ANNUAL REVIEWS

Annual Review of Linguistics

The Actuation Problem

Alan C.L. Yu

Phonology Laboratory, Department of Linguistics, University of Chicago, Chicago, Illinois, USA; email: aclyu@uchicago.edu

Annu. Rev. Linguist. 2023. 9:215–31

First published as a Review in Advance on October 13, 2022

The *Annual Review of Linguistics* is online at linguistics.annualreviews.org

https://doi.org/10.1146/annurev-linguistics-031120-101336

Keywords

actuation, incrementation, propagation, individual differences, sound change, cognitive processing style, agent-based modeling

Abstract

The actuation problem asks why a linguistic change occurs in a particular language at a particular time and space. Responses to this problem are multifaceted. This review approaches the problem of actuation through the lens of sound change, examining it from both individual and population perspectives. Linguistic changes ultimately actuate in the form of idiolectal differences. An understanding of language change actuation at the idiolectal level requires an understanding of (*a*) how individual speaker-listeners' different past linguistic experiences and physical, perceptual, cognitive, and social makeups affect the way they process and analyze the primary learning data and (*b*) how these factors lead to divergent representations and grammars across speakers-listeners. Population-level incrementation and propagation of linguistic innovation depend not only on the nature of contact between speakers with unique idiolects but also on individuals who have the wherewithal to take advantage of the linguistic innovations they encountered to achieve particular ideological projects at any given moment. Because of the vast number of contingencies that need to be aligned properly, the incrementation and propagation of linguistic innovation are predicted to be rare. Agent-based modeling promises to provide a controlled way to investigate the stochastic nature of language change propagation, but a comprehensive model of linguistic change actuation at the individual level remains elusive.

1. INTRODUCTION

"Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?" (Weinreich et al. 1968, p. 102). This question, which was first raised by Weinreich and colleagues in 1968, marked the earliest formulation of the actuation problem. Indeed, it was the first time "actuation" was presented as a "problem," which Weinreich et al. (1968) argued can be approached when four other problems are answered. That is, the answer to the actuation problem is contingent on understanding (*a*) the set of possible changes and the conditions for such changes (i.e., the so-called constraint problem); (*b*) the intervening stages that can be observed or posited between any two forms of a language defined for a language community at different times (i.e., the transition problem); (*c*) how a change is embedded in the linguistic and cultural/social contexts (i.e., the embedding problem); and (*d*) how a change is evaluated by speakers in terms of its effects upon linguistic structure and the communications between members of the community (i.e., the evaluation problem). As these problems suggest, the question of actuation is multifaceted, and it is not surprising that there are many different responses to it (see also Stevens & Harrington 2014). Some conceptualize the actuation problem as a two-part question, preferring to address the question of how a linguistic innovation comes into existence (i.e., the initiation of a change) separately from the question of how an innovation spreads (i.e., the propagation of a change). Often such discussions privilege community-level changes, relegating the question of the origins of innovative variants to matters of channel biases (i.e., biases attributable to universal factors shared by all speakers) or accidents. Indeed, for Weinreich et al. (1968, p. 186), the actuation of a linguistic change is the spreading of an innovative variant "through a specific subgroup of the speech community." Others prefer to conceptualize actuation not as a single event but as a process. Pinget (2015, p. 4), for example, sees actuation as "an iterative process at the individual level in which minimal changes incrementally accumulate in a speaker's system every time he speaks to a listener"; sound change is thus redefined as "a purely synchronic process, residing in the variability inherent to production and perception." In this sense, the propagation of a change, which can be discerned when comparing across age groups or diachronic time slices, is "merely a series of incremental acts of actuation in a defined direction. Actuation exists at every step. Change and propagation are the additive effects of such small steps." (Pinget 2015, p. 4). Echoing this sentiment, Harrington et al. (2018, p. 708) argue that the conditions that give rise to linguistic change and those that are responsible for its spread throughout the community might be more "artificial" than real.

In this article, I approach the problem of actuation through the lens of sound change, even though the actuation problem is decidedly not limited to the phonological domain (see, e.g., Stevens & Harrington 2014, Walkden 2017). I also address the actuation problem as it was posed by Weinreich et al. (1968), and not necessarily in the way they prefer to conceptualize it. How the actuation problem is addressed hinges on the way the nature of language change is conceptualized and on the emphasis individual researchers place on different aspects of language change. Broadly speaking, discourses related to the actuation problem can be dichotomized into theories that emphasize the contribution of individual-level factors and theories that focus on the societal and cultural forces at work. The rest of the article is broadly structured with this dichotomy in mind.

2. ACTUATION AT THE LEVEL OF THE INDIVIDUAL

The question of what constitutes a language is notoriously difficult to answer (Haugen 1966). Thus many theorists today settle on the view that language is best defined at the individual level, that is, in terms of the grammatical knowledge a speaker has. From this idiolect-centric view of language, language change stems from the fact that grammar transmission is never perfect. That is, no two individuals are likely to ever construct exactly the same grammatical system, because the type of linguistic inputs each person is exposed to and the order in which the linguistic inputs are encountered are never fully identical. Thus the ultimate shape of the constructed grammar and lexicon would differ across individuals even within the same speech community. From this point of view, then, the problem can be addressed by understanding what type of constraints—linguistic, cognitive, social, and otherwise—govern the way a grammar is constructed, the nature of the so-called primary learning data, and the order in which the learning data are presented in order to construct a grammar and its associated lexicon.

2.1. Idiolectal Models of Sound Change

All theories of language change accept the central role synchronic variations play in change at the community level. Theories, however, differ in how they conceptualize the nature of synchronic variations and their relationship with community-level changes. Building on his hyper- and hypoarticulation (H&H) theory of speaker-listener interaction, Lindblom and colleagues (1995), for example, argue that linguistic innovations emerge out of the functional-communicative nature of linguistic communication. For example, they see coarticulation in speech as reflecting the more economical realization the motor system defaults to when the communicative success of a situation does not hinge on the clarity of articulation. New phonetic variants can accumulate during the "how" mode of listening, which "puts the speakers-listeners in a state of readiness for phonetics and phonological innovations" (Lindblom et al. 1995, p. 13). The speakers-listeners are free to select from the pool of variants, a selection process that is governed by a host of different evaluative metrics involving articulatory, perceptual, structural, and social factors. Consider, for example, the emergence of vocalic nasalization. The historical change from VN to Ṽ is often assumed to emerge out of anticipatory vowel nasalization, that is, when the lowered velum gesture becomes a required configuration for the vowel rather than being coarticulatorily linked to the consonantal closure. The seeds of this change might stem from the variable realization of VN, particularly when VN is followed by a voiceless consonant such as a voiceless stop or fricative. Aerodynamically, the high intraoral air pressure needed to sustain a turbulent airstream for either a fricative or voiceless stop is incompatible with the low intraoral pressure when the velum is lowered. From an acoustic perspective, the presence of a nasal murmur is incompatible with the lack of energy during the closure phase of a voiceless stop. The need to resolve the so-called nasal repulsion (Carignan et al. 2021) might lead speakers to adjust the lowered velum gesture differently to preserve the nasal gesture. According to Carignan et al. (2021), successful repulsion of the nasal can be achieved (*a*) through an earlier phasing of the velum closing gesture with respect to the velum opening gesture that precedes it, (*b*) by shortening the onset phase of the velum lowering gesture (truncation), and/or (*c*) by rescaling the degree of velum lowering. The authors found evidence in German of gestural rescaling and temporal rephasing of the velum gesture associated with a nasal consonant preceding a voiceless obstruent relative to a nasal consonant preceding a voiced obstruent. Presumably due to the variable realization of VN sequences, some listeners might encounter difficulties with interpreting the production outputs of hyper- or hypoarticulating speakers. Ohala, for example, suggests that the VN > Ṽ could come about as a result of a listener's "hypocorrecting," that is, misattributing the nasal feature as part of the vowel if they fail to recognize the presence of the following nasal as the source of nasal coarticulation (Ohala 1981).[1]

[1]Listeners could also "hypercorrect," that is, misattribute an intended gesture as contextually determined. Sound changes canonically attributed to hypercorrection are metathesis and dissimilation. Readers are referred to Blevins (2004) for further discussions on other types of misperception and reanalysis.

