

RESEARCH

Pushes and pulls from below: Anatomical variation, articulation and sound change

Dan Dediu^{1,2,3,4} and Scott R. Moisiuk^{3,5}¹ Laboratoire Dynamique Du Langage UMR5596, Université Lumière Lyon 2, 14 avenue Berthelot, 69363 Lyon, FR² Collegium de Lyon, Institut d'Études Avancées, 24 rue Baldassini, 69007 Lyon, FR³ Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, NL⁴ Donders Institute for Brain, Cognition and Behaviour, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, NL⁵ Linguistics and Multilingual Studies, Nanyang Technological University, 50 Nanyang Avenue, SGCorresponding author: Dan Dediu (dan.dediu@univ-lyon2.fr)

This paper argues that inter-individual and inter-group variation in language acquisition, perception, processing and production, rooted in our biology, may play a largely neglected role in sound change. We begin by discussing the patterning of these differences, highlighting those related to vocal tract anatomy with a foundation in genetics and development. We use our *ArtiVarK* database, a large multi-ethnic sample comprising 3D intraoral optical scans, as well as structural, static and real-time MRI scans of vocal tract anatomy and speech articulation, to quantify the articulatory strategies used to produce the North American English /r/ and to statistically show that anatomical factors seem to influence these articulatory strategies. Building on work showing that these alternative articulatory strategies may have indirect coarticulatory effects, we propose two models for how biases due to variation in vocal tract anatomy may affect sound change. The first involves direct overt acoustic effects of such biases that are then reinterpreted by the hearers, while the second is based on indirect coarticulatory phenomena generated by acoustically covert biases that produce overt “at-a-distance” acoustic effects. This view implies that speaker communities might be “poised” for change because they always contain pools of “standing variation” of such biased speakers, and when factors such as the frequency of the biased speakers in the community, their positions in the communicative network or the topology of the network itself change, sound change may rapidly follow as a self-reinforcing network-level phenomenon, akin to a phase transition. Thus, inter-speaker variation in structured and dynamic communicative networks may couple the initiation and actuation of sound change.

Keywords: sound change; vocal tract anatomy; heterogeneous communicative networks

1 Introduction

Understanding *sound change* is key to a full account of how languages evolve and of the present-day distribution of linguistic diversity. While the last decades have witnessed a concerted effort across multiple disciplines, resulting in sustained progress towards a better model of sound change and its causes, there are still long-enduring puzzles as well as new questions that need to be addressed. Sound change can be usefully analyzed into its *initiation*, possibly followed by its *actuation* (Solé & Vives 2012; Yu 2013; Stevens & Harrington 2014). Initiation is traditionally located within an individual, where cognitive, perceptual or production processes, in a given context, produce a novel variant (say, a new articulatory realization of the trill /r/ as [ɾ] instead of the community norm of [r]), novelty that may or may not spread through the speech community, under the influence of many factors, to become actuated as the new norm. However, while apparently clear

and of great analytic help, this distinction is marred by the long-standing resilience of the “actuation problem” (Weinreich et al. 1968).

We suggest here that, at least in some cases, the initiation and the actuation of sound change are more intimately connected, and that conceptually severing this connection might hinder progress. More precisely, we argue that it might be more appropriate to picture sound change as happening not in a “neutral” speech community (i.e., one which does not “care” if [ɹ] or [r] is used), but in an already “poised” community, on the brink of a phase transition, where there are biases, preferences, constraints, affordances, factors that push and pull these variants, non-uniformities that are hidden below an apparently calm surface. While our arguments here will be focused on inter-individual and inter-group variation in the anatomy of the vocal tract and its impact on sound change (and, in particular, on the articulation of North American English /r/, which we show, using novel experimental evidence, is influenced by inter-individual variation in the anatomy of the anterior vocal tract), these ideas are more general and may apply equally well to other types of biases rooted in language perception, acquisition, processing, and production. As such, we suggest that variation between individuals and groups (not only sociolinguistic, but also anatomical, physiological, cognitive and psychological) must become central to linguistics as a whole, but we also warn against focusing on the isolated individual. Instead, language is a group-level phenomenon, emerging from structured interactions between individuals that simultaneously vary in myriad ways while also being embedded in dynamic and complex communication networks. We suggest that re-integrating inter-individual variation with network-level dynamics may dissolve the initiation/actuation dichotomy into a seamless process in which the two constantly interact.

The paper is structured as follows. Section 2 briefly discusses the patterning of inter-individual and inter-group variation, with a particular focus on the anatomy of the vocal tract. Section 3 reviews suggestions about the influence of vocal tract (and vocal tract variation) on sound change. In Section 4 we focus on the case of the North American English /r/, where we analyze our own *ArtiVarK* experimental sample, connecting MRI articulatory data with intra-oral anatomical information to show that the articulatory strategy employed is influenced by the anterior vocal tract anatomy; based on proposals about the overt and coarticulatory effects of alternative articulatory strategies, we suggest that this might be a case of anatomical biasing of sound change. We develop these ideas in a more abstract fashion in Section 5, where we suggest that linguistic communities contain vast amounts of “standing variation” due to inter-individual differences, and that even minor changes might trigger “phase transitions” in such “poised to change” communities, suggesting a way to unify the initiation and actuation of sound change. We close with a brief discussion and future directions in Section 6.

2 Vocal tract variation in context

Humans differ from each other, as individuals, in many ways, ranging from physical characteristics (such as height, weight and facial features; Durand & Rappold 2013) to psychological and cognitive ones (including temperament, memory capacity and vocabulary size; Furr & Bacharach 2007) and to subtle biochemical and genetic variation (e.g., drug metabolism; Tracy et al. 2016). It has usually been assumed, especially in the cognitive sciences, including the language sciences, that such inter-individual variation is epiphenomenal, merely noise that must be controlled for and abstracted away. Certainly, there is more that we share as humans than what sets us apart as individuals, and there are fundamental questions concerning the levels above the individual, addressed by scientific paradigms focused on universality. One such series of paradigms is represented by Noam Chomsky’s take on language, conceptualized as having a universal and invariable

core shared by all (normal) humans, whose interfacing with other systems allows cross-linguistic (and also presumably inter-individual) variation. This is a reflection of a wider social and scientific worldview rooted in the belief of an immutable and universal essence encapsulated in our genome. Scientific aspects aside, the horrors of the Second World War, and the self-questioning they generated, cast a long shadow on the study of differences, perceived as among the root causes of the mistreatment of individuals and groups seen as “different”. However, these paradigms are currently under strain from several mutually reinforcing directions.

Medicine deals with differences between individuals, ranging from variation in risk, to differential reactions to the same pathogen or treatment, to variation in the way people adhere to their prescriptions. There is increased attention devoted to individualized (or personalized) medicine (Topol 2014) and patient-centered health care (Mead & Bower 2000). The individual characteristics may have biological, cultural, social and economic components that may interact in complex ways (for example, the risk for developing lung cancer is increased by smoking but also by certain genetic variants).¹ Moreover, part of this variation between patients is explained by them belonging to certain groups that have cultural, socioeconomic, but also biological components. For example, the risk of developing diabetes is increased in groups with lower socioeconomic status (Agardh et al. 2011), in groups that do not follow healthy living guidelines (Lumb 2014), but also in groups of particular ancestry (Ferdinand & Nasser 2015), partly due to inter-group differences in the frequency of genetic risk factors (Wells 2009). Research in anthropology and genetics has uncovered several important facts about human variation. First, there is much more variation hidden in our genomes than anybody expected, most of the variants being restricted to a small number of people (The 1000 Genomes Project Consortium 2015), but we are still less diverse than, for example, the chimps (Bowden et al. 2012). Second, this variation is geographically structured, being relatively easy to assign individuals to broad regions based on their genomes (Novembre et al. 2008; Nassir et al. 2009; Paschou et al. 2010), but this information is spread across multiple genetic loci, with most common variants shared across many populations (The 1000 Genomes Project Consortium 2015) in gradual clines of continuously changing allele frequencies (Barbujani & Colonna 2010). Just approximately 8% of the variation is distributed between continents, about 8% between groups within the same continent, while the bulk of the variation is between individuals from the same group (Lewontin 1972; Barbujani & Colonna 2010). Third, this variation decreases gradually with distance from Africa (Barbujani & Colonna 2010), which hosts the most diverse human groups and also some of the oldest genetic lineages (Pickrell et al. 2012); this is largely reflected also by various phenotypic measures (Cramon-Taubadel & Lycett 2008; Betti et al. 2009). Fourth, present-day genetic diversity is the result of very complex histories of population differentiation, admixture and replacement (Reich 2018).

Thus, against a shared universal background, there are pervasive differences between individuals and between groups, geographically patterned, mostly continuous, gradual, and usually increasing with distance; moreover, they are usually statistical in nature, with discrete features being present at different frequencies in different groups, and with continuous ones having slightly different distributions. This patterning of variation is probably one of the strongest arguments against racism and discrimination, but we need to acknowledge and properly understand it (Reich 2018). Given that most of the variation

¹ Many cases of lung cancer (about 25%) are due to environmental factors (asbestos, radon, coal burning), infections (HPV), and genetics (Couraud et al. 2012; Torres-Durán et al. 2014). Genes might be involved even in the primary cause of lung cancer, smoking, by affecting the level of addiction and the difficulty of cessation (Benowitz 2009).

is distributed within groups, that groups are (and have always been) fluid entities, continuously merging into each other, always on the move and always interacting, any ideology built on categorical borders, on national or group “essences”, becomes untenable. Identities are dynamic, resulting from long and complex historical processes shaped by contingencies and by continuous interactions between biology and culture. However, even if differences are real, it matters what we make of them: racism, sexism and discrimination create and justify hierarchies based on perceived or invented differences (Schneider 2005; Rattansi 2007; Ridgeway 2011; Lippert-Rasmussen 2014; Safdar & Kosakowska-Berezecka 2015). Besides, denying the existence of variation leads (paradoxically) to the discrimination and oppression of those that are “different”.²

On this background, there is extensive variation in the *anatomy, physiology and control of vocal tract structures*, and while this was rather neglected by theoretical linguistics, more applied disciplines (such as speech therapy, education and experimental phonetics) have always been aware and interested in it. In what concerns the *pathological* extremes of variation, virtually every aspect of the vocal tract is known to be affected by specific conditions or by syndromes, some with a clear environmental etiology (e.g., accidents), but most involving an interplay between genetic and environmental risk factors. Research in this area, and especially in uncovering the genetic bases of pathological variation, is advancing fast, with, for example, *Online Mendelian Inheritance in Man* (OMIM; <https://omim.org>) and *PubMed* (<https://www.ncbi.nlm.nih.gov/pubmed>) being invaluable, free and up-to-date resources.³ The developmental pathologies affecting the *dentition* and the *jaw* have received particular attention due to their impact on feeding behavior and esthetics, with a good understanding of tooth development and the genetic bases of conditions such as *microdontia* (small teeth), *hypodontia* (missing teeth due to tooth agenesis) or *supernumerary teeth* (more teeth than normal), among others (Cobourne & Sharpe 2013; Klein et al. 2013; Brook et al. 2014). Another class of conditions that are well-studied and understood is represented by *cleft lip* with or without *cleft palate*, which can impact not only deglutition, but can have major deleterious effects on speech (Gibbon et al. 2008). These may be part of wider syndromes or occur isolated, and they helped shed light on genes involved in the development of the vocal tract (Dixon et al. 2011; Bush & Jiang 2012; Setó-Salvia & Stanier 2014). Much less well studied are pathologies affecting the *tongue* (such as micro-, macro- and ankylo-glossia; Reynoso et al. 1994; Hong 2013; Ounap 2016), the *velum* or the *larynx* (e.g., cleft larynx; Johnston et al. 2014). While probably less directly relevant to sound change than *normal variation*, such pathologies, coupled with animal models and evolutionary-developmental studies (Jernvall & Thesleff 2012), open the door to understanding the genetic and developmental mechanisms involved in the emergence of the observed patterns of variation between normal individuals.

While work towards understanding the genetic foundations of the *normal range of variation* is in its early stages, it is already clear that this range has been systematically underappreciated: for example, there is extensive variation in the size and shape of the *velum* (You et al. 2008; Kumar & Gopal 2011; Praveen et al. 2011) and the *hard palate* (Riquelme & Green 1970; Townsend et al. 1990; Lammert et al. 2013). However, there is more to variation than just its inter-individual dimension: while we expect that most of this

² One of the authors (DD) grew up in a nominally egalitarian society (Ceaușescu’s Communist Romania before 1989) which repressed differences while promoting uniformity. However, people continued being different (sometimes with deadly consequences) and these differences were sometimes recognized and even encouraged (e.g., school selection through exams) when perceived as useful to the society. This may have been an extreme Procrustes’ Bed, but “different” people everywhere tend to be forced into “universal” canons which do not fit them easily.

³ For example, a comprehensive search in OMIM for entries touching upon the evolution, development, pathology and genetics of vocal tract structures, returned as of February 2018 about 25,000 results.

diversity is distributed between members of the same population, some of this variation might still be distributed *between groups*. However, the currently available data are insufficient (in terms of both quantity and quality) to properly understand these patterns of variation, but we predict that such inter-group variation in aspects of the vocal tract to exist. We will discuss below several examples, after which we will turn to our own data showing that inter-individual variation in vocal tract anatomy affects the articulation of the North American English /r/.

3 Vocal tract variation and sound change

Phonetic variation is an essential precondition for sound change whatever the causes in particular instances may be. Ohala (1989) described this as the requisite “pool of synchronic variation” upon which sound change act(uate)s. The way that this variation is structured will (somehow) percolate through speech communities and crystalize into new phonetic and ultimately phonological norms. The question is what systematicity, if any, characterizes the variation. One intriguing possibility is that the phonetic variation behind sound change is, to some extent, governed by characteristic variation of the form and function of the vocal tract. This marks a finer grained characterization of what Weinreich et al. (1968) termed the “constraints problem”: what factors determine what sound changes are more expected or “natural” based on the nature of the speech and hearing mechanism. We propose here that the naturalness of sound change may not be universal, but rather, to a lesser or greater extent (depending on the sounds in question), *vary from one community to the next*. This would thus add another layer to the proposal that articulatory variation (in general) is important in the actuation of sound change (Baker et al. 2011; De Decker & Nycz 2012).

