. She found that syllables with /h-/ onsets are produced with greater velum opening than syllables with glottal stop onsets, and that nasal airflow during /hV/ syllables is greater during the /h/ than the following V. At least in Thai, then, velopharyngeal underspecification of /h/ – an articulatory bias – may be the source of spontaneous nasalization, with the nasal percept enhanced by the acoustic consequences of breathiness (Klatt and Klatt 1990).

2.4 Asymmetric sound changes involving labiovelar approximants

Ohala and Lorentz (1977) note that when /w/ becomes a nasal or when it functions as a target of place assimilation for a nasal, it nearly always behaves as a velar, rather than as a labial segment; but when it becomes a fricative or functions as a place assimilation target for a fricative, it nearly always behaves as a labial. Their article provides a large number of references; we provide just a few examples of /w/ behaving like a velar when alternating with a nasal (5) and /w/ behaving as a labial when alternating with a fricative (i.e., spirantization; 6a–c).

(5) Place assimilation of nasal in causative prefix /bɑN-/ in Khmer
(/N/ realized as [ŋ] before velars but also /h r ʔ w/)
ប៉ោង /paoŋ/ ‘to be swollen’ បំប៉ោង [bɑ*m*paoŋ] ‘to inflate’
កើត /kaət/ ‘to be born’ បង្កើត [bɑ*ŋ*kaət] ‘to give birth’
វិល /wəl/ ‘to revolve’ បង្វិល [bɑ*ŋ*wəl] ‘to turn something around’

(6) Examples of /w/ behaving as a labial when becoming or alternating with a fricative
a. Cham: /w/ alternates with /v/ [rəwəŋ – rəvəŋ] or with /β/ [βai̯ʔ ~ wai̯ʔ] (Blood 1967)
b. Jeh: /w/ alternates with /β/ syllable-initially (Gradin 1966)
c. Hungarian: bilabial *w became labiodental /v/ in Western dialects around 13c. (Kálmán 1972, 55): **wāšār* > *vāšār* <vásár> ‘market’, **köwek* > *kövek* <kövek> ‘stones’

As Ohala and Lorentz (1977) explain, both patterns have relatively straightforward acoustic explanations. The first case (/w/ covarying with velar, rather than bilabial nasals) has to do with nature of resonances and anti-resonances of the vocal tract (Fant 1960). Figure 1 illustrates the different schematic vocal tract configurations for /m n ŋ w/. The frequencies of the anti-resonances which are the primary acoustic feature distinguishing place of articulation in nasals depend on the length of the oral cavity which functions as a branching resonator of the pharyngeal-nasal airway. While labiovelar /w/ has two constrictions, it is only the back (velar) constriction which determines the length of the branching tube. As a result, the acoustic profile of (nasalized) /w/ is more similar to /ŋ/ than to any other nasal.

Ohala and Lorentz list several factors which might contribute to the tendency for /w/ to act as a labial rather than a velar fricative, the most important of which is that multiple constrictions are involved in the articulation of /w/. When /w/ is fricativized,

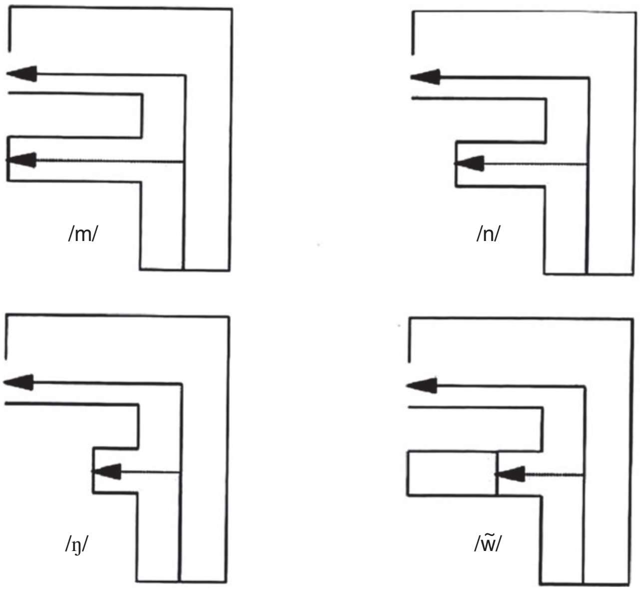

Figure 1 Schematic representation of vocal tract configurations for /m/, /n/, /ŋ/ and (nasalized) /w/. In each image, it is the lower arrows, showing the distance from the pharynx to the point of oral constriction, which contribute different resonances to the four sounds. From Ohala (2005, figure 3).

frication noise will be generated at two sources: the labial constriction and the velar constriction. It is the spectral qualities of the frication noise that provide the acoustic clues to the place of the constriction. In the case of fricativized /w/, the noise generated at the velar constriction will be attenuated by having to pass through the narrow labial constriction, but the noise generated at the downstream labial constriction will not be so attenuated, making these more perceptible. A corollary of this account is that labiovelar plosives should be more likely to develop into labial, rather than velar plosives, which finds support in sound changes such as PIE $*k^{w}$ > p in Greek *hippos* 'horse' (cf. Lat. *equus*).

2.5 Do changes ever have a purely acoustic-perceptual basis?

As noted in the preceding sections, Garrett and Johnson (2013) have questioned the existence of this entire category of sound changes (but cf. Recasens 2015 and especially Blevins 2019). Their primary argument centers on the lack of studies conclusively establishing the absence of articulatorily intermediate variants. Even setting aside this objection, however, as the preceding examples have made clear, it is not always possible to isolate acoustic-perceptual factors as the unambiguous and sole locus of a sound change. To take one final example, it has long been noted that in the Phnom Penh dialect of Khmer (Cambodian), words with /Cr/ onset clusters in the standard/literary language are produced without /r/ but with a low or falling pitch (Noss 1966; Huffman 1967; Wayland and Guion 2005; Kirby 2014a). Wayland and Guion (2005) showed that the loss of /r/ was accompanied by aspiration, a falling-rising pitch contour, and a change in vowel quality. Kirby (2014a) showed further that there were significant differences in spectral tilt, and that while aspiration, voice quality, and above all F0 differences contributed to the perception of the contrast, vowel quality (height) did not.

Wayland and Guion (2005) argued that the catalyst for this change was fundamentally aerodynamic. They proposed that devoicing of /r/ conditioned a drop in transglottal airflow and, as a result, reduced rate of vocal fold vibration and a resultant falling F0 contour, which may have subsequently been reanalyzed by listeners as a falling tone. Thus, at its core, this would be a non-perceptual account of actuation. However, Wayland and Guion's account leaves the vowel quality differences unexplained, and while they may not be parsed as cues to this contrast, they are extremely acoustically salient. On this basis, Kirby (2014a) argued that the explanation is fundamentally acoustic in nature: aspiration resulting from the devoicing of the trill may have created a percept of breathy voicing on the following vowel, which conditions both the percept of F1 lowering (resulting in vowel raising) and F0 (resulting in pitch lowering). That being said, the trajectory of this change in other Khmer dialects points to a more nuanced interaction between articulatory, aerodynamic, and acoustic-perceptual bias factors; see Kirby and Giang (2017) and "Tonogenesis and the Evolution of Tone Systems" (Section 3.1).

3 Perception, production, and gestural parsing

As defined in Section 1, sound change is initiated when a consistent change in a production norm is accompanied by realignment in perception. In Section 2, we have reviewed changes in which both phonetic cause and effect are arguably acoustic-perceptual, but in many more cases, regular sound change involves some form of *coarticulation* whereby the gestures involved in the production of two speech sounds influence one another in time, leading to changes in the *parsing* of those gestures. To understand how changes in parsing can come about, in Section 3.1 we first provide a brief overview of how phonetic variation and sound change are linked in a gestural model of speech production. We then consider in Section 3.2 and Section 3.3 how perceptual parsing (and perhaps mis-parsing) of gestures can lead to sound change.

3.1 Articulatory phonology: gestural overlap, blending, and spatial reduction

An account of coarticulation as the overlap and blending of articulators from sounds in sequence goes back to Joos (1948) and the pioneering empirical research by Öhman (1966, 1967). These ideas have contributed to the foundations of Articulatory Phonology (Browman and Goldstein 1986, 1989, 1992; Pouplier and Goldstein 2010) which seeks to provide a unified explanation of diverse types of synchronic variation in connected speech, including assimilation, deletion, and weakening.

In Articulatory Phonology (AP), gestures are autonomous and can vary in their relative phasing. Figure 2 shows a gestural model of speech in which the gestures of glottal opening and a stop closure are variably phased. This model has been applied to a sound change in Andalusian Spanish in which pre-aspirated stops, characteristic of older and East Andalusian speakers, are becoming post-aspirated, typically in younger and West Andalusian speakers (O'Neill 2010; Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016; Ruch 2018; Herrero de Haro and Hajek 2022): thus, /pa^{h}ta/ →/pat^{h}a/ or /pat^{s}a/ ('pasta') which contrasts with an unaspirated,

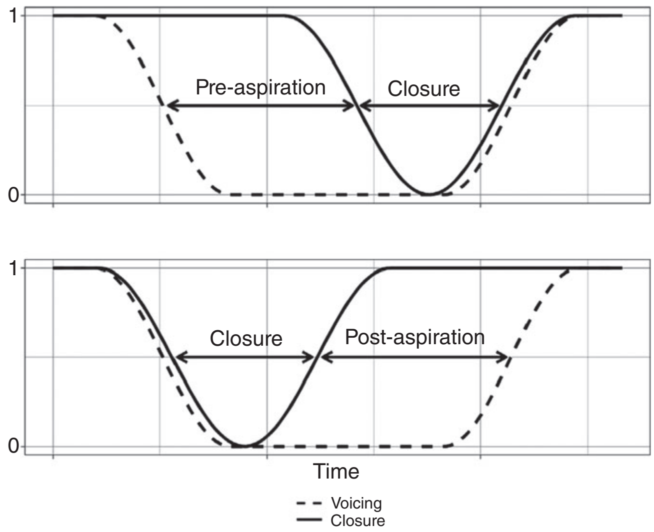

Figure 2 Idealized scheme of resynchronization of the closure with the voiceless interval in Andalusian Spanish /s/-aspiration. The solid line is the glottal gesture, where low values stand for an open glottis and hence voicelessness, the dashed line is the oral constriction gesture of the voiceless plosive, where the minimum of the curve indicates maximal closure. From Cronenberg et al. (2020, with permission of Elsevier).

intervocalic singleton stop /t/ (e.g., /pata/, 'paw'). There is also some evidence for a pre- to post-aspiration change synchronically in a faster speaking style (Parrell 2012). These synchronic and diachronic changes can be modeled as a rephasing of a closure gesture for /t/ relative to an unchanging open glottal gesture for voicelessness (Parrell 2012; Cronenberg et al. 2020). A progressively earlier phasing of the closure leads to less pre- and more post-aspiration. From this perspective, the articulatory prominences of pre- and post-aspiration stand in an inverse relationship to each other. This inverse relationship can then form the basis for a so-called perceptual trading relationship (Repp 1982) between these two cues (see Section 3.3) that are connected by gestural phasing.

Three further examples in which synchronic variation and sound change are linked by the variable phasing of different sets of articulator gestures include:

- *Increasing anticipatory coarticulatory nasalization* (Moll and Daniloff 1971; Kent, Carney, and Severeid 1974; Cohn 1990, 1993; Bell-Berti and Krakow 1991; Solé 1992, 1995; Delvaux, Metens, and Soquet 2002; Delvaux et al. 2008) in VN rhymes (e.g., /æn/ in *pan*, *pan's*, *pant*, *panned*) has been modeled as a velum gesture that is stable in both time and space but that is phased earlier with respect to the tongue dorsum gesture for the V (Beddor 2007, 2009, 2012; Beddor et al. 2013). This earlier phasing has been argued to form the basis of the sound change by which the V can be completely nasalized as the following N weakens and is lost (e.g., the development of French *son* 'sound', /sɔ̃/, from Latin *sonus;* cf. Italian *suono*).
- The *deletion or partial deletion of consonants in clusters* such as the /t/ in *perfect memory* has been modeled by an earlier phasing of the labial gesture for /m/ which overlaps and therefore hides the tongue tip gesture of /t/ (Browman and Goldstein

1990). Synchronically, there is evidence from Georgian (Chitoran, Goldstein, and Byrd 2008) that speakers avoid a too-close phasing of C_1C_2 clusters when the place of articulation of C_1 is further back than that of C_2 precisely because the release of C_1 would be masked perceptually by an early phasing of the C_2-closure. This gestural hiding in clusters is the likely physiological origin of sound changes such as Italian *notte* < Latin *noct-* 'night' or *immaculate*, Latin *immaculatus* < *in* + *maculatus* 'spotted, stained'.

- *Oral stop insertion* (Ohala 1997), for example, the variable production of *warmth* with or without a /p/ corresponds to variation in the gestural phasing of the velum and stop closure gestures: a voiceless bilabial stop /p/ is produced if the stop release is late relative to the time of velum lowering for /m/. Sound changes associated with this type of variation include labial insertions in, for example, *empty* (< Old English 'amtig'), *nimble* (< Old English 'næmel') and in French *chambre* (cf. Italian and Latin *camera*), *humble* (< Old French *umele*).

The preceding examples are of synchronic–diachronic connections for gestures involving *different* sets of articulators. When a sequence of sounds is produced with predominantly the *same* articulator, then they are in competition: the outcome can be some form of blending or averaging (Fowler and Saltzman 1993; De Jong, Beckman, and Edwards 1993; Romero 1996; Recasens 2019). Some examples of synchronic and diachronic variation due to gestural blending include:

- */u/-fronting* (Bauer 1985; Hawkins and Midgley 2005; Labov, Ash, and Boberg 2006; Fridland 2008; Harrington, Kleber, and Reubold 2008; Kataoka 2011; Alderton 2020), in which (for example, German *tut* 'does') opposing forces act on the tongue dorsum for its retraction due to the back vowel /u/ and its advancement due to the tongue tip articulations of the consonants (Öhman 1966; Flemming 2003, 2004). /u/-fronting as a sound change in progress is well documented and could be associated with the relatively high frequency of /u/ in a fronting context (*feud*, *tune*, *viewed*, *too*, etc.: e.g., Harrington 2007).
- */s/-retraction* in clusters in English varieties with a post-alveolar approximant (e.g., *street*) which derives from competing demands on the tongue tip between alveolar-dental and post-alveolar places of articulation. This synchronic variation has become a sound change in progress in some American (e.g. Shapiro 1995; Lawrence 2000; Baker, Archangeli, and Mielke 2011; Rutter 2011; Smith et al. 2019; Gunter, Vaughn, and Kendall 2021), Australian (Stevens and Harrington 2016), and New Zealand (Warren 2006) English varieties.
- Synchronic *transconsonantal (VCV) vowel coarticulation* is well documented (Öhman 1966; Cole et al. 2010; Hoole and Pouplier 2017). An associated sound change is metaphony in some varieties of Italian (Savoia and Maiden 1997; Loporcaro 2016) and umlaut in German (Twaddell 1938; Penzl 1949; Iverson and Salmons 2003; Kiparsky 2015) as in, for example, /mɛçtɪç/ *mächtig* 'mighty' derived from Proto-Germanic **mahtiga-*. The articulatory competition in this example from German is between the low and high jaw positions for /a/ and /i/ respectively resulting in an /ɛ/ mid vowel in the stem of the present-day form.