According to Lindblom and colleagues (1995, p. 15), "[f]or listeners to gain access to the surface value of a particular pronunciation variant, they do not necessarily have to make a perceptual error." They emphasize that the functional-communicative aspects of speech (e.g., reading a children's story, delivering a lecture, etc.) can give rise to innovative variants in different contexts, and listeners can engage in modes of listening that are not different from what phoneticians do when performing narrow transcription, without normalizing for potential contextual influences. To be sure, even if listeners heard the signal veridically (i.e., without normalization or contextual compensation), they might nonetheless analyze the signal differently, especially when the signal is consistent with multiple phonological analyses. Ultimately, speech perception is not the same as auditory phonetic analysis. For example, the rescaling and temporal rephasing of the lowered velum gesture in VN sequences could serve as the basis for the listeners to analyze the lowered velum gesture prosodically, anchoring the lowered velum gesture relative to the VN complex (or rime) rather than to the onset of tongue movement of the nasal stop. As suggested by Carignan et al. (2021), such a reanalysis could lead to the perceived equivalence scenario documented by Beddor (2009) whereby the duration of the nasal consonant is inversely related to the temporal extent of nasalization on the preceding vowel. Unlike the misperception/hypocorrective scenario discussed above, under the perceived equivalence account, the nasal consonant might still be present, but the variable duration profile of the nasal stop and the dissociation of the lowered velum gesture from the onset of the oral closure phase of nasal stop might lead the stop closure to be reanalyzed as part of the early phasing of the following oral stop. Thus, what was previously a VNT sequence may be analyzed by an innovative listener as ṼT, with no nasal stop present.

Returning to the question of actuation, the different models of sound change laid out thus far have in common the idea that linguistic innovation emerges at the level of the individual. Under both Lindblom's H&H and Ohala's misperception accounts, a linguistic innovation is actuated whenever a novel pronunciation variant is registered by the listener, either via the veridical how mode of listening or as a result of misperception. Both models are consistent with exemplar-based models of speech perception and production in which change is synonymous with memory update. That is, whenever a speaker-listener retains a new variant in their perceptual memories, conceptually a change has taken place. Their views essentially anticipated the notion that actuation is "an iterative process at the individual level in which minimal changes incrementally accumulate in a speaker's system every time he speaks to a listener" alluded to earlier (Pinget 2015, p. 4). Note that since both Lindblom's and Ohala's models see sound change as the accumulation of novel variants, the transition is lexically abrupt in the sense that change is due to the accumulation of word-specific innovative representations or exemplars. Sound change is lexically gradual in the sense that sound change does not target phonological classes per se but rather concerns the introduction of novel representations of individual words, one word at a time. It should be noted that, unlike Lindblom's H&H theory, the misperception scenario of change is restricted to inexperienced listeners or learners, since experienced listeners presumably have other means to reconstruct the presence of the coarticulatory source even if the contextual trigger is not clearly heard in the signal (Ohala 1993); that is, experienced listeners who have a good command of the lexicon and the phonotactics of the language are expected to be proficient at normalizing for context-specific distortions, as evidenced by robust lexical and phonotactic-guided effects in speech perception (e.g., Ganong 1980, Massaro & Cohen 1983).

While both these models assume innovation occurs at the level of individual words, Beddor's (2009) perceptual equivalence pathway, on the other hand, leads to the invention of novel phonological structures, resulting in change actuating at the level of the grammar. That is, to the extent that the perceptual analysis strategy from which the perceptual equivalence effect ultimately derives is part of the phonological grammar (or perceptual grammar, in Beddor's terminology), the innovative phonological analysis is predicted to apply to all relevant structures equally from the perspective of that idiolect. Thus, in this case, the transition is abrupt at the grammatical level as it involves a change in the phonological grammar.

2.2. Articulatory Origins of Idiolectal Variation and Change

The idea that individuals might arrive at different perceptual grammars raises questions about the cause of such differences. As noted earlier, a common assumption is that speakers-listeners might differ as a result of their past linguistic experiences. In his discussion of the life cycle of spirantization, Ramsammy (2018), for example, suggests that a change in status of a gradient phonetic process (i.e., unintended phonetic variation) may come about as a consequence of children's being exposed to adult speech models that include a sufficiently high number of fricated (i.e., changed and phonologized) intervocalic tokens. The linguistic exposure account of change, therefore, hinges on the existence of innovative variants to begin with, which raises a few questions: Where do innovative variants originate? How do linguistic innovations come to be regular and stable enough to affect change during child language acquisition? The traditional appeal to channel biases (i.e., the type of articulatory, acoustic, auditory, and perceptual constraints inherent to the vocal tract, along with the auditory and perceptual apparatus as alluded to above) is, perhaps paradoxically, not sufficient to answer these questions, since such biases are presumably always present and not unique to any particular individual. Phonetic precursors that are shared by all members of a community should not be expected to produce sound change because of their commonness (Baker et al. 2011). As argued by Yu (2021), the idiolectal approach to linguistic change necessitates the identification of innovative individuals and of the motivations underlying their exceptional behaviors.

Individual variability may come from differences in vocal tract physiology (Dediu & Moisik 2019), particularly related to the nature of sexual dimorphism of the vocal tract (Vorperian et al. 2011), vocal tract size and shape (Peterson & Barney 1952), and/or behavioral/etiological factors (Sachs et al. 1973, Ohala 1994). Inherent vocal tract anatomical differences across speakers could result in acoustically similar, but not identical, outputs. Such anatomical differences may lead the listeners to perceive as innovations the direct acoustic effects or the covert changing of neighboring sounds. The North American English /ɹ/ is one such example. The North American English rhotic may vary between being retroflexed, bunched, or some combinations of both tendencies (Baker et al. 2011, Mielke et al. 2016, Smith et al. 2019). Based on the examination of static MRI images of sustained /ɹ/ produced by 80 speakers from diverse ethnic and linguistic backgrounds who were trained to produce the North American /ɹ/, Dediu & Moisik (2019) concluded that anatomical aspects of the anterior vocal tract (e.g., hard palate width and height, overall size of the mouth, or the size/prominence of the alveolar ridge) may influence the articulation of North American English /ɹ/. The inherent articulatory variation of /ɹ/ may also exert covert influence on the articulation of neighboring sounds. In their investigation of /s/-retraction in American English (i.e., /s/ in words like *street* being pronounced more like /ʃ/), Baker et al. (2011) found that, among people who have not yet developed categorical /s/-retraction, the articulatory configuration of the /s/ in an *str-* sequence is predicted by the tongue shape of the following rhotic; all participants in Baker et al.'s (2011) study were bunchers. They hypothesize that the different bunched /ɹ/ variants correspond to different degrees of retraction of preceding /s/ and /z/ and the affrication of /tr/ and /dr/ sequences (e.g., [tʃɹ] and [dʒɹ] in words like *train* and *drain*, respectively; see also Smith et al. 2019). If listeners were exposed to individuals with strong brunched /ɹ/ that resulted in a high degree of /s/-retraction and affrication of /tr/ and /dr/, those listeners may develop /s/-retraction and /t, d/ affrication even if their own rhotic articulations would not have induced the same coarticulatory influence organically.