De Decker & Nycz (2012) point out that differences in articulatory strategy (arising from individual variation in the discovery of articulatory strategies to adequately approximate perceived acoustic targets) is critical for understanding sound change. They propose that such covert articulatory variation may manifest in distinct phonological patterns if learners converge on a particular strategy that differs from that used in previous generations. Thus, we are faced with two ways in which variation impacts sound change: variation in what is natural and variation in articulatory strategy. Both of these sources of variation, in turn, can be plausibly impacted by variation in vocal tract morphology.

The idea that human vocal tract variation biases phonological systems has been around for a long time (e.g., Vendryès 1902; Meillet 1937; Darlington 1947; Zipf 1949; Brosnahan 1961; Allott 1994), but, given the incredible causal complexity behind the organization of phonological systems, coupled with the perceived potential racist connotations, substantial proof has been wanting. Recent developments in vocal tract imaging, large scale typological data analysis, and biomechanical modeling suggest that a revisitation of this hypothesis is warranted (e.g., Stavness et al. 2013; Dediu et al. 2017), especially as more data are emerging which show that individual differences in phonetic behavior are linked with anatomical variation (Brunner et al. 2009; Weirich & Fuchs 2011; Stone et al. 2012).

Do any of these individual differences matter at the level of entire speech communities? Indeed, early suggestions for a speech-sound biasing theory were often rejected on the grounds of assumed universal vocal tract structure (e.g., de Saussure 1915: 147) or implicitly that compensatory mechanisms in production and perception of speech overcome any biases found within these systems given that anyone can learn to pronounce any language (e.g., Swadesh 1961; Lord 1966 or Davis 1983: 3). Brosnahan acknowledged humankind’s phonetic plasticity, but he saw this not as invalidating his argument but rather reflecting a “preoccupation” with the individual (Brosnahan 1961: 30). Nevertheless, such proposals have not been inducted into our modern view of sound change as arising from

the perceptually-driven phonologization of synchronic phonetic variation (Ohala 1989), propagated by lexical diffusion (Chen & Wang 1975), and constrained by phonological rules (Kiparsky 2003) and socio-phonetic factors (Aitchison 2001; Foulkes & Docherty 2006), such as language contact. The phonetic variation behind sound change is, however, structured such that some part comes from contextual factors (coarticulation and allophony) and another part from various stochastic effects acting on production and perception. Thus, it is a reasonable hypothesis that part of the phonetic variation present in a speech community reflects the aggregation of effects associated with organic (i.e., speaker specific) properties of voice quality⁴ and its associated influence on segmental and other suprasegmental phenomena.

For those who advocate for anatomico-physiological (and, perhaps, ultimately genetic) biasing of speech sound systems (Brosnahan 1961; Allott 1994; Dediu & Ladd 2007), this type of phonetic variation is viewed as reflecting the normal variation in form and function found in humans within and across populations. Such variations are posited to give rise to variation in the affordances or pressures governing the character or tendencies of speech production and perception, however subtle or profound they may be. A profound effect is readily observed when something goes wrong with vocal tract development or the motor system that results in speech problems, such as, in cases of cleft palate, hypernasality or the use of clicks as substitutions for oral stops (Gibbon et al. 2008). Individuals who have undergone partial laryngectomy need to resort to other means to generate voicing, such as aryepiglottic vibration (Crevier-Buchman et al. 2012). Less profound, but also highly relevant for speech, is the possibility that not everyone can perform a tongue tip trill even after substantial effort to learn to do so. Just as tongue curling “tricks” likely have a genetic basis (Ubbanowski & Wilson 1947), so too could tongue tip trilling (or at least the ease with which this sound is learned).

The potential for individual anatomical variation to influence speech production was acknowledged by Catford (1977: 21–23) citing the work of Brosnahan (1961). He gives some serious consideration to the possibility of population-wise differences in vocal tract anatomy being connected to differences in speech sound systems. Speculations that he notes include the variable presence of the *risorius muscles* leading to differential degrees of the spread lip configuration; variable *tongue length* possibly being related to aspects of Japanese phonology (the Japanese possessing relatively shorter tongues), although the exact details are not specified; and varying presence of certain *epilaryngeal muscles* (*thyroepiglottic inferior* and *thyromembranosus* muscles) possibly relating to increased tendency towards constricted voice qualities in German and Danish (compared against Japanese), and other phenomena (such as Danish *stød*). However, despite the potential importance of anatomical variation for understanding some phonetic variants, and despite the attention that individual phonetic variation has received in numerous studies, anatomical variables are rarely ever measured and analyzed, let alone discussed (some recent examples include Dalcher 2008; Beddor 2009; Lin et al. 2014). This neglect to include anatomical variation in studies of sound change, may be partly due to the stigma associated with such research, which is even condemned in some influential textbooks as irrevocably racist and obviously false (Campbell 1998: 284).

The large intra-group variation in the human form makes it difficult to track down speech biases operating at a level sufficient to have an influence on an entire phonological system. A good starting place would be populations which have had a long history

⁴ Not to be confused with the narrow sense referring to phonatory quality as generated by actions of the larynx, but rather expressing the quasi-permanent, holistic auditory quality arising from sustained muscular adjustment of the vocal apparatus *sensu* Abercrombie (1967: 91, 93) and Laver (1980).

of isolation; in such populations, it is expected that the relative influence of biases will be larger and therefore more readily observable. Australia may be a good example in this regard (Allott 1994: 135) and its phonological landscape is strongly suggestive that biasing effects are at work. Most striking is the “long, thin” (Butcher 2006) nature of the phonological inventory of the typical Australian language. The “long, thin” descriptor refers to the elaboration of place but reduction of manner contrasts within many, if not all (Gasser & Bower 2014), of the consonantal systems found on the continent. Butcher (2006) argues that the deeper causal factor for the “place-of-articulation imperative” (and associated phonetic and phonological properties) of Australian languages arises from a very high incidence of *otitis media* (middle ear infection) in Australian Aboriginal populations, and its possible long-term impact on speech perception. Otitis media, especially in its chronic form, can have long term consequences for hearing. Indeed, according to a recent World Health Organization report (World Health Organization 2004), Aboriginal populations show marked susceptibility to the condition. In Aboriginal communities (for Anangu Pitjantjatjara Yankunytjatjara, or APY lands, see Sánchez et al. 2010; for Elcho Island, Stoakes et al. 2011), the proportion of children who fail pure-tone audiometry testing (uni- or bilaterally) exceeds 60%. The hearing loss pattern in OM is *conductive* (via cholesteatoma, suppuration, tympanic perforations, and/or ossicular chain stiffening; see MacAndie & O’Reilly 1999): generally, frequencies below ≈ 500 Hz and those above ≈ 4000 Hz tend to be subject to considerable attenuation approaching -10 to -30 dB compared with normal hearing (Coates et al. 2002). Furthermore, Clarkson et al. (1989) provide evidence that categorical perception of voice onset time is hampered. Butcher (2006) and Butcher et al. (2012) propose several characteristics of the phonological inventory that reflect the pattern of hearing loss supposed to have prevailed in Aboriginal communities across the continent and extending deep into the past (e.g., first-contact records as early as 1788 exist which indicate Aboriginal children suffered from OM; see Butcher et al. 2012). Low frequency (≤ 500 Hz) attenuation would theoretically impact perceptibility of *obstruent voicing* (F_0 typically ranges well below 500 Hz) and *vowel height* (the higher the vowel, the worse the perceivability of F_1 , which decreases with height). High frequency (≥ 2000 – 4000 Hz) attenuation would affect *fricatives* (which typically have broad spectral regions of high frequency noise) and *noise bursts* (for similar reasons). The middle range, the least affected by the conductive hearing loss, happens to be the frequency zone within which spectral cues to place of articulation carried by formant transitions – particularly of F_2 – reside. The argument thus is that the structure and patterns of Australian Aboriginal phonologies reflect the affordances of an auditory system affected by conductive hearing loss. Avoidance of what are otherwise typologically typical processes, such as regressive nasalization of vowels in VN context, are rendered more likely by an increased pressure to optimize the phonological space around those acoustic cues (F_2 transitions) least impacted by otitis media. This possibly goes as far as putting a pressure on such phonologies to exhibit a strong preference for VC(V) prosodic structure.

“Ok, fine”, one might say, “there might well be all this variation in the anatomy of the vocal tract, but surely it is *compensated* by the speakers themselves, and what might still escape is *normalized* by the listeners (as speaker-specific non-linguistic noise), and anyway what would still survive gets *averaged out* at the speech community level”. This counter argument sounds very, very reasonable, but is it necessarily always true? Indeed, we know that speakers are capable of amazing feats of compensation, ranging from re-learning to speak without a tongue (Gerdeman & Fujimura 1990) or after resection of a significant portion of the larynx (Crevier-Buchman et al. 2012), to the more mundane accommodation to an artificial hard palate (Brunner et al. 2006) or to the normal developmental exploration and exploitation of one’s own anatomy (Brunner et al. 2009). However, this

type of articulatory accommodation and compensation need not be perfect; we propose that it can still have two main types of effects: a *direct acoustic signature* (discussed directly below), and an *indirect influence on coarticulation* (discussed in Section 4).

The first is probably the easiest to understand and study, as it rests on a relatively simple causal mechanism. This postulates that particular anatomical features of the vocal tract (categorical or metric) affect the acoustic properties of (possibly specific aspects of) speech, either directly or mediated by articulatory activity. An example of a direct influence could be the effect of a longer vocal tract on F_1 or of larger vocal folds on F_0 across vowels (Stevens 1998; Zemlin 1998). There are several illustrations of articulatory mediation, probably the best known being the influence of hard palate shape on articulatory variability when producing sounds such as /s/, /ʃ/ (Weirich & Fuchs 2013) and /j/ (Brunner et al. 2005; 2009), but we can also include here a large class of effects that involve variations on the *principle of least effort* (Zipf 1949). More precisely, certain anatomical configurations might reduce or increase the relative effort required to produce a desired acoustic output, for example, through increased muscle effort or more precise motor coordination. In general, such effects are expected to be extremely small, easily compensated by any particular speaker for any particular token, but strong enough to become non-negligible over everyday spoken interactions and potentially amplified by the repeated use and transmission of language (Dediu & Ladd 2007; Dediu et al. 2017; Moisiu & Dediu 2017).

4 The North American English /r/

The second type of effect – the *indirect influence* of vocal tract anatomical variation on coarticulation – may be nicely illustrated by the articulation of *North American English* /r/, where there are several articulatory configurations that result in acoustically similar, but not identical, outputs (Delattre & Freeman 1968; Tiede et al. 2004; Zhou et al. 2008). Two broad categories of configurations are generally recognized. One is with the *tongue tip up* (and often described as “retroflexed”, although true retroflexion, with sublingual stricture, does not always occur in these variants; for example, in the Delattre & Freeman 1968 taxonomy, compare type 7, which is tip up but not truly retroflex, and type 8, which is truly retroflex).⁵ The other is with the tongue tip down and the anterior tongue showing varying degrees of “*bunching*”⁶ (although note that it is possible for the bunched variant to be produced with the tip up, Derrick & Gick 2011). Delattre & Freeman’s (1968) types 5 and 6 respectively represent retracted and fronted variants with the tip elevated and the tongue bunched to the point of producing a primary or even secondary laminal constriction). These might in reality represent attractors in a wider space of possible configurations (Tiede et al. 2004). The distribution of tongue shapes is complex and varies in relation to a number of factors, including segmental (Guenther et al. 1999; Mielke et al. 2010) and prosodic context and individual preference (Boyce et al. 2015), with some speakers adopting (fairly) consistent configurations across contexts. Mielke et al. (2016) provide a comprehensive summary of factors influencing intra-speaker variation between the two broad types, noting that retroflexion is favored adjacent to back/low (vs. front) vowels, labial (vs. dorsal and less so coronal) consonants, prevocalic (vs. postvocalic) position and word-initially (vs. in onset clusters). Acoustically, the retroflex and bunched configurations result in very similar F_1 – F_3 but different F_4 and F_5 patterns (Zhou et al. 2008), but it is unclear if these are perceived by listeners (Twist et al. 2007). If listen-

⁵ Please note that the main difference between these types in the (Delattre & Freeman 1968) taxonomy is the degree of pharyngeal constriction, but we point out the difference in degree of tongue tip retroflexion in the X-ray traces there as a useful point of reference.

⁶ Another name for this “bunched” /r/ variant is “molar” /r/ (Uldall 1958).

ers could hear these differences and control the articulatory configuration of their own /r/ production, then this might result in a “classic” scenario of sound change where, for example, one variant would become dominant and replace the others, or where different variants would become phonetically conditioned allophones, perhaps eventually diverging into different phonemes.

While such overt acoustic signatures of anatomical variation can in principle be heard and reinterpreted as phonetic variation, leading to sound change across repeated episodes of interactions and language acquisition, they might not be the main pathway for the amplification of such weak biases. We suggest here, building on earlier proposals in the same vein by Baker et al. (2011), De Decker & Nycz (2012), and Mielke et al. (2017), a more indirect causal mechanism that does not require direct overt acoustic effects and may therefore amplify covert, articulatorily-mediated anatomical variation. In a nutshell, we suggest that alternative articulatory strategies (partly due to anatomical variation, but not necessarily so) that produce indistinguishable acoustic output⁷ can nevertheless have coarticulatory effects. These may result in overt acoustic effects on *other* segments and be therefore potentially reinterpreted by hearers and amplified by repeated language use and learning. Such *covert at-a-distance effects* make the inference of observed sound changes even more complex, as they break the direct causal connection between articulatory variability, acoustic variation, and its phonologization, interposing articulatory interactions that may depend on particular anatomical contexts.