AP also models spatial reduction through various mechanisms such as gestural rescaling and truncation (Harrington, Fletcher, and Roberts 1995; Carignan et al. 2021). Synchronically, spatial reduction is shown by the weakening and lenition of consonants often in prosodically weak and intervocalic positions (Beckman et al. 1992) and diachronically by the lenition of intervocalic stops (Hualde, Simonet, and Nadeu 2011; Torreira and Ernestus 2011) and fricatives (Hualde and Prieto 2014) in Spanish, and of voiced stops in English (e.g., English *daisy* < Old English *dæges eage* 'day's eye').

3.2 Gestural parsing: from articulatory variation to categorical change

There is, however, more to sound change than variation in gestural phasing and magnitude: at some point there must be a change in how gestures are categorized. This point is explicitly made by Browman and Goldstein (1991, 327) in their analyses of some sound changes that "cannot be completely accounted for by the principles of gestural reduction and increase in overlap." In one of their examples by which /luːs/ changed to /lyː/ in Tibetan, they comment that "[a]s the alveolar fricative gesture is deleted, the constriction location of the tongue body gesture is recategorized as palatal rather than velar." Browman and Goldstein (1991) suggest that recategorization in such cases comes about because listeners fail to recover the speech production gestures from the speech signal. There is a similar perceptual interpretation in Beckman et al.'s (1992) analysis of changes due to prosodic weakening: they suggest, for example, that the diachronic reduction of 'chocolate' to two syllables may have come about because the medial weak vowel comes to be misinterpreted by the listener as part of the release of the /k/.

This idea that sound change is driven by the listener's parsing error is central to Ohala's (1981, 1993b, 2005, 2012) conception of sound change. As discussed above, in some varieties of English the /s/ in /str/ clusters tends toward post-alveolar because of the anticipatory coarticulatory influence of the retracted tongue-tip for /r/. Acoustically, tongue tip retraction causes a lowering of the spectral center of gravity of /s/ (Stevens 1971; Jongman, Wayland, and Wong 2000; Koenig et al. 2013). This acoustic shift has the consequence that when an /s/ is decontextualized by excising it acoustically from words like *string*, then it sounds slightly more like /ʃ/ (Stevens and Harrington 2016) given that spectral center of gravity lowering is also one of the main cues for the /s/ ~ /ʃ/ distinction (Stevens 1971). But listeners according to Ohala's (1993b) model would typically perceive the same /s/ in *string* and in words like *sting* for which there is no post-alveolar context. This is because in perceiving *string*, it is assumed that listeners perceptually normalize for context (Mann and Repp 1980): that is, they attribute the spectral center of gravity lowering to the anticipatory influence of the /r/ and factor perceptually its contextual influence from /s/ (see Fowler and Smith 1986; Fowler and Thompson 2010 for an analogous explanation in terms of the direct perception of gestures). If, according to Ohala's model, listeners fail to normalize for context, then the coarticulatory influence is no longer attributed to the post-alveolar approximant and so becomes attached to the /s/. In AP terms, the gestural score under such circumstances defining the place of articulation changes categorically from alveolar to post-alveolar.

The categorical change in the above example of the /s/ → /ʃ/ sound change in *string* comes about because a listener *under*-compensates, that is, insufficiently normalizes for context. By contrast, in sound changes due to dissimilation (Alderete and Frisch 2007; see Abrego-Collier 2013; Harrington, Kleber, and Stevens 2016; Jatteau and Hejná 2016 for related experimental analyses; see also "Assimilation and Dissimilation Processes in Sound Change"), listeners might *over*-compensate for coarticulation: that is, they attribute too much of the variation in the speech signal to the coarticulatory source. For example, in the sound change from pre-Shona **kumwa* to Shona /kumɣa/ (Guthrie 1967; Ohala 1981; Browman and Goldstein 1991), the lip-rounding in the second consonant of the medial cluster is, according to Ohala's model, incorrectly attributed by the listener to a coarticulatory carryover effect from /m/ and so is factored out from the labial-velar /w/ leaving a plain velar /ɣ/.

The emergence of sound change from phonetic variation requires not only the listener to make a mistake in gestural parsing (which according to Ohala happens only rarely), but also for the listener to subsequently reproduce the perceived error (e.g., to produce *string* with /ʃ/) and for this to be copied by others. The model therefore correctly predicts that while phonetic variation is ubiquitous, sound change – requiring a succession of improbable stages from (i) a parsing error by the listener error to (ii) the replication of the error in the production of the same listener to (iii) the copying of this error by others – is very unlikely to occur.

3.3 Sound change and listener variation in parsing strategies

A Necker cube and Rubin's vase are visually bistable because in both the viewer flips categorically between two possibilities (e.g., different 3D-perspectives in the former; either a vase or two faces in the latter). Just this type of categorical switch is suggested to the listener on the brink of parsing correctly or incorrectly the speaker's gestures in ambiguous speech production signals in Ohala's model: thus, the listener flips between hearing a non-retracted or retracted /s/ in *string*. From this it follows that the emergence of sound change from phonetic variation is *abrupt* in Ohala's model (see Ohala 1993b): in the *string* example, one variant [s] is exchanged for a categorically different variant [s̠] and is not *gradual* (in which [s] within the same speaker morphs over time into [s̠]).

In other models by contrast, notably those of Beddor and colleagues (Beddor et al. 2013, 2018), as well as those based on exemplar theory (Pierrehumbert 2001; Todd, Pierrehumbert, and Hay 2019), sound change is predicted to advance incrementally rather than due to a switch between categorically different variants. In exemplar models, the possibility of sound change exists because speech production samples from the same exemplars (or a subset thereof) that are memorized and updated in perception (Wedel 2004, 2006, 2007; Todd, Pierrehumbert, and Hay 2019). Thus, if an individual is exposed to a new dialect over a prolonged period of time, a certain degree of sound change within that individual is very likely because exposure to the new dialect updates the (stochastic) relationship between words and phonological categories that are also input to speech production (see, e.g., Sancier and Fowler 1997; Sankoff 2004; Harrington, Palethorpe, and Watson 2000; Leeuw 2019; Riverin-Coutlée and Harrington 2022 for compatible evidence). In contrast to Ohala (1993b), sound change in exemplar models is also not caused by any kind of perceptual error but is instead a byproduct of experience, that is, of absorbing new exemplars in memory.

As in exemplar theory, sound change in Beddor's model is incremental and not driven by any kind of perceptual error. However, in contrast to exemplar theory, the progression from continuous phonetic variation to a categorical change in Beddor's model is more directly grounded in the perceptual processing of coarticulation and overlapping articulatory gestures. Beddor's model is especially relevant for a particular class of sound change whose outcome is *transphonologization* (Hyman

1976, 2013; Kiparsky 2015). These are sound changes in which cues are transferred from a coarticulatory source to a coarticulatory effect, culminating in the loss or near-loss of the source. They include umlaut and metaphony, which have their origins in trans-consonantal vowel coarticulation (see Section 3.1); the compensatory lengthening of vowels due to loss of following consonants (Kavitskaya 2001); and (the motivating example for Beddor's model) the phonologization of nasalization in VN sequences in which an oral-nasal contrast develops in the vowel with potentially the complete loss of the N (see Section 3.1).

The perceptual mechanisms in Beddor's model that provide the foundations for phonetic variation to develop into sound change are firstly that listeners are attuned to, and make use of, coarticulatory information for processing the speech signal into categories (Martin and Bunnell 1981; Alfonso and Baer 1982; Connine and Darnieder 2009; Beddor et al. 2013); and secondly that listeners often only compensate for coarticulation *partially* (Fowler and Brown 2000; Beddor, Harnsberger, and Lindemann 2002; Beddor 2009; Yu and Lee 2014; Zellou 2017; Zellou, Barreda, and Ferenc Segedin 2020; Zellou, Cohn, and Block 2021). This idea that listeners compensate partially for coarticulation is a departure from Ohala's model in which compensation for coarticulation takes place completely (in which case there is no sound change) or not at all (leading potentially to sound change). In their analysis of perceived nasalization of vowels in an NVN context, Beddor and Krakow (1999) reasoned that if listeners compensated completely for coarticulation, then a nasalized V perceived in an N_N context should sound completely oral (if listeners attribute all the nasalization to the N context). But Beddor and Krakow (1999) shows that some listeners nevertheless hear the vowel as slightly nasal in the N_N context, which points to a partial compensation: that is, not all the nasalization is parsed with the N contexts that gave rise to it.

Partial compensation on its own need not lead to sound change. The conditions for phonetic variation to turn into a categorical change are instead more likely to be met when a perceptual *trading relationship* (Repp 1982; Pisoni and Luce 1987) begins to develop between the coarticulatory source and effect. A trading relationship is a mechanism in speech perception by which listeners can flexibly weight perceptual cues in relation to their strength in the speech signal (Chandrasekaran, Sampath, and Wong 2010; Beddor 2012, 2015; Schertz et al. 2015; Clayards 2018; Kim and Clayards 2019). With regard to a stop voicing contrast for example, studies show that if the voice onset time (VOT) cue is compromised, then listeners pay increasing attention to F0 cues at the onset of the vowel for identifying the distinction between voiced and voiceless stops (Whalen et al. 1993). The perceptual catalyst for a categorical sound change is that listeners pay greater attention to one of the cues (e.g., F0) and, because of the trade-off relationship, increasingly less to the other cue (VOT) with which it trades. Applied to nasalization, an increasingly skewed trading relationship of this kind means that listeners pay almost exclusive attention to coarticulatory nasalization in the V and virtually none to the (inherent) nasalization in the following N.

For the sound change to progress, a correspondingly skewed relationship must obtain not only in perception but also in production. The development of an inverse relationship between coarticulatory effect and source is demonstrated in the various analyses of Beddor and colleagues by comparing contexts in which this sound change is more or less likely to occur. As various studies have shown, the phonologization of nasalization is more likely when the nasal precedes a voiceless consonant in VNC̥ contexts like *bent* than a voiced VNC̬ context like *bend* (Ruhlen 1978; Ohala and Ohala 1991; Tuttle 1991; Hajek 1997; Sampson 1999; Busà 2007). This is because nasalization and a following voiceless consonant are less compatible: nasalization causes the presence of low-frequency energy (House and Stevens 1956; Fujimura 1962), whereas the perception of a voiceless consonant requires its absence (Ohala and Ohala 1991, 1993). This acoustic incompatibly is likely to be one of the factors that contributes both to the relative scarcity of NC̥ clusters in the languages of the world (Itô and Mester 1986; Pater 1999; see Carignan et al. 2021 for a more detailed review), as well as to the diachronic loss of nasal consonants before voiceless consonants, especially voiceless fricatives with high-frequency energy (e.g., English *institute*, Italian *istituto*; German *Insel*, Latin *insula*, but English *island* and Italian *isola*; German *Wunsch,* English *wish*, etc.). Beddor's studies show that there is a greater skewed relationship toward the coarticulatory effect in VN preceding voiceless than voiced consonants in both perception and production (Beddor 2012; Beddor et al. 2018). Thus, an eye-tracking study (Beddor et al. 2013) shows that listeners pay much more attention to nasalization in the vowel for distinguishing for example, *sent* from *set* than *send* from *said*. Compatibly, the duration of vowel nasalization in production is greater while that of the N is less in words of the 'sent' than the 'send' type (see also Carignan et al. 2021 showing that the size of the N gesture is less in German voiceless NC̥ than voiced NC̬ clusters).

Some studies have also investigated whether an analogous relationship between perception and production exists at the level of the individual. There is some evidence that 'innovative' individuals at the forefront of sound change are both more sensitive perceptually to the presence of nasalization in the vowel and have more extensive vowel nasalization in their own productions. This relationship at the level of the individual has been demonstrated for (i) coarticulatory nasalization by Beddor et al. (2013, 2018) and by Zellou (2017), (ii) the extent to which F0 and Voice Onset Time co-vary with respect to an ongoing tonogenetic sound change in two varieties of Afrikaans by Coetzee et al. (2018, 2022), and (iii) in a study by Yu (2019) concerned with the relationship between categorizing /s/ and /ʃ/ in perception in two vowel contexts and measuring the contextual influence of same two vowels on the sibilants in production. On the other hand, this type of production-perception relationship at the level of the individual was found neither in Grosvald's (2009) study of long-distance anticipatory vowel-to-vowel coarticulation nor in Kataoka's (2011) analysis of the coarticulatory influence of /t/ on /u/-fronting. In a recent comprehensive survey of cue weighting in perception and production, Schertz and Clare (2020, 18) conclude that, while the sensitivity in perceiving cues is well matched to the extent to which they are manifested in production at a *group* level, "it has proven difficult to understand the relationship between them on an individual level. Even those studies that have found evidence of a direct relationship did not find it in all of the expected places" (see also the review of sound change in Harrington, Kleber et al. 2019 for a similar conclusion). Thus, evidence for production-perception parity is at best mixed. That having been said, detecting individual-level correlations requires extremely high statistical power; if the link between the modalities is weak, it may not be detectable by small, underpowered studies.

4 Enhancement

A persistent puzzle in the study of sound change is explaining why phonetic biases, acoustic-perceptual or otherwise, do not inevitably or inexorably lead to sound change.

At some stage, some intrinsic phonetic biases become *phonologized*, that is, "exaggerated to such an extent that [they] cannot be entirely predicted on the basis of the universal effect" of the bias (Hyman 1976, 408). For instance, in many languages, F0 is perturbed upward following syllable-initial voiceless obstruents (House and Fairbanks 1953; Lehiste and Peterson 1961; Kohler 1982; Hanson 2009; Chen 2011; Dmitrieva et al. 2015; Xu and Xu 2021). In languages where this intrinsic effect has been phonologized and where it constitutes the sole or primary cue to the laryngeal contrast, such as Afrikaans (Coetzee et al. 2018), Eastern Kmhmu' (Svantesson and House 2006; Kirby, Pittayaporn, and Brunelle 2023), or Malagasy (Howe 2017), the magnitude and temporal extent of the effect is far greater than in languages where it has not been phonologized. What is less clear is what, exactly, underlies this exaggeration. One line of explanation which is frequently invoked is that of *enhancement*, but this term is used to mean slightly different things by different researchers. In particular, the target of enhancement is not always made explicit: Are speakers aiming to exaggerate particular gestures, acoustic targets which are diagnostic of a particular category, or something more abstract, like the contrast itself? In this section, we review several understandings of enhancement (featural, auditory, articulatory, adaptive), with the aim of orienting readers toward the different uses of the term in the literature.