2.3. Perceptual Sources of Idiolectal Variation and Change

Just as speaker-specific articulation might stem from inherent anatomical differences across individuals or idiosyncratic articulatory habits (Klatt 1986), listeners might also have inherent differences in speech-processing strategies that could lead to variation in perceptual analysis (Yu 2010, 2013, 2021). For example, as noted above, Beddor (2009) found production and perceptual evidence that some American English speakers-listeners exhibit covariation between vocalic nasalization and the duration of the postvocalic nasal stop in VN(C) sequences. Beddor et al. (2018) further established that speakers-listeners who attended more to the presence of nasalization during the vowel interval, as evidenced by their eye movement in a perception task, also produced earlier onset vocalic nasalization in their production. What factors contribute to such individual variability in speech processing remains a largely understudied puzzle. According to Beddor (2009) and Beddor et al. (2018), their findings are reminiscent of individual differences in cue weighting that have been frequently reported in the literature (e.g., Escudero & Boersma 2004, Shultz et al. 2012, Kong & Edwards 2016, Kapnoula et al. 2017, Ou et al. 2021). That is, not only are speech categories defined by multiple acoustic dimensions, but speakers and listeners also show differential weighting to those cue dimensions in production and in perception.

The processing of stop voicing contrasts offers a particularly instructive example of differential cue weighting and its relation to sound change. Voice onset time (VOT) and fundamental frequency (F0) at the onset of the following vowel often covary in languages. In English, for example, phonologically voiceless plosives are followed by raised F0 at the vocalic onset, while phonologically voiced plosives (which are canonically realized with zero to weakly positive VOT) are followed by lowered F0 at onset (e.g., Hombert et al. 1979, Shultz et al. 2012, Dmitrieva et al. 2015). Similar covariation between VOT and F0 has been found in many other languages [e.g., Cantonese (Francis et al. 2006), Spanish (Dmitrieva et al. 2015), French and Italian (Kirby & Ladd 2016), and Khmer, Central Thai, and Northern Vietnamese (Kirby 2018)]. Understanding the nature of the covariation between speech cues is critical from the perspective of this type of sound change, especially since some scholars have posited that certain typologically common sound changes stem from the restructuring (or "transphonologization") of the cue weight relationship between covarying cues (e.g., Hyman 1976, Kirby 2013). For example, tonal distinctions in some languages might have emerged from the type of F0 differences associated with the production of voicing distinctions in consonants, via what might be called the cue reweighting pathway to sound change (Hyman 1976, Kang 2014, Coetzee et al. 2018).

Consider, for example, the case of Kammu, a Mon-Khmer language spoken in northern Laos. It has three main dialects, Eastern, Northern, and Western. As described by Svantesson & House (2006), Eastern Kammu is nontonal, and it retains the original contrasts between voiceless and voiced stops and sonorants. In Northern and Western Kammu, syllables with historically voiceless and voiced initials have developed high and low tones, respectively. F0 perturbations have been argued to be a reflex of aerodynamic (Ladefoged 1967) and/or articulatory (Halle & Stevens 1971, Ohala 1973, Löfqvist et al. 1989) byproducts of stop voicing production. When F0 perturbations come to be actively controlled by speakers, perhaps to enhance a phonological contrast (Kingston & Diehl 1994, Keyser & Stevens 2006, Kingston 2007, Solé 2007, Hanson 2009), F0 perturbation is phonologized in the sense that it has become part of the phonetic knowledge (i.e., part of the grammatical system) that is language-specific and learned (Cohn 1993, Kingston & Diehl 1994).

In the case of English, various studies have shown that the F0 perturbation effect not only is language-specific (Hombert et al. 1979, Dmitrieva et al. 2015) but also varies extensively across individuals (Shultz et al. 2012, Chodroff & Wilson 2018, Clayards 2018b). This means that F0 perturbation in English should not be considered a channel bias effect. Instead, English speakers appear to have already developed idiolectal differences in VOT/F0 covariation. Consistent with this idiolectal view is the fact that the perceptual weightings of the VOT and F0 cues also exhibit a trading relationship that is listener-specific. That is, English listeners who rely more on the VOT cue are found to rely less on the onset F0 cue (Kapnoula 2016, Kapnoula et al. 2017, Ou et al. 2021), but the specific weights they assign for each cue differ between listeners.

While cue weight variability might emerge from differences in individual perceptual experience (e.g., Francis et al. 2008, Lehet & Holt 2017, Zhang & Holt 2018), crucially, it can also stem from individual variation in speech-processing strategies. Similar to the processing variability reported in Beddor et al. (2018), Ou et al. (2021) found, for example, that individuals who integrate secondary cues more extensively during processing are more likely to utilize a buffer processing strategy. That is, rather than activating words by the earlier arriving cue without waiting for the later ones (i.e., a cascade processing strategy), listeners who exhibit greater integration of multiple cues in the signal show a delayed reaction to the early-arriving cue until other relevant cues are available. Individual variability in the weighting of VOT and onset F0 cue has also been linked to individual variability in categorization gradience (Kong & Edwards 2016, Kapnoula et al. 2017, Ou et al. 2021), which, in turn, is linked to individual differences in neural encoding of the speech signal at the subcortical and cortical levels (Ou & Yu 2021). These studies suggest that individual variation in cognitive processing style could influence how the primary learning data are processed and analyzed, which would lead to the actuation of significant idiolectal differences (e.g., Yu 2010, 2013).

To be sure, the existence of between-individual variation alone does not necessarily imply language change. However, when individual differences in cue weighting can be shown to be stable across time (Idemaru et al. 2012, Schertz et al. 2015, Kapnoula 2016)—or can be shown to exhibit speaker control, such as controlled adjustments relative to global duration conditions such as differences in speaking rate or prosodic contexts (Solé 2007)—those speakers-listeners can be said to have phonologized the relationship between covarying cues differently (see also Yu 2021). In the case of the covariation between VOT and onset F0 cue in American English, for example, listeners not only show consistent cue weighting strategies across contrasts (Clayards 2018a, Ou et al. 2021) but also are sensitive to socio-indexical factors, such as gender (both the perceived gender of the talker and the gender of the listener) and perceived talker personality traits, when evaluating the VOT and F0 production of a talker (Yu 2022). This type of socio-indexical and affective linkage with these specific voicing cues further supports the idea that the covariation between VOT and onset F0 in North American English might be on a path to develop tonal distinctions similar to what has been reported recently in Afrikaans. Coetzee et al. (2018) investigated the production and perception of phonologically voiced and voiceless plosives in older and younger female speakers of Afrikaans. While there was significant individual variability, older speakers were more likely than younger speakers to produce prevoicing and to rely on prevoicing perceptually, even though all speakers produced large and systematic F0 differences after phonologically voiced and voiceless plosives and used F0 (especially in the absence of prevoicing) to perceptually differentiate those plosives. Like speakers of Afrikaans, older speakers of Seoul Korean convey the contrast between lenis and aspirated plosives in phrase-initial position primarily by VOT (with longer VOT for aspirated plosives), but this pattern has been replaced for younger speakers by an F0 contrast such that the two plosive categories are no longer robustly differentiated by VOT but by low F0 after historically lenis plosives and by high F0 after historically aspirated plosives (e.g., Oh 2011, Kang 2014, Bang et al. 2018). In particular, females make larger F0 distinctions than male speakers, and, correspondingly, females are ahead of males in the VOT merger process (Oh 2011, Kang 2014).