While the two “canonical” variants of /r/ have very similar acoustic characteristics, and, as noted above, are probably not perceptually differentiated by speakers (Twist et al. 2007), the possibility of a different tongue posture may have consequences for coarticulatory patterns. For instance, Derrick & Gick (2011) observe sub-phonemic kinematic variation in the production of the English flap/tap allophone of /t/ and /d/ that interacts with the lingual posture of neighboring vowels, including the rhotic vowel /ɚ/. They found that speakers employ kinematic variants of the flap/tap allophone that minimize articulatory conflict with contextual vowels. An example is the word ‘saturday’ /sætədeɪ/, which contains two potential flap/tap allophones flanking the /ɚ/; the occurrence of a tip-up variant of the vowel, i.e., [ɪ̥], induces the first stop (/t/) to be an up-flap and the second to be a down-flap (in preparation for the upcoming /e/).

Along similar lines, Baker et al. (2011) investigated North American English s-retraction whereby speakers realize /s/ as a sound closer to [ʃ] (particularly common in the /str/ sequence, as in ‘stretch’, but also occurring word-medially, as in ‘grocery’). Some speakers, “retractors”, consistently employ an allophone that is acoustically very close to [ʃ]; “non-retractors” vary in degree of retraction. Baker et al. (2011) found that, for the non-retractors, variation in /r/ tongue shape seemed to cause variability in the degree of s-retraction, with greater retraction occurring in proportion to how similar the speakers produced their /s/ and /r/. Most of this variation involved bunched /r/ (since retroflex varieties are uncommon in clusters with coronal consonants), thus fine differences in production style can have measurably different coarticulatory effects. Jeff Mielke and colleagues (2017, this special issue), using a large sample of 175 native speakers from Raleigh, NC (Dodsworth & Kohn 2012) supplemented by original ultrasound and (external) video articulatory data from 29 demographically-matched speakers, are investigating whether the type of rhotic variant, acting as covert (i.e., not audible) articulatory variation, can produce differential coarticulatory effects. They hypothesize that people

⁷ Indistinguishability might or might not mean acoustic identity, but represents the situation where a hearer does not interpret the acoustic differences as potentially linguistic; this is probably context-dependent and subtle.

employing bunched /r/ variants will exhibit retraction of preceding /s/ and /z/ and affrication in /tr/ and /dr/ sequences (as [tʃɹ] and [dʒɹ] in ‘train’ and ‘drain’, respectively). In both cases, it is argued that actuation of sound change requires enough variability within coarticulatory processes to produce extreme phonetic variants that become reinterpreted as a different articulatory target (even by those speakers who may not produce the coarticulatory pattern because they lack the necessary phonetic bias).

Irrespective of the effects (direct overt acoustic ones, or indirect covert coarticulatory ones) that this type of variation in the articulation of North American English /r/ has, it is interesting to inquire into its probable causes. While some limited data indicates that even identical twins may adopt different articulations in acquiring the sound (Magloughlin 2016), these articulations seem to also be influenced by palate shape, as suggested by an artificial palate speech-perturbation study (Tiede et al. 2010). Although their participants showed individual variation in the form and onset of their adaptation strategies, the artificial palate, which enlarged the alveolar ridge prominence and uniformly lowered the hard palate, resulted in a change in articulation in most participants. We may imagine that this is probably due to forcing the re-exploration of articulatory space and the discovery of a “secondary strategy”, or, as Tiede et al. (2010) suggest, perhaps through the rediscovery of prior articulatory strategies used in acquisition. Interestingly, the only participant which did not change articulation was the one with the largest vocal tract, for which the standardized prosthetics made the smallest difference, but who did exhibit an increased tendency towards bunching nevertheless.

In this context, our own *ArtiVarK* data might offer new insights into the factors influencing the articulatory configuration used for producing the North American English /r/. The *ArtiVarK* project is currently the largest multi-ethnic dataset designed to explore the relationships between vocal tract anatomy and speech production (see Supplementary file 1 for details). Briefly, our 90 participants (35 female) come from four umbrella ethnic backgrounds (“Chinese”, “North Indian”, “South Indian” and “European”) and have little prior training in phonetics. After a standardized *phonetic training*, we acquired *structural MRI scans*, *intraoral optical scans*, *static MRI scans of sustained articulations* and *real-time MRI scans of speech*. For quantifying the anatomy of the vocal tract structures, we designed a set of *landmarks* and *semi-landmarks*, from which we further derived *classical measures* (essentially angles and distances between landmarks) and *geometric morphometric analyses* that capture more global aspects of shape (Zelditch et al. 2012; Bosman et al. 2017). We extracted 52 variables capturing variation in the “hard” components of the vocal tract. We then used three ways of dealing with the fact that these anatomical variables are not statistically independent, resulting in three sets of variables (for details see Supplementary file 3): *AN* (we removed variables that are strongly inter-correlated, keeping 35 weakly correlated ones), *CL* (we used hierarchical clustering, resulting in 8 clusters), and *PC* (we conducted Principal Component Analysis, resulting in 13 PCs that together explain 91% of the variance). Given that each of these three sets compresses and represents the variance in the original 52 variables in slightly different ways, we ran all analyses separately for each set.

As detailed in Supplementary file 3, for this study we used a subsample of 80 participants from which we acquired static MRI images of the sustained articulation of the North American English /r/ (1 trial per participant), as well as real-time MRI videos of the articulation of the same sound in the isolated /a_a/ context (5 trials) and in the /a_a/ context in the sentence “I said ara for him” (1 trial), resulting in non-missing data for 548 trials. Participants were given systematic phonetic training prior to conducting the MRI. This training explained what it means to sustain a speech sound, providing several examples where different sounds were gradually extended (e.g., [asa], [as:a], [as::a]). In the static

sequence portion of the MRI session, as was the case for all sounds we examined, participants listened to an example (as many times as desired) of the target ([ara] as produced by SRM) prior to scan acquisition. During the scan, we presented a text prompt (“ara”) to remind them of the target. We allowed retakes for participants that failed to sustain the sound correctly (for example, accidentally producing [a::] instead). Participants were instructed (and frequently reminded) to speak slowly and carefully for the real-time scans because of the relatively low frame rate. For each such trial, a trained phonetician (SRM) manually coded the articulations on *four dimensions* based on visual appearance:

- *type* (shortened in the following to *rtype*), with possible values “tip-up” (shortened to “tu”), “bunched” (“bn”), “retroflex” (“rtr”) and “tap/trill” (“tt”);
- *place* (*poa*), with values “dental/alveolar” (“da”) and “palatal” (“p”);⁸
- *dorsum*, with values “concave” (“cv”), “flat” (“f”) and “convex” (“cv”); and
- *lip rounding*, with values “yes” or “no”.

From these, we defined the binary variables: *tip-up/retroflex vs bunched* (coded “yes” = tip-up/retroflex), *bunched vs anything else* (coded “yes” = bunched) and *tip-up/retroflex vs anything else* (coded “yes” = tip-up/retroflex). Figure 1 shows the *rtype* and *lip rounding* per trial within participants, as well as other information related to the participant (*sex*, *phonetic expertise* and *English proficiency*) and the participant’s native *language subfamily* (for details, see below).

The categorization of tongue type (into bunched, tip-up, etc.) can only crudely capture the complexity of tongue shape. Therefore, for the static MRI samples only, we also manually traced the shape of the tongue, including the posterior contour of the mandibular symphysis (shortened to “ms”; where the two halves of the mandible fuse in infancy) for consistency purposes, as can be seen in Figure 2. We then performed a Principal Component Analysis (PCA) on these traces, resulting in 6 PCs that explain together 84% of the variance, as seen in Figures 3–6 (to avoid confusion with other Principal Components described below, we will denote these as *tongue PCs*, *tPC1* to *tPC6*). The PCA reveals the nuanced modes of variation in tongue shape that our speakers used, visualized either as *warps* to the mean shape (Figure 3) or in terms of where the individual participants fall within the *tPC space* (visualized as two components at a time up to the 6th *tPC*, in Figures 4–6). We employ this characterization of continuous variation in tongue shape in our statistical modeling to examine whether anatomical variation has an effect on the production of /r/. The visualizations aid in interpreting the results of these models (linking anatomical variation to production variation). Qualitatively, the *tPCs* seem to capture the following patterns in shape variation. *tPC1* presents the most recognizable differentiation between the canonical “bunched” (red line in Figure 3) and “retroflex” (blue line) variants. However, it should be noted that for *tPC1* and the other *tPCs*, this variation is entangled to a greater or lesser extent with shape variation (mainly of height) of the mandibular symphysis (which was included to provide a better registration among the traces as it is a clearly identifiable structure). *tPC2* could also be said to capture the “bunched” (red line) and “retroflex” (blue line) variants, but it also seems to show a correlation of the tongue as either tall and narrow (in its anteroposterior dimension; red line) or squat and wide (blue line). While the variants in *tPC3* could be classified as tip-up, this *tPC* also shows variation in the contouring of the dorsum, either being rather convex (red line) or rather flat (blue line). We might consider the former a type of bunching with the tip-up,

⁸ These two broad place categories are defined with reference to where the primary oral constriction occurs in relation to the postalveolar zone (posterior slope of the alveolar ridge leading to the roof of the palate). Those anterior of this are grouped under dental/alveolar; those posterior are classified as palatal.

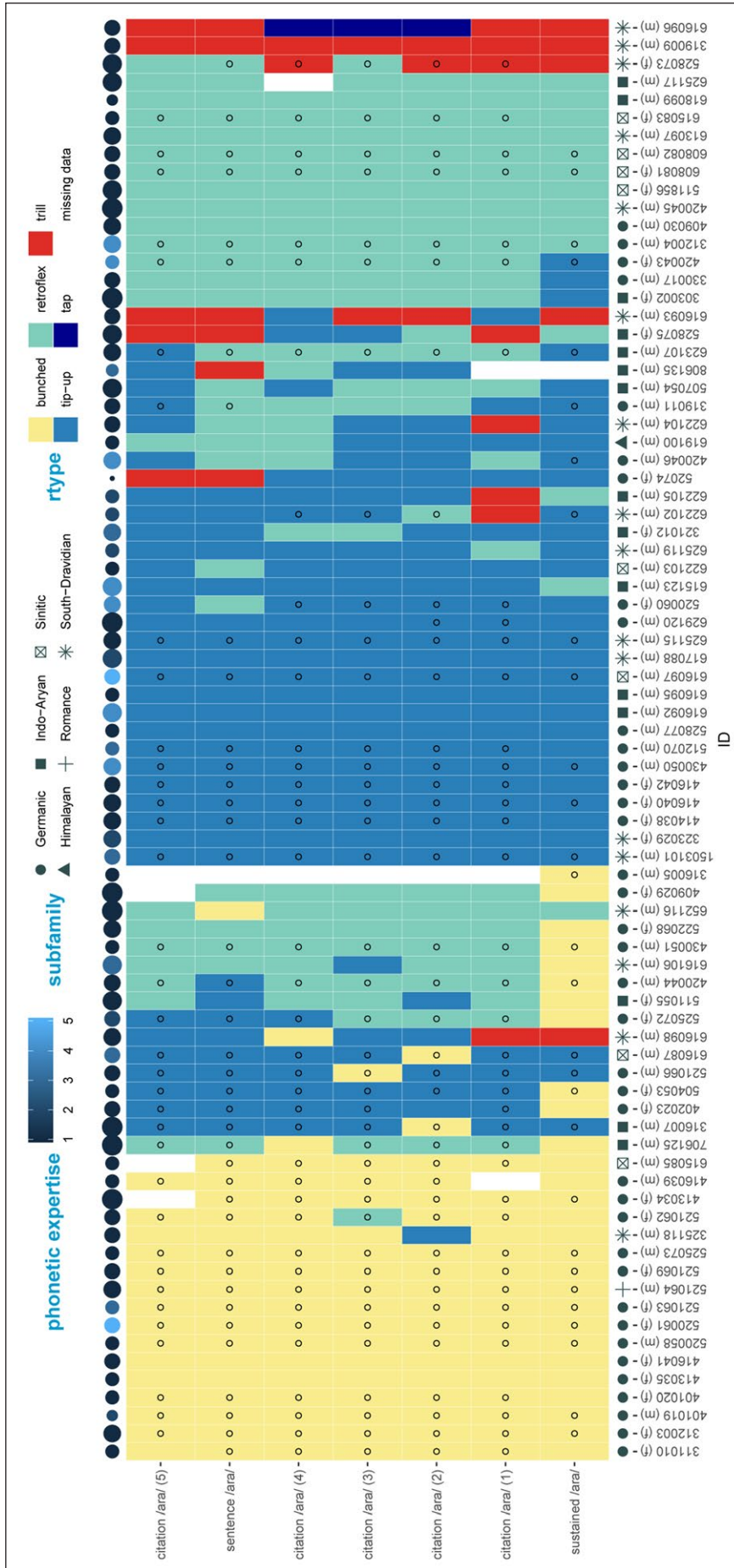


Figure 1: Matrix plot of rtype for each participant (horizontal axis) across trials (vertical axis). The trials are (from bottom to top): sustained articulation of /r/ in /ara/; four isolated /ara/ (trials 1 to 4); /ara/ in the sentence "I said ara for him"; and one more isolated /ara/ (trial 5). The participants are sorted by decreasing number of bunched, then retroflex, and then tip-up responses. Missing data is represented by white cells. Rounding is represented by small black circles. Participant IDs are followed by the participant's sex (in parentheses). The primary participant's language subfamily is represented by the bottom symbols. The top circles represent the participant's English proficiency (circle size: bigger is better) and phonetic expertise (circle color: lighter is better).