4.1 Phonetic (featural) enhancement

In many contexts, enhancement refers to enhancement of a contrast: when the phonetic cues to a phonological contrast are for whatever reason not sufficiently perceptually recoverable, additional cues to the contrast may be produced. Stevens and Keyser (1989) use the term *phonetic enhancement* to refer to changes in pronunciation which highlight the 'phonetic essence' of a distinctive feature. This typically takes the form of a secondary, enhancing gesture being superimposed on the primary, defining gesture. For example, lip rounding is regarded as an enhancing feature applied to the production of /ʃ/, because it contributes to a lowering of F3, thereby increasing the acoustic distance between /s/ and /ʃ/ (what Stevens and Keyser regard as an enhancement of the feature [-anterior]). Phonetic enhancement can also involve the introduction of a new, potentially context-sensitive acoustic property unrelated to the acoustic property of the defining feature. On this view, the co-intrinsic pitch perturbations which accompany the production of [-voice] consonants in many languages are an enhancement of the defining feature [+stiff vocal folds] (Stevens and Keyser 2010). Prenasalization of [+voice] consonants may be similarly regarded as an enhancement of the [voice] contrast (Keyser and Stevens 2006).

The essential aspects of this type of enhancement are (i) that it targets a *contrast*, not a particular articulatory or acoustic dimension and (ii) it presupposes a dichotomy on which all features are either universally primary or universally secondary. Other approaches, such as Hall's (2007, 2011) *contrast and enhancement theory*, relax this second assumption. On Hall's approach, phonological features in a given language are organized hierarchically (Dresher 2009), meaning that the same phoneme may be specified for a given feature in one language but not another. For example, there are two possible feature hierarchies for the inventory /i a u/: in one, /a/ is specified as [+low, +back], but in the other, /a/ has no specification for backness (Hall 2011, 13). Accordingly, [a] is only predicted to be a target of enhancing gestures if it is specified for [+back] in a particular language. In other words, enhancement is still conceived of as targeting phonological features, but their language-specific hierarchical organization provides a mechanism for predicting which feature will be enhanced in a particular language (cf. Section 4.4).

A related conception is that of Avery and Idsardi (2001), who regard enhancement as the introduction of a non-contrastive privative feature. For instance, the plosives /b d ɡ/ in a language like Japanese would be specified for something like a feature [voice], with /p t k/ left unspecified; in a language like English, it would be /p t k/ that are specified for something like [spread], with /b d ɡ/ left unspecified. Enhancement would thus involve the introduction of (non-contrastive) glottal spreading to /p t k/ in Japanese, but glottal tension to /b d ɡ/ in English. Iverson and Salmons (2003) show how this notion of enhancement could explain the introduction of aspiration into the laryngeal contrast of early Germanic, which itself is thought to have provided the phonetic conditions for Grimm's Law (Iverson and Salmons 2003, 53 ff.).

4.2 Articulatory enhancement

Garrett and Johnson (2013) use enhancement to refer broadly to "processes by which a relatively small initial bias effect is amplified to its eventual categorical result" (2013, 78). They distinguish two types of enhancement, *articulatory* and *auditory*. Auditory enhancement, discussed in Section 4.3 below, involves the addition of a new gesture; Garrett and Johnson restrict articulatory enhancement to mean a shift in gestural magnitude and/or temporal realignment of an *existing* gesture. They give the example of a shift in the magnitude by which /u/ is fronted to [y] in a /uCi/ sequence (umlaut); the phonologization of co-intrinsic pitch in languages like Afrikaans (Coetzee et al. 2018) could also be classified as articulatory enhancement, because it targets the magnitude of an existing gesture (or gestural complex).

The Andalusian Spanish sound change mentioned in Section 3.1, whereby pre-aspirated plosives become post-aspirated, could be regarded as an example of articulatory enhancement by way of temporal realignment, because the aspiration feature has shifted to a more perceptually salient position (e.g., Ruch and Harrington 2014). The centralization of /aɪ/ > /ʌɪ/ before voiceless consonants characteristic of many varieties of English (Moreton 2021) provides another example of enhancement by means of temporal realignment: as argued by Moreton and Thomas (2007), this effect began as a phonetic bias affecting the offglide and only later spread to nucleus.

Although a contrastive distinction may be thus perceptually enhanced, the Garrett and Johnson conception of articulatory enhancement differs subtly from that of Keyser and Stevens: for Keyser and Stevens, enhancement can be change in pronunciation that aids in the recovery of a 'primary property', whereas for Garrett and Johnson enhancement is seen as applying to a particular gesture.

4.3 Auditory enhancement

The term *auditory enhancement* seems to be used in (at least) two senses. The first sense refers to the *Auditory Enhancement Hypothesis* of Kingston, Diehl, and colleagues (Diehl and Kluender 1989; Kingston and Diehl 1994; Kingston et al. 2008; Diehl 2008, 2011). The Auditory Enhancement Hypothesis proposes that sound systems tend to be characterized by combinations of articulatory gestures having mutually reinforcing effects on particular acoustic-auditory properties. Cues which reinforce a particular auditory effect are referred to as *enhancing*. For instance, the high back vowel /u/ is characterized acoustically by both a low F1 and low F2. Hyperarticulated tokens of /u/ will tend to be produced with a range of seemingly unrelated articulatory accompaniments, including lip protrusion, larynx lowering, and greater velar constriction, all of which have the effect of lowering F1 and/or F2 (Fant 1960). However, the observed range of seemingly disparate articulatory maneuvers deployed can be explained by the fact that they all contribute to increasing the salience of the acoustic properties most characteristic of the category. A similar example is provided by the various cues to the [voice] contrast. Low F1 and F0 near the constriction interval and glottal pulsing during the closure phase are all typical acoustic correlates of [+voice] segments (Stevens and Blumstein 1981), and have been demonstrated to trade with one another in cueing the perception of [+voice] segments (Kingston et al. 2008). On this approach, enhancement is fundamentally listener-oriented, in the sense that auditory distinctiveness is taken to be a desirable property of communicative signaling systems, and thus that sound systems which display such properties will be more common than those that do not.

Garrett and Johnson (2013) use the term auditory enhancement to refer to the introduction of a new articulatory feature with the aim of increasing the auditory salience of an existing contrast (2013, sec. 3.5.1). It differs from their articulatory enhancement (Section 3.2) in that it involves the introduction of a *new* articulatory feature, rather than exaggeration or realignment of an existing one (cf. Stevens and Keyser 2010, sec. 3). In addition to the classic example of lip rounding enhancing the salience of the [back] contrast in vowels (Liljencrants and Lindblom 1972), they mention several other possible examples, such as prenasalization as an enhancement to voicing and labialization enhancing properties relevant to the perception of [θ] (cf. Section 4.1 and Section 2.2, respectively). While closely related to Kingston and Diehl's notion of auditory enhancement, Garrett and Johnson remain agnostic as to whether auditory enhancements involve a functional, listener-oriented calculation on the part of the speaker, or whether they arise by chance but are more likely to be propagated because they are more likely to be retained in listeners' memories (cf. Section 4.5).

4.4 Probabilistic enhancement

Kirby (2010, 2013, 2014b), building on previous functional approaches including Lindblom (1990), Boersma (1998), and Flemming (2001), develops a listener-oriented notion of *probabilistic enhancement*, which predicts that talkers enhance particular acoustic dimensions based on their assessment of the listener's informational needs. This proposal bears many similarities to Lindblom's H&H theory (1990), which predicts that talkers hyperarticulate in contexts where the listener's informational needs are deemed to be high, but extends this to make (probabilistic) predictions about precisely *which* acoustic dimension(s) the talker will target based on a cue's distributional *informativeness* in signaling the phonetic contrast (see also Clayards 2008). Like Garrett and Johnson's notion of articulatory enhancement (Section 4.2), this involves strengthening of a contrast through emphasizing a redundant cue, but also provides a mechanism for predicting when actuation/initiation is likely: namely, in settings under which (due, e.g., to bias factors of the types discussed in Section 2) the perceptibility of the contrast is threatened. Examples of sound changes for which probabilistic enhancement provides a compelling account include the phonologization of F0 in Seoul Korean (Kirby 2013; Bang et al. 2018) and the phonologization of vowel duration and nasalized /æ/ in Australian English (Cox and Palethorpe 2014).

In terms of helping to understand how redundant cues are exaggerated, probabilistic enhancement posits that production biases provide the initial impetus which drives subsequent enhancement of secondary/redundant cues. However, the actual locus of enhancement is the speaker. The primary innovation of probabilistic enhancement is that it provides a mechanism for identifying which acoustic dimensions will be targeted for enhancement (in the sense of Section 4.2) and by how much they are predicted to be enhanced, as well as why other cues typically *aren't* targeted for enhancement (see also Garrett 2015). In the case of Seoul Korean, for example, probabilistic enhancement offers an explanation for why it is F0, and not some other cue to the laryngeal contrast, which is exaggerated when the VOT contrast is threatened: because F0 is the most distributionally informative cue dimension after VOT (cf. the idea that the directionality of mergers may be driven by differences between them in terms of the orientation of their variance components: Harrington et al. 2018).

4.5 Passive enhancement

It has also been proposed that listener-driven enhancements need not be 'listener-oriented' in the sense of being directly attributable to the speaker's assessment of the listener's communicative needs, but can instead be passive in nature. Silverman (2006) discusses the phonetic underpinnings of labial spreading in Trique, whereby historical *uka > uk^{w}a, but *uta > uta. He argues that the spreading of labiality increases the acoustic distinction between velars and alveolars: the F2 transition is sharply falling out of the alveolar plosive and only shallowly falling following velars, while both /k^{w}/ and /t^{w}/ have rising F2 transitions. Therefore, the acoustic distance between /k^{w}/ ~ /t/ is greater than /k/ ~ /t^{w}/ would be, which Silverman argues explains the targeting/restriction of spreading to the velar context. The results of an accompanying identification experiment indicate that the overall extent of confusability correlates with F2 similarity: while the presence vs. absence of a glide was easily perceived by listeners, confusion in terms of place of articulation was greatest when labiality was present.

Silverman argues that this enhancement does not originate with a listener-oriented enhancement on the part of the speaker, but rather that because "labiality was already loitering in the neighborhood," it might instead "be enhanced passively, evolving over generations of speakers, due to the communicative success of some tokens, and the communicative failure of others" (2006, 141). That is, all else being equal, exemplars of more acoustically distinct tokens are less likely to be rejected/more likely to be retained by the listener/learner, so the sound change will 'naturally' take place over generations without needing to attribute any functional/teleological goal to the speaker (cf. Section 4.3 on auditory enhancement).

4.6 Quantal theory and enhancement

The *quantal theory* of speech (Stevens 1972, 1989) proposes that the relationship between change in an articulatory parameter and its acoustic consequences may be non-linear. Within so-called quantal regions, articulatory variability will be tolerated because the acoustic effect of that variation is minimal; but at certain critical points, a small articulatory difference results in a phase shift between different quantal regions. Stevens proposed that this non-linearity in the mapping between articulation and acoustics leads to natural classes of speech sounds.

Quantal theory also provides a way of understanding how incrementation in production may condition a sudden perceptual 'jump' when a critical acoustic threshold is crossed. For example, Lin, Beddor, and Coetzee (2014) studied the phonetic underpinnings of sound changes involving loss or post-vocalic vocalization of /l/, as found in many varieties of English (e.g., SSBE *folk* /fəʊk/, *palm* /pɑːm/, *talk* /tɔːk/) or French *animaux* < Latin *animales* (Müller 2011). Lin, Beddor, and Coetzee (2014) showed that a small, incremental articulatory change (lowering of the tongue tip) can cause a comparatively large increase or approximation of F1 and F2, which could lead to its reinterpretation as a back vowel. /s/-retraction is also likely to have a quantal component given that, as Stevens (1972, 1989) shows, there is a sudden perceptual jump from [s] to [ʃ] as the tongue tip slides back toward the hard palate. The ongoing fronting of /u/ in British English (Harrington, Kleber, and Reubold 2008) may also be quantal in nature, as the relationship between tongue position and the second formant frequency, perhaps its most salient acoustic consequence, is non-linear due to acoustic coupling between the sub- and supraglottal systems (Stevens 1989; Sonderegger 2004; Chi and Sonderegger 2007).

5 Perceptual learning

(Mis)perception-based accounts of the initiation of change such as Ohala's (Section 3.2) are challenged by findings in psycholinguistics that speech perception is highly flexible and adaptable, and that much of this flexibility and adaptability is *because*, not in spite, of the variability inherent in the speech signal. In fact, it is questionable whether it makes sense at all to think of listeners as 'compensating' for phonetic biases, because it seems clear that listeners make active use of this variability to assist in speech processing (Cutler 2012).

Not only do listeners have considerable experience with how articulatory overlap and gestural co-activation engender different acoustic outcomes and are able to use this information in perception, this adaptation is also extremely rapid: listeners are very adept at adjusting to non-local accents and contexts, as demonstrated from a large literature on *perceptual learning* (Gibson 1963; Goldstone 1998; for overviews, see Samuel and Kraljic 2009; Cutler 2012). In speech, perceptual learning refers to the general phenomenon that, when making use of contextual information to resolve ambiguous acoustic speech input, listeners appear to learn in a manner that influences how they later categorize ambiguous speech sounds in the absence of the disambiguating context. This literature is quite broad and includes work studying the effects of exposure to unfamiliar (non-native, accented, degraded) speech (Logan, Lively, and Pisoni 1991; Lively, Logan, and Pisoni 1993; Bradlow et al. 1997, 1999; Bradlow and Bent 2008; Winn, Chatterjee, and Idsardi 2013). Here, we briefly review two strands of this work focusing on the retuning of phonetic categories in L1 speech, as well as related work on phonetic accommodation, and discuss the implications of these findings for sound change research.

5.1 Lexically guided perceptual learning

Lexically guided perceptual learning refers to a particular paradigm in perceptual learning research which focuses on how phonetic retuning is guided by context. The classic study in this paradigm is that of Norris, McQueen, and Cutler (2003). In this study, Dutch listeners heard words in which a final fricative (/f/ or /s/) was replaced by an ambiguous sound [?] acoustically halfway between [f] and [s]. One group of listeners heard words in which a final [f] of a real word, like *witlof* ‘chicory’ was replaced by the ambiguous sound, while another heard words in which the final [s] of a word like *naaldbos* ‘pine forest’ was replaced with the ambiguous sound. Crucially, in both conditions, there existed no Dutch word with the alternative fricative: that is, **witlos* and **naaldbof* do not exist. Norris, McQueen, and Cutler found that, after exposure, the two groups of listeners differed in their categorization patterns along an /ɛf/–/ɛs/ continuum: listeners who heard [?] in contexts like *witlo?* were more likely to respond with /f/, whereas listeners who heard [?] in contexts like *naaldbo?* were more likely to respond with /s/. An important finding of this study was that this perceptual learning was lexically driven, because the same post-exposure categorization shift was *not* observed in listeners exposed solely to non-words with the ambiguous sound in final position. Subsequent studies using this paradigm have explored the extent to which lexically guided perceptual learning is persistent over time, as well as to what degree it generalizes beyond the specific items, contexts, and even languages presented in the experiment (Eisner and McQueen 2005; Kraljic and Samuel 2005, 2006; Reinisch, Weber, and Mitterer 2013).

The results from lexically guided perceptual learning studies have been argued to be consistent with exemplar-theoretic conceptions of sound change in which the category boundary between two sounds varies depending on asymmetries between them in the direction of variance (Harrington et al. 2018). If a listener is presented with an exemplar containing an ambiguous phone and the identity of that phone distinguishes between lexical items, the exemplar will likely be categorized as belonging to one or the other lexical item. If there is no competing lexical item distinguished by this phone, on the other hand, the exemplar (if not discarded outright) can only be assigned to one lexical item, which over time would lead to a more variable phonetic representation of the phone.

5.2 Dimension-based statistical learning

Dimension-based statistical learning describes the phenomenon whereby the perceptual system flexibly adjusts to short-term deviations from its learned cue weights. In dimension-based statistical learning studies (e.g., Francis and Nusbaum 2002; Idemaru and Holt 2011, 2012, 2014, 2020; Yang and Sundara 2019; Kim, Clayards, and Kong 2020; Lehet and Holt 2020), listeners are exposed to speech input in which cue distributions differ from their canonical, long-term regularities. For example, Idemaru and Holt (2011) exposed English listeners to an artificial ‘accent’ where the canonical correlation between VOT and onset F0 was neutralized or reversed. Accordingly, listeners rapidly learned to down-weight reliance on the F0 dimension in response to stimuli in which it was less reliable (i.e., less distributionally informative). This paradigm differs from lexically guided perceptual learning in that, in the experimental context, lexical status never serves to disambiguate the ambiguous acoustics: all the response possibilities are always real words of the target language. On this basis, Idemaru and Holt (2011) suggest that a consistent and unchanging primary cue dimension (e.g., VOT) can serve a similar function to lexical status in helping to orient the listener to the function of the secondary cue dimension.

Studies using the dimension-based statistical learning paradigm typically employ an exposure condition where the canonical cue correlation is neutralized or reversed, demonstrating that listeners will down-weight cue dimensions which they learn to be unreliable. However, in the context of sound change, a more likely scenario would seem to involve perceptual up-weighting of a previously redundant cue, either due to enhancement (Section 4), or to the decrease in informativeness of a primary cue. This latter scenario is especially relevant for sound changes for which phonetic bias factors bring about the masking or complete loss of a primary cue, such as spontaneous consonant devoicing, which could redirect attention to a secondary acoustic property. A study by Gao and Kirby (2023) explored this possibility by exposing French listeners to monosyllabic CV stimuli in which either the informativeness of F0 (a secondary cue to voicing) was enhanced, or the informativeness of VOT (the primary cue) was decreased. An increase in F0 informativeness led listeners to upweight the use of high F0, making them overall more likely to give voiceless responses. However, when lexical feedback was given, decreased informativeness of VOT led to increased use of F0 to identify voiced plosives. These findings suggest that the effect of changes in distribution of cues may not be uniform in all regions of the phonetic space, and that distributional changes alone may not always be sufficient to trigger perceptual adaptation, *pace* Idemaru and Holt (2011). If true, this could provide a partial account of cases of transphonologization that cannot easily be explained by gestural rephasing, such as the phonologization of onset F0 (Kingston 2011; Ratliff 2015; “Tonogenesis and the Evolution of Tone Systems”).

5.3 Perceptual learning, phonetic convergence, and speech production

The findings of the perceptual learning literature suggest that, at least under some circumstances and for at least some period of time, listeners are able to adjust their cue weights and/or phonetic category boundaries. Nevertheless, it remains unclear how, or whether, these adjustments extend to production, a critical consideration for sound change. Perhaps surprisingly, there seems to be little work in the perceptual learning tradition, dimension-based or lexically guided, on production. One study which addresses this question is Kraljic, Brennan, and Samuel (2008), who examined subjects’ productions of /s/ and /ʃ/ before and after a classic lexically guided perceptual learning manipulation. While they found evidence of robust perceptual learning, they failed to observe evidence of a corresponding shift in production. Lehet and Holt (2017) studied whether rapid adjustments in listeners’ pre-lexical perceptual cue weight also affected those same listeners’ own speech productions by exposing English listeners to an ‘artificial accent’ which reversed the natural, canonical correlation between F0 and plosive voicing. They found that exposure not only caused listeners to decrease their perceptual reliance on F0, but also elicited a decrease in the use of F0 in the same individuals’ productions.

However, there does exist a related body of work on spontaneous phonetic imitation or *convergence* (also *accommodation*: e.g., Goldinger 1998; Pardo 2006; Nielsen 2011; Babel 2012; Yu, Abrego-Collier, and Sonderegger 2013; Zellou, Scarborough, and Nielsen 2016; Kim and Clayards 2019; Kwon 2019, 2021) which has studied unintentional changes or ‘drift’ in speech patterns in the direction of stimuli or an interlocutor. For example, English listeners who hear a model talker producing phonologically voiceless plosives with artificially lengthened VOT will tend to lengthen their own

VOTs, even in words that were not part of the exposure phase (Nielsen 2011). If phonetic imitation in the laboratory is a reasonable proxy for conversational interaction 'in the wild', imitation and convergence could be an important locus of sound change propagation (Babel 2012). However, despite its intuitive appeal, the past two decades of research have found the link between imitation and change to be tenuous at best (Pardo 2012; Sonderegger, Bane, and Graff 2017; Harrington, Gubian et al. 2019).

While phonetic imitation/accommodation does not strictly speaking imply that perceptual learning has taken place, some degree of perceptual flexibility is presumably a prerequisite for one speaker to shift their productions in the direction of another. The relationship between perceptual learning, cue reweighting, and phonetic accommodation is an area in need of further investigation, but clearly of great importance for our understanding of sound change. An outstanding challenge for future models of sound change will be to reconcile the fact of contextual phonetic variation with the flexible and adaptable nature of speech perception.

6 Concluding remarks

This entry has shown that it is scarcely possible to understand the phonetic conditions that give rise to sound change without taking into account acoustic-perceptual factors and their connection to speech production. One of the most enduring issues in sound-change research from this perspective that has been analyzed in numerous landmark studies by Ohala (1981, 1993b, 2005) is that certain sound sequences can be perceptually ambiguous: that is, a listener is sometimes confronted with more than one way of parsing an acoustic speech signal into articulatory gestures. Various examples of such ambiguities and the sound changes they can give rise to have been discussed in Sections 2 and 3.

Another important conclusion of this review is that although coarticulation may be responsible for considerable acoustic variability, this variability itself does not automatically lead to change. There is considerable evidence that listeners track the progression of coarticulation in perception (Martin and Bunnell 1982; Fowler and Thompson 2010): that is, coarticulatory cues inform phonological and lexical decisions as soon as the coarticulatory information becomes available (Beddor et al. 2013). Because speakers coarticulate to different degrees and in different ways (Grosvald 2009), and also because coarticulation and its possible enhancement (Section 4) depend on factors such as speaking situation, style, and discourse context, a byproduct of coarticulatory tracking in perception is likely to be that listeners encounter and memorize many different types of coarticulatory patterns for the same sequence of speech sounds. This accumulation of coarticulatory variation through experience of listening to many different types of speakers and speech styles is another important way in which perception and sound change are connected. However, as reviewed in Section 5, there is a large body of work showing that perceptual adaptation to novel accents and speech contexts can be both rapid and robust, without necessarily having a long-term impact on speech production (Section 5.3).

Finally, we have suggested that at least two types of acoustic-perceptual factors are involved in the progression from incremental phonetic variation toward a categorical change. One of these is the perceptual trading relationship (Section 3.3) that is at the heart of Beddor's account of cue transfer from coarticulatory source to effect (with possible loss of the source) in phonologization-based sound changes. Another – one that is to date largely unexplored in relation to sound change – is that categorical change could be associated with a progression of incremental articulatory change that initially has minimal acoustic consequences but a dramatic acoustic-perceptual effect once a quantal boundary (Section 4.6) is crossed.

Acknowledgments

This research was partly funded by the European Union (ERC SoundAct, project number 101053194 and ERC EVOTONE, project number 758605). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

References and Suggested Readings

Abrego-Collier, Carissa. 2013. “Liquid Dissimilation as Listener Hypocorrection.” *Annual Meeting of the Berkeley Linguistics Society* 37 (1): 3–17. DOI: 10.3765/bls.v37i1.3195.

Alderete, John, and Stefan A. Frisch. 2007. “Dissimilation in Grammar and the Lexicon.” In *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 379–398. Cambridge: Cambridge University Press.

Alderton, Roy. 2020. “Speaker Gender and Salience in Sociolinguistic Speech Perception: Goose-Fronting in Standard Southern British English.” *Journal of English Linguistics* 48 (1): 72–96. DOI: 10.1177/0075424219896400.

Alfonso, Peter J., and Thomas Baer. 1982. “Dynamics of Vowel Articulation.” *Language and Speech* 25 (2): 151–173. DOI: 10.1177/002383098202500203.

Avery, Peter, and William J. Idsardi. 2001. “Laryngeal Dimensions, Completion and Enhancement.” In *Distinctive Feature Theory*, edited by Tracy Alan Hall and Ursula Kleinhanz, 41–70. Berlin: Walter de Gruyter.

Babel, Molly. 2012. “Evidence for Phonetic and Social Selectivity in Spontaneous Phonetic Imitation.” *Journal of Phonetics* 40 (1): 177–189. DOI: 10.1016/j.wocn.2011.09.001.

Baker, Adam, Diana Archangeli, and Jeff Mielke. 2011. “Variability in American English S-Retraction Suggests a Solution to the Actuation Problem.” *Language Variation and Change* 23 (3): 347–374. DOI: 10.1017/S0954394511000135.

Bang, Hye-Young, Morgan Sonderegger, Yoonjung Kang, Meghan Clayards, and Tae-Jin Yoon. 2018. “The Emergence, Progress, and Impact of Sound Change in Progress in Seoul Korean: Implications for Mechanisms of Tonogenesis.” *Journal of Phonetics* 66 (January): 120–144. DOI: 10.1016/j.wocn.2017.09.005.

Bauer, Laurie. 1985. “Tracing Phonetic Change in the Received Pronunciation of British English.” *Journal of Phonetics* 13 (1): 61–81. DOI: 10.1016/S0095-4470(19)30726-0.

Beckman, Mary E., Kenneth De Jong, Sun-Ah Jun, and Sook-Hyang Lee. 1992. “The Interaction of Coarticulation and Prosody in Sound Change.” *Language and Speech* 35 (1–2): 45–58. DOI: 10.1177/002383099203500205.

Beddor, Patrice Speeter. 2007. “Nasals and Nasalization: The Relation between Segmental and Coarticulatory Timing.” In *Proceedings of the 16th International Congress of Phonetic Sciences*, edited by Jürgen Trouvain and William J. Barry, 249–254.

Beddor, Patrice Speeter. 2009. “A Coarticulatory Path to Sound Change.” *Language* 85 (4): 785–821.

Beddor, Patrice Speeter. 2012. “Perception Grammars and Sound Change.” In *The Initiation of Sound Change: Perception, Production, and Social Factors*, edited by Maria-Josep Solé and Daniel Recasens, 37–56. Amsterdam: John Benjamins.

Beddor, Patrice Speeter. 2015. “The Relation between Language Users’ Perception and Production Repertoires.” In *Proceedings of the 18th International Congress of Phonetic Sciences*, edited by The Scottish Consortium for ICPhS 2015, paper number 1041. https://www.internationalphoneticassociation.org/icphs/icphs2015.

Beddor, Patrice Speeter, Andries W. Coetzee, Will Styler, Kevin B. McGowan, and Julie E. Boland. 2018. “The Time Course of Individuals’ Perception of Coarticulatory Information Is Linked to Their Production: Implications for Sound Change.” *Language* 94 (4): 931–968. DOI: 10.1353/lan.2018.0051.

Beddor, Patrice Speeter, James D. Harnsberger, and Stephanie Lindemann. 2002. “Language-Specific Patterns of Vowel-to-Vowel Coarticulation: Acoustic Structures and Their Perceptual Correlates.” *Journal of Phonetics* 30 (4): 591–627. DOI: 10.1006/jpho.2002.0177.

Beddor, Patrice Speeter, and Rena Arens Krakow. 1999. “Perception of Coarticulatory Nasalization by Speakers of English and Thai: Evidence for Partial Compensation.” *The Journal of the Acoustical Society of America* 106 (5): 2868–2887. DOI: 10.1121/1.428111.

Beddor, Patrice Speeter, Kevin B. McGowan, Julie E. Boland, Andries W. Coetzee, and Anthony Brasher. 2013. “The Time Course of Perception of Coarticulation.” *The Journal of the Acoustical Society of America* 133 (4): 2350–2366.

Bell-Berti, Fredericka, and Rena Arens Krakow. 1991. “Anticipatory Velar Lowering: A Coproduction Account.” *The Journal of the Acoustical Society of America* 90 (1): 112–123. DOI: 10.1121/1.401304.

Bermúdez-Otero, Ricardo. 2015. “Amphichronic Explanation and the Life Cycle of Phonological Processes.” In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph C. Salmons. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo, and Graeme Trousdale. 2012. “Cycles and Continua: On Unidirectionality and Gradualness in Language Change.” In *The Oxford Handbook of the History of English*, 1st ed., edited by Terttu Nevalainen and Elizabeth Closs Traugott, 691–720. Oxford University Press. DOI: 10.1093/oxfordhb/9780199922765.013.0059.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press. http://www.cambridge.org/gb/academic/subjects/languages-linguistics/phonetics-and-phonology/evolutionary-phonology-emergence-sound-patterns?format=PB.

Blevins, Juliette. 2006. “New Perspectives on English Sound Patterns: ‘Natural’ and ‘Unnatural’ in Evolutionary Phonology.” *Journal of English Linguistics* 34 (1): 6–25. DOI: 10.1177/0075424206287585.

Blevins, Juliette. 2015. “Evolutionary Phonology: A Holistic Approach to Sound Change Typology.” In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons, 485–500. Oxford: Oxford University Press.

Blevins, Juliette. 2019. “Notes on θ > f and f > θ: Deconstructing Markedness in Sound Change Typology.” In *Perspectives on Language Structure and Language Change: Studies in Honor of Henning Andersen*, edited by Lars Heltoft, Iván Igartua, Brian D. Joseph, Kirsten Jeppesen Kragh, and Lene Schøsler, 107–122. Amsterdam: John Benjamins. https://benjamins.com/catalog/cilt.345.04ble.

Blood, David L. 1967. “Phonological Units in Cham.” *Anthropological Linguistics* 9 (8): 15–32.

Blust, Robert A. 2009. *The Austronesian Languages*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Boersma, Paul. 1998. “Functional Phonology.” PhD diss., University of Amsterdam.

Bradley, David, ed. 1985. “Arakanese Vowels.” In *Linguistics of the Sino-Tibetan Area: The State of the Art*, edited by Graham Thurgood, James A. Matisoff, and David Bradley, 180–200. Canberra: The Australian National University, Research School of Pacific Studies.

Bradlow, Ann R., and Tessa Bent. 2008. “Perceptual Adaptation to Non-Native Speech.” *Cognition* 106 (2): 707–729. DOI: 10.1016/j.cognition.2007.04.005.

Bradlow, Ann, David Pisoni, Reiko Akahane-Yamada, and Yoh'ichi Tohkura. 1997. "Training Japanese Listeners to Identify English /r/ and /l/: IV. Some Effects of Perceptual Learning on Speech Production." *The Journal of the Acoustical Society of America* 101 (May): 2299–2310. DOI: 10.1121/1.418276.

Bradlow, Ann R., Reiko Akahane-Yamada, David B. Pisoni, and Yoh'ichi Tohkura. 1999. "Training Japanese Listeners to Identify English /r/and /l/: Long-Term Retention of Learning in Perception and Production." *Perception & Psychophysics* 61 (5): 977–985. DOI: 10.3758/BF03206911.

Britain, David J. 2005. "The Dying Dialects of England?" In *Historical Linguistic Studies of Spoken English*, edited by Antonio Bertacca, 35–46. Pisa: Edizioni Plus.

Browman, Catherine P., and Louis M. Goldstein. 1986. "Towards an Articulatory Phonology." *Phonology* 3 (May): 219–252. DOI: 10.1017/S0952675700000658.

Browman, Catherine P., and Louis Goldstein. 1989. "Articulatory Gestures as Phonological Units." *Phonology* 6 (2): 201–251. DOI: 10.1017/S0952675700001019.

Browman, Catherine P., and Louis Goldstein. 1990. "Tiers in Articulatory Phonology, with Some Implications for Casual Speech." In *Papers in Laboratory Phonology*, vol. 1: *Between the Grammar and Physics of Speech*, edited by John Kingston and Mary E. Beckman, 341–376. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511627736.019.

Browman, Catherine P., and Louis M. Goldstein. 1991. "Gestural Structures: Distinctiveness, Phonological Processes, and Historical Change." In *Modularity and the Motor Theory of Speech Perception*, edited by Ignatius G. Mattingly and Michael Studdert-Kennedy, 313–338. Hillsdale, NJ: Lawrence Erlbaum.

Browman, Catherine P., and Louis Goldstein. 1992. "Articulatory Phonology: An Overview." *Phonetica* 49 (3–4): 155–180. DOI: 10.1159/000261913.

Busà, Maria Grazia. 2007. "Coarticulatory Nasalization and Phonological Developments: Data from Italian and English Nasal-Fricative Sequences." In *Experimental Approaches to Phonology*, edited by Maria-Josep Solé, Patrice S. Beddor, and Manjari Ohala, 155–174. Oxford: Oxford University Press.

Carignan, Christopher, Stefano Coretta, Jens Frahm, Jonathan Harrington, Phil Hoole, Arun Joseph, Esther Kunay, and Dirk Voit. 2021. "Planting the Seed for Sound Change: Evidence from Real-Time MRI of Velum Kinematics in German." *Language* 97 (2): 333–364. DOI: 10.1353/lan.2021.0020.

Catford, J.C. 2001. "On Rs, Rhotacism and Paleophony." *Journal of the International Phonetic Association* 31 (2): 171–185. DOI: 10.1017/S0025100301002018.

Chandrasekaran, Bharath, Padma D. Sampath, and Patrick C.M. Wong. 2010. "Individual Variability in Cue-Weighting and Lexical Tone Learning." *The Journal of the Acoustical Society of America* 128 (1): 456–465. DOI: 10.1121/1.3445785.

Chang, Steve S., Madelaine Plauché, and John J. Ohala. 2001. "Markedness and Consonant Confusion Asymmetries." In *The Role of Speech Perception in Phonology*, edited by Elisabeth Hume and Keith Johnson, 79–101. San Diego: Academic Press.

Chen, Yiya. 2011. "How Does Phonology Guide Phonetics in Segment–F0 Interaction?" *Journal of Phonetics* 39 (4): 612–625. DOI: 10.1016/j.wocn.2011.04.001.

Chi, Xuemin, and Morgan Sonderegger. 2007. "Subglottal Coupling and Its Influence on Vowel Formants." *The Journal of the Acoustical Society of America* 122 (3): 1735–1745. DOI: 10.1121/1.2756793.

Chitoran, Ioana, Louis Goldstein, and Dani Byrd. 2008. "Gestural Overlap and Recoverability: Articulatory Evidence from Georgian." In *Laboratory Phonology 7*, edited by Carlos Gussenhoven and Natasha Warner, 419–448. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110197105.2.419.

Clayards, Meghan Alison. 2008. "The Ideal Listener: Making Optimal Use of Acoustic-Phonetic Cues for Word Recognition." PhD diss., University of Rochester.

Clayards, Meghan. 2018. “Differences in Cue Weights for Speech Perception Are Correlated for Individuals within and across Contrasts.” *The Journal of the Acoustical Society of America* 144 (3): EL172–177. DOI: 10.1121/1.5052025.

Coetzee, Andries W., Patrice Speeter Beddor, Kerby Shedden, Will Styler, and Daan Wissing. 2018. “Plosive Voicing in Afrikaans: Differential Cue Weighting and Tonogenesis.” *Journal of Phonetics* 66 (January): 185–216. DOI: 10.1016/j.wocn.2017.09.009.

Coetzee, Andries W., Patrice Speeter Beddor, Will Styler, Stephen Tobin, Ian Bekker, and Daan Wissing. 2022. “Producing and Perceiving Socially Structured Coarticulation: Coarticulatory Nasalization in Afrikaans.” *Laboratory Phonology* 13 (1). DOI: 10.16995/labphon.6450.

Cohn, Abigail C. 1990. “Phonetic and Phonological Rules of Nasalization.” PhD diss., University of California at Los Angeles.

Cohn, Abigail C. 1993. “Nasalisation in English: Phonology or Phonetics.” *Phonology* 10 (1): 43–81.

Cole, Jennifer, Gary Linebaugh, Cheyenne Munson, and Bob McMurray. 2010. “Unmasking the Acoustic Effects of Vowel-to-Vowel Coarticulation: A Statistical Modeling Approach.” *Journal of Phonetics* 38 (2): 167–184. DOI: 10.1016/j.wocn.2009.08.004.

Connine, Cynthia M., and Laura M. Darnieder. 2009. “Perceptual Learning of Co-Articulation in Speech.” *Journal of Memory and Language* 61 (3): 412–422. DOI: 10.1016/j.jml.2009.07.003.

Cox, Felicity, and Sallyanne Palethorpe. 2014. “Phonologisation of Vowel Duration and Nasalised /æ/ in Australian English.” In *Proceedings of the 15th Australasian Speech Science & Technology Conference*, edited by Jennifer Hay and Emma Parnell, 33–36.

Cronenberg, Johanna, Michele Gubian, Jonathan Harrington, and Hanna Ruch. 2020. “A Dynamic Model of the Change from Pre- to Post-Aspiration in Andalusian Spanish.” *Journal of Phonetics* 83: 101016. DOI: 10.1016/j.wocn.2020.101016.

Cutler, Anne. 2012. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge, MA: The MIT Press.

Delvaux, Véronique, Didier Demolin, Bernard Harmegnies, and Alain Soquet. 2008. “The Aerodynamics of Nasalization in French.” *Journal of Phonetics* 36 (4): 578–606. DOI: 10.1016/j.wocn.2008.02.002.

Delvaux, Véronique, Thierry Metens, and Alain Soquet. 2002. “French Nasal Vowels: Acoustic and Articulatory Properties.” In *7th International Conference on Spoken Language Processing (ICSLP 2002)*, 53–56. DOI: 10.21437/ICSLP.2002-51.

De Jong, Kenneth, Mary E. Beckman, and Jan Edwards. 1993. “The Interplay between Prosodic Structure and Coarticulation.” *Language and Speech* 36 (2–3): 197–212. DOI: 10.1177/002383099303600305.

Diehl, R.L. 2008. “Acoustic and Auditory Phonetics: The Adaptive Design of Speech Sound Systems.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1493): 965–978. DOI: 10.1098/rstb.2007.2153.

Diehl, Randy L. 2011. “On the Robustness of Speech Perception.” In *Proceedings of the 17th International Congress of Phonetic Sciences*, edited by Wai Sum Lee and Eric Zee, 1–8.

Diehl, Randy L., and Keith R. Kluender. 1989. “On the Objects of Speech Perception.” *Ecological Psychology* 1 (2): 121–144. DOI: 10.1207/s15326969eco0102_2.

Dmitrieva, Olga, Fernando Llanos, Amanda A. Shultz, and Alexander L. Francis. 2015. “Phonological Status, Not Voice Onset Time, Determines the Acoustic Realization of Onset F0 as a Secondary Voicing Cue in Spanish and English.” *Journal of Phonetics* 49 (March): 77–95. DOI: 10.1016/j.wocn.2014.12.005.

Dresher, B. Elan. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511642005.

Eisner, Frank, and James M. McQueen. 2005. “The Specificity of Perceptual Learning in Speech Processing.” *Perception & Psychophysics* 67 (2): 224–238.

Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton & Co.

Flemming, Edward S. 2001. *Auditory Representations in Phonology*. New York: Routledge.

Flemming, Edward S. 2003. “The Relationship between Coronal Place and Vowel Backness.” *Phonology* 20 (3): 335–373.

Flemming, Edward. 2004. “Contrast and Perceptual Distinctiveness.” In *Phonetically Based Phonology*, edited by Bruce Hayes, Robert Kirchner, and Donca Steriade, 232–276. Cambridge: Cambridge University Press. http://ebooks.cambridge.org/ref/id/CBO9780511486401A016.

Flynn, Darin, and Sean Fulop. 2014. “Acoustic Phonetic Features in Athabaskan Sound Change.” *Diachronica* 31 (2): 192–222. DOI: 10.1075/dia.31.2.02fly.

Fowler, Carol A., and Julie M. Brown. 2000. “Perceptual Parsing of Acoustic Consequences of Velum Lowering from Information for Vowels.” *Perception & Psychophysics* 62 (1): 21–32. DOI: 10.3758/BF03212058.

Fowler, Carol A., and Elliot Saltzman. 1993. “Coordination and Coarticulation in Speech Production.” *Language and Speech* 36 (2–3): 171–195. DOI: 10.1177/002383099303600304.

Fowler, Carol A., and Mary R. Smith. 1986. “Speech Perception as ‘Vector Analysis’: An Approach to the Problems of Invariance and Segmentation.” In *Invariance and Variability in Speech Processes*, 123–139. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Fowler, Carol A., and Jaqueline M. Thompson. 2010. “Listeners’ Perception of Compensatory Shortening.” *Attention, Perception, & Psychophysics* 72: 481–491. DOI: 10.3758/APP.72.2.481.

Francis, Alexander L., and Howard C. Nusbaum. 2002. “Selective Attention and the Acquisition of New Phonetic Categories.” *Journal of Experimental Psychology: Human Perception and Performance* 28 (2): 349–366. DOI: 10.1037//0096-1523.28.2.349.

Fridland, Valerie. 2008. “Patterns of /Uw/, /ʊ/, and /Ow Fronting in Reno, Nevada.” *American Speech* 83 (4): 432–454. DOI: 10.1215/00031283-2008-030.

Fujimura, Osamu. 1962. “Analysis of Nasal Consonants.” *The Journal of the Acoustical Society of America* 34 (12): 1865–1875. DOI: 10.1121/1.1909142.

Gao, Jiayin, and James Kirby. 2023. “Perceptual Adaptation to Altered Cue Informativeness: Distributional, Auditory, and Lexical Factors.” In *Proceedings of the 20th International Congress of Phonetic Sciences*, edited by Radek Skarnitzl and Jan Volín, 137–141.

Garrett, Andrew. 2015. “Sound Change.” In *The Routledge Handbook of Historical Linguistics*, edited by Claire Bowern and Bethwyn Evans, 227–248. London and New York: Routledge.

Garrett, Andrew, and Keith Johnson. 2013. “Phonetic Bias in Sound Change.” In *Origins of Sound Change: Approaches to Phonologization*, edited by Alan Yu, 51–97. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199573745.003.0003.

Gibson, E.J. 1963. “Perceptual Learning.” *Annual Review of Psychology* 14 (1): 29–56. DOI: 10.1146/annurev.ps.14.020163.000333.

Goldinger, Stephen D. 1998. “Echoes of Echoes? An Episodic Theory of Lexical Access.” *Psychological Review* 105 (2): 251–279.

Goldstone, Robert L. 1998. “Perceptual Learning.” *Annual Review of Psychology* 49 (1): 585–612.

Gradin, Dwight. 1966. “Consonantal Tone in Jeh Phonemics.” *Mon-Khmer Studies* 2: 41–53.

Grierson, George A. 1922. “Spontaneous Nasalization in the Indo-Aryan Languages.” *The Journal of the Royal Asiatic Society of Great Britain and Ireland* 3: 381–388.

Grosvald, Michael. 2009. “Interspeaker Variation in the Extent and Perception of Long-Distance Vowel-to-Vowel Coarticulation.” *Journal of Phonetics* 37 (2): 173–188. DOI: 10.1016/j.wocn.2009.01.002.

Grosvald, Michael, and David Corina. 2012. “Perception of Long-Distance Coarticulation: An Event-Related Potential and Behavioral Study.” *Applied Psycholinguistics* 33: 55–82. DOI: 10.1017/S0142716411000105.

Guion, Susan Guignard. 1998. “The Role of Perception in the Sound Change of Velar Palatalization.” *Phonetica* 55 (1–2): 18–52. DOI: 10.1159/000028423.

Gunter, Kaylynn, Charlotte Vaughn, and Tyler Kendall. 2021. “Contextualizing /s/ Retraction: Sibilant Variation and Change in Washington DC African American Language.” *Language Variation and Change* 33 (3): 331–357. DOI: 10.1017/S095439452100020X.

Guthrie, Malcolm. 1967. *Comparative Bantu*, vol. 1. Farnborough: Gregg.

Hajek, John. 1997. *Universals of Sound Change in Nasalization*. Oxford: Blackwell.

Hall, Daniel Currie. 2007. "The Role and Representation of Contrast in Phonological Theory." PhD diss., University of Toronto. https://twpl.library.utoronto.ca/index.php/twpl/article/view/6497.

Hall, Daniel Currie. 2011. "Phonological Contrast and Its Phonetic Enhancement: Dispersedness without Dispersion." *Phonology* 28 (1): 1–54. DOI: 10.1017/S0952675711000029.

Hanson, Helen M. 2009. "Effects of Obstruent Consonants on Fundamental Frequency at Vowel Onset in English." *The Journal of the Acoustical Society of America* 125 (1): 425–441. DOI: 10.1121/1.3021306.

Harrington, Jonathan. 2007. "Evidence for a Relationship between Synchronic Variability and Diachronic Change in the Queen's Annual Christmas Broadcasts." In *Laboratory Phonology 9*, edited by Jennifer Cole and José Ignacio Hualde, 125–143. Berlin: Mouton de Gruyter.

Harrington, Jonathan, Janet Fletcher, and Corinne Roberts. 1995. "Coarticulation and the Accented/Unaccented Distinction: Evidence from Jaw Movement Data." *Journal of Phonetics* 23 (3): 305–322. DOI: 10.1016/S0095-4470(95)80163-4.

Harrington, Jonathan, Michele Gubian, Mary Stevens, and Florian Schiel. 2019. "Phonetic Change in an Antarctic Winter." *The Journal of the Acoustical Society of America* 146 (5): 3327–3332. DOI: 10.1121/1.5130709.

Harrington, Jonathan, Felicitas Kleber, and Ulrich Reubold. 2008. "Compensation for Coarticulation, /u/-Fronting, and Sound Change in Standard Southern British: An Acoustic and Perceptual Study." *The Journal of the Acoustical Society of America* 123 (5): 2825–2835. DOI: 10.1121/1.2897042.

Harrington, Jonathan, Felicitas Kleber, Ulrich Reubold, Florian Schiel, and Mary Stevens. 2018. "Linking Cognitive and Social Aspects of Sound Change Using Agent-Based Modeling." *Topics in Cognitive Science* 10 (4): 707–728. DOI: 10.1111/tops.12329.

Harrington, Jonathan, Felicitas Kleber, Ulrich Reubold, Florian Schiel, and Mary Stevens. 2019. "The Phonetic Basis of the Origin and Spread of Sound Change." In *The Routledge Handbook of Phonetics*, edited by William F. Katz and Peter F. Assmann, 401–426. Abingdon and New York: Routledge. DOI: 10.4324/9780429056253-15.

Harrington, Jonathan, Felicitas Kleber, and Mary Stevens. 2016. "The Relationship between the (Mis)-Parsing of Coarticulation in Perception and Sound Change: Evidence from Dissimilation and Language Acquisition." In *Recent Advances in Nonlinear Speech Processing*, edited by Anna Esposito, Marcos Faundez-Zanuy, Antonietta M. Esposito, Gennaro Cordasco, Thomas Drugman, Jordi Solé-Casals, and Francesco Carlo Morabito, 15–34. Cham: Springer.

Harrington, Jonathan, Sallyanne Palethorpe, and Catherine I. Watson. 2000. "Does the Queen Speak the Queen's English?" *Nature* 408 (6815): 927–928.

Harris, Katherine Safford. 1958. "Cues for the Discrimination of American English Fricatives in Spoken Syllables." *Language and Speech* 1 (1): 1–7. DOI: 10.1177/002383095800100101.

Hawkins, Sarah, and Jonathan Midgley. 2005. "Formant Frequencies of RP Monophthongs in Four Age Groups of Speakers." *Journal of the International Phonetic Association* 35 (2): 183–199.

Heinz, John M., and Kenneth N. Stevens. 1961. "On the Properties of Voiceless Fricative Consonants." *The Journal of the Acoustical Society of America* 33 (5): 589–596. DOI: 10.1121/1.1908734.

Herrero de Haro, Alfredo, and John Hajek. 2022. "Eastern Andalusian Spanish." *Journal of the International Phonetic Association* 52 (1): 135–156. DOI: 10.1017/S0025100320000146.

Hill, Nathan W. 2007. "Aspirated and Unaspirated Voiceless Consonants in Old Tibetan." *Language and Linguistics* 8 (2): 471–493.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. "Phonetic Explanations for the Development of Tones." *Language* 55 (1): 37–58. DOI: 10.2307/412518.

Hoole, Philip, and Marianne Pouplier. 2017. "Öhman Returns: New Horizons in the Collection and Analysis of Imaging Data in Speech Production Research." *Computer Speech & Language* 45 (September): 253–277. DOI: 10.1016/j.csl.2017.03.002.

House, Arthur S., and Grant Fairbanks. 1953. "The Influence of Consonant Environment upon the Secondary Acoustic Characteristics of Vowels." *The Journal of the Acoustical Society of America* 25 (1): 105–113.

House, Arthur S., and Kenneth N. Stevens. 1956. "Analog Studies of the Nasalization of Vowels." *Journal of Speech and Hearing Disorders* 21 (2): 218–232. DOI: 10.1044/jshd.2102.218.

Howe, Penelope Jane. 2017. "Tonogenesis in Central Dialects of Malagasy: Acoustic and Perceptual Evidence with Implications for Synchronic Mechanisms of Sound Change." PhD diss., Rice University, Houston.

Hualde, José Ignacio, and Pilar Prieto. 2014. "Lenition of Intervocalic Alveolar Fricatives in Catalan and Spanish." *Phonetica* 71 (2): 109–127. DOI: 10.1159/000368197.

Hualde, José Ignacio, Miquel Simonet, and Marianna Nadeu. 2011. "Consonant Lenition and Phonological Recategorization." *Laboratory Phonology* 2 (2). DOI: 10.1515/labphon.2011.011.

Huffman, Franklin E. 1967. "An Outline of Cambodian Grammar." PhD diss., Cornell University.

Hughes, George W., and Morris Halle. 1956. "Spectral Properties of Fricative Consonants." *The Journal of the Acoustical Society of America* 28 (2): 303–310. DOI: 10.1121/1.1908271.

Hyman, Larry M. 1976. "Phonologization." In *Linguistic Studies Offered to Joseph Greenberg on the Occasion of His Sixtieth Birthday*, edited by Alphonse Julliand, Andrew M. Devine, and Laurence D. Stephens, 470–418. Saratoga, CA: Anma Libri.

Hyman, Larry M. 2013. "Enlarging the Scope of Phonologization." In *Origins of Sound Change: Approaches to Phonologization*, edited by Alan C.L. Yu, 3–28. Oxford: Oxford University Press.

Idemaru, Kaori, and Lori L. Holt. 2011. "Word Recognition Reflects Dimension-Based Statistical Learning." *Journal of Experimental Psychology: Human Perception and Performance* 37 (6): 1939–1956. DOI: 10.1037/a0025641.

Idemaru, Kaori, and Lori L. Holt. 2012. "Examining Talker and Phoneme Generalization of Dimension-Based Statistical Learning in Speech Perception." In *Proceedings of the 14th Australasian International Conference on Speech Science and Technology 2012 (SST 2012)*, edited by Felicity Cox, Katherine Demuth, Susan Lin, Kelly Miles, Sallyanne Palethorpe et al., 165–168.

Idemaru, Kaori, and Lori L. Holt. 2014. "Specificity of Dimension-Based Statistical Learning in Word Recognition." *Journal of Experimental Psychology: Human Perception and Performance* 40 (3): 1009–1021. DOI: 10.1037/a0035269.

Idemaru, Kaori, and Lori L. Holt. 2020. "Generalization of Dimension-Based Statistical Learning." *Attention, Perception, & Psychophysics* 82 (4): 1744–1762. DOI: 10.3758/s13414-019-01956-5.

Itô, Junko, and Ralf-Armin Mester. 1986. "The Phonology of Voicing in Japanese: Theoretical Consequences for Morphological Accessibility." *Linguistic Inquiry* 17 (1): 49–73.

Iverson, Gregory K., and Joseph C. Salmons. 2003. "Laryngeal Enhancement in Early Germanic." *Phonology* 20 (1): 43–74.

Jatteau, Adèle, and Michaela Hejná. 2016. "Dissimilation Can Be Gradient: Evidence from Aberystwyth English." *Papers in Historical Phonology* 1 (December): 359–386. DOI: 10.2218/pihph.1.2016.1737.

Johnson, Sarah E. 2019. "Spontaneous Nasalization: An Articulatory Investigation of Glottal Consonants in Thai." PhD diss., University of Illinois at Urbana-Champaign.

Jones, Mark J. 2002. "More on the 'Instability' of Interdental Fricatives: Gothic *þliuhan* 'flee' and Old English *flēon* 'flee' Revisited." *Word* 53 (1): 1–8. DOI: 10.1080/00437956.2002.11432521.

Jongman, Allard, Ratree Wayland, and Serena Wong. 2000. "Acoustic Characteristics of English Fricatives." *The Journal of the Acoustical Society of America* 108 (3): 1252–1263.

Joos, Martin. 1948. "Acoustic Phonetics." *Language Monographs* 23: 1–136.

Kálmán, Béla. 1972. "Hungarian Historical Phonology." In *The Hungarian Language*, edited by Loránd Benkö and Samu Imre, 49–84. The Hague: De Gruyter. DOI: 10.1515/9783110880236-003.

Kataoka, Reiko. 2011. "Phonetic and Cognitive Bases of Sound Change." PhD diss., University of California, Berkeley.

Kavitskaya, Darya. 2001. “Compensatory Lengthening: Phonetics, Phonology, Diachrony.” PhD diss., University of California, Berkeley.

Kawasaki, Haruko. 1986. “Phonetic Explanation for Phonological Universals: The Case of Distinctive Vowel Nasalization.” In *Experimental Phonology*, edited by John J. Ohala and Jeri J. Jaeger, 239–252. Orlando: Academic Press.

Kent, Raymond D., Patrick J. Carney, and Larry R. Severeid. 1974. “Velar Movement and Timing: Evaluation of a Model for Binary Control.” *Journal of Speech and Hearing Research* 17 (3): 470–488. DOI: 10.1044/jshr.1703.470.

Keyser, Samuel Jay, and Kenneth N. Stevens. 2006. “Enhancement and Overlap in the Speech Chain.” *Language* 82 (1): 33–63. DOI: 10.1353/lan.2006.0051.

Kim, Donghyun, and Meghan Clayards. 2019. “Individual Differences in the Link between Perception and Production and the Mechanisms of Phonetic Imitation.” *Language, Cognition and Neuroscience* 34 (6): 769–786. DOI: 10.1080/23273798.2019.1582787.

Kim, Donghyun, Meghan Clayards, and Eun Jong Kong. 2020. “Individual Differences in Perceptual Adaptation to Unfamiliar Phonetic Categories.” *Journal of Phonetics* 81 (July): 100984. DOI: 10.1016/j.wocn.2020.100984.

Kingston, John. 2011. “Tonogenesis.” In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewan, Elisabeth V. Hume, and Keren Rice, 4: 2304–2333. Hoboken, NJ: Wiley Blackwell.

Kingston, John, and Randy L. Diehl. 1994. “Phonetic Knowledge.” *Language* 70 (3): 419–454. DOI: 10.2307/416481.

Kingston, John, Randy L. Diehl, Cecilia J. Kirk, and Wendy A. Castleman. 2008. “On the Internal Perceptual Structure of Distinctive Features: The [Voice] Contrast.” *Journal of Phonetics* 36 (1): 28–54. DOI: 10.1016/j.wocn.2007.02.001.

Kiparsky, Paul. 1995. “The Phonological Basis of Sound Change.” In *The Handbook of Phonological Theory*, edited by John A. Goldsmith, 640–670. Oxford: Blackwell.

Kiparsky, Paul. 2015. “Phonologization.” In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph C. Salmons, 563–582. Oxford: Oxford University Press.

Kirby, James. 2010. “Cue Selection and Category Restructuring in Sound Change.” PhD diss., University of Chicago.

Kirby, James. 2013. “The Role of Probabilistic Enhancement in Phonologization.” In *Origins of Sound Patterns: Approaches to Phonologization*, edited by Alan C.L. Yu, 228–246. Oxford: Oxford University Press.

Kirby, James. 2014a. “Incipient Tonogenesis in Phnom Penh Khmer: Acoustic and Perceptual Studies.” *Journal of Phonetics* 43: 69–85. DOI: 10.1016/j.wocn.2014.02.001.

Kirby, James. 2014b. “Incipient Tonogenesis in Phnom Penh Khmer: Computational Studies.” *Laboratory Phonology* 5 (1): 195–230. DOI: 10.1515/lp-2014-0008.

Kirby, James and Đinh Lư Giang. 2017. “On the R>h Shift in Kiên Giang Khmer.” *Journal of the Southeast Asian Linguistics Society* 10 (2): 66–85.

Kirby, James, Pittayawat Pittayaporn, and Marc Brunelle. 2023. “Transphonologization of Onset Voicing: Revisiting Northern and Eastern Kmhmu’.” *Phonetica* 79 (6): 591–629. DOI: 10.1515/phon-2022-0029.

Klatt, Dennis H., and Laura C. Klatt. 1990. “Analysis, Synthesis, and Voice Quality Variations among Female and Male Talkers.” *The Journal of the Acoustical Society of America* 87 (2): 820–857.

Koenig, Laura L., Christine H. Shadle, Jonathan L. Preston, and Christine R. Mooshammer. 2013. “Toward Improved Spectral Measures of /s/: Results from Adolescents.” *Journal of Speech, Language, and Hearing Research* 56 (4): 1175–1189. DOI: 10.1044/1092-4388(2012/12-0038).

Kohler, Klaus J. 1982. “F0 in the Production of Fortis and Lenis Plosives.” *Phonetica* 39: 199–218.

Kraljic, Tanya, Susan E. Brennan, and Arthur G. Samuel. 2008. “Accommodating Variation: Dialects, Idiolects, and Speech Processing.” *Cognition* 107 (1): 54–81. DOI: 10.1016/j.cognition.2007.07.013.

Kraljic, Tanya, and Arthur G. Samuel. 2005. “Perceptual Learning for Speech: Is There a Return to Normal?” *Cognitive Psychology* 51 (2): 141–178. DOI: 10.1016/j.cogpsych.2005.05.001.

Kraljic, Tanya, and Arthur G. Samuel. 2006. “Generalization in Perceptual Learning for Speech.” *Psychonomic Bulletin & Review* 13 (2): 262–268.

Kwon, Harim. 2019. “The Role of Native Phonology in Spontaneous Imitation: Evidence from Seoul Korean.” *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10 (1): 1–24. DOI: 10.5334/labphon.83.

Kwon, Harim. 2021. “A Non-Contrastive Cue in Spontaneous Imitation: Comparing Mono- and Bilingual Imitators.” *Journal of Phonetics* 88 (September): 101083. DOI: 10.1016/j.wocn.2021.101083.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology, and Sound Change: A Multimedia Reference Tool*. Berlin and New York: Mouton de Gruyter.

Lawrence, Wayne P. 2000. “/Str → /ʃtr/: Assimilation at a Distance?” *American Speech* 75 (1): 82–87. DOI: 10.1215/00031283-75-1-82.

Leeuw, Esther de. 2019. “Native Speech Plasticity in the German-English Late Bilingual Stefanie Graf: A Longitudinal Study over Four Decades.” *Journal of Phonetics* 73 (March): 24–39. DOI: 10.1016/j.wocn.2018.12.002.

Lehet, Matthew, and Lori L. Holt. 2017. “Dimension-Based Statistical Learning Affects Both Speech Perception and Production.” *Cognitive Science* 41 (April): 885–912. DOI: 10.1111/cogs.12413.

Lehet, Matthew, and Lori L. Holt. 2020. “Nevertheless, It Persists: Dimension-Based Statistical Learning and Normalization of Speech Impact Different Levels of Perceptual Processing.” *Cognition* 202 (September): 104328. DOI: 10.1016/j.cognition.2020.104328.

Lehiste, Ilse, and Gordon E. Peterson. 1961. “Some Basic Considerations in the Analysis of Intonation.” *The Journal of the Acoustical Society of America* 33 (4): 419–425.

Lehnert-LeHouillier, Heike. 2010. “A Cross-Linguistic Investigation of Cues to Vowel Length Perception.” *Journal of Phonetics* 38 (3): 472–482. DOI: 10.1016/j.wocn.2010.05.003.

Lehnert-LeHouillier, Heike. 2013. “From Long to Short and from Short to Long: Perceptual Motivations for Changes in Vocalic Length.” In *Origins of Sound Change*, edited by Alan C.L. Yu, 98–111. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199573745.003.0004.

Liljencrants, Johan, and Björn Lindblom. 1972. “Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast.” *Language* 48 (4): 839–862.

Lin, Susan, Patrice Speeter Beddor, and Andries W. Coetzee. 2014. “Gestural Reduction, Lexical Frequency, and Sound Change: A Study of Post-Vocalic /l/.” *Laboratory Phonology* 5 (1): 9–36. DOI: 10.1515/lp-2014-0002.

Lindblom, Björn. 1990. “Explaining Phonetic Variation: A Sketch of the H&H Theory.” In *Speech Production and Speech Modelling*, edited by William J. Hardcastle and Alain Marchal, 403–439. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-009-2037-8_16.

Lindblom, Björn, Susan Guion, Susan Hura, Seung-Jae Moon, and Raquel Willerman. 1995. “Is Sound Change Adaptive?” *Rivista Di Linguistica* 7 (1): 5–37.

Lively, Scott E., John S. Logan, and David B. Pisoni. 1993. “Training Japanese Listeners to Identify English /r/ and /l/. II: The Role of Phonetic Environment and Talker Variability in Learning New Perceptual Categories.” *The Journal of the Acoustical Society of America* 94 (3 Pt 1): 1242–1255.

Logan, John S., Scott E. Lively, and David B. Pisoni. 1991. “Training Japanese Listeners to Identify English /r/ and /l/: A First Report.” *The Journal of the Acoustical Society of America* 89 (2): 874–886.

Lombardi, Linda. 1991. “Laryngeal Features and Laryngeal Neutralization.” PhD diss., University of Massachusetts, Amherst.

Loporcaro, Michele. 2016. “Metaphony and Diphthongization in Southern Italy: Reconstructive Implications for Sound Change in Early Romance.” In *Metaphony and Diphthongization in Southern Italy: Reconstructive Implications for Sound Change in Early Romance*, 55–88. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110366310-004.

Mackay, Carolyn J. 1995. *A Veneto Lexicon: The Dialect of Segusino and Chipilo*. Treviso: Grafiche Antiga.

Mann, Virginia A., and Bruno H. Repp. 1980. “Influence of Vocalic Context on Perception of the [ʃ]-[s] Distinction.” *Perception & Psychophysics* 28 (3): 213–228. DOI: 10.3758/BF03204377.

Martin, James G., and H. Timothy Bunnell. 1981. “Perception of Anticipatory Coarticulation Effects.” *The Journal of the Acoustical Society of America* 69 (2): 559–567. DOI: 10.1121/1.385484.

Martin, James G., and H. Timothy Bunnell. 1982. “Perception of Anticipatory Coarticulation Effects in Vowel–Stop Consonant–Vowel Sequences.” *Journal of Experimental Psychology: Human Perception and Performance* 8 (3): 473–488. DOI: 10.1037/0096-1523.8.3.473.

Matisoff, James A. 1975. “Rhinoglottophilia: The Mysterious Connection between Nasality and Glottality.” In *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, edited by Charles A. Ferguson, Larry M. Hyman, and John J. Ohala, 265–287. Stanford, CA: Language Universals Project.

McGuire, Grant, and Molly Babel. 2012. “A Cross-Modal Account for Synchronic and Diachronic Patterns of /f/ and /θ/ in English.” *Laboratory Phonology* 3 (2): 251–272. DOI: 10.1515/lp-2012-0014.

Michailovsky, Boyd. 1975. “A Case of Rhinoglottophilia in Hayu.” *Linguistics of the Tibeto-Burman Area* 2 (2): 293.

Miller, George A., and Patricia E. Nicely. 1955. “An Analysis of Perceptual Confusions among Some English Consonants.” *The Journal of the Acoustical Society of America* 27 (2): 338–352. DOI: 10.1121/1.1907526.

Moll, K.L., and R.G. Daniloff. 1971. “Investigation of the Timing of Velar Movements during Speech.” *The Journal of the Acoustical Society of America* 50 (2): 678–684. DOI: 10.1121/1.1912683.

Moreton, Elliott. 2021. “Phonological Abstractness in English Diphthong Raising.” *The Publication of the American Dialect Society* 106 (1): 13–44. DOI: 10.1215/00031283-9551267.

Moreton, E., and Erik Thomas. 2007. “Origins of Canadian Raising in Voiceless-Coda Effects: A Case Study in Phonologization.” In *Laboratory Phonology 9*, edited by Jennifer Cole and José Ignacio Hualde, 37–64. Berlin: Mouton de Gruyter.

Müller, Daniela. 2011. “Developments of the Lateral in Occitan Dialects and Their Romance and Cross-Linguistic Context.” PhD diss., Université Toulouse le Mirail – Toulouse II; Ruprecht-Karls-Universität.

Nielsen, K. 2011. “Specificity and Abstractness of VOT Imitation.” *Journal of Phonetics* 39 (2): 132–142. DOI: 10.1016/j.wocn.2010.12.007.

Norris, Dennis, James M. McQueen, and Anne Cutler. 2003. “Perceptual Learning in Speech.” *Cognitive Psychology* 47 (2): 204–238. DOI: 10.1016/S0010-0285(03)00006-9.

Noss, Richard B. 1966. “The Treatment of */r/ in Two Modern Khmer Dialects.” In *Studies in Comparative Austroasiatic Linguistics*, edited by Norman H. Zide, 89–95. London: Mouton & Co.

Ohala, John J. 1975. “Phonetic Explanations for Nasal Sound Patterns.” In *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, edited by C.A. Ferguson, Larry M. Hyman, and John J. Ohala, 289–316. Stanford, CA: Language Universals Project.

Ohala, John J. 1981. “The Listener as a Source of Sound Change.” In *Proceedings from the 17th Annual Meeting of the Chicago Linguistic Society*, vol. 2: *Papers from the Parasession on Language and Behavior*, edited by Carrie S. Masek, Roberta A. Hendrick, and Mary Frances Miller, 178–203.

Ohala, John J. 1983. “The Origin of Sound Patterns in Vocal Tract Constraints.” In *The Production of Speech*, edited by Peter F. MacNeilage, 189–216. New York: Springer.

Ohala, John J. 1989. “Sound Change Is Drawn from a Pool of Synchronic Variation.” In *Language Change: Contributions to the Study of Its Causes*, edited by Leiv E. Breivik and Ernst H. Jahr, 173–198. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110853063.173.

Ohala, John J. 1992. “What’s Cognitive, What’s Not, in Sound Change.” In *Diachrony within Synchrony: Language History and Cognition*, edited by Günter Kellermann and Michael D. Morrissey, 309–355. Frankfurt: Peter Lang.

Ohala, John J. 1993a. “Coarticulation and Phonology.” *Language and Speech* 36 (2–3): 155–170. DOI: 10.1177/002383099303600303.

Ohala, John J. 1993b. “The Phonetics of Sound Change.” In *Historical Linguistics: Problems and Perspectives*, edited by Charles Jones, 237–278. London: Longman.

Ohala, John J. 1994. “Towards a Universal, Phonetically-Based, Theory of Vowel Harmony.” In *3rd International Conference on Spoken Language Processing (ICSLP 1994)*, 491–494. DOI: 10.21437/ICSLP.1994-113.

Ohala, John J. 1997. “Emergent Stops.” In *Proceedings of the 4th Seoul International Conference on Linguistics [SICOL]*, 84–91.

Ohala, John J. 2005. “Phonetic Explanations for Sound Patterns: Implications for Grammars of Competence.” In *A Figure of Speech: A Festschrift for John Laver*, edited by William J. Hardcastle and Janet M. Beck, 23–38. London: Erlbaum. https://escholarship.org/uc/item/5m77t155.

Ohala, John J. 2011. “Accommodation to the Aerodynamic Voicing Constraint and Its Phonological Relevance.” In *Proceedings of the 16th International Congress of Phonetic Sciences*, edited by Jürgen Trouvain and William J. Barry, 64–67.

Ohala, John J. 2012. “The Listener as a Source of Sound Change: An Update.” In *The Initiation of Sound Change: Perception, Production, and Social Factors*, edited by Maria-Josep Solé and Daniel Recasens, 21–36. Amsterdam: John Benjamins.

Ohala, John J., and Maria Grazia Busà. 1995. “Nasal Loss before Voiceless Fricatives: A Perceptually-Based Sound Change.” *Rivista Di Linguistica* 7 (1): 125–144.

Ohala, John J., and James Lorentz. 1977. “The Story of [w]: An Exercise in the Phonetic Explanation for Sound Patterns.” In *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society (1977)*, 577–599. DOI: 10.3765/bls.v3i0.2264.

Ohala, John J., and Manjari Ohala. 1993. “The Phonetics of Nasal Phonology: Theorems and Data.” In *Nasals, Nasalization, and the Velum*, edited by Marie K. Huffman and Rena A. Krakow, 5: 225–249. San Diego: Academic Press. DOI: 10.1016/B978-0-12-360380-7.50013-2.

Ohala, Manjari, and John J. Ohala. 1991. “Nasal Epenthesis in Hindi.” *Phonetica* 48 (2–4): 207–220. DOI: 10.1159/000261885.

Öhman, Sven E.G. 1966. “Coarticulation in VCV Utterances: Spectrographic Measurements.” *The Journal of the Acoustical Society of America* 39 (1): 151–168. DOI: 10.1121/1.1909864.

Öhman, Sven E.G. 1967. “Numerical Model of Coarticulation.” *The Journal of the Acoustical Society of America* 41 (2): 310–320. DOI: 10.1121/1.1910340.

O’Neill, Paul. 2010. “Variación y cambio en las consonantes oclusivas del español de Andalucía.” *Estudios de Fonética Experimental XIX*: 11–41.

Pardo, Jennifer S. 2006. “On Phonetic Convergence during Conversational Interaction.” *The Journal of the Acoustical Society of America* 119 (4): 2382–2393. DOI: 10.1121/1.2178720.

Pardo, Jennifer S. 2012. “Reflections on Phonetic Convergence: Speech Perception Does Not Mirror Speech Production.” *Language and Linguistics Compass* 6 (12): 753–767. DOI: 10.1002/lnc3.367.

Parrell, Benjamin. 2012. “The Role of Gestural Phasing in Western Andalusian Spanish Aspiration.” *Journal of Phonetics* 40 (1): 37–45. DOI: 10.1016/j.wocn.2011.08.004.

Pater, Joe. 1999. “Austronesian Nasal Substitution and Other NC Effects.” In *The Prosody–Morphology Interface*, edited by René Kager, Harry van der Hulst, and Wim Zonneveld, 310–343. Cambridge: Cambridge University Press. http://roa.rutgers.edu/article/view/171.

Penzl, Herbert. 1949. “Umlaut and Secondary Umlaut in Old High German.” *Language* 25 (3): 223–240. DOI: 10.2307/410084.

Pierrehumbert, Janet B. 2001. "Exemplar Dynamics: Word Frequency, Lenition and Contrast." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan L. Bybee and Paul J. Hopper, 137–158. Amsterdam: John Benjamins. DOI: 10.1075/tsl.45.08pie.

Pisoni, David B., and Paul A. Luce. 1987. "Acoustic-Phonetic Representations in Word Recognition." *Cognition* 25 (1–2): 21–52. DOI: 10.1016/0010-0277(87)90003-5.

Plauché, Madelaine, Cristina Delogu, and John J. Ohala. 1997. "Asymmetries in Consonant Confusion." In *EUROSPEECH-1997*, 2187–2190. DOI: 10.21437/Eurospeech.1997-577.

Pouplier, M., and L. Goldstein. 2010. "Intention in Articulation: Articulatory Timing in Alternating Consonant Sequences and Its Implications for Models of Speech Production." *Language and Cognitive Processes* 25 (5): 616–649.

Ramsammy, Michael. 2015. "The Life Cycle of Phonological Processes: Accounting for Dialectal Microtypologies: The Life Cycle of Phonological Processes." *Language and Linguistics Compass* 9 (1): 33–54. DOI: 10.1111/lnc3.12102.

Ratliff, Martha. 2015. "Tonoexodus, Tonogenesis, and Tone Change." In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons, 245–261. Oxford: Oxford University Press.

Recasens, Daniel. 2015. "Phonetic Factors Contributing to the Inception and Evolution of Sound Change." In *Proceedings of the XI Convegno Nazionale Dell'Associazione Italiana Di Scienze Della Voce*, 23–40. DOI: 10.17469/O2101AISV000002.

Recasens, Daniel. 2018. "Coarticulation." In *Oxford Research Encyclopedia of Linguistics.* Oxford: Oxford University Press. DOI: 10.1093/acrefore/9780199384655.013.416.

Recasens, Daniel. 2019. "Typology of Mixing Articulatory Gestures in Phonetics and Phonology." *Loquens* 6 (1): e057–e057. DOI: 10.3989/loquens.2019.057.

Reinisch, Eva, Andrea Weber, and Holger Mitterer. 2013. "Listeners Retune Phoneme Categories across Languages." *Journal of Experimental Psychology: Human Perception and Performance* 39 (1): 75–86. DOI: 10.1037/a0027979.

Repp, Bruno H. 1982. "Phonetic Trading Relations and Context Effects: New Experimental Evidence for a Speech Mode of Perception." *Psychological Bulletin* 92 (1): 81–110.

Riverin-Coutlée, Josiane, and Jonathan Harrington. 2022. "Phonetic Change over the Career: A Case Study." *Linguistics Vanguard* 8 (1): 41–52. DOI: 10.1515/lingvan-2021-0122.

Romero, Joaquín. 1996. "Articulatory Blending of Lingual Gestures." *Journal of Phonetics* 24 (1): 99–111. DOI: 10.1006/jpho.1996.0007.

Ruch, Hanna. 2018. "Perception of Speaker Age and Speaker Origin in a Sound Change in Progress: The Case of /s/-Aspiration in Andalusian Spanish." *Journal of Linguistic Geography* 6 (1): 40–55. DOI: 10.1017/jlg.2018.4.

Ruch, Hanna, and Jonathan Harrington. 2014. "Synchronic and Diachronic Factors in the Change from Pre-Aspiration to Post-Aspiration in Andalusian Spanish." *Journal of Phonetics* 45 (July): 12–25. DOI: 10.1016/j.wocn.2014.02.009.

Ruch, Hanna, and Sandra Peters. 2016. "On the Origin of Post-Aspirated Stops: Production and Perception of /s/ + Voiceless Stop Sequences in Andalusian Spanish." *Laboratory Phonology* 7 (1): 1–36. DOI: 10.5334/labphon.2.

Ruhlen, Merrit. 1978. "Nasal Vowels." In *Universals of Human Language: Phonology*, edited by Joseph H. Greenberg, 203–241. Stanford, CA: Stanford University Press.

Rutter, Ben. 2011. "Acoustic Analysis of a Sound Change in Progress: The Consonant Cluster /stɹ/ in English." *Journal of the International Phonetic Association* 41 (1): 27–40.

Sampson, Rodney. 1999. *Nasal Vowel Evolution in Romance*. Oxford and New York: Oxford University Press.

Samuel, Arthur G., and Tanya Kraljic. 2009. "Perceptual Learning for Speech." *Attention, Perception, & Psychophysics* 71 (6): 1207–1218. DOI: 10.3758/APP.71.6.1207.

Sancier, Michele L., and Carol A. Fowler. 1997. "Gestural Drift in a Bilingual Speaker of Brazilian Portuguese and English." *Journal of Phonetics* 25 (4): 421–436. DOI: 10.1006/jpho.1997.0051.

Sankoff, Gillian. 2004. “Adolescents, Young Adults and the Critical Period: Two Case Studies from ‘Seven Up.’” In *Sociolinguistic Variation: Critical Reflections*, edited by Carmen Fought, 121–139. Oxford: Oxford University Press.

Savoia, Leonardo, and Martin Maiden. 1997. “Metaphony.” In *The Dialects of Italy*, edited by Martin Maiden and Mair Parry, 15–25. London: Routledge.

Schertz, Jessamyn, Taehong Cho, Andrew Lotto, and Natasha Warner. 2015. “Individual Differences in Phonetic Cue Use in Production and Perception of a Non-Native Sound Contrast.” *Journal of Phonetics* 52 (September): 183–204. DOI: 10.1016/j.wocn.2015.07.003.

Schertz, Jessamyn, and Emily J. Clare. 2020. “Phonetic Cue Weighting in Perception and Production.” *WIREs Cognitive Science* 11 (2): e1521. DOI: 10.1002/wcs.1521.

Schleef, Erik, and Michael Ramsammy. 2013. “Labiodental Fronting of /θ/ in London and Edinburgh: A Cross-Dialectal Study.” *English Language and Linguistics* 17 (1): 25–54. DOI: 10.1017/S1360674312000317.

Shapiro, Michael. 1995. “A Case of Distant Assimilation: /Str/ → /Σtr/.” *American Speech* 70 (1): 101–107. DOI: 10.2307/455876.

Silverman, Daniel. 2006. “The Diachrony of Labiality in Trique, and the Functional Relevance of Gradience and Variation.” In *Laboratory Phonology 8*, edited by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 133–154. Berlin: De Gruyter Mouton.

Smith, Bridget J., Jeff Mielke, Lyra Magloughlin, and Eric Wilbanks. 2019. “Sound Change and Coarticulatory Variability Involving English /ɹ/.” *Glossa: A Journal of General Linguistics* 4 (1). DOI: 10.5334/gjgl.650.

Solé, Maria-Josep. 1992. “Phonetic and Phonological Processes: The Case of Nasalization.” *Language and Speech* 35 (1–2): 29–43. DOI: 10.1177/002383099203500204.

Solé, Maria-Josep. 1995. “Spatio-Temporal Patterns of Velopharyngeal Action in Phonetic and Phonological Nasalization.” *Language and Speech* 38 (1): 1–23. DOI: 10.1177/002383099503800101.

Solé, Maria-Josep. 2007. “Controlled and Mechanical Properties in Speech: A Review of the Literature.” In *Experimental Approaches to Phonology*, edited by Maria-Josep Solé, Patrice Speeter Beddor, and Manjari Ohala, 302–321. Oxford: Oxford University Press.

Sonderegger, Morgan A. 2004. “Subglottal Coupling and Vowel Space: An Investigation in Quantal Theory.” BA thesis, Massachusetts Institute of Technology.

Sonderegger, Morgan, Max Bane, and Peter Graff. 2017. “The Medium-Term Dynamics of Accents on Reality Television.” *Language* 93 (3): 598–640. DOI: 10.1353/lan.2017.0038.

Sprigg, R.K. 1987. “‘Rhinoglottophilia’ Revisited: Observations on ‘the Mysterious Connection between Nasality and Glottality.’” *Linguistics of the Tibeto-Burman Area* 10 (1): 44–62.

Stevens, Kenneth N. 1971. “Airflow and Turbulence Noise for Fricative and Stop Consonants: Static Considerations.” *The Journal of the Acoustical Society of America* 50 (4B): 1180–1192. DOI: 10.1121/1.1912751.

Stevens, Kenneth N. 1972. “The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data.” In *Human Communication: A Unified View*, edited by Edward E. David Jr. and Peter B. Denes, 51–56. New York: McGraw-Hill.

Stevens, Kenneth N. 1989. “On the Quantal Nature of Speech.” *Journal of Phonetics* 17: 3–45.

Stevens, Kenneth, and Sheila Blumstein. 1981. “The Search for Invariant Acoustic Correlates of Phonetic Features.” In *Perspectives on the Study of Speech*, edited by Peter D. Eimas and Joanne L. Miller, 1–38. Hillsdale, NJ: Erlbaum.

Stevens, Kenneth N., and Samuel Jay Keyser. 1989. “Primary Features and Their Enhancement in Consonants.” *Language* 65 (1): 81–106. DOI: 10.2307/414843.

Stevens, Kenneth Noble, and Samuel Jay Keyser. 2010. “Quantal Theory, Enhancement and Overlap.” *Journal of Phonetics* 38 (1): 10–19. DOI: 10.1016/j.wocn.2008.10.004.

Stevens, Mary, and Jonathan Harrington. 2016. “The Phonetic Origins of /s/-Retraction: Acoustic and Perceptual Evidence from Australian English.” *Journal of Phonetics* 58 (September): 118–134. DOI: 10.1016/j.wocn.2016.08.003.

Stuart-Smith, J., and C. Timmins. 2006. "'Tell Her to Shut Her Moof': The Role of the Lexicon in TH-Fronting in Glaswegian." In *The Power of Words: Essays in Lexicography, Lexicology and Semantics*: *In Honour of Christian J. Kay*, edited by C.J. Kay, G.D. Caie, C. Hough, and I. Wotherspoon, 171–183. Amsterdam: Rodopi.

Svantesson, Jan-Olof, and David House. 2006. "Tone Production, Tone Perception and Kammu Tonogenesis." *Phonology* 23 (2): 309–333. DOI: 10.1017/S0952675706000923.

Sweet, Herry. 1874. "The History of English Sounds." *Transactions of the Philological Society* 15 (1): 461–623. DOI: 10.1111/j.1467-968X.1874.tb00881.x.

Tabain, Marija. 1998. "Non-Sibilant Fricatives in English: Spectral Information above 10 KHz." *Phonetica* 55 (3): 107–130. DOI: 10.1159/000028427.

Todd, Simon, Janet B. Pierrehumbert, and Jennifer Hay. 2019. "Word Frequency Effects in Sound Change as a Consequence of Perceptual Asymmetries: An Exemplar-Based Model." *Cognition* 185 (April): 1–20. DOI: 10.1016/j.cognition.2019.01.004.

Torreira, Francisco. 2012. "Investigating the Nature of Aspirated Stops in Western Andalusian Spanish." *Journal of the International Phonetic Association* 42 (1): 49–63.

Torreira, Francisco, and Mirjam Ernestus. 2011. "Realization of Voiceless Stops and Vowels in Conversational French and Spanish." *Laboratory Phonology* 2 (2): 331–353. DOI: 10.1515/labphon.2011.012.

Trouvain, Jürgen. 2021. "Aerodynamic Voicing Constraint in Stops: From Wolfgang von Kempelen to John Ohala." In *Proceedings of Fourth International Workshop on the History of Speech Communication Research (HSCR 2021)*, 75–81. DOI: 10.21437/HSCR.2021-7.

Tuttle, Edward F. 1991. "Nasalization in Northern Italy: Syllabic Constraints and Strength Scales as Developmental Parameters." *Rivista di Linguistica* 3 (1): 33–92.

Twaddell, William F. 1938. "A Note on Old High German Umlaut." *Monatshefte für Deutschen Unterricht* 30 (3/4): 177–181.

Warren, Paul. 2006. "/S/-Retraction, /t/-Deletion and Regional Variation in New Zealand English /Str/ and /Stj/ Clusters." In *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*, edited by Paul Warren and Catherine I. Watson, 466–471.

Wayland, Ratree P., and Susan G. Guion. 2005. "Sound Change Following the Loss of /r/ in Khmer: A New Tonogenetic Mechanism?" *Mon-Khmer Studies* 35: 55–82.

Wedel, Andrew. 2004. "Category Competition Drives Contrast Maintenance within an Exemplar-Based Production/Perception Loop." In *Proceedings of the 7th Meeting of the ACL Special Interest Group in Computational Phonology: Current Themes in Computational Phonology and Morphology*, 1–10.

Wedel, Andrew B. 2006. "Exemplar Models, Evolution and Language Change." *The Linguistic Review* 23 (3). DOI: 10.1515/TLR.2006.010.

Wedel, Andrew B. 2007. "Feedback and Regularity in the Lexicon." *Phonology* 24 (1): 147–185. DOI: 10.1017/S0952675707001145.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Whalen, Douglas H., Leigh Lisker, Arthur S. Abramson, and Maria Mody. 1993. "F0 Gives Voicing Information Even with Unambiguous Voice Onset Times." *The Journal of the Acoustical Society of America* 93 (4): 2152–2159.

Winitz, Harris, M.E. Scheib, and James A. Reeds. 1972. "Identification of Stops and Vowels for the Burst Portion of /p, t, k/ Isolated from Conversational Speech." *The Journal of the Acoustical Society of America* 51 (4B): 1309–1317. DOI: 10.1121/1.1912976.

Winn, Matthew B., Monita Chatterjee, and William J. Idsardi. 2013. "The Roles of Voice Onset Time and F0 in Stop Consonant Voicing Perception: Effects of Masking Noise and Low-Pass Filtering." *Journal of Speech, Language, and Hearing Research: JSLHR* 56 (4): 1097–1107. DOI: 10.1044/1092-4388(2012/12-0086).

Xu, Yi, and Anqi Xu. 2021. "Consonantal F0 Perturbation in American English Involves Multiple Mechanisms." *The Journal of the Acoustical Society of America* 149 (4): 2877–2895. DOI: 10.1121/10.0004239.

Yang, Meng, and Megha Sundara. 2019. "Cue-Shifting between Acoustic Cues: Evidence for Directional Asymmetry." *Journal of Phonetics* 75 (July): 27–42. DOI: 10.1016/j.wocn.2019.04.002.

Yu, Alan C.L. 2019. "On the Nature of the Perception-Production Link: Individual Variability in English Sibilant-Vowel Coarticulation." *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10 (1): 2. DOI: 10.5334/labphon.97.

Yu, Alan C.L., Carissa Abrego-Collier, and Morgan Sonderegger. 2013. "Phonetic Imitation from an Individual-Difference Perspective: Subjective Attitude, Personality and 'Autistic' Traits." *PLoS ONE* 8 (9): e74746. DOI: 10.1371/journal.pone.0074746.

Yu, Alan C.L., and Hyunjung Lee. 2014. "The Stability of Perceptual Compensation for Coarticulation within and across Individuals: A Cross-Validation Study." *The Journal of the Acoustical Society of America* 136 (1): 382–388. DOI: 10.1121/1.4883380.

Zellou, Georgia. 2017. "Individual Differences in the Production of Nasal Coarticulation and Perceptual Compensation." *Journal of Phonetics* 61 (March): 13–29. DOI: 10.1016/j.wocn.2016.12.002.

Zellou, Georgia. 2022. *Coarticulation in Phonology*. Cambridge: Cambridge University Press. DOI: 10.1017/9781009082488.

Zellou, Georgia, Santiago Barreda, and Bruno Ferenc Segedin. 2020. "Partial Perceptual Compensation for Nasal Coarticulation Is Robust to Fundamental Frequency Variation." *The Journal of the Acoustical Society of America* 147 (3): EL271–276. DOI: 10.1121/10.0000951.

Zellou, Georgia, Michelle Cohn, and Aleese Block. 2021. "Partial Compensation for Coarticulatory Vowel Nasalization across Concatenative and Neural Text-to-Speech." *The Journal of the Acoustical Society of America* 149 (5): 3424–3436. DOI: 10.1121/10.0004989.

Zellou, Georgia, Rebecca Scarborough, and Kuniko Nielsen. 2016. "Phonetic Imitation of Coarticulatory Vowel Nasalization." *The Journal of the Acoustical Society of America* 140 (5): 3560–3575. DOI: 10.1121/1.4966232.