The covariation in VOT and F0 across English, Afrikaans, and Seoul Korean points to a cline of individual variability in different language communities. All three language communities (or

Language Customs, see further discussion below) exhibit greater-than-expected F0 differences after the relevant phonation contexts, but they vary in terms of the magnitude and temporal extent of the F0 difference. Also variable is the degree of VOT merger; it is most prevalent in Seoul Korean and less extensive in Afrikaans. While there is no indication that VOT is merging in American English generally, the VOT difference between voiced and voiceless stops is smaller in males than in females, and the gender difference might be the result of the amplification of an intrinsic variation due to anatomical differences between males and females (Whiteside & Marshall 2001). Note that this type of gender difference in VOT is not universal across languages (Oh 2011, Lundeborg et al. 2012, Li 2013, Reddy et al. 2013, Peng et al. 2014), further supporting the potential socio-indexical relevance of this gender-based VOT difference in English and disassociating it from any potential physiological motivations. The broad differences between language communities in VOT/F0 covariation and different degrees of transphonologization from a primarily VOT-based contrast to an F0-based contrast suggest that additional forces are at work that constrain idiolectal variation.

Before moving on, it is worth noting that many analysts euphemistically refer to such idiolectal changes as mini sound changes (e.g., Ohala 1993, Lindblom et al. 1995). However, from the idiolectal perspective of language change, such changes follow from the basic modus operandi of language transmission; thus, diminuting such changes risks obfuscating the centrality of these outcomes in language change. Rather than being called "mini," they are also sometimes referred to as "local" (Eckert 2019) or "under-the-counter" (Milroy 2002) changes. In the next section, we look at how local changes come to be supralocal or over the counter.

3. ACTUATION AT THE POPULATION LEVEL

The recognition of sporadic and localized changes within individuals notwithstanding, many historical linguists see the holy grail of explaining language change as lying in an understanding of the population-level implementation of a change. According to Weinreich et al. (1968), there are two ways within an idiolectal view of language change to conceptualize population-level changes, or a change in Language Custom (LC; *Sprachusus* in German) in the terminology of Paul (1880). Change in LC could come about as a result of a change in idiolects that define the LC or "through additions or subtractions of idiolects from the set of idiolects over which a Language Custom is defined" (Weinreich et al. 1968, p. 107). That is, for a given LC that is composed of the set of idiolects {A, B, C, D}, a change of idiolect B at time t to B_i at time $t+1$ would lead to a change in LC to LC_i at $t+1$, consisting of the set {A, B_i, C, D}. LC could also change as result of the addition of an idiolect (e.g., due to the birth of a child or immigration into the community) such that LC_i consists of the set {A, B, C, D, E} or as a result of the reduction of the set (e.g., due to death or migration away from the community) such that LC_i includes only the set {A, C, D}. Changes in idiolects (i.e., from idiolect B to B_i), in turn, can come about in one of two ways: spontaneous change or adoption of features from idiolects of other speakers. As reviewed earlier, spontaneous changes could come about as a result of Lindblom-style how-mode accumulation of novel pronunciations or structures, or via misperception à la Ohala. Spontaneous changes could also come about as a result of grammar updates, if the learner encounters new learning data that prompt them to adopt a novel grammatical analysis of the input data. The adoption of features from idiolects of other speakers, or what Labov (2007) calls transmission (i.e., parent-to-child) and diffusion (i.e., adult-to-adult), assumes that speakers engage in selective adoption of features from other speakers who have idiolectal features that differ in the relevant way.[2] Much work within sociolinguistics, and

[2]Note that, unlike the Labovian notions of transmission and diffusion, even though the type of reanalysis assumed in the grammar update scenario requires linguistic exchanges between speakers with different idiolects,

especially within the literature under the rubric of language variation and change, is dedicated to understanding the mechanisms behind such feature adoption. The question of actuation, from this population-level perspective, becomes a question of how linguistic innovation spreads across a population—specifically, the conditions under which a speaker would take up a feature from idiolects of other speakers.

The simplest model of the spreading of linguistic innovation relies on the notion of density of communication (e.g., Bloomfield 1933, Labov 2001). That is, the likelihood of a change being adopted is proportional to the rate of exposure to the innovative variant as a result of the interaction between individuals. Learners are assumed to engage in probability matching relative to the exposure data (Labov 1994, pp. 580–83). This conception of innovation propagation assumes that language users adopt new variants indiscriminately. Yet, imitation and convergence studies have found that linguistic accommodation and convergence are not only individual-specific (e.g., Wade et al. 2021) but also highly selective, with sensitivity toward both the specificity of the linguistic system (Nielsen 2011) and the subjective evaluations between interlocutors (e.g., Babel 2012, Yu et al. 2013). A less mechanistic take on this idiolectal contact view of change comes from sociolinguistic work on networks (e.g., Milroy 1980, Dodsworth & Benton 2020), which emphasizes the role of local social networks in the adoption of change. Specifically, people within a tighter network with fewer connections with other networks would spread a novel feature within the network faster and would be less likely to be affected by changes coming from outside the network. But even with more sophisticated network models, linguistic innovation propagation is still based on frequency of exposure, and exceptional individual behaviors within the same network are not uncommon (Dodsworth 2019).

Bermúdez-Otero (2020) argues that, in conjunction with density of communication, the incrementation of linguistic innovation spread can be modeled as a result of what he calls community-oriented momentum-sensitive learning. The language learners are community oriented in the sense that they internalize and follow a mental representation of the collective linguistic norm of the speech community and reject individual idiosyncrasies. Learning is momentum sensitive in that the learner's mental representation of the community norm incorporates an age vector encoding differences in variable use between age groups. While he acknowledges the role of idiolectal innovation as an important source of "mutation" within the population, Bermúdez-Otero (2020) argues that the community-oriented learners would reject most innovations as "randomly scattered idiosyncratic deviations from the community norm." Innovative variants would only get adopted if there is an "inverse correlation between the frequency of a mutation and speaker age." In particular, it is assumed that, if the correlation is strong, the learners who adopt an age vector would increase their usage of the new variant until late adolescence, yielding the "adolescence peak" pattern commonly observed in apparent time studies (Labov 2001, chap. 14). This approach, however, is silent with respect to how the "inverse correlation between the frequency of a mutation and speaker age" would come to be in the first place.

To this end, Eckert (2019) emphasizes the importance of the social embedding of linguistic innovations in the incrementation and propagation of change. As she noted, linguistic changes are not autonomous and do not occur in isolation; they are always accompanied by many other variables, linguistic or otherwise, all embedded within the so-called semiotic landscape, which is an "imagined array of social types, distinguished on the basis of social issues and grounding linguistic

no mismatch in outputs needs to be assumed. That is, the two differing speakers might very well be producing similar outputs, but their respective idiolects nonetheless are assuming different underlying grammatical analyses.

variability in ideology" (Eckert 2019, p. 4). From this perspective, the adolescent peak, which is associated with the acceleration of linguistic change during preadolescence and adolescence, follows from the children's "opening up of the social landscape and its concomitant structuring and elaboration of the semiotic landscape" (p. 6). The developmental imperative (Eckert 2000)—that is, the desire to be older and to move on to the next stage of life—encourages adolescents to differentiate themselves from their elders within the context of ongoing social changes. Rather than a decline in linguistic plasticity (Bermúdez-Otero 2020, p. 2), the deceleration of change after adolescence may reflect a change to a life stage that encourages conformity rather than differentiation.

The discussion thus far suggests that, while individual-to-individual contact is a necessary condition for change to spread at the local level, it is not sufficient. For population-level spreading of an innovation to actuate, it requires the "joint indexical process that establishes a change as locally meaningful" (Eckert 2019, p. 2). The association of a variable with an age vector can be seen as a part of this general process of associating language variability with differentiations in the social world. Variation and change in idiolects offer the raw material that "puts the speakers-listeners in a state of readiness for phonetics and phonological innovations" (Lindblom et al. 1995, p. 13). Ultimately, however, it is ideological projects that "go looking for linguistic material" (Eckert 2019, p. 3). After all, just as the mapping between sound and lexical meaning is arbitrary, so the mapping between linguistic innovation and social meaning is fundamentally random, even if the mapping makes sense within the locally and historically specific framing in which a speaker's ideological project is embedded. For example, many studies have documented variation in /s/ production. As already discussed above, /s/ in /str/ sequences could be retracted (i.e., a retracted /s/ is realized with a lower spectral mean than nonretracted /s/). The sibilant could also be fronted, commonly realized with a higher spectral mean than nonfronted /s/. Different communities recruit /s/ variation for different ideological projects. In North America, fronted /s/ is often associated with gayness (Munson et al. 2006), but in rural northern California, town-oriented speakers generally have more fronted /s/ than country-oriented speakers. However, the younger country-oriented men are fronting their /s/, which Podesva & Hofwegen (2016) argued reflects a move away from hypermasculinity. Pharao et al. (2014) likewise found an association of fronted /s/ with gayness in Copenhagen Danish, but the perception of gayness is only evident when the talker has a modern guise [i.e., a man speaking in a (white) urban Copenhagen accent] and not when the talker has a street guise (i.e., Danish that is associated with Copenhagen's immigrant population and with a so-called gangster lifestyle). Maegaard & Pharao (2021) further found that the fronted /s/ effects on the modern voices are evident only in male voices and not in female ones. In other words, the connection between linguistic variables on the one hand and social correlates and social meaning on the other can vary quite drastically even among speakers who arguably speak the same language. Complicating the picture even further, such connection is not static either. Levon (2014), for example, found that, for listeners who endorse normative stereotypes of masculinity and male gender roles, fronted /s/ serves as a salient cue of nonmasculinity and gayness, in contrast to those who reject those stereotypes, who show no such association with fronted /s/. Likewise, Phillips & Resnick (2019) showed that the perception of /s/ variation in /s/-retraction contexts is affected by indicators of masculine toughness, both in terms of the talker's voice and face and in terms of the listener's endorsement of masculine stereotypes. These findings point to the significant impact listeners' individual beliefs have on the types of social meanings they associate with linguistic variation. That is, while innovative linguistic variants may have semiotic potential, the pairing of a linguistic sign with indexical meaning is nonetheless dependent on an individual's noticing, or what Gal & Irvine (2019) called uptake. Crucially, uptakes not only are "embedded in the flow of social life" (Gal & Irvine 2019, p. 16)—that is, they take place in a particular time and place—but also hinge on the receptiveness of the individual. That is, speakers-listeners might differ in sociolinguistic awareness (Garrett & Johnson 2013) or what Labov et al. (2011, p. 435) called the sociolinguistic monitor, a cognitive mechanism that "tracks, stores and processes" socially salient quantitative linguistic distributions. In other words, despite the fact that the social meaning attached to a linguistic variable emerges out of interactions between individuals, ultimately, it is the speaker-listener who has to decide what to make of the pairing between a linguistic variable and what it might index.

4. ACTUATION IN THE VIRTUAL WORLD

Many students of language change found the idiolectal view of actuation conceptually unsatisfying. Weinreich et al. (1968, p. 112) critique it as follows:

> For even when the course of a language change has been fully described and its ability explained, the question always remains as to why the change was not actuated sooner, or why it was not simultaneously actuated wherever identical functional properties prevailed. The unresolved actuation riddle is the price paid by any facile and individualistic explanation of language change. It creates the opposite problem—of explaining why language fails to change.

The attitude reflected in the quote above is endemic in the language change literature. As explained above, however, actuation takes place whenever a listener/language acquirer constructs a representation or grammar that does not match that of the person who provides the input, and since there are many possible factors that would lead to imperfect transmission, change ought to be the norm rather than the exception, just as Weinreich et al. (1968) predicted. But unlike Weinreich et al. (1968), we should not be dismayed about this conclusion. To be sure, change at the idiolectal level is not likely to be what they are concerned about; but change at the community level is not a straightforward matter, as we have seen. As noted above, a model of spread by communication density is likely to be far too simplistic. To the extent that the spread of linguistic change must be mediated by social meaning (Eckert 2019), which must be negotiated locally between language users with diverse backgrounds, idiolectal peculiarities, and ideological concerns, the incrementation and propagation of change, the level of actuation that Weinreich et al. (1968) were most concerned about, are predicted to be rare due to the need for multiple factors to converge in the right way. As noted by Ohala, "it is probably not possible to explain why a given sound change happens in one language but not another and it is generally not a fruitful question to pursue" (Ohala & Ohala 1991, p. 271). Just as it is not possible to predict precisely the movement of individual molecules, as molecules are constantly in motion, it is not productive to seek a deterministic solution to the incrementation and propagation aspects of the actuation riddle. As already alluded to above, Baker et al. (2011) emphasized that phonetic precursors that are shared by all members of a community should not be expected to produce sound change because of their commonness. Instead, it is speaker-specific linguistic changes that are favored as novel targets, and it is these novel targets that spread across individuals. The results of speaker-specific responses to phonological pressures "offer up changes, which are then recruited into ideological projects" (Eckert 2019). Inter-speaker variation, as reflected in changes observed across idiolects, is par for the course within any linguistic community. The spread of a particular speaker-specific innovation happens when there is a "spurious correlation between a phonetic variable (as produced by a particular speaker or subset of speakers) and a social variable, and a speaker adopts the change into his/her own speech" (Baker et al. 2011, p. 365).

Must we be resigned to the fact that the actuation riddle is an unresolvable problem? While "deductive-nomological explanations of particular instances of sound change" (Bermúdez-Otero 2020) are unlikely, the random nature of linguistic change actuation suggests that it is better modeled using stochastic methods. To this end, agent-based models (ABMs) have been deployed to investigate how communication density, in concert with different linguistic factors, drives sound change. ABMs vary greatly in terms of the assumptions made about the nature of the input distributions, the perceptual and production systems, the phonological and lexical organizations, memory retention and loss, and the interactional relationship between agents (e.g., Blevins & Wedel 2009; Kirby 2013, 2014; Sóskuthy 2015; Harrington & Schiel 2017; Stevens et al. 2019). Models are set up quite differently, and their results also vary quite drastically depending on how specific parameters are set. As such, it can be difficult to assess the relative successes across models and what they can tell us about actuation. Consider, for example, the work of Kirby (2013), who modeled the emergence of contrastive F0 in Seoul Korean using an ABM in which the agents may enhance a cue in production to improve the precision of a contrast or bias a cue to reduce contrast in production. The simulation results show that agents equipped with enhancement alone produce no noticeable changes, whereas agents who are subject to bias alone lead to complete mergers for all cues. The type of transphonologization between VOT and F0 found in Seoul Korean is only evident when bias and enhancement are allowed to work together. Kirby (2014) modeled the incipient tonogenesis found in Phnom Penh Khmer, where the fortition of /r/ endangered the distinction between /CrV/, /CV/, and /C^hV/ forms. He found that the loss of /r/ drove the emergence of F0 drop or breathy phonation even when probablistic enhancement was not part of the production architecture of the agents. Kirby (2014) suggested that the fact that the ABM for Phnom Penh Khmer was able to yield transphonologization results without the probablistic enhancement mechanism might have to do with the use of batch learning (contra online learning in the Seoul Korean simulation) or the number of cue dimensions involved in each simulation (i.e., five in Kirby 2013 versus three in Kirby 2014). Many ABM studies assume that the agents share the same initial state, and channel biases (in production or perception) are introduced during the learning process between agents who experience those biases with the same likelihood. A scenario more in line with the idiolectal model of actuation laid out above is an ABM in which the agents are initialized with speaker-specific cue/lexical distributions. Stevens et al. (2019), for example, examine the likelihood of /s/-retraction by initializing the agents with phonetic and phonological distributions drawn from 19 of the 20 Australian English speakers studied by Stevens & Harrington (2016). That is, each agent is initialized with the cue distributions of one of 19 Australian speakers. A talker agent would randomly select a word class from the lexicon and generate a signal consistent with the distribution of acoustic parameters associated with that word class. The selected word class and derived signal are transmitted to the agent listener, who then incorporates the input into the same word class if it is probabilistically closest to one of the agent listener's phonological classes. Unlike in other ABMs, words were allowed to be reclassified into different phonological classes after 100 iterations of talker-listener interactions. After 100 runs, the authors examined the difference between the baseline and the post-runs and found that /s/ became more retracted, but it did not get reclassified as /ʃ/ completely. Their findings suggest that the agents created an intermediate category of /str/ that is closer to /s/ rather than merging /str/ with /ʃ/.

These and other ABM-based studies offer potential constructive models of how likely community-level shifts would actuate. Kirby & Sonderegger (2015) found that unlike learners in ABMs with a simple prior or no prior at all, learners in ABMs who were equipped with a complex prior (i.e., a categoricity bias toward certain categories) showed a bifurcation, such as a sudden change from a stable contextual variation (i.e., the F1 of /a/ lowers gradually before /i/) to a stable umlaut (i.e., /a/ → /i/ before /i/). They also found that, in addition to the categoricity bias, population structure itself (i.e., whether agents learn from a single or multiple teachers) can play a role in promoting stability of existing phonetic categories. Harrington & Schiel (2017) investigated the effects of dialect contact on the likelihood of /u/-fronting between two dialects of agents. Agents from one group were initialized with cue distributions taken from the younger British English speakers studied by Harrington et al. (2008), while agents of the other group were initialized with cue distributions taken from the older speakers from the same study. The younger speakers exhibited completely fronted /u/, while the older speakers exhibited only mild evidence of /u/-fronting. Harrington & Schiel (2017) used an ABM to understand the effect of asymmetric cue distribution on the direction of vowel shifts between the two groups of agents. While they found a larger shift of the older speakers' /u/ toward the front of the vowel space under most simulation runs, 12 of the 100 runs (with 50,000 iterations of agent interactions each) show the two groups converging to an intermediate vowel between the starting positions of each group. These and other findings (Garrett & Johnson 2013, Stanford & Kenny 2013) suggest that community-level changes are not inevitable, or at least they do not always result in the same outcomes. While ABMs are often highly idealized, they nonetheless provide investigators with a means to examine how the propagation of linguistic innovation can be "stochastically influenced by factors such as which speakers come into contact with each other, how often they do so, and whether upon contact a produced item is absorbed and then subsequently retained in the perceiver's memory" (Harrington & Schiel 2017, p. 437).

5. CONCLUSION

Actuation is a multifaceted phenomenon that has been examined at both individual and community levels. Ultimately, linguistic changes actuate in the form of idiolectal differences. This idiolectal view of actuation raises questions about the traditional appeal to channel biases to provide insights into the seeds of sound change. In the end, if the physical and perceptual principles assumed to be the phonetic precursors to change are truly universal and thus always present, they cannot in principle offer the sufficient conditions for explaining why individuals would nonetheless arrive at divergent representations and grammars. An understanding of language change actuation at the idiolectal level thus requires an understanding of how individual speaker-listeners' different past linguistic experiences and physical, perceptual, and cognitive makeups affect the way they process and analyze the primary learning data and how these factors lead to divergent representations and grammars. Population-level incrementation and propagation of linguistic innovation depend not only on the nature of contact between speakers with unique idiolects but also on individuals who have the wherewithals to take advantage of the linguistic innovations they encounter to achieve particular ideological projects they are engaging in at any given moment. Because of the vast number of contingencies that need to be aligned properly, the incrementation and propagation of linguistic innovation are predicted to be rare. ABMs promise to provide a controlled way to investigate the stochastic nature of language change propagation, but a comprehensive model of linguistic change actuation at the individual level remains elusive.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Many thanks to Stephen Anderson, Andrew Garrett, and James Kirby for their feedback on earlier versions of this review. Any errors are of course my own. This work was supported in part by the National Science Foundation (BCS1827409).

LITERATURE CITED

Babel M. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *J. Phonet.* 40(1):177–89

Baker A, Archangeli D, Mielke J. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Lang. Var. Change* 23(3):347–74

Bang H-Y, Sonderegger M, Kang Y, Clayards M, Yoon T-J. 2011. The emergence, progress, and impact of sound change in progress in Seoul Korean: implications for mechanisms of tonogenesis. *J. Phonet.* 66:120–44

Beddor PS. 2009. A coarticulatory path to sound change. *Language* 85(4):785–832

Beddor PS, Coetzee A, Styler W, McGowan K, Boland J. 2018. The time course of individuals' perception of coarticulatory information is linked to their production: implications for sound change. *Language* 94:931–68

Bermúdez-Otero R. 2020. The initiation and incrementation of sound change: community-oriented momentum-sensitive learning. *Glossa J. Gen. Linguist.* 5(1):121

Blevins J. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge, UK: Cambridge Univ. Press

Blevins J, Wedel A. 2009. Inhibited sound change: an evolutionary approach to lexical competition. *Diachronica* 26:143–83

Bloomfield L. 1933. *Language*. New York: H. Holt & Co.

Carignan C, Coretta S, Frahm J, Harrington J, Hoole P, et al. 2021. Planting the seed for sound change: evidence from real-time MRI of velum kinematics in German. *Language* 97(2):333–64

Chodroff E, Wilson C. 2018. Predictability of stop consonant phonetics across talkers: between-category and within-category dependencies among cues for place and voice. *Linguist. Vanguard* 4(S2):20170047

Clayards M. 2018a. Differences in cue weights for speech perception are correlated for individuals within and across contrasts. *J. Acoust. Soc. Am.* 144:EL172

Clayards M. 2018b. Individual talker and token variability in multiple cues to stop voicing. *Phonetica* 75(1):1–23

Coetzee AW, Beddor PS, Shedden K, Styler W, Wissing D. 2018. Plosive voicing in Afrikaans: differential cue weighting and sound change. *J. Phonet.* 66:185–216

Cohn A. 1993. Nasalization in English: phonology or phonetics. *Phonology* 10(1):43–81

Dediu D, Moisik SR. 2019. Pushes and pulls from below: anatomical variation, articulation and sound change. *Glossa J. Gen. Linguist.* 4(1):7

Dmitrieva O, Llanos F, Shultz AA, Francis AL. 2015. Phonological status, not voice onset time, determines the acoustic realization of onset $f0$ as a secondary voicing cue in Spanish and English. *J. Phonet.* 49(1):77–95

Dodsworth R. 2019. Bipartite network structures and individual differences in sound change. *Glossa J. Gen. Linguist.* 4(1):61

Dodsworth R, Benton RA. 2020. *Language Variation and Change in Social Networks: A Bipartite Approach*. London: Routledge

Eckert P. 2000. *Linguistic Variation as Social Practice*. Oxford, UK: Blackwell

Eckert P. 2019. The individual in the semiotic landscape. *Glossa J. Gen. Linguist.* 4(1):14

Escudero P, Boersma P. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Stud. Second Lang. Acquis.* 26:551–85

Francis AL, Ciocca V, Wong VK, Chan JK. 2006. Is fundamental frequency a cue to aspiration in initial stops? *J. Acoust. Soc. Am.* 120(5):2884–95

Francis AL, Kaganovich N, Discoll-Huber C. 2008. Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *J. Acoust. Soc. Am.* 124(2):1234–51

Gal S, Irvine JT. 2019. *Signs of Difference: Language and Ideology in Social Life*. Cambridge, UK: Cambridge Univ. Press

Ganong WF. 1980. Phonetic categorization in auditory word perception. *J. Exp. Psychol. Hum. Percept. Perform.* 6:110–25

Garrett A, Johnson K. 2013. Phonetic biases in sound change. In *Origins of Sound Change: Approaches to Phonologization*, ed. ACL Yu, pp. 51–97. Oxford, UK: Oxford Univ. Press

Halle M, Stevens K. 1971. A note on laryngeal features. *MIT Res. Lab. Electron. Q. Prog. Rep.* 101:198–213

Hanson HM. 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *J. Acoust. Soc. Am.* 125(1):425–41

Harrington J, Kleber F, Reubold U. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: an acoustic and perceptual study. *J. Acoust. Soc. Am.* 123(5):2825–35

Harrington J, Kleber F, Reubold U, Schiel F, Stevens M. 2018. Linking cognitive and social aspects of sound change using agent-based modeling. *Top. Cogn. Sci.* 10(4):707–28

Harrington J, Schiel F. 2017. /u/-fronting and agent-based modeling: the relationship between the origin and spread of sound change. *Language* 93(2):414–45

Haugen E. 1966. Dialect, language, nation. *Am. Anthropol.* 68(4):922–35

Hombert J-M, Ohala JJ, Ewan WG. 1979. Phonetic explanations for the development of tones. *Language* 55(1):37–58

Hyman L. 1976. Phonologization. In *Linguistic Studies Presented to Joseph H. Greenberg*, ed. A Juilland, pp. 407–18. Saratoga, CA: Anma Libri

Idemaru K, Holt LL, Seltman H. 2012. Individual differences in cue weight are stable across time: the case of Japanese stop lengths. *J. Acoust. Soc. Am.* 132(6):3950–64

Kang Y. 2014. Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: a corpus study. *J. Phonet.* 45(1):76–90

Kapnoula EC, Winn MB, Kong EJ, Edwards J, McMurray B. 2017. Evaluating the sources and functions of gradiency in phoneme categorization: an individual differences approach. *J. Exp. Psychol. Hum. Percept. Perform.* 43:1594–611

Kapnoula EE. 2016. *Individual differences in speech perception: sources, functions, and consequences of phoneme categorization gradiency*. PhD Thesis, Univ. Iowa, Iowa City

Keyser SJ, Stevens KN. 2006. Enhancement and overlap in the speech chain. *Language* 82:33–63

Kingston J. 2007. Segmental influences on F_0: automatic or controlled? In *Tones and Tunes*, Vol. 2: *Experimental Studies in Word and Sentence Prosody*, ed. C Gussenhoven, T Riad, pp. 171–201. Berlin: Mouton de Gruyte

Kingston J, Diehl RL. 1994. Phonetic knowledge. *Language* 70(3):419–54

Kirby JP. 2013. The role of probabilistic enhancement in phonologization. In *Origins of Sound Change: Approaches to Phonologization*, ed. ACL Yu, pp. 228–46. Oxford, UK: Oxford Univ. Press

Kirby JP. 2014. Incipient tonogenesis in Phnom Penh Khmer: computational studies. *Lab. Phonol.* 5:195–230

Kirby JP. 2018. Onset pitch perturbations and the cross-linguistic implementation of voicing: evidence from tonal and non-tonal languages. *J. Phonet.* 71:326–54

Kirby JP, Ladd DR. 2016. Effects of obstruent voicing on vowel *F*0: evidence from "true voicing" languages. *J. Acoust. Soc. Am.* 140:2400–11

Kirby JP, Sonderegger M. 2015. Bias and population dynamics in the actuation of sound change. arXiV:1507.04420 [cs.CL]

Klatt DH. 1986. The problem of variability in speech recognition and in models of speech perception. In *Invariance and Variability in Speech Processes*, ed. JS Perkell, DH Klatt, pp. 300–19. Hillsdale, NJ: Erlbaum

Kong EJ, Edwards J. 2016. Individual differences in categorical perception of speech: cue weighting and executive function. *J. Phonet.* 59(1):40–57

Labov W. 1994. *Principles of Linguistic Change*, Vol. 1: *Internal Factors*. Oxford, UK: Blackwell

Labov W. 2001. *Principles of Linguistic Change*, Vol. 2: *Social Factors*. Malden, MA: Blackwell

Labov W. 2007. Transmission and diffusion. *Language* 83(2):344–87

Labov W, Ash S, Ravindranath M, Weldon T, Nagy N. 2011. Properties of the sociolinguistic monitor. *J. Sociolinguist.* 15(4):431–63

Ladefoged P. 1967. *Three Areas of Experimental Phonetics*. Oxford, UK: Oxford Univ. Press

Lehet M, Holt LL. 2017. Dimension-based statistical learning affects both speech perception and production. *Cogn. Sci.* 41(S4):885–912

Levon E. 2014. Categories, stereotypes, and the linguistic perception of sexuality. *Lang. Soc.* 43:539–66

Li FF. 2013. The effect of speakers' sex on voice onset time in Mandarin stops. *J. Acoust. Soc. Am.* 133:142–47

Lindblom B, Guion S, Hura S, Moon S-J, Willerman R. 1995. Is sound change adaptive? *Riv. Linguist.* 7:5–36

Löfqvist A, Baer T, McGarr NS, Story RS. 1989. The cricothryoid muscle in voicing control. *J. Acoust. Soc. Am.* 85:1314–21

Lundeborg I, Larsson M, Wiman S, McAllister AM. 2012. Voice onset time in Swedish children and adults. *Logop. Phoniatr. Vocology* 37(3):117–22

Maegaard M, Pharao N. 2021. Features, meanings, and indexical fields. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*, ed. L Hall-Lew, E Moore, R Podesva, pp. 203–21. Cambridge, UK: Cambridge Univ. Press

Massaro DW, Cohen MM. 1983. Phonological constraints in speech perception. *Percept. Psychophys.* 34:338–48

Mielke J, Baker A, Archangeli D. 2016. Individual-level contact limits phonological complexity: evidence from bunched and retroflex /ɹ/. *Language* 92(1):101–40

Milroy L. 1980. *Language and Social Networks*. Oxford, UK: Blackwell

Milroy L. 2002. Off the shelf or under the counter? On the social dynamics of sound changes. In *Studies in the History of the English Language*, Vol. III: *Managing Chaos: Strategies for Identifying Change in English*, ed. CM Cain, G Russom, pp. 149–72. Berlin: Mouton de Gruyter

Munson B, McDonald EC, DeBoe NL, White AR. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *J. Phonet.* 34(2):202–40

Nielsen K. 2011. Specificity and abstractness of VOT imitation. *J. Phonet.* 39(2):132–42

Oh E. 2011. Effects of speaker gender on voice onset time in Korean stops. *J. Phonet.* 39:59–67

Ohala JJ. 1973. *Explanations for the intrinsic pitch of vowels*. Mon. Intern. Memo., Phonol. Lab., Univ. Calif., Berkeley

Ohala JJ. 1981. The listener as a source of sound change. In *Papers from the Parasession on Language and Behavior*, pp. 178–203. Chicago: Chicago Linguist. Soc.

Ohala JJ. 1993. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, ed. C Jones, pp. 237–78. London: Longman

Ohala JJ. 1994. The frequency code underlies the sound symbolic use of voice pitch. In *Sound Symbolism*, ed. L Hinton, J Nichols, JJ Ohala, pp. 325–47. Cambridge, UK: Cambridge Univ. Press

Ohala M, Ohala JJ. 1991. Reply to commentators. *Phonetica* 48:271–74

Ou J, Yu ACL. 2021. Neural correlates of individual differences in speech categorization: evidence from subcortical, cortical, and behavioral measures. *Lang. Cogn. Neurosci.* 37(3):269–84

Ou J, Yu ACL, Xiang M. 2021. Individual differences in categorization gradience as predicted by online processing of phonetic cues during spoken word recognition: evidence from eye movements. *Cogn. Sci.* 45(3):e12948

Paul H. 1880. *Prinzipien der Sprachgeschichte*. Halle, Ger.: M. Niemayer

Peng J-F, Mei Chen L, Lee C-C. 2014. Voice onset time of initial stops in Mandarin and Hakka: effect of gender. *Taiwan J. Linguist.* 21(1):63–80

Peterson GE, Barney HL. 1952. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24:175–84

Pharao N, Maegaard M, Møller J, Kristiansen T. 2014. Indexical meanings of [s^+] among Copenhagen youth: social perception of a phonetic variant in different linguistic contexts. *Lang. Soc.* 43(1):1–31

Phillips J, Resnick P. 2019. Masculine toughness and the categorical perception of onset sibilant clusters. *J. Acoust. Soc. Am.* 145:EL574

Pinget A-F. 2015. *The Actuation of Sound Change*. Utrecht, Neth.: LOT

Podesva RJ, Hofwegen JV. 2016. /s/exuality in small-town California: gender normativity and the acoustic realization of /s/. In *Language, Sexuality, and Power: Studies in Intersectional Linguistics*, ed. E Levon, RB Mendes, pp. 16–88. Oxford, UK: Oxford Univ. Press

Ramsammy M. 2018. The phonology-phonetics interface in constraint-based grammar: gradience, variability, and phonological change. In *The Routledge Handbook of Phonological Theory*, ed. S Hannahs, ARK Bosch. Abingdon, UK: Routledge. **https://doi.org/10.4324/9781315675428**

Reddy BMS, Kumar NM, Sreedevi N. 2013. Voice onset time across gender and different vowel contexts in Telugu. *Lang. India* 14(12):252–63

Sachs J, Lieberman P, Erickson D. 1973. Anatomical and cultural determinants of male and female speech. In *Language Attitudes: Current Trends and Prospects*, ed. RW Shuy, RW Fasold, pp. 74–84. Washington, DC: Georgetown Univ. Press

Schertz J, Cho T, Lotto A, Warner N. 2015. Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *J. Phonet.* 52:183–204

Shultz AA, Francis AL, Llanos F. 2012. Differential cue weighting in perception and production of consonanat voicing. *J. Acoust. Soc. Am.* 132:EL95

Smith BJ, Mielke J, Magloughlin L, Wilbanks E. 2019. Sound change and coarticulatory variability involving English /ɹ/. *Glossa J. Gen. Linguist.* 4(1):63

Solé M-J. 2007. Controlled and mechanical properties in speech: a review of the literature. In *Experimental Approaches to Phonology*, ed. M-J Solé, PS Beddor, M Ohala, pp. 302–21. Oxford, UK: Oxford Univ. Press

Sóskuthy M. 2015. Understanding change through stability: a computational study of sound change actuation. *Lingua* 163:40–60

Stanford JN, Kenny LA. 2013. Revisiting transmission and diffusion: an agent-based model of vowel chain shifts across large communities. *Lang. Var. Change* 25(2):119–53

Stevens M, Harrington J. 2014. The individual and the actuation of sound change. *Loquens* 1(1):e003

Stevens M, Harrington J. 2016. The phonetic origins of /s/-retraction: acoustic and perceptual evidence from Australian English. *J. Phonet.* 58:118–34

Stevens M, Harrington J, Schiel F. 2019. Associating the origin and spread of sound change using agent-based modelling applied to /s/-retraction in English. *Glossa J. Gen. Linguist.* 4(1):8

Svantesson J-O, House D. 2006. Tone production, tone perception and Kammu tonogenesis. *Phonology* 23:309–33

Vorperian HK, Wang S, Schimek EM, Durtschi RB, Kent RD, et al. 2011. Developmental sexual dimorphism of the oral and pharyngeal portions of the vocal tract: an imaging study. *J. Speech Lang. Hear. Res.* 54:995–1010

Wade L, Lai W, Tamminga M. 2021. The reliability of individual differences in VOT imitation. *Lang. Speech* 64(3):576–93

Walkden G. 2017. The actuation problem. In *The Cambridge Handbook of Historical Syntax*, ed. A Ledgeway, I Roberts, pp. 403–24. Cambridge, UK: Cambridge Univ. Press

Weinreich U, Labov W, Herzog M. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics*, ed. W Lehmann, Y Malkiel, pp. 95–188. Austin: Univ. Tex. Press

Whiteside SP, Marshall J. 2001. Developmental trends in voice onset time: some evidence for sex differences. *Phonetica* 58:196–210

Yu ACL. 2010. Perceptual compensation is correlated with individuals' "autistic" traits: implications for models of sound change. *PLOS ONE* 5(8):e11950

Yu ACL. 2013. Individual differences in socio-cognitive processing and the actuation of sound change. In *Origins of Sound Change: Approaches to Phonologization*, ed. ACL Yu, pp. 201–27. Oxford, UK: Oxford Univ. Press

Yu ACL. 2021. Toward an individual difference perspective on phonologization. *Glossa J. Gen. Linguist.* 6(1):14

Yu ACL. 2022. Perceptual cue weighting is influenced by the listener's gender and subjective evaluations of the speaker: the case of English stop voicing. *Front. Psychol. Lang. Sci.* 13:840291

Yu ACL, Abrego-Collier C, Sonderegger M. 2013. Phonetic imitation from an individual-difference perspective: subjective attitude, personality, and "autistic" traits. *PLOS ONE* 8(9):e74746

Zhang X, Holt LL. 2018. Simultaneous tracking of coevolving distributional regularities in speech. *J. Exp. Psychol. Hum. Percept. Perform.* 44(11):1760–79