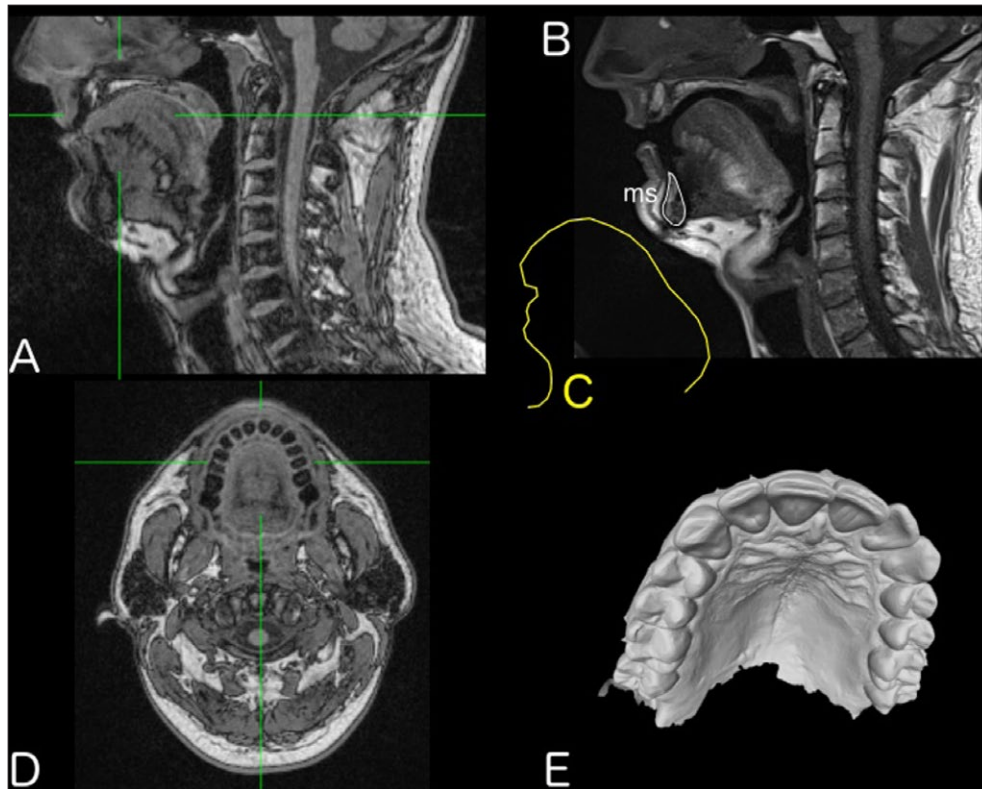


Figure 2: Example of *ArtiVarK* data (author DD, 41 years old male). **Panels A and D** are, respectively, the midsagittal and transverse sections through the 3D T1 structural MRI scan. **Panel B** is the T2 midsagittal static MRI of sustained articulation of North American English /r/ (the mandibular symphysis has been outlined with a thin white line and indicated with “ms”), with **Panel C** showing (in yellow) the manual tracing of the tongue shape. **Panel E** is a projection of the 3D intraoral scan of the hard palate and upper dentition. Images in panels A, B and D were created using Horos v2.4.0, panel E with MeshLab v2016.12, and the whole image assembled with GIMP 2.8.22, on macOS 10.13 High Sierra.

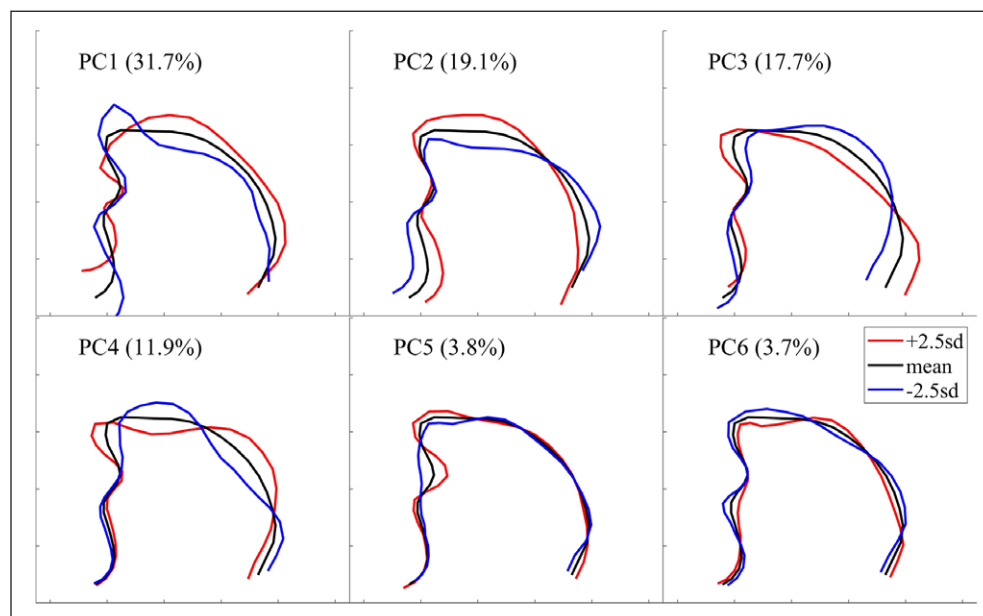


Figure 3: Principal Component Analysis of tongue shape showing the mean shape (black contour) and ± 2.5 standard deviations (red and blue contours, respectively) for the first 6 tongue PCs (*tPCs*).

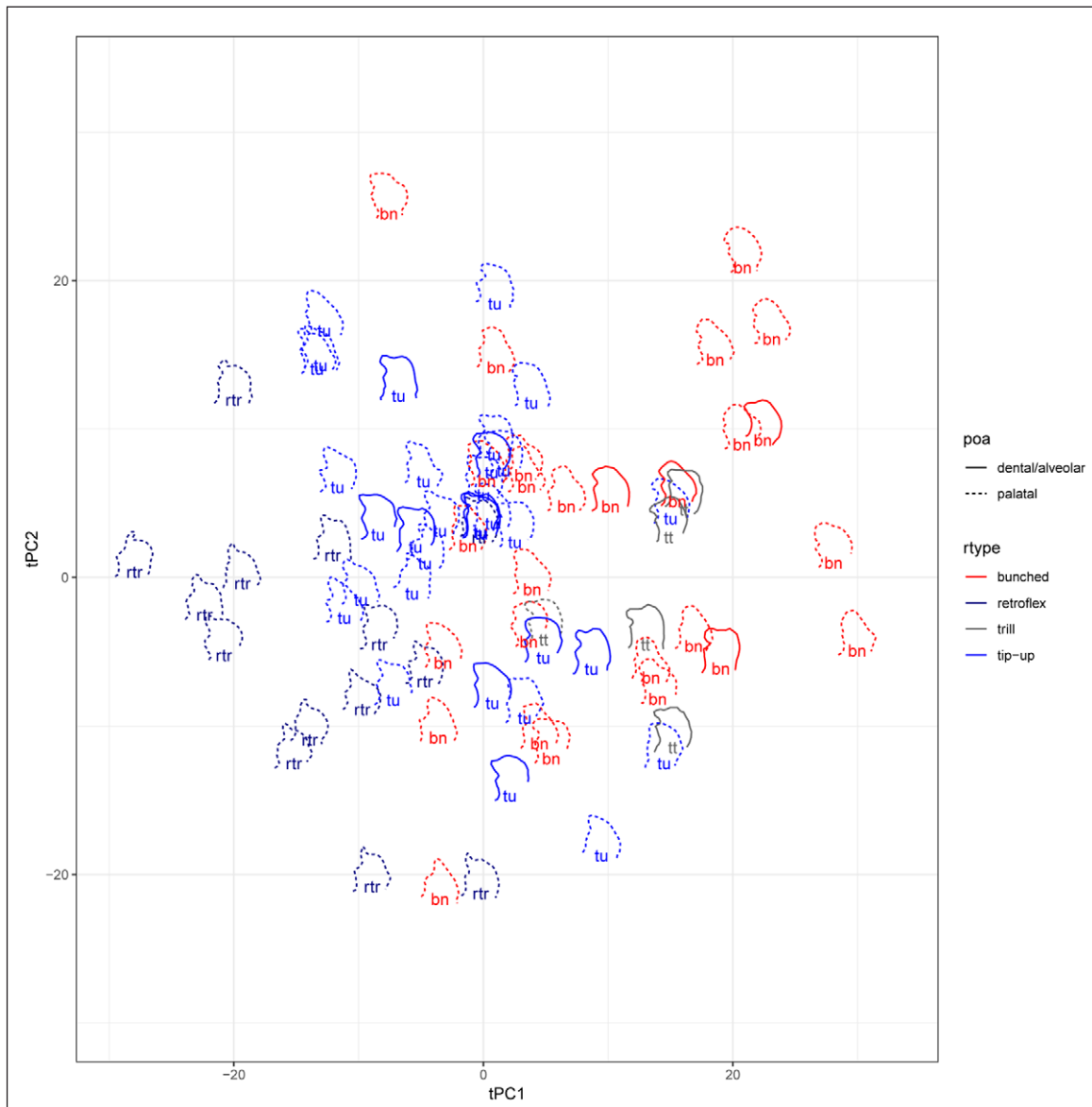


Figure 4: Actual tongue shapes and discrete judgments of type (*rtype*) and place (*poa*) of articulation in the *tPC1* × *tPC2* plane.

while the latter is more adequately described as simply tip-up (with possible pharyngealization). The shape variation captured by *tPC4* opposes an apical stricture (used especially by those with tongue tip trilling; red line; a *t*-test between tongue trill and other types is highly significant: $t(6) = 7.0, p = 3.10^{-4}$) with what could be described as a “double bunched” (blue line) posture (Catford 1983: 349), which is, incidentally, reminiscent of what tends to happen in the production of pharyngealized vowels (for discussion and lingual ultrasound imaging data, see Moisiu 2013, Section 5.4). The double bunched posture is so named because the tongue is bunched up in two areas: one that is directed towards the anterior palate and another that is the bulging of the tongue root into the lower pharyngeal region. There is also a conspicuous volume left in the area of the back dorsum near the posterior velum and uvula. *tPC5* and *tPC6* show finer modes of variation (with “higher spatial frequency”), in addition to accounting for relatively little variation. *tPC5* mainly identifies shaping of the sublingual space, either larger (red line) or smaller (blue line), as the key factor. *tPC6* appears to vary between rather more apical (red line) and laminal (blue line) forms of tip-up stricture, with the tongue appearing quite bunched in all of

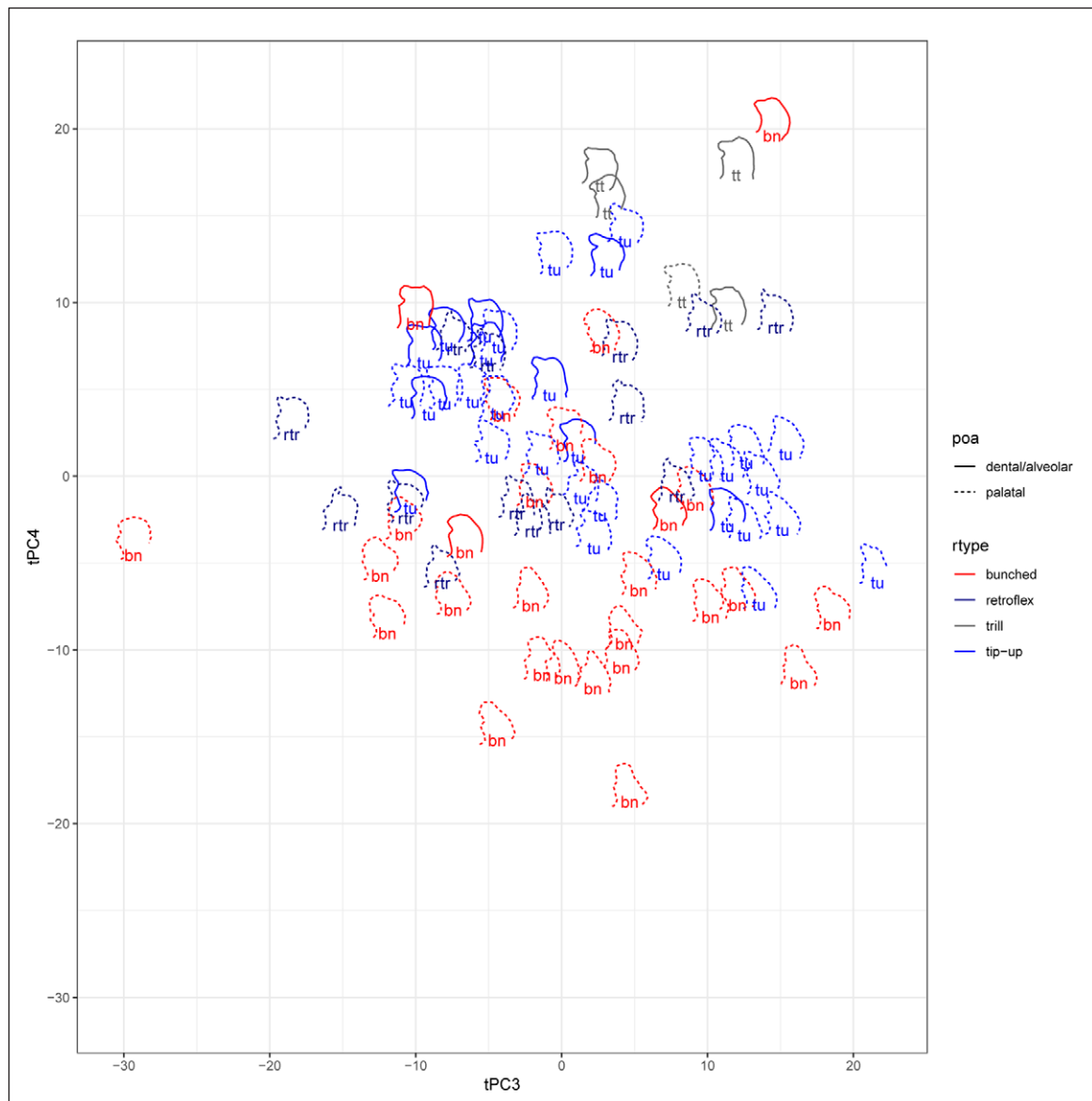


Figure 5: Actual tongue shapes and discrete judgments of type (*rtype*) and place (*poa*) of articulation in the *tPC3* × *tPC4* plane.

these cases. It should be noted that mandibular symphysis variation is largely absent from *tPC4*–*tPC6*. So while it is possible to spot the classic shapes in the PCA, it is clear that the variation is more nuanced than just the simple possibilities of (tip-up) retroflex and (tip-down) bunched.

The articulatory strategies used are very stable within participants across trials: we compared the average number of switches in consecutive trials across participants to 10,000 permutations of the data, and for all variables tested there was not a single permutation that had a lower average number of switches (the observed and permuted averages are 1.07 vs 3.98 for *rtype*, 0.57 vs 1.89 for *poa*, 0.42 vs 2.41 for *dorsum*, 0.35 vs 2.88 for *lip rounding*, and 0.57 vs 2.60 for *tip-up/retroflex* vs *bunched*).

Then, to better understand tongue shape, we performed separate linear regressions of the 6 *tPCs* using, as predictors, the participants' *sex* (male vs female), *age*, *group* (an umbrella ethnic background with four categories: “European”, “Chinese”, “North Indian” and “South Indian”), *phonexp* (self-declared level of formal phonetic training on a Likert scale from “none” to “expert”), self-declared *English proficiency* (Likert scale from “none”

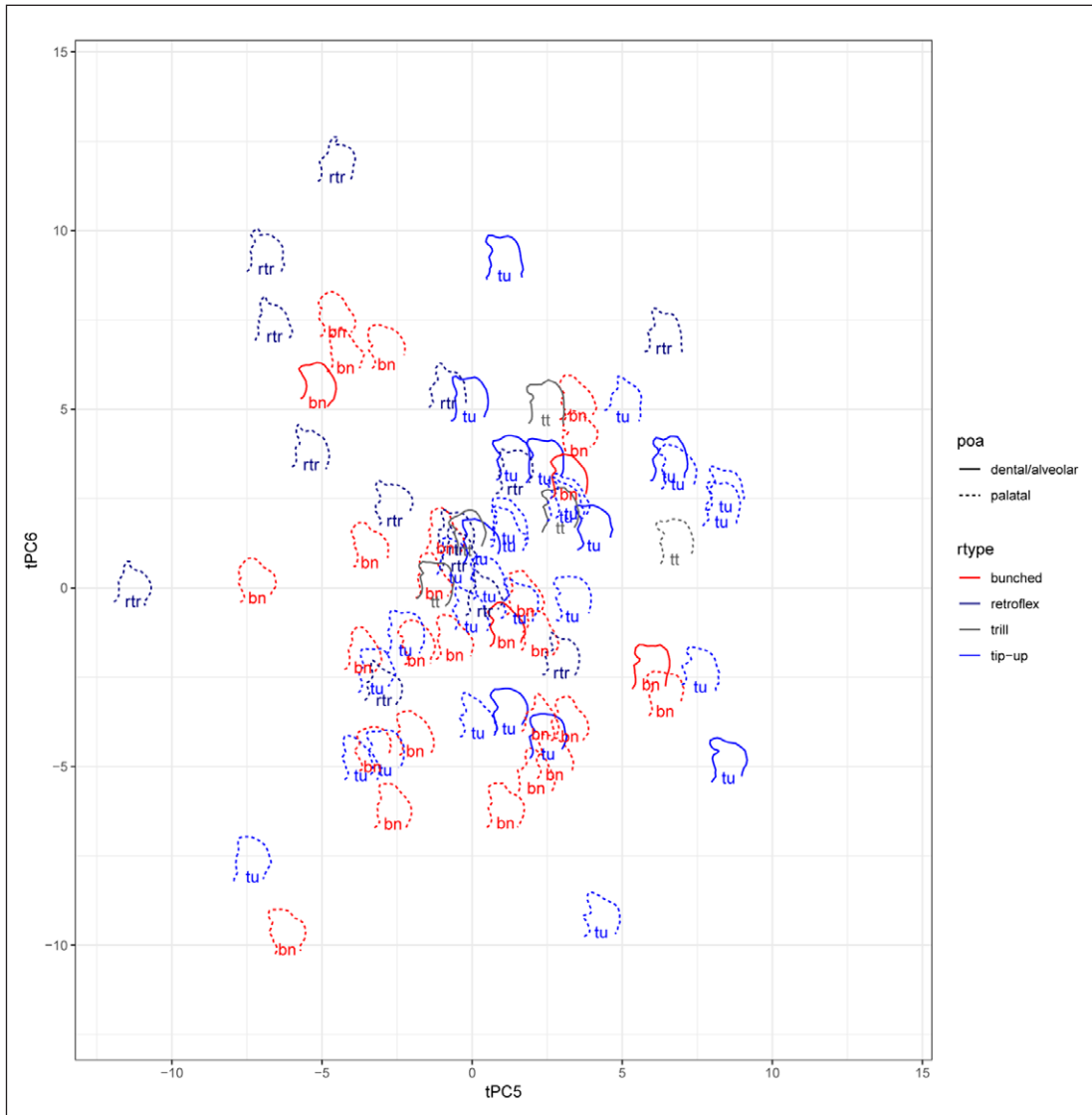


Figure 6: Actual tongue shapes and discrete judgments of type (*rtype*) and place (*poa*) of articulation in the *tPC5* × *tPC6* plane.

to “native speaker”) and a set of binary variables related to the phonetics and phonology of their native language:⁹ does the language have *retroflex stops* /ʈ ɖ/, *alveolar trill* /r/, *retroflex approximant* /ɻ/, *alveolar flap or tap* /ɾ/, *alveolar approximant R* /ɹ/, or *uvular R* (fricative and/or trill) /ʁ R/? We used an α -level of 0.01 throughout and all continuous predictors were z-scored. We simplified the regression models as follows (which we implemented as a fully automatic procedure in R; R Core Team 2017):

- we started with the full model (containing all the relevant predictors),
- followed by the iterative elimination of predictors with a variance inflation factor (VIF) greater than 2,
- possibly followed by an Akaike Information Criterion (AIC)-based model simplification (as implemented by R’s `step()` function), and
- ending with the iterative elimination of predictors with p-values greater than the considered α -level.

⁹ Data on these segment inventories was collected from PHOIBLE (Moran et al. 2014; Dediu & Moisik 2016) and Wikipedia.

Model simplification has advantages, but is also criticized, and should be applied carefully and appropriately to each research question. Here, we are not primarily interested in whether a set of IVs (i.e., *independent variables* or *predictors*) predict a given DV (i.e., the *dependent variable*) – in which case one would rather use powerful machine learning techniques such as SVMs or neural networks – but in the question of which particular IVs significantly improve our understanding of the DV. Therefore, we opted for an unbiased (i.e., purely statistical) model simplification approach which iteratively removes predictors that do not contribute significantly, but which is affected by noise in the data, potentially resulting in slightly different predictors being retained due to minor fluctuations in the data or procedure. Nevertheless, we found that here, these alternatively retained predictors tend to paint similar stories, and we tried to interpret them in rather general terms and to avoid capitalizing on details that might not be robust to noise. The same worries, of unduly capitalizing on noise, coupled with the relatively low statistical power of our sample (despite its size), prompted us to perform the various analyses using the multiple methods that we present here, but focusing on their mutual agreements and commonalities.

The full results are shown in Table 1, and, in summary, they suggest that indeed, tongue shape is affected by anatomy, especially *tPC2* (positively by a higher hard palate, and a larger, “V”-shaped mouth with overjet) and *tPC4* (positively by a longer and wider mouth, with a bigger alveolar ridge¹⁰ and angled lower incisors). We can suggest that a higher palate dome (as captured by *m2height*) would require a higher tongue position, as positive values of *tPC2* indicate. Overjet (anterior-posterior distance between the maxillary and mandibular incisors) favoring positive *tPC2* shapes may reflect the less anterior positioning of the mandibular symphysis (and hence the lower incisors) compared to negative values of this *tPC*. Overall scale (reflected inversely in *PC1*) and maxillary dental arch shape (with low values of *C.P2w* having a more “V”-shaped arch) may impact how the participant exploits lingual-palatal bracing to stabilize the articulation (see Gick et al. 2017). Larger overall dimensions of the mouth seem consistent with the positive shape indicated for *tPC2*, but it is not exactly clear why an anteriorly narrower dental arch would favor such a posture. Acoustic factors may also be relevant, but would require a modeling study to discern. *tPC3* seems mostly affected (negatively) by the presence of the alveolar trill /r/ in the participant’s native language. Positive values of this *tPC* capture the shape associated with trilling, but note that positive values of *tPC4* also capture some of these cases of trilling (however, this variable was not a significant predictor for that *tPC*). *tPC1*, *tPC5* and *tPC6* do not seem reliably influenced by any of the variables in our dataset. Interestingly, *sex* only seems to affect *tPC4*, *age* *tPC2*, with *group* and most *language-related characteristics* having no significant effects.

¹⁰ The fact that *aRBulge* appears as positively predictive in the *tPC4* formula here may be partially consistent with the results of the palate perturbation experiment in Tiede et al. (2010). A major feature of the artificial palate in that study was a considerable enlargement to the alveolar ridge prominence. Participants who used at least some retroflexion in the unperturbed/BASE setting (F1 and M1) showed even more retroflex variants when their palates were perturbed. Participant F2, who did not have any retroflex variants in the unperturbed condition, used them in the perturbed /ara/ context. The only participant who did not switch to retroflexion, but rather showed an increase in bunching, was also noted to have the largest vocal tract and hence was perhaps least susceptible to the influence of the palate perturbation. Returning to our own study, positive values of *tPC4* are associated with the tip-up/retroflex posture (in addition to, at the extreme positive values, the tongue tip trill, but such a production is considered a failed attempt), thus, the more bulgy the alveolar ridge was judged to be, the more likely the participant would use retroflexion. So, despite the highly variable findings reported in Tiede et al. (2010), there is some congruence between that study and our own. Of course, this result requires replication and further testing to confirm that there is indeed a true effect of the alveolar ridge on tendency towards tip-up/retroflex variants as the presence of this variable in the regression equation is possibly an artifact of the model simplification procedure.

Table 1: The simplified linear regression equations of the tongue shape PCs ($tPCn$) on all predictors, separately for each of the three sets of anatomical variables (AN , CL and PC). We show only those cases that did not fully simplify to the null model (thus that retain more than just the intercept). The numbers in parentheses represent the p -values. Sex represents the effect of being male versus being female. $Adj. R^2$ is the adjusted percent of explained variance. The anatomical predictors are (with meaning in parentheses; followed by what a larger value represents): *overjet* (overjet; greater anterior-posterior distance between the maxillary and mandibular incisors), *m2Height* (height at the second molars; taller hard palate), *C.P2w* (ratio of inter-canine to inter-second premolar widths; more “U”-shaped than “V”-shaped maxillary dental arch), *lowIA* (lower incisor angle; more vertical lower incisors), *C.P2l* (as *C.P2w* for length; longer anterior mouth), *aRBulge* (alveolar ridge bulge subjective judgment on a 1–5 Likert scale; bigger alveolar ridge), *C44* (width of the mouth; narrower mouth), *PC1* (overall dimensions of the mouth; higher values mean a smaller mouth), *PC2* (front of the lower mouth; less overbite and more sublingual space), *PC3* (width of the mouth; narrow mouth), *PC8* (low incisors; less vertical) and *PC13* (teeth and alveolar ridge). For example, the entry in row 2 represents the simplified regression model of $tPC2$ on the covariates and all the clusters CL , which simplifies to $tPC2 = 0.36 * age$, where the intercept α is not significantly different from 0, the slope of age is (with $p = 0.001$), and the $adjusted R^2 = 11.9\%$. Directly comparing the slopes and p -values across models and methods is difficult, but we expect consistent estimates of the slopes (absolute values of similar magnitudes and especially with the same sign) for those predictors that are significant.

TPC	SET	REGRESSION EQUATION (WITH P-VALUES)	ADJ. R ²
$tPC2$	AN	$0.33*overjet (0.002) + 0.32*m2Height (0.002) - 0.30*C.P2w (0.005)$	22.3%
	CL	$0.36*age (0.001)$	11.9%
	PC	$0.35*age (0.001) - 0.28*PC1 (0.007)$	18.9%
$tPC3$	$AN \& CL$	$-1.35*trill (0.003)$	9.7%
	PC	$0.30*PC2 (0.004) - 1.48*trill (0.001)$	17.9%
$tPC4$	AN	$0.32 - 0.85*sex (0.000) - 0.27*lowIA (0.003) + 0.35*C.P2l (0.000) + 0.36*aRBulge (0.000)$	42.9%
	CL	$0.28 - 0.73*sex (0.001) - 0.30*C44 (0.005)$	26.6%
	PC	$-0.26*PC1 (0.007) - 0.25*PC3 (0.010) - 0.34*PC8 (0.001) - 0.29*PC13 (0.003)$	31.5%
$tPC6$	PC	$-0.30*PC5 (0.005) + 0.31*PC10 (0.004)$	15.2%

Finally, for each of the binary variables *bunched vs. tip-up/retroflex* (denoted as $b:r$), *bunched vs. everything else* ($b:*$), *tip-up/retroflex vs. everything else* ($r:*$), and *lip rounding* (lr), we conducted logistic regressions with model simplification. For the *real-time MRI* data, we performed multi-level logistic regressions (with the *participants* as random effect), as well as “flat” regressions (by collapsing each participant’s observations to her/his most frequent value). For the *sustained articulation* (i.e., static) MRI data, we only performed “flat” logistic regressions as there is a single observation per participant. In all cases, we used the three sets of anatomical variables (AN , CL and PC), the covariates (sex , $group$, $phonexp$ and age), the properties of the native language (*Retroflexes*, *R.trill*, *R.retroflex*, *R.flap.tap*, *R.approximant* & *R.uvular*), and *English proficiency* as IVs; to these, for the “static” case only, we also added the *tongue shape PCs* ($tPC1 - tPC6$). In all cases we used both *Maximum Likelihood* and *Markov Chain Monte Carlo* estimation (Supplementary file 3). The results are in Tables 2–4 (for full details please see the Supplementary file 3)¹¹ Due to the very high within-participant consistency, the Intraclass correlations coefficients

¹¹ Here, we focus on the *agreement* between methods as we are relatively underpowered, there is probably coding noise in the data, and there might be a less-than-perfect fit with the methods’ assumptions. Nevertheless, as each method has its own “strengths” and “weaknesses”, a robust finding across them has a lower probability of being a false positive or a method artifact.

Table 2: Simplified logistic regressions for the sustained articulations (i.e., the static MRI data). DVs: *b:r* = *bunched* vs. *tip-up/retroflex* (1 = tip-up/retroflex), *b:** = *bunched* vs. *everything else* (1 = bunched), *r:** = *tip-up/retroflex* vs. *everything else* (1 = tip-up/retroflex), and *lr* = *lip rounding* (1 = lip rounding is present). *Set* = which one of the three IV sets is used. *Meth* = method (Maximum Likelihood or MCMC – shortened to MC). The numbers in parentheses represent the *p*-values. *IC* is AIC for ML and DIC for MCMC. Please note that for MCMC the numerical estimates might differ between runs. The predictors are: *tPC1* = canonical “bunched” vs “retroflex”; *tPC4* = apical stricture vs “double bunched”; *tPC6* = more apical vs laminal forms of tip-up stricture with tongue bunching; *jLowPC2* = larger values reflect a narrower and more anterior lower jaw; *sex* = difference between males and females; *R.uvular* = does the participants’ native language use uvular R? *R.flap.tap* = does the participants’ native language use flap or tap? *PC10* = width and slope of the anterior palate roof. The conventions are as in Table 1.

DV	SET	METH	REGRESSION EQUATION (WITH P-VALUES)	IC
<i>b:r</i>	AN	ML	-0.60 + 0.79*jLowPC2 (0.0058)	93.1
		MC	-0.98 + 2.28*tPC1 (<6·10 ⁻⁵) - 0.99*tPC6 (0.0079)	62.0
	CL	ML	-0.82 - 1.38*tPC4 (0.00034)	83.8
		MC	-0.82 + 2.12*tPC1 (<6·10 ⁻⁵) - 1.48*tPC4 (<6·10 ⁻⁵)	54.0
	PC	MC	-0.98 + 2.28*tPC1 (<6·10 ⁻⁵) - 0.98*tPC6 (0.0092)	62.0
	<i>b:*</i>	AN	ML	-0.70 + 0.85*jLowPC2 (0.0037)
MC			-2.05 + 1.80*sex (0.0086) + 2.05*tPC1 (<6·10 ⁻⁵) - 1.27*tPC6 (0.00078)	66.0
CL		ML	-0.70 + 0.91*C44 (0.0033)	95.7
		MC	-0.96 + 2.02*tPC1 (<6·10 ⁻⁵) - 1.68*tPC4 (<6·10 ⁻⁵)	56.0
PC		ML	-1.36 + 1.36*R.uvular (0.00757)	99.0
		MC	-1.24 + 1.99*tPC1 (<6·10 ⁻⁵) - 1.06*tPC6 (0.00322)	72.0
<i>r:*</i>	AN	MC	0.53 - 2.14*tPC1 (<6·10 ⁻⁵)	72.0
	CL	MC	0.53 - 2.14*tPC1 (<6·10 ⁻⁵)	72.0
	PC	MC	0.53 - 2.14*tPC1 (<6·10 ⁻⁵)	72.0
<i>lr</i>	AN	MC	-0.21 - 1.66*R.flap.tap (0.0082)	99.0
		CL	ML	-0.15 - 1.51*R.flap.tap (0.013)
		MC	-0.21 - 1.65*R.flap.tap (0.0093)	99.0
	PC	ML	-0.15 - 1.51*R.flap.tap (0.013)	100.5
		MC	-0.71 - 0.78*PC10 (0.0079)	99.0

(ICCs) are extremely high (> 98%; Table 3), strongly suggesting that a multi-level logistic regression approach may not be appropriate (in fact, most of these regressions either do not find any significant predictor or suffer from convergence issues; Table 3). Please note that future studies are planned to look quantitatively at the actual articulatory variation in the real-time MRI data in a way comparable to what was done for the static MRI data, which we presented above. For now, we must content ourselves with this much more simplistic characterization of the articulatory variation (based on subjective visual phonetic categorization), but this should be kept in mind when interpreting the results. The “flat” logistic regressions of the real-time MRI data (Table 4) found that participants speaking languages using a flap/tap tend to produce less “bunching” and lip rounding, but also that *C44* (representing the width of the mouth, with larger values standing for a narrower palate and mouth) has a positive influence on “bunching”. The static productions (Table 2) are influenced by tongue shape (as quantified by the *tPC* variables), with “bunching” being reflected by *tPC1* (positively), *tPC4* (negatively) and *tPC6* (negatively), but there are also hints of an effect of anatomy, with positive influences of *jLowPC2* (larger values reflect a narrower and more anterior lower jaw) and *C44*. Interestingly, there is also a

Table 3: Simplified mixed-effects logistic regressions for the real-time articulations (i.e., the participant is the random effect). *ICC* = Intraclass Correlation (how similar are responses within participants); the other notations as in Table 2. Empty rows represent cases where no predictor was significant after simplification or where converge errors occurred. *PC1* = overall dimensions of the mouth with higher values meaning a smaller mouth.

DV	ICC	SET	METH	REGRESSION EQUATION (WITH P-VALUES)	IC
<i>b:r</i>	99.2%			[No significant IVs]	
<i>b:*</i>	99.2%			[No significant IVs]	
<i>r:*</i>	98.2%			[No significant IVs]	
<i>lr</i>	99.5%	CL	ML	$10.4 - 21.9 * R.flap.tap$ ($8.6 \cdot 10^{-14}$)	182.0
	99.5%	PC	ML	$7.47 * PC1$ (0.0013) - $7.97 * PC10$ (0.00081)	220.0

Table 4: Simplified “flat” logistic regressions for the real-time articulations (collapsing a participant to her/his most frequent response). Conventions as in Tables 2 and 3. *Retroflexes* = does the participants’ native language use retroflex sounds.

DV	SET	METH	REGRESSION EQUATION (WITH P-VALUES)	IC
<i>b:r</i>	AN	ML	$-0.66 - 2.67 * Retroflexes$ (0.012)	73.0
	CL	ML	$-0.84 - 2.25 * R.flap.tap$ (0.0344)	77.2
		MC	$-1.65 + 0.88 * C44$ (0.0077)	76.0
	PC	ML	$-0.84 - 2.25 * R.flap.tap$ (0.0344)	77.2
		MC	$-1.05 - 2.11 * R.flap.tap$ (0.0098)	76.0
<i>b:*</i>	AN	MC	$-1.06 - 2.23 * R.flap.tap$ (0.0064)	76.0
	CL	ML	$-0.84 - 2.38 * R.flap.tap$ (0.0251)	77.4
		MC	$-1.70 + 0.93 * C44$ (0.0058)	77.0
	PC	ML	$-0.84 - 2.38 * R.flap.tap$ (0.0251)	77.4
		MC	$-1.06 - 2.25 * R.flap.tap$ (0.0041)	76.0
<i>r:*</i>			[No significant IVs]	
<i>lr</i>	AN	MC	$0.65 - 1.99 * R.flap.tap$ (0.00067)	100.0
	CL	ML	$0.58 - 1.79 * R.flap.tap$ (0.0011)	101.3
		MC	$0.65 - 2.00 * R.flap.tap$ (0.00078)	100.0
	PC	ML	$0.58 - 1.79 * R.flap.tap$ (0.0011)	101.3
		MC	$0.66 - 2.00 * R.flap.tap$ (0.00067)	100.0

hint of a negative influence of *PC10* (width and slope of the anterior palate roof) on lip rounding.

The static productions (Table 2) are influenced by tongue shape (as quantified by the *tPC* variables), with “bunching” being reflected by *tPC1* (positively), *tPC4* and *tPC6* (both negatively), but there are also hints of an effect of anatomy, with positive influences of *jLowPC2* (larger values reflect a narrower and more anterior lower jaw) and *C44* (representing the width of the mouth, with larger values standing for a narrower palate and mouth). Interestingly, there is also a hint of a negative influence of *PC10* (representing the width and slope of the anterior palate roof) on lip rounding. The “flat” logistic regressions of the real-time MRI data (Table 4) found that participants speaking languages using a flap/tap tend to produce less “bunching” and lip rounding, but also that *C44* has a positive influence on “bunching”. Thus, the anatomical influences found tend to be consistent between methods, but stronger for sustained articulation than for real-time; this could

be due to the better capacity to suppress native language motor patterns during the first compared to the second condition.¹²

We tentatively suggest that this influence of anatomy may relate to bracing (as pointed out above). Because the sides of the tongue need to be elevated to produce the bunched configuration, it might benefit more in terms of biomechanical stability than the retroflex configuration from a narrow mouth (and hence narrow dental arch).¹³ Retroflex and non-bunched tip-up configurations may instead rely more on pharyngeal wall bracing (which we see more of in back and especially low vowels; Gauffin & Sundberg 1978). The preference of the retroflex variant in low back vowel contexts (see Mielke et al. 2016: 1106) probably reflects this (and retroflex sounds in general share much in common with the articulation of back, especially non-high vowels; see Hamann 2003). Additionally, there once again could be acoustic reasons for the preference, although these are not immediately clear. This interpretation is confounded by the problem that, because muscle and bone are so tightly coupled in growth and development (DiGirolamo et al. 2013), the tongue itself should reflect the dimensions of the skeletal structures (wider mouth, wider tongue). Individuals settle on their production strategies during childhood, when the structures are still growing, although the greatest changes are observed in the first 18 months of life (Vorperian et al. 2005), after which point dental arch dimensions (mouth width) are relatively stable (Tsujino & Machida 1998). Continuous growth, however, would pose additional problems to establishing motor control, possibly leading to different solutions, even in (presumably) the same vocal tract (as in Magloughlin 2016). Most characteristics of the native language do not seem to strongly affect the articulation of /r/ in our data, with the exception of the presence of a flap/tap /ɾ/; likewise, *English proficiency*, *phonetic expertise*, *sex*, *group* and *age* have no significant influences, suggesting that our results might truly capture relatively weak effects of vocal tract anatomy (and compensatory strategies) on articulation.

Coupled with the results of the artificial alteration of the hard palate (Tiede et al. 2010), our findings from the *ArtiVarK* sample suggest that, besides linguistic explanations, *anatomical aspects of the anterior vocal tract may influence the articulation of the North American English /r/*, in particular *hard palate width and height*, *the overall size of the mouth*, and *the size (i.e., prominence) of the alveolar ridge*. However, it is important to point out several limitations of our study, including the fact that the vast majority of our participants were not native English speakers (and among those that were, most spoke varieties of British English) and that while our standardized training was based entirely on acoustic stimuli (recorded by John Esling, former president of the International Phonetic Association from 2011 to 2015) without any visual or articulatory feedback, it took place on a background of quite high proficiency in English as a second language (excluding the native speakers, on a scale from 0 to 10, the minimum is 5, maximum is 10, and mean and median 8). This is reflected in the general impression (from SRM, who is a native speaker of Canadian English) that the /r/s produced sounded quite natural (excepting cases where they were trilled or tapped). Part of this may reflect the fact that many of our participants are Dutch speakers who possess the alveolar approximant allophone of /r/ in coda position (e.g., see Scobbie et al. 2006). A small set of our participants are native speakers of Mandarin, which possesses a retroflex /ɻ/, and this might contribute to the impression of good accuracy of the productions. We have not made any attempt, however, to objectively characterize just how “native-like” the /r/s are judged to be by native speakers. Moreover,

¹² We wish to thank an anonymous reviewer for this suggestion.

¹³ Jeff Mielke (personal communication) points out that studies may have overlooked this relationship because they are restricted to articulatory information available in the midsagittal plane.

despite its relatively large sample size, our *ArtiVarK* study may suffer from low statistical power to reliably identify the small effect sizes we expect such anatomical biases to have; also, we want to highlight that the actual predictors retained after model simplification should not be taken too literally but rather as representatives of quite general anatomical features. Nevertheless, the fact that demographic characteristics, group of origin, amount of formal phonetic training and most native language characteristics do not contribute significantly, suggests that these results are not artifacts. If, indeed, these influences of vocal tract anatomy on the articulatory strategies used to produce the North American English /r/ survive replication in samples of native speakers of the concerned dialects of English, they might provide a fascinating example of (at least in part) anatomically-grounded covert articulatory variation with coarticulatory effects, potentially affecting language change.

5 Initiation and actuation of sound change in biased populations

We argued here that variation in the anatomy of the vocal tract might matter for sound change, by either producing overt, direct acoustic effects which the hearers may perceive and reinterpret as innovations, or by covertly changing other speech sounds, producing indirect effects that may be heard and reinterpreted. Anatomy is but one dimension on which people differ, and our choice to focus on it is mostly pragmatic (Dediu et al. 2017), but these “first steps” should offer a methodological springboard for the investigation of biases affecting other levels of language, such as morpho-syntax, semantics and pragmatics. Therefore, we can generalize to *abstract biases* that *probabilistically* affect certain aspects of language, and we need not assume that this bias has a genetic basis, but only that it is relatively stable within individuals. For simplicity, we consider a binary bias, that is either present or absent for a given individual (or that varies continuously in strength and direction among individuals). If this bias is shared by all “normal” humans, its effects are (probabilistically) universal cross-linguistically, but if it varies between individuals, it may lead to (probabilistic) patterns of linguistic diversity. If the frequency of biased individuals is relatively low and similar across populations, we would expect its effects to not lead to sound changes, but to be perceived as individual idiosyncrasies with varying degrees of associated social stigma or prestige. However, there may be circumstances where its effects are amplified, leading to sound change, and, if these circumstances vary between populations and/or across time, patterns of linguistic diversity may emerge.

A growing literature explores the conditions under which a universally-shared *weak bias* may be amplified by the repeated use and transmission of language, and the general findings are that this amplification process is very *sensitive* to the strength, type and effects of the bias, and that it is *non-linear* (for example, see Kirby et al. 2007; 2008; Dediu 2008; 2009; Thompson et al. 2016; Janssen & Dediu 2018). However, there is relatively little work on the amplification of biases that are not universally shared (Dediu 2008; 2009), and we must study the crucial importance of the *communication network* in which the individuals (biased or not) act. We must therefore consider not only the *topology*, *structural properties* and *connectivity patterns* of these networks (Newman 2010; Aggarwal 2011; Kadushin 2012), but also the *individual properties* of their nodes. For example, a few biased nodes placed in the hubs of a communication network might be able to “nudge” language towards change, while many more such nodes at the periphery may not have any effect on the network-wide language; the status of the nodes may amplify (or dampen) the effects of the nodes’ intrinsic biases; and keeping the same nodes (with their biases, status, and approximate connectivity properties) but changing the topology of the network should affect the spread and amplification of the bias.

This view of language change, potentially triggered by biases that vary between individuals within communicative networks, may open a new way of conceptualizing not only the initiation of sound change, but also its spread (or not) throughout the community. Because such biases are not uniformly distributed and their amplification depends non-linearly on the frequency and network properties of the biased individuals, in some cases a bias-initiated sound change might spread to fixation, in others it might stall midway (or even reverse), while in others it may completely fail to go beyond the stage of individual idiosyncrasy. However, in our view, such biases do not emerge suddenly in a single individual (a “hopeful monster” or “Prometheus”; a “mutation”), but are present at various frequencies in a population (“*standing variation*” in population genetics) for various reasons (“drift”, “founder effects”, “selective pressures”, “byproducts”, etc.) that may (but usually do not) have anything to do with their linguistic effects. Thus, part of the population is “*poised*” for change, in the sense that there already exist individuals that manifest the bias to a certain degree, but their biases simply are not yet amplified enough to expand beyond idiosyncrasies. This “almost there” or “poised” state of affairs may persist effectively hidden in the speech community for a while, but when conditions change, it may catastrophically coalesce into a network-wide state change. We might metaphorically think of the speech community as a system close to a *phase transition*, which can be triggered by minor changes in the system, such as the addition of a nucleation center to a supercooled liquid. If our suggestion is on the right track, we may predict that there are communicative networks with heterogeneous nodes where the addition of just a few edges or biased nodes will push the network towards a self-reinforcing positive feedback loop, where the biased nodes suddenly form an “echo chamber” that not only amplifies their biases, but also exposes the rest of the network to this new variant enough to make its uptake by the whole community possible¹⁴ (see similar ideas in, for example, Kirby & Sonderegger 2015).

6 Conclusions

In this paper we briefly discussed the origins, patterning and significance of inter-individual and inter-group differences in the perception and production of speech and language, and in particular biases that are rooted in the anatomy of the vocal tract. We explored the articulatory strategies that are used to produce the North American English /r/ in our own *ArtiVarK* sample, showing that the two “canonical” strategies (“retroflex” vs. “bunched”) are attractors on a multi-dimensional continuum, and seem to be influenced by individual anatomical characteristics of the anterior vocal tract. Moreover, based on previous work, we suggest that such biases may affect sound change either through “direct” acoustic signatures that can be reinterpreted by the hearers, or through “indirect” coarticulatory effects “at a distance”. While there is still some resistance to such proposals, we argue that they open the way to a full understanding of language as a complex phenomenon at the intersection of culture and biology. In many ways, our ideas are less radical than they might seem, being just a refinement of established paradigms, highlighting that, on a universally shared background, particular individuals and groups might create their own constraints and affordances.¹⁵ We propose that to properly understand the features of sound change, we must seriously consider that individuals are intrinsically variable,

¹⁴ Arguably, such phenomena might govern, for example, the runaway polarization of social media (Zollo et al. 2015; Bessi et al. 2016).

¹⁵ As an anecdote, when one of the authors (DD) was asked, after a talk, a question about these views, somebody else in the audience replied along the lines of “this is really not that new, but ideas that have been around in laboratory phonology for ages – it’s just that they are applied to smaller groups than the whole human species!”

and that such varying biases must be viewed as acting within dynamic and structured communicative networks with heterogeneous nodes. This opens the possibility of seeing linguistic communities as constantly poised towards sound changes, and that relatively minor changes in the frequency of biased individuals, in their position within the network, or to network structure, might trigger rapid phase changes that may or may not reach completion, depending on further changes and feedback loops within the network. This view may allow us to unify the initiation and actuation of sound change as resulting from a background of pre-existing (but invisible) standing variation of potential sound changes which may be triggered by multiple factors.

However, these are just ideas and very preliminary results that should be followed by better data collection designs in larger samples, more motivated case studies and more informed hypotheses about the dimensions of vocal tract variation that might affect sound change and their coarticulatory effects. We should also move beyond anatomy and consider variation in the physiology and motor control of articulation, on our way towards biases that are more cognitive in nature. Finally, we must build more realistic models of language use in heterogeneous communicative networks, and study the network-level changes induced by relatively small alterations to network structure, dynamics, and the frequency and network location of the biased nodes.

Additional Files

The additional files for this article can be found as follows:

- **Supplementary file 1.** Description of the *ArtiVarK* sample. DOI: <https://doi.org/10.5281/zenodo.1480427>
- **Supplementary file 2.** The *data* (as TAB-separated .TSV files and .RData standard R data files) and Rmarkdown *script* needed to reproduce the results reported here as a single .TAR.XZ archive. Please note that the actual results obtained when compiling this script may differ slightly from those reported in the paper, especially for Markov-Chain Monte Carlo (due to the heavy use of random numbers), but also for the Maximum Likelihood (due to differences between particular BLAS/LAPACK implementations). The first compilation of this script is quite expensive computationally (on an Intel Core i7-3770 3.40 GHz with 32 Gb RAM this took about 25 hours), but the next ones are much faster as the most expensive results are cached locally. DOI: <https://doi.org/10.5281/zenodo.1481941>
- **Supplementary file 3.** Full description of the *sample, analysis and results* (HTML report obtained by compiling the Rmarkdown script and data contained in Supplementary file 2, compressed with XZ). DOI: <https://doi.org/10.5281/zenodo.1481943>

Ethics and Consent

The *ArtiVarK* study is covered by the ethics approval 45659.091.14 (1 June 2015), Donders Center for Brain, Cognition and Behaviour, Nijmegen.

Acknowledgements

We wish to thank our *ArtiVarK* participants; David Norris and Paul Gaalman for access to and piloting on the Avanto MRI scanner; Thomas Maal, Frans Delfos and Cees Kreulen for access to and help with the TRIOS intraoral scanner; Carly Jaques for participant recruitment and management; John Esling for recording the phonetic training materials; Sabine Kooijman for assistance with ethics; Alexandra Dima for advice on statistics. Special thanks to the organizers of the 4th edition of the Workshop for Sound Change (2007)

in Edinburgh (who also edited this Special Issue), and to one anonymous reviewer and to Jeff Mielke for invaluable feedback on earlier drafts.

Funding Information

This work was funded by the Netherlands Organisation for Scientific Research (NWO) VIDI grant 276-70-022 to DD. During the writing of this paper, DD was funded by an European Institutes for Advanced Study (EURIAS) Fellowship Program, and an IDEX Lyon Fellowship from the Université de Lyon, France.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

Conceptualization: DD, SRM. Formal Analysis: DD, SRM. Funding Acquisition: DD. Investigation: DD, SRM. Methodology: DD, SRM. Resources: DD, SRM. Software: DD, SRM. Visualization: DD, SRM. Writing: DD, SRM.

References

- Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Agardh, Emilie, Peter Allebeck, Johan Hallqvist, Tahereh Moradi & Anna Sidorchuk. 2011. Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. *International Journal of Epidemiology* 40(3). 804–818. DOI: <https://doi.org/10.1093/ije/dyr029>
- Aggarwal, Charu C. 2011. Social network data analytics. Springer Science & Business Media. DOI: <https://doi.org/10.1007/978-1-4419-8462-3>
- Aitchison, Jean. 2001. *Language change: Progress or decay?* Cambridge University Press.
- Allott, Robin. 1994. Motor theory of language: The diversity of languages. In Jan Wind, Abraham Jonker, Robin Allott & Leonard Rolfe (eds.), *Studies in Language Origins* 3. 125–160. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- 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. DOI: <https://doi.org/10.1017/S0954394511000135>
- Barbujani, Guido & Vincenza Colonna. 2010. Human genome diversity: Frequently asked questions. *Trends in Genetics* 26(7). 285–295. DOI: <https://doi.org/10.1016/j.tig.2010.04.002>
- Beddor, Patrice Speeter. 2009. A coarticulatory path to sound change. *Language* 85(4). 785–821. DOI: <https://doi.org/10.1353/lan.0.0165>
- Benowitz, Neal L. 2009. Pharmacology of nicotine: Addiction, smoking-induced disease, and therapeutics. *Annual Review of Pharmacology and Toxicology* 49(1). 57–71. DOI: <https://doi.org/10.1146/annurev.pharmtox.48.113006.094742>
- Bessi, Alessandro, Fabiana Zollo, Michela Del Vicario, Michelangelo Puliga, Antonio Scala, Guido Caldarelli, Brian Uzzi & Walter Quattrociocchi. 2016. Users polarization on facebook and youtube. *PLoS One* 11(8). e0159641. DOI: <https://doi.org/10.1371/journal.pone.0159641>
- Betti, Lia, François Balloux, William Amos, Tsunehiko Hanihara & Andrea Manica. 2009. Distance from Africa, not climate, explains within-population phenotypic diversity in humans. *Proceedings of the Royal Society B Biological Sciences* 276(1658). 809–814. DOI: <https://doi.org/10.1098/rspb.2008.1563>

- Bosman, Abel M., Scott R. Moisiu, Dan Dediu & Andrea Waters-Rist. 2017. Talking heads: Morphological variation in the human mandible over the last 500 years in the Netherlands. *HOMO – Journal of Comparative Human Biology* 68(5). 329–342. DOI: <https://doi.org/10.1016/j.jchb.2017.08.002>
- Bowden, Rory, Tammie S. MacFie, Simon Myers, Garrett Hellenthal, Eric Nerrienet, Ronald E. Bontrop, Colin Freeman, Peter Donnelly & Nicholas I. Mundy. 2012. Genomic tools for evolution and conservation in the chimpanzee: *Pan troglodytes ellioti* is a genetically distinct population. *PLoS Genetics* 8(3). e1002504. DOI: <https://doi.org/10.1371/journal.pgen.1002504>
- Boyce, Suzanne, Mark Tiede, Carol Espy-Wilson & Kathy Groves-Wright. 2015. Diversity of tongue shapes for the American English rhotic liquid. In The Scottish Consortium for ICPHS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Paper no. 1041. Glasgow, UK: The University of Glasgow. <https://www.international-phoneticassociation.org/icphsproceedings/ICPHS2015/Papers/ICPHS0847.pdf>.
- Brook, Alan H., Jukka Jernvall, Richard N. Smith, Toby E. Hughes & Grant C. Townsend. 2014. The dentition: the outcomes of morphogenesis leading to variations of tooth number, size and shape. *Australian Dental Journal* 59. 131–142. DOI: <https://doi.org/10.1111/adj.12160>
- Brosnahan, Leonard Francis. 1961. *The sounds of language: An inquiry into the role of genetic factors in the development of sound systems*. Cambridge: W. Heffer and Sons.
- Brunner, Jana, Phil Hoole, Pascal Perrier & Susanne Fuchs. 2006. Temporal development of compensation strategies for perturbed palate shape in German/sch/-production. In *Proceedings of the 7th International Seminar on Speech Production*, 247–254. <http://halshs.archives-ouvertes.fr/hal-00403289/>.
- Brunner, Jana, Susanne Fuchs & Pascal Perrier. 2005. The influence of the palate shape on articulatory token-to-token variability. *ZAS Papers in Linguistics* 42. 43–67. http://www.zas.gwz-berlin.de/fileadmin/material/ZASPiL_Volltexte/zp42/zaspil42-fuchs-perrier-brunner.pdf.
- Brunner, Jana, Susanne Fuchs & Pascal Perrier. 2009. On the relationship between palate shape and articulatory behavior. *The Journal of the Acoustical Society of America* 125(6). 3936–3949. DOI: <https://doi.org/10.1121/1.3125313>
- Bush, Jeffrey O. & Rulang Jiang. 2012. Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development. *Development* 139(2). 231–243. DOI: <https://doi.org/10.1242/dev.067082>
- Butcher, Andrew. 2006. Australian aboriginal languages: Consonant salient phonologies and the ‘place-of-articulation imperative’. In John Harrington & Marija Tabain (eds.), *Speech production: Models, phonetic processes, and techniques* (Macquarie Monographs in Cognitive Science), 187–210. Psychology Press: NY. https://www.researchgate.net/publication/251509897_Australian_Aboriginal_languages_Consonant_salient_phonologies_and_the_place-of-articulation_imperative.
- Butcher, Andrew, Janet Fletcher, Hywel Stoakes & Marija Tabain. 2012. Australian phonologies and aboriginal hearing: Is there a link? <https://www.jcu.edu.au/the-cairns-institute/publications-of-the-cairnsinstitute/e-lectures/australian-phonologies-and-aboriginal-hearing>.
- Campbell, Lyle. 1998. *Historical linguistics: An introduction*. Cambridge, Mass: MIT Press.
- Catford, John C. 1977. *Fundamental problems in phonetics*. London: Indiana University Press.
- Catford, John C. 1983. Pharyngeal and laryngeal sounds in Caucasian languages. In Diane Bless & James Abbs (eds.), *Vocal fold physiology: Contemporary research and clinical issues*, 344–350. San Diego: College Hill Press.

- Chen, Matthew Y. & William S.-Y. Wang. 1975. Sound change: Actuation and implementation. *Language* 51(2). 255–281. DOI: <https://doi.org/10.2307/412854>
- Clarkson, Richard L., Peter D. Eimas & G. Cameron Marean. 1989. Speech perception in children with histories of recurrent otitis media. *The Journal of the Acoustical Society of America* 85(2). 926–933. DOI: <https://doi.org/10.1121/1.397989>
- Coates, Harvey L., Peter S. Morris, Amanda J. Leach & Sophie Couzos. 2002. Otitis media in Aboriginal children: tackling a major health problem. *The Medical Journal of Australia* 177(4). 177–178. <https://www.mja.com.au/journal/2002/177/4/otitis-mediaaboriginal-children-tackling-major-health-problem>.
- Cobourne, Martyn T. & Paul T. Sharpe. 2013. Diseases of the tooth: the genetic and molecular basis of inherited anomalies affecting the dentition. *Wiley Interdisciplinary Reviews: Developmental Biology* 2(2). 183–212. DOI: <https://doi.org/10.1002/wdev.66>
- Couraud, Sébastien, Gérard Zalcman, Bernard Milleron, Franck Morin & Pierre-Jean Souquet. 2012. Lung cancer in never smokers – a review. *European Journal of Cancer* 48(9). 1299–1311. DOI: <https://doi.org/10.1016/j.ejca.2012.03.007>
- Cramon-Taubadel, Noreen von & Stephen J. Lycett. 2008. Human cranial variation fits iterative founder effect model with African origin. *American Journal of Physical Anthropology* 136(1). 108–113. DOI: <https://doi.org/10.1002/ajpa.20775>
- Crevier-Buchman, Lise, Claire Pillot-Loiseau, Annie Rialland, Narantuya, Coralie Vincent & Alain Desjacques. 2012. Analogy between laryngeal gesture in Mongolian Long Song and supracricoid partial laryngectomy. *Clinical Linguistics & Phonetics* 26(1). 86–99. DOI: <https://doi.org/10.3109/02699206.2011.590920>
- Dalcher, Christina Villafaña. 2008. Consonant weakening in Florentine Italian: A cross-disciplinary approach to gradient and variable sound change. *Language Variation and Change* 20(2). 275–316. DOI: <https://doi.org/10.1017/S0954394508000021>
- Darlington, Cyril Dean. 1947. The genetic component of language. *Heredity* 1. 269–286. DOI: <https://doi.org/10.1038/175178b0>
- Davis, Lawrence M. 1983. *English dialectology*. Alabama: Alabama University Press.
- De Decker, Paul M. & Jennifer R. Nycz. 2012. Are tense [æ]s really tense? The mapping between articulation and acoustics. *Lingua* 122(7). 810–821. DOI: <https://doi.org/10.1016/j.lingua.2012.01.003>
- de Saussure, Ferdinand. 1915. *Course in general linguistics*. New York: McGraw-Hill.
- Dediu, Dan. 2008. The role of genetic biases in shaping language-genes correlations. *Journal of Theoretical Biology* 254. 400–407. DOI: <https://doi.org/10.1016/j.jtbi.2008.05.028>
- Dediu, Dan. 2009. Genetic biasing through cultural transmission: do simple Bayesian models of language evolution generalize? *Journal of Theoretical Biology* 259(3). 552–561. DOI: <https://doi.org/10.1016/j.jtbi.2009.04.004>
- Dediu, Dan & D. Robert Ladd. 2007. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences of the United States of America* 104(26). 10944–10949. DOI: <https://doi.org/10.1073/pnas.0610848104>
- Dediu, Dan, Rick Janssen & Scott R. Moisiik. 2017. Language is not isolated from its wider environment: vocal tract influences on the evolution of speech and language. *Language and Communication* 54. 9–20. DOI: <https://doi.org/10.1016/j.lang-com.2016.10.002>
- Dediu, Dan & Scott R. Moisiik. 2016. Defining and counting phonological classes in cross-linguistic segment databases. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Paris, France: European

- Language Resources Association (ELRA). http://www.lrecconf.org/proceedings/lrec2016/pdf/151_Paper.pdf.
- Delattre, Pierre & Donald C. Freeman. 1968. A dialect study of American R's by X-ray motion picture. *Linguistics* 6(44). 29–68. DOI: <https://doi.org/10.1515/ling.1968.6.44.29>
- Derrick, Donald & Bryan Gick. 2011. Individual variation in English flaps and taps: A case of categorical phonetics. *The Canadian Journal of Linguistics/La revue canadienne de linguistique* 56(3). 307–319. DOI: <https://doi.org/10.1353/cjl.2011.0024>
- DiGirolamo, Douglas J., Douglas P. Kiel & Karyn A. Esser. 2013. Bone and skeletal muscle: neighbors with close ties. *Journal of Bone and Mineral Research* 28(7). 1509–1518. DOI: <https://doi.org/10.1002/jbmr.1969>
- Dixon, Michael J., Mary L. Marazita, Terri H. Beaty & Jeffrey C. Murray. 2011. Cleft lip and palate: synthesizing genetic and environmental influences. *Nature reviews. Genetics* 12(3). 167–178. DOI: <https://doi.org/10.1038/nrg2933>
- Dodsworth, Robin & Mary Kohn. 2012. Urban rejection of the vernacular: The SVS undone. *Language Variation and Change* 24(2). 221–245. DOI: <https://doi.org/10.1017/S0954394512000105>
- Durand, Claudia & Gudrun A. Rappold. 2013. Height matters – from monogenic disorders to normal variation. *Nature Reviews Endocrinology* 9(3). 171–177. DOI: <https://doi.org/10.1038/nrendo.2012.251>
- Ferdinand, Keith C. & Samar A. Nasser. 2015. Racial/ethnic disparities in prevalence and care of patients with type 2 diabetes mellitus. *Current Medical Research and Opinion* 31(5). 913–923. DOI: <https://doi.org/10.1185/03007995.2015.1029894>
- Foulkes, Paul & Gerard Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34(4). 409–438. DOI: <https://doi.org/10.1016/j.wocn.2005.08.002>
- Furr, R. Michael & Verne R. Bacharach. 2007. *Psychometrics: An introduction*. Sage Publications, Inc. 1st edn.
- Gasser, Emily & Claire Bowern. 2014. Revisiting phonotactic generalizations in Australian languages. In *Proceedings of the annual meetings on phonology* 1. DOI: <https://doi.org/10.3765/amp.v1i1.17>
- Gauffin, Jan & Johan Sundberg. 1978. Pharyngeal constrictions. *Phonetica* 35(3). 157–168. DOI: <https://doi.org/10.1159/000259927>
- Gerdeman, Bernice & Osamu Fujimura. 1990. Speaking without a tongue. *The Journal of the Acoustical Society of America* 87(S1). S89–S89. DOI: <https://doi.org/10.1121/1.2028415>
- Gibbon, Fiona, Alice Lee, Ivan Yuen & Lisa Crampin. 2008. Clicks produced as compensatory articulations in two adolescents with velocardiofacial syndrome. *The Cleft Palate-Craniofacial Journal* 45(4). 381–392. DOI: <https://doi.org/10.1597/06-232.1>
- Gick, Bryan, Blake Allen, François Roewer-Després & Ian Stavness. 2017. Speaking tongues are actively braced. *Journal of Speech, Language, and Hearing Research* 60(3). 494–506. DOI: https://doi.org/10.1044/2016_JSLHR-S-15-0141
- Guenther, Frank H., Fatima T. Husain, Michael A. Cohen & Barbara G. Shinn-Cunningham. 1999. Effects of categorization and discrimination training on auditory perceptual space. *The Journal of the Acoustical Society of America* 106(5). 2900–2912. DOI: <https://doi.org/10.1121/1.428112>
- Hamann, Silke. 2003. *The phonetics and phonology of retroflexes*. Utrecht, The Netherlands: LOT dissertation. <https://dspace.library.uu.nl/handle/1874/627>.
- Hong, Paul. 2013. Ankyloglossia (tongue-tie). *CMAJ: Canadian Medical Association Journal* 185(2). E128. DOI: <https://doi.org/10.1503/cmaj.120785>
- Janssen, Rick & Dan Dediu. 2018. Genetic biases in language: computer models and experimental approaches. In Thierry Poibeau & Aline Villavicencio (eds.), *Language, cognition, and computational models* (Studies in Natural Language Processing),

- 256–288. Cambridge, UK: Cambridge University Press 1st edn. DOI: <https://doi.org/10.1017/9781316676974.010>
- Jernvall, Jukka & Irma Thesleff. 2012. Tooth shape formation and tooth renewal: evolving with the same signals. *Development* 139(19). 3487–3497. DOI: <https://doi.org/10.1242/dev.085084>
- Johnston, Douglas R., Karen Watters, Lynne R. Ferrari & Reza Rahbar. 2014. Laryngeal cleft: Evaluation and management. *International Journal of Pediatric Otorhinolaryngology* 78(6). 905–911. DOI: <https://doi.org/10.1016/j.ijporl.2014.03.015>
- Kadushin, Charles. 2012. *Understanding social networks: Theories, concepts, and findings*. Oxford University Press, USA.
- Kiparsky, Paul. 2003. The phonological basis of sound change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 311–342. Blackwell Publishing Ltd. DOI: <https://doi.org/10.1002/9780470756393.ch6>
- Kirby, James & Morgan Sonderegger. 2015. Bias and population structure in the actuation of sound change. arXiv: 1507.04420. <http://arxiv.org/abs/1507.04420>.
- Kirby, Simon, Hannah Cornish & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America* 105(31). 10681–10686. DOI: <https://doi.org/10.1073/pnas.0707835105>
- Kirby, Simon, Mike Dowman & Thomas L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of the United States of America* 104(12). 5241–5245. DOI: <https://doi.org/10.1073/pnas.0608222104>
- Klein, Ophir D., Snehlata Oberoi, Ann Huysseune, Maria Hovorakova, Miroslav Peterka & Renata Peterkova. 2013. Developmental disorders of the dentition: An update. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 163(4). 318–332. DOI: <https://doi.org/10.1002/ajmg.c.31382>
- Kumar, D. Kalyan & K. Saraswathi Gopal. 2011. Morphological variants of soft palate in normal individuals: A digital cephalometric study. *Journal of Clinical and Diagnostic Research* 5(6). 1310–1313. http://www.jcdr.net/article_fulltext.asp?issn=0973-709x&year=2011&month=November&volume=5&issue=6&page=1310-1314&id=1646.
- Lammert, Adam, Michael Proctor & Shrikanth Narayanan. 2013. Morphological variation in the adult hard palate and posterior pharyngeal wall. *Journal of Speech, Language and Hearing Research* 56. 521–530. DOI: [https://doi.org/10.1044/1092-4388\(2012/12-0059\)](https://doi.org/10.1044/1092-4388(2012/12-0059))
- Laver, John. 1980. *The phonetic description of voice quality*. Cambridge University Press.
- Lewontin, Richard C. 1972. The apportionment of human diversity. In Theodosius Dobzhansky, Max K. Hecht & William C. Steere (eds.), *Evolutionary biology* 6. 381–398. New York: Springer. https://link.springer.com/chapter/10.1007/978-1-4684-9063-3_14.
- Lin, Susan, Patrice Speeter Beddor & Andries W. Coetzee. 2014. Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/. *Laboratory Phonology* 5(1). 9. DOI: <https://doi.org/10.1515/lp-2014-0002>
- Lippert-Rasmussen, Kasper. 2014. *Born free and equal?: A philosophical inquiry into the nature of discrimination*. Oxford, UK: Oxford University Press.
- Lord, Robert. 1966. *Comparative linguistics*. London, UK: English Universities Press.
- Lumb, Alistair. 2014. Diabetes and exercise. *Clinical Medicine* 14(6). 673–676. DOI: <https://doi.org/10.7861/clinmedicine.14-6-673>
- MacAndie, Christine & Brian F. O'Reilly. 1999. Sensorineural hearing loss in chronic otitis media. *Clinical Otolaryngology and Allied Sciences* 24(3). 220–222. DOI: <https://doi.org/10.1046/j.1365-2273.1999.00237.x>

- Magloughlin, Lyra V. 2016. Accounting for variability in North American English/ɪ: Evidence from children's articulation. *Journal of Phonetics* 54. 51-67. DOI: <https://doi.org/10.1016/j.wocn.2015.07.007>
- Mead, Nicola & Peter Bower. 2000. Patient-centredness: a conceptual framework and review of the empirical literature. *Social Science & Medicine* 51(7). 1087–1110. DOI: [https://doi.org/10.1016/S0277-9536\(00\)00098-8](https://doi.org/10.1016/S0277-9536(00)00098-8)
- Meillet, Antoine. 1937. *Introduction à l'étude comparative des langues Indoeuropéens*. Buck, Alabama: University of Alabama Press 8th edn.
- Mielke, Jeff, Adam Baker & Diana Archangeli. 2010. Variability and homogeneity in American English /r/ allophony and /s/ retraction. *Laboratory phonology* 10. 699–730. DOI: <https://doi.org/10.1017/S0954394511000135>
- Mielke, Jeff, Adam Baker & Diana Archangeli. 2016. Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language* 92(1). 101–140. DOI: <https://doi.org/10.1353/lan.2016.0019>
- Mielke, Jeff, Bridget Smith & Michael J. Fox. 2017. Implications of covert articulatory variation for several phonetic variables in Raleigh, North Carolina English. *The Journal of the Acoustical Society of America* 141(5). 3981–3981. DOI: <https://doi.org/10.1121/1.4989092>
- Moisiu, Scott R. 2013. *The Epilarynx in Speech*. Victoria, British Columbia, Canada: University of Victoria PhD Thesis. <https://dspace.library.uvic.ca:8443/handle/1828/4690>.
- Moisiu, Scott R. & Dan Dediu. 2017. Anatomical biasing and clicks: Evidence from bio-mechanical modeling. *Journal of Language Evolution* 2(1). 37–51. DOI: <https://doi.org/10.1093/jole/lzx004>
- Moran, Steven, Daniel McCloy & Richard Wright (eds.). 2014. *PHOIBLE online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://phoible.org>.
- Nassir, Rami, Roman Kosoy, