WILEY THE WILEY BLACKWELL COMPANION TO DIACHRONIC LINGUISTICS

Acoustic-perceptual factors in the actuation of sound change

Journal:	*The Wiley Blackwell Companion to Diachronic Linguistics*
Manuscript ID	Draft
Wiley - Manuscript type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Kirby, James; Ludwig-Maximilians-Universität München Institut für Phonetik und Sprachverarbeitung Harrington, Jonathan; Ludwig-Maximilians-Universität München Institut für Phonetik und Sprachverarbeitung
Keywords:	phonetics, enhancement, speech perception, sound change
Abstract:	Speech is characterized by ubiquitous variation. Although variation is often quite stable, both within speech communities and over long spans of historical time, it is widely accepted that it is this systemic phonetic variation that provides the 'seeds' of sound change (Kiparsky 1995; Ohala 1989; Blevins 2004; Lindblom et al. 1995). But how does variation become change? The focus of this chapter is on changes characterized primarily by the first stage of the life cycle model (see Bermúdez-Otero 2015; Bermúdez-Otero & Trousdale 2012; Ramsammy 2015; WBCDL018), i.e. 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.

Acoustic-perceptual factors in the actuation of sound change

1. Introduction: acoustic-perceptual factors in sound change

Speech is characterized by ubiquitous variation. Although variation is often quite stable, both within speech communities and over long spans of historical time, it is widely accepted that it is this systemic phonetic variation that provides the 'seeds' of sound change (Kiparsky 1995; Ohala 1989; Blevins 2004; Lindblom et al. 1995). But how does variation become change? The focus of this chapter is on changes characterized primarily by the first stage of the life cycle model (see Bermúdez-Otero 2015; Bermúdez-Otero & Trousdale 2012; Ramsammy 2015; WBCDL018), i.e. 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, i.e. contextually-conditioned variation in the acoustic realization of a speech sound (Ohala 1993b; Recasens 2018; Zellou 2022). Commonly cited examples include the emergence of umlaut due to vowel-to-vowel coarticulation (Ohala 1994; Grosvald & Corina 2012), the development of pitch-based (tonal) contrasts grounded in consonant-to-vowel coarticulation (Hombert, Ohala & Ewan 1979; WBCDL076), or the emergence of contrastive vowel nasalization due to retiming of oral and velic gestures (Kawasaki 1986; Cohn 1993; Solé 1992; 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 §2, we discuss a number of these changes and consider different accounts of their actuation.

While phonetic variation is ubiquitous in speech, it can persist over long periods of time. As a result, 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 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 & Busà 1995: 3). For the purposes of this chapter when we talk about 'sound change' we mean it in this narrow sense: a consistent change in production accompanied by some type of realignment in perception. What 'parsed together' in perception means is a topic we will cover in more detail in §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 §4, we discuss several definitions of enhancement and how they relate to sound change. Finally, in §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

Scholars of sound change have often made a distinction between changes having a clear phonetic conditioning *of some type*, and those that have no such obvious phonetic conditioning (Garrett & Johnson 2013; Garrett 2015). The set of changes with clear phonetic conditioning can be further subdivided by their putative phonetic origin, namely in the articulatory, aerodynamic, or acoustic-perceptual properties. The development of contrastive vowel nasalization from historical VN sequences (Beddor 2009; Ohala 1975) 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; Ohala 2011; Trouvain 2021).

However, changes may also 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 some type of reconfiguration at some point, whether due to error, cue reweighting, or gestural reparsing (see §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 (Garrett & 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), Garrett & Johnson (2013), Recasens (2015), Lehnert-Lehouillier (2010; 2013) 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] > ME *chin*
   OE *ciese* [kiːese] > ME *cheese*

2. *Palatalization in Romance*
   Latin *centum* ['kɛntum] > Ital. *cento* ['tʃɛnto]
   Latin *cantare* [kan'taːre] > Old French *chanter* ['tʃan'ter] (cf. Engl *chant*)

3. *Slavic ‘First Regressive Palatalization’*
   PIE **kers-* > OCS *černý*, Cz *černý,* Ru черный ‘black’ (cf. Skst. *kṛṣṇa* ‘Krishna’)
   PIE **kʷetwer-* > OCS *četyre*, Cz čtyři, Ru четыре ‘four’ (cf. Latin *quadra-*)
   PIE **gʷḗn* > OCS *žena*, Cz *žena*, Sb жена (cf. ME *queen* < OE *cwēn*, Ger *Königin*)

PIE **g^{w}iH_3wo-* > OCS *životŭ* 'life', *živŭ* 'alive', Cz *život*, Ru жизнь
(cf. ME *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 (i.e., articulatory overlap) is implicated in velar palatalization inasmuch as the production of [k] will be affected by the front vowel target. Still, articulatorily fronted [k^ji] 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 et al. (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 et al. (1997) and Chang et al. (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 aspiration noise as a fricative (/ki/ > [k^hi] > [tʃi], cf. §2.5).

Garrett and Johnson (2013) problematize the acoustic-perceptual account by (a) noting the absence of studies of ongoing changes that document a clear [k^ji] > [tʃi] shift without any intermediate variants, and (b) suggesting that the affrication step could be the result of listener-oriented enhancements on the part of the speaker (see §5). Recasens (2015, §1.2.3) also suggests 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 Scottish English varieties (Stuart-Smith & Timmins 2006; Schleef & Ramsammy 2013), many British English varieties (Britain 2005)

c. South Slavey *θa* > Mountain Slavey *fa* 'sand', *θaj* ~ *fa* 'tent pole', *θɛ-* ~ *fɛ-* PERFECTIVE (Flynn & Fulop 2014)

d. Veneto Italian θ*émena* ~ f*emena* (Ital. 'femmina'; Eng. 'woman'); θ*onc* ~ f*onc* (cf. Ital. 'fungo', Eng. 'mushroom') (Mackay 1995:xvii)

Observers as early as Sweet (1874: 470) cast doubt on incremental articulatory explanations for these correspondences[1]. 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 & Nicely 1955; McGuire & Babel 2012; Harris 1958) points towards an acoustic-perceptual trigger. However, unlike velar palatalization, most phonetic studies have found relatively symmetric spectral similarities between the two sounds (Tabain 1998; Hughes & Halle 1956; Heinz & Stevens 1961; Jongman, Wayland & 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 & Nicely 1955; Jones 2002; McGuire & 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. Results showed that 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/.[2]

Garrett & Johnson (2013: 71-2) 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 (§4.4), rather than the cause being fundamentally acoustic in nature. For further discussion see Garrett & Johnson (2013) and Blevins (2019).

2.3 Nasalization and aspiration

[1] "The not unfrequent change of *th* into *f* is no doubt purely imitative".
[2] As noted by Blevins (2019), this account has a further advantage in that it is only available for languages in which /f/ already exists as a (visible) articulatory target. This would explain why context-free *θ > f changes are nearly always mergers: if there is no /f/ in a language, then there is nothing to cue the articulatory-acoustic association, and (correctly) no sound change is predicted.

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, e.g. 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 the nasal(ization) entirely (see §3). Experimental evidence supporting this account comes from Ohala & 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 makes 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 also 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 thus be the ultimate origin of spontaneous nasalization, with the nasal percept enhanced by the acoustic consequences of breathiness (Klatt & Klatt 1990).

2.4 Asymmetric sound changes involving labiovelar approximants

Ohala & 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 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 targeted by a nasal (5) and /w/ behaving as a labial when targeted by a fricative (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 targeted by 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 & 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.

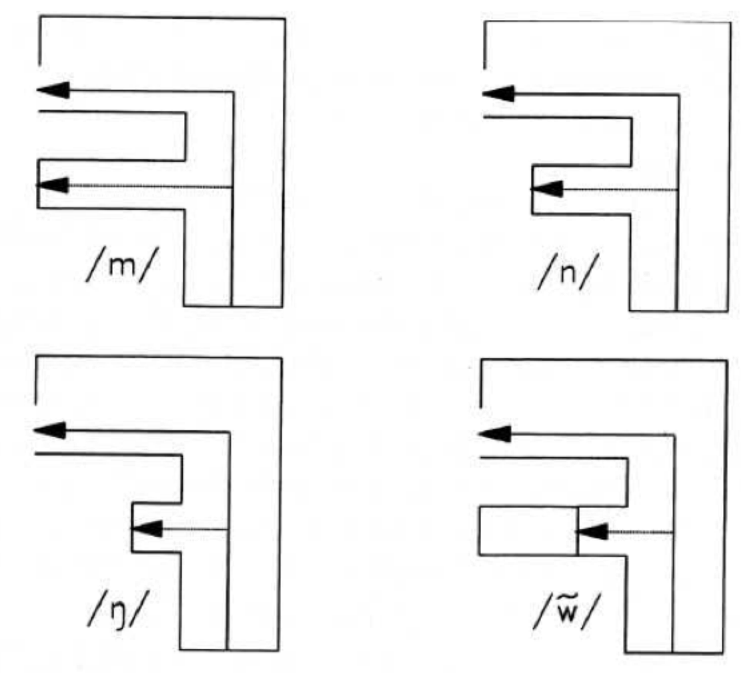

Figure 1: schematic representation of vocal tract configurations for /m/, /n/, /ŋ/ and (nasalized) /w/. Solid arrows indicate airflow contributing resonances;dotted arrows indicate anti-resonances. After Ohala and Lorentz (1977), Figure 2.

Ohala & Lorentz list a number of 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, 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 will be more likely to develop into labial, rather than velar plosives, which finds support in sound changes such as $*k^w$ > p in e.g. Latin *equus* > Greek *hippos*.

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

As noted in the preceding sections, Garrett & 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. However, as all the preceding examples have made clear, it is rarely 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; Kirby 2014a; Wayland & Guion 2005). 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 & Guion (2005) argued that the catalyst for this change was fundamentally aerodynamic: the propose 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 then 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 & 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 F1 lowering (resulting in vowel raising) and f0 (resulting in pitch lowering). However, the trajectory of this change in other dialects points to a more nuanced interaction between articulatory, aerodynamic, and acoustic-perceptual bias factors; see Kirby & Giang (2017) and WBCDL076 (§3.1).

3. Perception, production, and gestural parsing

As defined in §1, sound change – a change in a pronunciation norm – is initiated when listeners cease to parse together phonetic cause and effect. In §2, we have reviewed changes in which both phonetic cause and effect are 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. To understand how changes in parsing can come about, in §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 §3.2 and §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 & Goldstein 1986; Browman & Goldstein 1989; Browman & Goldstein 1992; Pouplier & Goldstein 2010) which seeks to provide a unified explanation of diverse types of synchronic variation in connected speech including assimilation, deletion, and weakening.

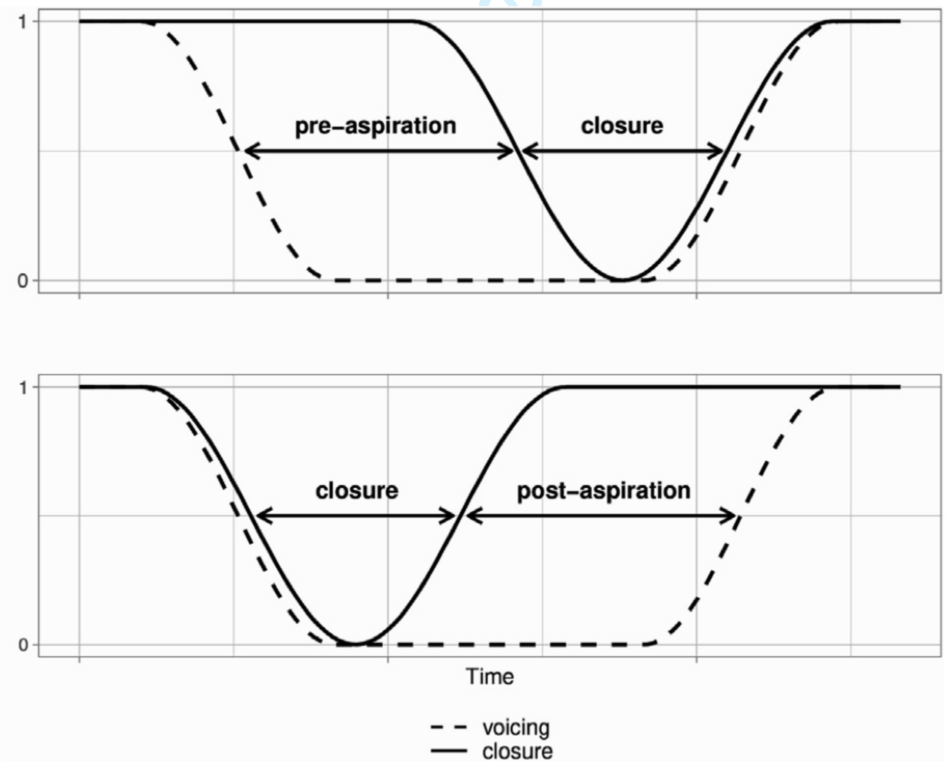

Figure 2: Idealized scheme of resynchronization of the closure with the voiceless interval in Andalusian Spanish /s/-aspiration. The solid line is the glottal gesture, where low values stand for an open glottis and hence voicelessness, the dashed line is the oral constriction gesture of the voiceless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

In articulatory phonology (AP), gestures are autonomous and can vary in their relative phasing. Fig. 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 (Herrero de Haro & Hajek 2022; O'Neill 2010; Ruch 2018; Ruch & Harrington 2014; Ruch & Peters 2016; Torreira 2012): thus, /pa^{h}ta/ →/pat^{h}a/ or /pat^{s}a/ (*pasta*) which contrasts with an unaspirated, 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 modelled as a re-phasing of a closure gesture for /t/ relative to an unchanging open glottal gesture for voicelessness (Cronenberg et al. 2020; Parrell 2012). 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 §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* (Bell-Berti & Krakow 1991; Cohn 1990; Cohn 1993; Delvaux, Metens & Soquet 2002; Delvaux et al. 2008; Kent, Carney & Severeid 1974; Solé 1992; Solé 1995; Moll & Daniloff 1971) in VN rhymes (e.g., /æn/ in *pan*, *pan's, pant*, *panned*) has been modelled 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; Beddor et al. 2013; Beddor 2009; Beddor 2012). 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', /sõ/, from Latin *sonus;* cf. Italian *suono*).

- The *deletion or partial deletion of consonants in clusters* such as the /t/ in *perfect memory* has been modelled by an earlier phasing of the labial gesture for /m/ which overlaps and therefore hides the tongue tip gesture of /t/ (Browman & Goldstein 1990). Synchronically, there is evidence from Georgian (Chitoran, Goldstein & 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), e.g., 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 e.g., *empty* (< Old English 'amtig'), *humble* (cf. *humility*), *nimble* (< Old English 'næmel') and in French *chambre* (cf. Italian and Latin *camera*).

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 (de Jong, Beckman & Edwards 1993; Recasens 2019; Fowler & Saltzman 1993; Romero 1996). Some examples of synchronic and diachronic variation due to gestural blending include:

- */u/-fronting* (Alderton 2020; Bauer 1985; Fridland 2008; Harrington, Kleber & Reubold 2008; Hawkins & Midgley 2005; Kataoka 2011; Labov, Ash & Boberg 2006) in which for e.g., German *tut* (*does*), opposing forces act on the tongue dorsum for its retraction due the back vowel /u/ and its advancement due to the tongue tip articulations of the consonants (Flemming 2003; Flemming 2004; Öhman 1966). /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., Baker, Archangeli & Mielke 2011; Lawrence 2000; Rutter 2011; Shapiro 1995), Australian (Stevens & Harrington 2016) and New Zealand (Warren 2006) English varieties.

- Synchronic *transconsonantal (VCV) vowel coarticulation* is well documented (Cole et al. 2010; Hoole & Pouplier 2017; Öhman 1966). An associated sound change is metaphony in some varieties of Italian (Loporcaro 2016; Savoia & Maiden 1997) and umlaut in German (Iverson & Salmons 2003; Kiparsky 2015; Penzl 1949; Twaddell 1938) as in e.g., /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 (Carignan et al. 2021; Harrington, Fletcher & Roberts 1995). 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 & Nadeu 2011; Torreira & Ernestus 2011) and fricatives (Hualde & 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 /lu:s/ changed to /ly:/ 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 re-categorization 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; 1993a; 2005; 2012) conception of sound change. As discussed above, in some varieties of English the /s/ in /str/ clusters tends towards 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/ (Jongman, Wayland & Wong 2000; Koenig et al. 2013; Stevens 1971). 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 & 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 (1993a) 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 normalize perceptually for context (Mann & 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 & Smith 1986; Fowler & 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, i.e. normalizes insufficiently for context. By contrast, in sound changes due to dissimilation (Alderete & Frisch 2007; see Abrego-Collier 2013; Harrington, Kleber & Stevens 2016; Jatteau & Hejná 2016 for related experimental analyses), 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/ (Browman & Goldstein 1991; Guthrie 1967; Ohala 1981), 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 produce 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) the application of a 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 1993a): 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; Beddor et al. 2018) as well as those based on exemplar theory (Pierrehumbert 2001; Todd, Pierrehumbert & Hay 2019), sound change is predicted to emerge continually within the same individual rather than due to a switch between categorically different variants. Moreover, both models distance themselves from Ohala's idea that sound charge is caused by any kind of error.

The progression from continuous phonetic variation to a categorical change in Beddor's model of sound change also has its origins in speech perception and its relationship to gestures in articulatory phonology. Beddor's model is especially relevant for a particular class of sound change whose outcome is TRANSPHONOLOGIZATION (Hyman 2013; Hyman 1976; Kiparsky 2015). These are sound changes in which the cues are transferred from a coarticulatory source to a coarticulatory effect often leading to the loss or near loss of the source. They include umlaut and metaphony that have their origins in trans-consonantal vowel coarticulation (see §3.1); the compensatory lengthening of vowels due to loss of following consonants (Kavitskaya 2001); and (the main focus of 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 §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 (Alfonso & Baer 1982; Beddor et al. 2013; Connine & Darnieder 2009; Martin & Bunnell 1981); and secondly that listeners often only compensate for coarticulation *partially* (Beddor, Harnsberger & Lindemann 2002; Beddor 2009; Fowler & Brown 2000; Yu & Lee 2014; Zellou 2017; Zellou, Cohn & Block 2021; Zellou, Barreda & Ferenc Segedin 2020). 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 & 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 & 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 (Pisoni & Luce 1987; Repp 1982) 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 (Beddor 2012; Beddor 2015; Chandrasekaran, Sampath & Wong 2010; Clayards 2018; Kim & Clayards 2019; Schertz et al. 2015). 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 VNÇ contexts like *bent* than a voiced VNÇ contexts like *bend* (Busà 2007; Hajek 1997; Ohala & Ohala 1991; Ruhlen 1978; Tuttle 1991; Sampson 1999). This is because nasalization and a following voiceless consonant are less compatible. Nasalization causes the presence of low frequency energy (Fujimura 1962; House & Stevens 1956) whereas the perception of a voiceless consonant requires its absence (Ohala & Ohala 1991; Ohala & Ohala 1993). This acoustic incompatibly is likely to be one the factors that contributes both to the relative scarcity of NÇ clusters in the languages of the world (Itô & 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*; English *wish*, German *Wunsch* etc.). Beddor's studies show that there is a greater skewed relationship towards 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 e.g., *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 NÇ̥ than voiced NÇ̬ clusters).

Some studies are also beginning to show an analogous relationship between perception and production at the level of the individual. Thus, there is now 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 tonogenesis sound change in two varieties of Afrikaans by Coetzee et al (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 recent comprehensive survey of cue weighting in perception and production, Scherz & 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 et al. 2019 for a similar conclusion).

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, i.e. "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 & Fairbanks 1953; Lehiste & Peterson 1961; Kohler 1982; Hanson 2009; Chen 2011; Dmitrieva et al. 2015; Xu & 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 & House 2006; Kirby, Pittayaporn & 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 (Hombert, Ohala & Ewan 1979; Solé 1992; Solé 1995; Solé 2007). 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 towards 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 & 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 Keyser & Stevens regard as an enhancement of the feature [-anterior]). Phonetic enhancement can also involve the introduction 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 accompanying the production of [-voice] consonants in many languages are an enhancement of the defining feature [+stiff vocal folds] (Stevens & Keyser 2010). Prenasalization of [+voice] consonants may be similarly regarded as an enhancement of the [voice] contrast (Keyser & Stevens 2006).

The essential aspects of this type of enhancement are (a) that it enhances a *contrast*, not a particular articulatory or acoustic dimension and (b) it presupposes a dichotomy on which all features are either universally primary or universally secondary. Other approaches, such as Hall's (2011; 2007) 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 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, under this system, this phoneme would only be predicted to be a target of a phonetic enhancement such as lip rounding 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. §4.4).

4.2. Articulatory enhancement

Garrett & 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 §4.3 below, involves the addition of a new gesture; Garrett & 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 §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 & 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 & Johnson conception of articulatory enhancement differs subtly from that of Keyser & Stevens: for Keyser & Stevens, enhancement can be change in pronunciation that aids in the recovery of a 'primary property', whereas for Garrett & 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 & Kluender 1989; Kingston & Diehl 1994; Kingston et al. 2008; Diehl 2008; Diehl 2011). The Auditory Enhancement Hypothesis proposes that sound systems tend to be characterized by combinations of articulatory properties that have 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 & 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 & 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 (§3.2) in that it involves the introduction of a new feature, rather than exaggeration or realignment of an existing one (cf. Stevens & Keyser 2010, Sec. 3). In addition to the classic example of lip rounding enhancing the salience of the [back] contrast in vowels (Liljencrants & 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. §4.1 and §2.2, respectively). While closely related to Kingston & Diehl's notion of auditory enhancement, Garrett & 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 listener's memories (cf. §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 the talker will target based on a cue's distributional INFORMATIVENESS in terms of signaling the phonetic contrast (see also Clayards 2008). Like Garrett & Johnson's notion of articulatory enhancement (§4.2), this involves strengthening of a contrast through emphasizing a redundant cue, but it 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 §2) the perceptibility of the contrast is threatened. Examples of sound changes for which adaptive enhancement provides a compelling account include the phonologization of F0 in Seoul Korean (Kirby 2013; Bang et al. 2018) and the phonologisation of vowel duration and nasalised /æ/ in Australian English (Cox & Palethorpe 2014).

In terms of helping to understand how redundant cues are exaggerated, probabilistic enhancement predicts that bias affecting production are fundamentally 'responsible' for enhancement of the secondary/redundant cues which ultimately become phonologized. 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 §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). While not intended as a general theory of sound change, probabilistic enhancement suggests a more fine-grained understanding of how changes might unfold over time.

4.5. Passive enhancement

It has also been proposed that listener-driven enhancements need not be 'listener-oriented', i.e. attributable to an assessment of communicative need, but can instead be passive in nature. Silverman (2006) discusses the phonetic underpinnings of labial spreading in Trique, whereby historical *uka > uk^wa, 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 were easily perceived, 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. §4.3 on auditory enhancement). While many of the other accounts reviewed above also make reference to an existing phonetic dimension, this account differs in terms of how it proposes that dimension comes to be enhanced: via the natural selection of more extreme variants over time, rather than through any explicit targeting of those variants by speakers.

4.6. Quantal theory and enhancement

The QUANTAL THEORY of speech (Stevens 1972; Stevens 1989) proposes that the relationship between change in an articulatory parameter and its acoustic consequences may be nonlinear. 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 change 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 et al. (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 Bavarian German (e.g. *vui* vs. SG *viel* 'much', *Abrui* vs. SG *April*, *Gäid* vs. SG *Geld* 'money'). Lin et al. (2014) show 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 of as a back vowel. /s/-retraction is also likely to have a quantal component give that, as Stevens (1972; 1989) shows there is a sudden perceptual jump from [s] to [ʃ] as the tongue tip slides back towards the hard palate. The ongoing fronting of /u/ in British English (Harrington, Kleber & 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 (Sonderegger 2004; Chi & Sonderegger 2007; Stevens 1989).

5. Perceptual learning

(Mis)-perception based accounts of the initiation of change such as Ohala's (§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 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 extremely rapid: listeners are very adept at adjusting to nonlocal accents and contexts, as demonstrated from a large literature on PERCEPTUAL LEARNING (Gibson 1963; Goldstone 1998; for overviews, see Samuel & Kraljic 2009; Cutler 2012). In speech, perceptual learning refers to the general phenomenon that, when making use of contextual information to resolve ambiguous speech acoustic 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 (Lively, Logan & Pisoni 1993; Logan, Lively & Pisoni 1991; Bradlow et al. 1997; Bradlow et al. 1999; Bradlow & Bent 2008; Winn, Chatterjee & 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 et al. (2003). In this study, Dutch listeners heard words in which a final fricative (/f/ or /s/) was replaced by an ambiguous sound [?] acoustically halfway been [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 *naadlbos* 'pine forest' was replaced with the ambiguous sound. Crucially, in both conditions, there existed no Dutch word with the alternative fricative: i.e., **witlos* and **naadlbof* are not possible Dutch words. Norris et al. 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 *naadlbo?* 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, 2006; Kraljic and Samuel 2005, 2006, 2007; Reinisch, Weber, & 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. Idemaru & Holt 2011; Idemaru & Holt 2012; Idemaru & Holt 2014; Idemaru & Holt 2020; Yang & Sundara 2019; Lehet & Holt 2020; Kim, Clayards & Kong 2020; Francis & Nusbaum 2002), listeners are exposed to speech input in which cue distributions differ from their canonical, long-term regularities. For example, Idemaru & 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 & 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 (§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 abrupt loss of a primary cue, such as spontaneous consonant devoicing, which could redirect attention to a secondary acoustic property. A study by Gao & 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 in identifying voiceless plosives, but did not affect their responses to voiced plosives. However, when lexical feedback was given, decreased informativeness of VOT did lead 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 & 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; WBCDL076).

5.3 Perceptual learning, phonetic convergence, and speech production

While the findings of the perceptual learning literature suggests 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, it remains unclear how, or whether, they 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. More recently, 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; Babel 2012; Kwon

2019; Kwon 2021; Zellou, Scarborough & Nielsen 2016; Nielsen 2011; Yu, Abrego-Collier & Sonderegger 2013; Kim & Clayards 2019) which has documented 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). For scholars who assume that phonetic imitation in the laboratory is a reasonable proxy for conversational interaction "in the wild", imitation and convergence are hypothesized to be an important locus of sound change propagation (Babel 2012).

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 re-weighting, and phonetic accommodation is an area in need of further investigation, but are clearly of great importance for our understanding 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.

References

Abrego-Collier, Carissa. 2013. Liquid dissimilation as listener hypocorrection. *Annual Meeting of the Berkeley Linguistics Society* 37(1). 3–17. https://doi.org/10.3765/bls.v37i1.3195.

Alderete, John & Stefan A. Frisch. 2007. Dissimilation in grammar and the lexicon. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, 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*. SAGE Publications Inc 48(1). 72–96. https://doi.org/10.1177/0075424219896400.

Alfonso, Peter J. & Thomas Baer. 1982. Dynamics of vowel articulation. *Language and Speech*. SAGE Publications Ltd 25(2). 151–173. https://doi.org/10.1177/002383098202500203.

Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40(1). 177–189. https://doi.org/10.1016/j.wocn.2011.09.001.

Baker, Adam, Diana Archangeli & Jeff Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23(3). 347–374. https://doi.org/10.1017/S0954394511000135.

Bang, Hye-Young, Morgan Sonderegger, Yoonjung Kang, Meghan Clayards & 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. 120–144. https://doi.org/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. https://doi.org/10.1016/S0095-4470(19)30726-0.

Beckman, Mary E., Kenneth De Jong, Sun-Ah Jun & Sook-Hyang Lee. 1992. The interaction of coarticulation and prosody in sound change. *Language and Speech* 35(1–2). 45–58. https://doi.org/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*, 249–254. Saarbrücken.

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 Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change: Perception, production, and social factors*, 37–56. Amsterdam: John Benjamins. https://www.jbe-platform.com/content/books/9789027273666-cilt.323.06bed. (8 November, 2021).

Beddor, Patrice Speeter. 2015. The relation between language users' perception and production repertoires. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow.

Beddor, Patrice Speeter, Andries W. Coetzee, Will Styler, Kevin B. McGowan & 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. https://doi.org/10.1353/lan.2018.0051.

Beddor, Patrice Speeter, James D. Harnsberger & Stephanie Lindemann. 2002. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* 30(4). 591–627. https://doi.org/10.1006/jpho.2002.0177.

Beddor, Patrice Speeter & 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. https://doi.org/10.1121/1.428111.

Beddor, Patrice Speeter, Kevin B. McGowan, Julie E. Boland, Andries W. Coetzee & 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 & Rena Arens Krakow. 1991. Anticipatory velar lowering: A coproduction account. *The Journal of the Acoustical Society of America* 90(1). 112–123. https://doi.org/10.1121/1.401304.

Bermúdez-Otero, Ricardo. 2015. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford Handbook of the History of English*, 691–720. 1st edn. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0059.

Blevins, Juliette. 2004. *Evolutionary phonology*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2006. New Perspectives on English Sound Patterns: "Natural" and "Unnatural" in Evolutionary Phonology. *Journal of English Linguistics* 34(1). 6–25. https://doi.org/10.1177/0075424206287585.

Blevins, Juliette. 2015. Evolutionary Phonology: a holistic approach to sound change typology. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford Handbook of Historical Phonology*, 485–500. Oxford: Oxford University Press.

Blevins, Juliette. 2019. Notes on θ > f and f > θ: Deconstructing markedness in sound change typology. In Lars Heltoft, Iván Igartua, Brian D. Joseph, Kirsten Jeppesen Kragh &

Lene Schøsler (eds.), *Perspectives on Language Structure and Language Change: Studies in honor of Henning Andersen* (Current Issues in Linguistic Theory), 107–122. John Benjamins Publishing Company.
Blood, David L. 1967. Phonological units in Cham. *Anthropological Linguistics*. [Anthropological Linguistics, Trustees of Indiana University] 9(8). 15–32.
Blust, Robert A. 2009. *The Austronesian languages* (Pacific Linguistics 602). Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
Boersma, Paul. 1998. *Functional phonology*. University of Amsterdam PhD dissertation.
Bradley, David (Ed). 1985. Arakanese vowels. PDF. In Graham Thurgood, James A. Matisoff & David Bradley (eds.), *Linguistics of the Sino-Tibetan area: the state of the art*, 180–200. The Australian National University: Research School of Pacific Studies.
Bradlow, Ann, David Pisoni, Reiko Akahane-Yamada & 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. 2299–310. https://doi.org/10.1121/1.418276.
Bradlow, Ann R., Reiko Akahane-Yamada, David B. Pisoni & 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. https://doi.org/10.3758/BF03206911.
Bradlow, Ann R. & Tessa Bent. 2008. Perceptual adaptation to non-native speech. *Cognition* 106(2). 707–729. https://doi.org/10.1016/j.cognition.2007.04.005.
Britain, David J. 2005. The dying dialects of England? In Antonio Bertacca (ed.), *Historical linguistic studies of spoken English*, 35–46. Pisa: Edizioni Plus.
Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology*. Cambridge University Press 6(2). 201–251. https://doi.org/10.1017/S0952675700001019.
Browman, Catherine P. & Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary E. Beckman (eds.), *Papers in Laboratory Phonology: Volume 1: Between the Grammar and Physics of Speech* (Papers in Laboratory Phonology), vol. 1, 341–376. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511627736.019.
Browman, Catherine P. & Louis Goldstein. 1992. Articulatory Phonology: An overview. *Phonetica* 49(3–4). 155–180. https://doi.org/10.1159/000261913.
Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology* 3. 219–252. https://doi.org/10.1017/S0952675700000658.
Browman, Catherine P. & Louis M. Goldstein. 1991. Gestural structures: distinctiveness, phonological processes, and historical change. In Ignatius G. Mattingly & Michael Studdert-Kennedy (eds.), *Modularity and the motor theory of speech perception*, 313–338. Hillsdale, NJ: Lawrence Erlbaum.
Busà, Maria Grazia. 2007. Coarticulatory nasalization and phonological developments: Data from Italian and English nasal-fricative sequences. In Maria-Josep Solé, Patrice S. Beddor & Manjari Ohala (eds.), *Experimental approaches to phonology*, 155–174. Oxford: Oxford University Press.
Carignan, Christopher, Stefano Coretta, Jens Frahm, Jonathan Harrington, Phil Hoole, Arun Joseph, Esther Kunay & Dirk Voit. 2021. Planting the seed for sound change: Evidence from real-time MRI of velum kinematics in German. *Language*. Linguistic Society of America 97(2). 333–364. https://doi.org/10.1353/lan.2021.0020.
Catford, J. C. 2001. On Rs, rhotacism and paleophony. *Journal of the International Phonetic Association* 31(2). 171–185. https://doi.org/10.1017/S0025100301002018.

Chandrasekaran, Bharath, Padma D. Sampath & 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. https://doi.org/10.1121/1.3445785.
Chang, Steve S., Madelaine Plauché & John J. Ohala. 2001. Markedness and consonant confusion asymmetries. In Elisabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 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. https://doi.org/10.1016/j.wocn.2011.04.001.
Chi, Xuemin & Morgan Sonderegger. 2007. Subglottal coupling and its influence on vowel formants. *The Journal of the Acoustical Society of America* 122(3). 1735–1745. https://doi.org/10.1121/1.2756793.
Chitoran, Ioana, Louis Goldstein & Dani Byrd. 2008. Gestural overlap and recoverability: Articulatory evidence from Georgian. In *Gestural overlap and recoverability: Articulatory*, 419–448. De Gruyter Mouton. https://doi.org/10.1515/9783110197105.2.419.
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–EL177. https://doi.org/10.1121/1.5052025.
Clayards, Meghan Alison. 2008. *The ideal listener: making optimal use of acoustic-phonetic cues for word recognition*. University of Rochester PhD dissertation.
Coetzee, Andries W., Patrice Speeter Beddor, Kerby Shedden, Will Styler & Daan Wissing. 2018. Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis. *Journal of Phonetics* 66. 185–216. https://doi.org/10.1016/j.wocn.2017.09.009.
Coetzee, Andries W., Patrice Speeter Beddor, Will Styler, Stephen Tobin, Ian Bekker & Daan Wissing. 2022. Producing and perceiving socially structured coarticulation: Coarticulatory nasalization in Afrikaans. *Laboratory Phonology*. Open Library of Humanities 13(1). https://doi.org/10.16995/labphon.6450.
Cohn, Abigail C. 1990. *Phonetic and phonological rules of nasalization*. University of California at Los Angeles PhD dissertation.
Cohn, Abigail C. 1993. Nasalisation in English: phonology or phonetics. *Phonology* 10(1). 43–81.
Cole, Jennifer, Gary Linebaugh, Cheyenne Munson & Bob McMurray. 2010. Unmasking the acoustic effects of vowel-to-vowel coarticulation: A statistical modeling approach. *Journal of phonetics* 38(2). 167–184. https://doi.org/10.1016/j.wocn.2009.08.004.
Connine, Cynthia M. & Laura M. Darnieder. 2009. Perceptual learning of co-articulation in speech. *Journal of Memory and Language* 61(3). 412–422. https://doi.org/10.1016/j.jml.2009.07.003.
Cox, Felicity & Sallyanne Palethorpe. 2014. Phonologisation of vowel duration and nasalised /æ/ in Australian English. In Jennifer Hay & Emma Parnell (eds.), *Proceedings of the 15th Australasian Speech Science & Technology Conference*, 33–36. Christchurch.
Cronenberg, Johanna, Michele Gubian, Jonathan Harrington & Hanna Ruch. 2020. A dynamic model of the change from pre- to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 83. 101016. https://doi.org/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 & Alain Soquet. 2008. The aerodynamics of nasalization in French. *Journal of Phonetics* 36(4). 578–606. https://doi.org/10.1016/j.wocn.2008.02.002.

Delvaux, Véronique, Thierry Metens & Alain Soquet. 2002. French nasal vowels: acoustic and articulatory properties. In *7th International Conference on Spoken Language Processing (ICSLP 2002)*, 53–56. ISCA. https://doi.org/10.21437/ICSLP.2002-51.

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. https://doi.org/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*. Hong Kong.

Diehl, Randy L. & Keith R. Kluender. 1989. On the objects of speech perception. *Ecological Psychology* 1(2). 121–144. https://doi.org/10.1207/s15326969eco0102_2.

Dmitrieva, Olga, Fernando Llanos, Amanda A. Shultz & 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. 77–95. https://doi.org/10.1016/j.wocn.2014.12.005.

Dresher, B. Elan. 2009. *The Contrastive Hierarchy in Phonology* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511642005.

Fant, Gunnar. 1960. *Acoustic theory of speech production*. The Hague: Mouton & Co.

Flemming, Edward. 2003. The relationship between coronal place and vowel backness. *Phonology*. Cambridge University Press 20(3). 335–373.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, 232–276. Cambridge: Cambridge University Press. http://ebooks.cambridge.org/ref/id/CBO9780511486401A016. (30 December, 2014).

Flemming, Edward S. 2001. *Auditory representations in phonology*. Routledge.

Flynn, Darin & Sean Fulop. 2014. Acoustic phonetic features in Athabaskan sound change. *Diachronica*. John Benjamins 31(2). 192–222. https://doi.org/10.1075/dia.31.2.02fly.

Fowler, Carol A. & Julie M. Brown. 2000. Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Perception & Psychophysics* 62(1). 21–32. https://doi.org/10.3758/BF03212058.

Fowler, Carol A. & Elliot Saltzman. 1993. Coordination and coarticulation in speech production. *Language and Speech* 36(2–3). 171–195. https://doi.org/10.1177/002383099303600304.

Fowler, Carol A. & 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, US: Lawrence Erlbaum Associates, Inc.

Fowler, Carol A. & Jaqueline M. Thompson. 2010. Listeners' perception of compensatory shortening. *Attention, Perception, & Psychophysics*. US: Psychonomic Society 72. 481–491. https://doi.org/10.3758/APP.72.2.481.

Francis, Alexander L. & 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. https://doi.org/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. https://doi.org/10.1215/00031283-2008-030.

Fujimura, Osamu. 1962. Analysis of nasal consonants. *The Journal of the Acoustical Society of America* 34(12). 1865–1875. https://doi.org/10.1121/1.1909142.

Gao, Jiayin & James Kirby. 2023. Perceptual adaptation to altered cue informativeness: distributional, auditory, and lexical factors. In *Proceedings of the 20th International Congress of Phonetic Sciences*. Prague.

Garrett, Andrew. 2015. Sound change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 227–248. London and New York: Routledge.
Garrett, Andrew & Keith Johnson. 2013. Phonetic bias in sound change. In Alan Yu (ed.), *Origins of sound change: Approaches to phonologization*, 51–97. Oxford: Oxford University Press.
Gibson, E J. 1963. Perceptual Learning. *Annual Review of Psychology* 14(1). 29–56. https://doi.org/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*. [Cambridge University Press, 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. https://doi.org/10.1016/j.wocn.2009.01.002.
Grosvald, Michael & David Corina. 2012. Perception of long-distance coarticulation: An event-related potential and behavioral study. *Applied Psycholinguistics*. United Kingdom: Cambridge University Press 33. 55–82. https://doi.org/10.1017/S0142716411000105.
Guion, Susan Guignard. 1998. The Role of Perception in the Sound Change of Velar Palatalization. *Phonetica* 55(1–2). 18–52. https://doi.org/10.1159/000028423.
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*. University of Toronto PhD dissertation.
Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: dispersedness without dispersion. *Phonology* 28(1). 1–54. https://doi.org/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. https://doi.org/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 Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory Phonology 9*, 125–143. Berlin: Mouton de Gruyter.
Harrington, Jonathan, Janet Fletcher & Corinne Roberts. 1995. Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics* 23(3). 305–322. https://doi.org/10.1016/S0095-4470(95)80163-4.
Harrington, Jonathan, Felicitas Kleber & 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. https://doi.org/10.1121/1.2897042.
Harrington, Jonathan, Felicitas Kleber, Ulrich Reubold, Florian Schiel & Mary Stevens. 2018. Linking cognitive and social aspects of sound change using agent-based modeling. *Topics in Cognitive Science* 10(4). 707–728. https://doi.org/10.1111/tops.12329.

Harrington, Jonathan, Felicitas Kleber, Ulrich Reubold, Florian Schiel & Mary Stevens. 2019. The phonetic basis of the origin and spread of sound change. In William F. Katz & Peter F. Assmann (eds.), *The Routledge Handbook of Phonetics*, 401–426. 1st edn. Abingdon, Oxon ; New York, NY : Routledge, 2019. | Series: Routledge handbooks in linguistics: Routledge. https://doi.org/10.4324/9780429056253-15.
Harrington, Jonathan, Felicitas Kleber & Mary Stevens. 2016. The relationship between the (mis)-parsing of coarticulation in perception and sound change: evidence from dissimilation and language acquisition. In Anna Esposito, Marcos Faundez-Zanuy, Antonietta M. Esposito, Gennaro Cordasco, Thomas Drugman, Jordi Solé-Casals & Francesco Carlo Morabito (eds.), *Recent Advances in Nonlinear Speech Processing* (Smart Innovations,Systems and Technologies 48), 15–34.
Harris, Katherine Safford. 1958. Cues for the Discrimination of American English Fricatives in Spoken Syllables. *Language and Speech*. SAGE Publications Ltd 1(1). 1–7. https://doi.org/10.1177/002383095800100101.
Hawkins, Sarah & Jonathan Midgley. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*. Cambridge University Press 35(2). 183–199.
Heinz, John M. & Kenneth N. Stevens. 1961. On the properties of voiceless fricative consonants. *The Journal of the Acoustical Society of America*. Acoustical Society of America 33(5). 589–596. https://doi.org/10.1121/1.1908734.
Herrero de Haro, Alfredo & John Hajek. 2022. Eastern Andalusian Spanish. *Journal of the International Phonetic Association*. Cambridge University Press 52(1). 135–156. https://doi.org/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 & William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55(1). 37–58. https://doi.org/10.2307/412518.
Hoole, Philip & Marianne Pouplier. 2017. Öhman returns: New horizons in the collection and analysis of imaging data in speech production research. *Computer Speech & Language* 45. 253–277. https://doi.org/10.1016/j.csl.2017.03.002.
House, Arthur S. & 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. & Kenneth N. Stevens. 1956. Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders*. American Speech-Language-Hearing Association 21(2). 218–232. https://doi.org/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*. Houston: Rice University PhD dissertation.
Hualde, José Ignacio & Pilar Prieto. 2014. Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica* 71(2). 109–127. https://doi.org/10.1159/000368197.
Hualde, José Ignacio, Miquel Simonet & Marianna Nadeu. 2011. Consonant lenition and phonological recategorization. *Laboratory Phonology* 2(2). https://doi.org/10.1515/labphon.2011.011.
Huffman, Franklin E. 1967. *An outline of Cambodian grammar*. Cornell University PhD dissertation.
Hughes, George W. & Morris Halle. 1956. Spectral properties of fricative consonants. *The Journal of the Acoustical Society of America* 28(2). 303–310. https://doi.org/10.1121/1.1908271.

Hyman, Larry M. 1976. Phonologization. In Alphonse Julliand, Andrew M. Devine & Laurence D. Stephens (eds.), *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*, 470–418. Saratoga,CA: Anma Libri.

Hyman, Larry M. 2013. Enlarging the scope of phonologization. (Ed.) Alan C. L. Yu. *Origins of sound change: Approaches to phonologization* 3–28.

Idemaru, Kaori & Lori L. Holt. 2011. Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance*. US: American Psychological Association 37(6). 1939–1956. https://doi.org/10.1037/a0025641.

Idemaru, Kaori & Lori L. Holt. 2012. Examining talker and phoneme generalization of dimension-based statistical learning in speech perception. In *Proc. SST 2012*, 165–168. Sydney.

Idemaru, Kaori & 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. https://doi.org/10.1037/a0035269.

Idemaru, Kaori & Lori L. Holt. 2020. Generalization of dimension-based statistical learning. *Attention, Perception, & Psychophysics* 82(4). 1744–1762. https://doi.org/10.3758/s13414-019-01956-5.

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

Iverson, Gregory K. & Joseph C. Salmons. 2003. Laryngeal enhancement in early Germanic. *Phonology*. Cambridge University Press 20(1). 43–74.

Jatteau, Adèle & Michaela Hejná. 2016. Dissimilation can be gradient: evidence from Aberystwyth English. *Papers in Historical Phonology* 1. 359–386. https://doi.org/10.2218/pihph.1.2016.1737.

Johnson, Sarah E. 2019. *Spontaneous nasalization: an articulatory investigation of glottal consonants in Thai*. University of Illinois at Urbana-Champaign PhD dissertation.

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. https://doi.org/10.1080/00437956.2002.11432521.

Jong, Kenneth de, Mary E. Beckman & Jan Edwards. 1993. The interplay between prosodic structure and coarticulation. *Language and Speech* 36(2–3). 197–212. https://doi.org/10.1177/002383099303600305.

Jongman, Allard, Ratree Wayland & 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 Loránd Benkö & Samu Imre (eds.), *The Hungarian Language*, 49–84. De Gruyter. https://doi.org/10.1515/9783110880236-003.

Kataoka, Reiko. 2011. *Phonetic and cognitive bases of sound change*. University of California, Berkeley PhD dissertation.

Kavitskaya, Darya. 2001. *Compensatory lengthening: phonetics, phonology, diachrony*. University of California, Berkeley PhD dissertation.

Kawasaki, Haruko. 1986. Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. In John J. Ohala & Jeri J. Jaeger (eds.), *Experimental phonology*, 239–252. Orlando, FL: Academic Press.

Kent, Raymond D., Patrick J. Carney & Larry R. Severeid. 1974. Velar movement and timing: evaluation of a model for binary control. *Journal of Speech and Hearing Research*. American Speech-Language-Hearing Association 17(3). 470–488. https://doi.org/10.1044/jshr.1703.470.

Keyser, Samuel Jay & Kenneth N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82(1). 33–63. https://doi.org/10.1353/lan.2006.0051.
Kim, Donghyun & 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. https://doi.org/10.1080/23273798.2019.1582787.
Kim, Donghyun, Meghan Clayards & Eun Jong Kong. 2020. Individual differences in perceptual adaptation to unfamiliar phonetic categories. *Journal of Phonetics* 81. 100984. https://doi.org/10.1016/j.wocn.2020.100984.
Kingston, John. 2011. Tonogenesis. In Marc van Oostendorp, Colin J. Ewan, Elisabeth V. Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, vol. 4, 2304--2333. Hoboken: Wiley-Blackwell.
Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70(3). 419–454. https://doi.org/10.2307/416481.
Kingston, John, Randy L. Diehl, Cecilia J. Kirk & Wendy A. Castleman. 2008. On the internal perceptual structure of distinctive features: The [voice] contrast. *Journal of Phonetics* 36(1). 28–54. https://doi.org/10.1016/j.wocn.2007.02.001.
Kiparsky, Paul. 1995. The phonological basis of sound change. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 640–670. Oxford: Blackwell.
Kiparsky, Paul. 2015. Phonologization. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*, 563–582. Oxford: Oxford University Press.
Kirby, James. 2010. *Cue selection and category restructuring in sound change*. Chicago: University of Chicago PhD dissertation.
Kirby, James. 2013. The role of probabilistic enhancement in phonologization. In Alan C. L. Yu (ed.), *Origins of sound patterns: approaches to phonologization*, 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. https://doi.org/10.1016/j.wocn.2014.02.001.
Kirby, James. 2014b. Incipient tonogenesis in Phnom Penh Khmer: Computational studies. *Laboratory Phonology* 5(1). 195–230. https://doi.org/10.1515/lp-2014-0008.
Kirby, James & 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 & Marc Brunelle. 2023. Transphonologization of onset voicing: revisiting Northern and Eastern Kmhmu'. *Phonetica* 79(6). 591–629. https://doi.org/10.1515/phon-2022-0029.
Klatt, Dennis H. & 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 & Christine R. Mooshammer. 2013. Toward Improved Spectral Measures of /s/: Results From Adolescents. *Journal of Speech, Language, and Hearing Research* 56(4). 1175–1189. https://doi.org/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 & Arthur G. Samuel. 2008. Accommodating variation: Dialects, idiolects, and speech processing. *Cognition* 107(1). 54–81. https://doi.org/10.1016/j.cognition.2007.07.013.

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. https://doi.org/10.5334/labphon.83.
Kwon, Harim. 2021. A non-contrastive cue in spontaneous imitation: Comparing mono- and bilingual imitators. *Journal of Phonetics* 88. 101083. https://doi.org/10.1016/j.wocn.2021.101083.
Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: phonetics, phonology, and sound change: a multimedia reference tool*. Berlin ; New York: Mouton de Gruyter.
Lawrence, Wayne P. 2000. /str → /ʃtr/: Assimilation at a distance? *American Speech* 75(1). 82–87. https://doi.org/10.1215/00031283-75-1-82.
Lehet, Matthew & Lori L. Holt. 2017. Dimension-based statistical learning affects both speech perception and production. *Cognitive Science* 41. 885–912. https://doi.org/10.1111/cogs.12413.
Lehet, Matthew & Lori L. Holt. 2020. Nevertheless, it persists: Dimension-based statistical learning and normalization of speech impact different levels of perceptual processing. *Cognition* 202. 104328. https://doi.org/10.1016/j.cognition.2020.104328.
Lehiste, Ilse & 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. https://doi.org/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 Alan C. L. Yu (ed.), *Origins of Sound Change*, 98–111. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573745.003.0004.
Liljencrants, Johan & Björn Lindblom. 1972. Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast. *Language* 48(4). 839–862.
Lin, Susan, Patrice Speeter Beddor & Andries W. Coetzee. 2014. Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/. *Laboratory Phonology*. De Gruyter Mouton 5(1). 9–36. https://doi.org/10.1515/lp-2014-0002.
Lindblom, Björn. 1990. Explaining phonetic variation: a sketch of the H&H theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech Production and Speech Modelling*, 403–439. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-2037-8_16.
Lindblom, Björn, Susan Guion, Susan Hura, Seung-Jae Moon & Raquel Willerman. 1995. Is sound change adaptive? *Rivista di Linguistica* 7(1). 5–37.
Lively, Scott E., John S. Logan & 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 & 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*. University of Massachusetts, Amherst PhD dissertation.
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. De Gruyter Mouton. https://doi.org/10.1515/9783110366310-004.

Mackay, Carolyn J. 1995. *A Veneto lexicon: the dialect of Segusino and Chipilo*. Treviso: Grafiche Antiga.
Mann, Virginia A. & Bruno H. Repp. 1980. Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception & Psychophysics* 28(3). 213–228. https://doi.org/10.3758/BF03204377.
Martin, James G. & H. Timothy Bunnell. 1981. Perception of anticipatory coarticulation effects. *The Journal of the Acoustical Society of America* 69(2). 559–567. https://doi.org/10.1121/1.385484.
Matisoff, James A. 1975. Rhinoglottophilia: the mysterious connection between nasality and glottality. In Charles A. Ferguson, Larry M. Hyman & John J. Ohala (eds.), *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, 265–287. Stanford: Language Universals Project.
McGuire, Grant & Molly Babel. 2012. A cross-modal account for synchronic and diachronic patterns of /f/ and /θ/ in English. *Laboratory Phonology*. De Gruyter Mouton 3(2). 251–272. https://doi.org/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. & 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. https://doi.org/10.1121/1.1907526.
Moll, K. L. & 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. https://doi.org/10.1121/1.1912683.
Moreton, E. & Erik Thomas. 2007. Origins of Canadian raising in voiceless-coda effects: A case study in phonologization. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory Phonology 9*, 37–64. Berlin: Mouton de Gruyter.
Moreton, Elliott. 2021. Phonological Abstractness In English Diphthong Raising. *The Publication of the American Dialect Society* 106(1). 13–44. https://doi.org/10.1215/00031283-9551267.
Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39(2). 132–142. https://doi.org/10.1016/j.wocn.2010.12.007.
Norris, Dennis, James M. McQueen & Anne Cutler. 2003. Perceptual learning in speech. *Cognitive Psychology* 47(2). 204–238. https://doi.org/10.1016/S0010-0285(03)00006-9.
Noss, Richard B. 1966. The treatment of */r/ in two modern Khmer dialects. In Norman H. Zide (ed.), *Studies in comparative Austroasiatic linguistics*, 89–95. London: Mouton & Co.
Ohala, John J. 1975. Phonetic explanations for nasal sound patterns. In C. A. Ferguson, Larry M. Hyman & John J. Ohala (eds.), *Nasalfest: Papers from a symposium on nasals and nasalization*, 289–316. Stanford: Language Universals Project.
Ohala, John J. 1981. The listener as a source of sound change. In Carrie S. Masek, Roberta A. Hendrick & Mary Frances Miller (eds.), *Proceedings from the 17th Annual Meeting of the Chicago Linguistic Society, Volume 2: Papers from the Parasession on Language and Behavior*, vol. 2, 178–203. Chicago: Chicago Linguistic Society.
Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.
Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In Leiv E. Breivik & Ernst H. Jahr (eds.), *Language change: contributions to the study of its causes* (Trends in Linguistics. Studies and Monographs [TiLSM] 43), 173–198. De Gruyter Mouton. https://doi.org/10.1515/9783110853063.173.

Ohala, John J. 1992. What's cognitive, what's not, in sound change. In Günter Kellermann & Michael D. Morrissey (eds.), *Diachrony within synchrony: language history and cognition* (Duisburger Arbeiten Zur Sprach- Und Kulturwissenschaft 14), 309–355. Frankfurt: Peter Lang.

Ohala, John J. 1993a. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 237–278. London: Longman.

Ohala, John J. 1993b. Coarticulation and phonology. *Language and Speech* 36(2–3). 155–170. https://doi.org/10.1177/002383099303600303.

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. ISCA. https://doi.org/10.21437/ICSLP.1994-113.

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

Ohala, John J. 2005. Phonetic explanations for sound patterns: implications for grammars of competence. In William J. Hardcastle & Janet M. Beck (eds.), *A figure of speech: a festschrift for John Laver*, 23–38. London: Erlbaum.

Ohala, John J. 2011. Accommodation to the aerodynamic voicing constraint and its phonological relevance. In *Proceedings of The 16th International Congress of Phonetic Sciences*, 64–67. Hong Kong.

Ohala, John J. 2012. The listener as a source of sound change: An update. In Maria-Josep Solé & Daniel Recasens (eds.), *The Initiation of Sound Change: Perception, production, and social factors* (Current Issues in Linguistic Theory 323), 21–36. Amsterdam: John Benjamins Publishing Company.

Ohala, John J. & Maria Grazia Busà. 1995. Nasal loss before voiceless fricatives: A perceptually-based sound change. *Rivista di Linguistica* 7(1). 125–144.

Ohala, John J. & Manjari Ohala. 1993. The phonetics of nasal phonology: theorems and data. In Marie K. Huffman & Rena A. Krakow (eds.), *Nasals, Nasalization, and the Velum* (Phonetics and Phonology), vol. 5, 225–249. San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-360380-7.50013-2.

Ohala, John & James Lorentz. 1977. The story of [w]: an exercise in the phonetic explanation for sound patterns. *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society (1977)* 577–599. https://doi.org/10.3765/bls.v3i0.2264.

Ohala, Manjari & John J. Ohala. 1991. Nasal epenthesis in Hindi. *Phonetica*. De Gruyter Mouton 48(2–4). 207–220. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1121/1.2178720.

Parrell, Benjamin. 2012. The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of phonetics* 40(1). 37–45. https://doi.org/10.1016/j.wocn.2011.08.004.

Pater, Joe. 1999. Austronesian Nasal Substitution and other NC Effects. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody-morphology interface*,

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*. Linguistic Society of America 25(3). 223–240. https://doi.org/10.2307/410084.
Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Bybee & Paul J. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure* (Typological Studies in Language 45), 137–158. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.45.08pie.
Pisoni, David B. & Paul A. Luce. 1987. Acoustic-phonetic representations in word recognition. *Cognition* 25(1–2). 21–52. https://doi.org/10.1016/0010-0277(87)90003-5.
Plauché, Madelaine, Cristina Delogu & John J. Ohala. 1997. Asymmetries in consonant confusion. In *EUROSPEECH-1997*, 2187–2190. Rhodes, Greece. https://doi.org/10.1121/1.417051.
Pouplier, M. & 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. https://doi.org/10.1111/lnc3.12102.
Ratliff, Martha. 2015. Tonoexodus, tonogenesis, and tone change. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 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. https://doi.org/10.17469/O2101AISV000002.
Recasens, Daniel. 2018. Coarticulation. *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.416.
Recasens, Daniel. 2019. Typology of mixing articulatory gestures in phonetics and phonology. *Loquens* 6(1). e057–e057. https://doi.org/10.3989/loquens.2019.057.
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.
Romero, Joaquín. 1996. Articulatory blending of lingual gestures. *Journal of Phonetics* 24(1). 99–111. https://doi.org/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*. Cambridge University Press 6(1). 40–55. https://doi.org/10.1017/jlg.2018.4.
Ruch, Hanna & Jonathan Harrington. 2014. Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 45. 12–25. https://doi.org/10.1016/j.wocn.2014.02.009.
Ruch, Hanna & Sandra Peters. 2016. On the origin of post-aspirated stops: production and perception of /s/ + voiceless stop sequences in Andalusian Spanish. *Laboratory Phonology*. Open Library of Humanities 7(1). 1–36. https://doi.org/10.5334/labphon.2.
Ruhlen, Merrit. 1978. Nasal vowels. In Joseph H. Greenberg (ed.), *Universals of human language: phonology*, 203–241. Stanford: 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*. Cambridge University Press 41(1). 27–40.

Sampson, Rodney. 1999. *Nasal vowel evolution in Romance*. Oxford, New York: Oxford University Press.

Samuel, Arthur G. & Tanya Kraljic. 2009. Perceptual learning for speech. *Attention, Perception, & Psychophysics* 71(6). 1207–1218. https://doi.org/10.3758/APP.71.6.1207.

Savoia, Leonardo & Martin Maiden. 1997. Metaphony. In Martin Maiden & Mair Parry (eds.), *The Dialects of Italy*, 15–25. London: Routledge.

Schertz, Jessamyn, Taehong Cho, Andrew Lotto & Natasha Warner. 2015. Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics* 52. 183–204. https://doi.org/10.1016/j.wocn.2015.07.003.

Schertz, Jessamyn & Emily J. Clare. 2020. Phonetic cue weighting in perception and production. *WIREs Cognitive Science* 11(2). e1521. https://doi.org/10.1002/wcs.1521.

Schleef, Erik & Michael Ramsammy. 2013. Labiodental fronting of /θ/ in London and Edinburgh: a cross-dialectal study. *English Language and Linguistics* 17(1). 25–54. https://doi.org/10.1017/S1360674312000317.

Shapiro, Michael. 1995. A case of distant assimilation: /str/ → /ʃtr/. *American Speech*. [Duke University Press, American Dialect Society] 70(1). 101–107. https://doi.org/10.2307/455876.

Silverman, Daniel. 2006. The diachrony of labiality in Trique, and the functional relevance of gradience and variation. In Louis Goldstein, Douglas H. Whalen & Catherine T. Best (eds.), *Laboratory phonology 8*, 133–154. De Gruyter Mouton.

Solé, Maria-Josep. 1992. Phonetic and phonological processes: the case of nasalization. *Language and Speech* 35(1–2). 29–43. https://doi.org/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. https://doi.org/10.1177/002383099503800101.

Solé, Maria-Josep. 2007. Controlled and mechanical properties in speech: a review of the literature. In Maria-Josep Solé, Patrice Speeter Beddor & Manjari Ohala (eds.), *Experimental approaches to phonology*, 302–321. Oxford: Oxford University Press.

Sonderegger, Morgan A. 2004. *Subglottal coupling and vowel space: an investigation in quantal theory*. Cambridge, MA: Massachusetts Institute of Technology BA thesis.

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 & Sheila Blumstein. 1981. The search for invariant acoustic correlates of phonetic features. In Peter D. Eimas & Joanne L. Miller (eds.), *Perspectives on the study of speech*, 1–38. Hillsdale, NJ: Erlbaum.

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. https://doi.org/10.1121/1.1912751.

Stevens, Kenneth N. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In Edward E. David, Jr. & Peter B. Denes (eds.), *Human Communication: A Unified View*, 51–56. New York: McGraw-Hill.

Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17. 3–45.

Stevens, Kenneth N. & Samuel Jay Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65(1). 81–106. https://doi.org/10.2307/414843.

Stevens, Kenneth Noble & Samuel Jay Keyser. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38(1). 10–19. https://doi.org/10.1016/j.wocn.2008.10.004.

Stevens, Mary & Jonathan Harrington. 2016. The phonetic origins of /s/-retraction: Acoustic and perceptual evidence from Australian English. *Journal of Phonetics* 58. 118–134. https://doi.org/10.1016/j.wocn.2016.08.003.

Stuart-Smith, J. & C. Timmins. 2006. "Tell her to shut her moof": the role of the lexicon in TH-fronting in Glaswegian. In C. J. Kay, G. D. Caie, C. Hough & I. Wotherspoon (eds.), 171–183. Amsterdam, Netherlands: Rodopi.

Svantesson, Jan-Olof & David House. 2006. Tone production, tone perception and Kammu tonogenesis. *Phonology* 23(02). 309–333. https://doi.org/10.1017/S0952675706000923.

Sweet, Herry. 1874. The history of English sounds. *Transactions of the Philological Society* 15(1). 461–623. https://doi.org/10.1111/j.1467-968X.1874.tb00881.x.

Tabain, Marija. 1998. Non-Sibilant Fricatives in English: Spectral Information above 10 kHz. *Phonetica*. De Gruyter Mouton 55(3). 107–130. https://doi.org/10.1159/000028427.

Todd, Simon, Janet B. Pierrehumbert & Jennifer Hay. 2019. Word frequency effects in sound change as a consequence of perceptual asymmetries: An exemplar-based model. *Cognition* 185. 1–20. https://doi.org/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*. Cambridge University Press 42(1). 49–63.

Torreira, Francisco & Mirjam Ernestus. 2011. Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology* 2(2). 331–353. https://doi.org/10.1515/labphon.2011.012.

Trouvain, Jürgen. 2021. Aerodynamic voicing constraint in stops: From Wolfgang von Kempelen to John Ohala. In *Fourth International Workshop on the History of Speech Communication Research (HSCR 2021)*, 75–81. ISCA. https://doi.org/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*. University of Wisconsin Press 30(3/4). 177–181.

Warren, Paul. 2006. /s/-retraction, /t/-deletion and regional variation in New Zealand English /str/ and /stj/ clusters. In Paul Warren & Catherine I. Watson (eds.), *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*, 466–471. Auckland, New Zealand: Australasian Speech Science and Technology Association.

Wayland, Ratree P. & Susan G. Guion. 2005. Sound change following the loss of /r/ in Khmer: a new tonogenetic mechanism? *Mon-Khmer Studies* 35. 55–82.

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

Whalen, Douglas H., Leigh Lisker, Arthur S. Abramson & 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 & 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. https://doi.org/10.1121/1.1912976.

Winn, Matthew B., Monita Chatterjee & 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. https://doi.org/10.1044/1092-4388(2012/12-0086).

Xu, Yi & Anqi Xu. 2021. Consonantal F0 perturbation in American English involves multiple mechanisms. *The Journal of the Acoustical Society of America* 149(4). 2877–2895. https://doi.org/10.1121/10.0004239.

Yang, Meng & Megha Sundara. 2019. Cue-shifting between acoustic cues: Evidence for directional asymmetry. *Journal of Phonetics* 75. 27–42. https://doi.org/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. https://doi.org/10.5334/labphon.97.

Yu, Alan C. L., Carissa Abrego-Collier & Morgan Sonderegger. 2013. Phonetic imitation from an individual-difference perspective: subjective attitude, personality and “autistic” traits. *PLoS ONE* 8(9). e74746. https://doi.org/10.1371/journal.pone.0074746.

Yu, Alan C. L. & 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. https://doi.org/10.1121/1.4883380.

Zellou, Georgia. 2017. Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics* 61. 13–29. https://doi.org/10.1016/j.wocn.2016.12.002.

Zellou, Georgia. 2022. *Coarticulation in phonology*. 1st edn. Cambridge University Press. https://doi.org/10.1017/9781009082488.

Zellou, Georgia, Santiago Barreda & 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–EL276. https://doi.org/10.1121/10.0000951.

Zellou, Georgia, Michelle Cohn & 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. https://doi.org/10.1121/10.0004989.

Zellou, Georgia, Rebecca Scarborough & Kuniko Nielsen. 2016. Phonetic imitation of coarticulatory vowel nasalization. *The Journal of the Acoustical Society of America* 140(5). 3560–3575. https://doi.org/10.1121/1.4966232.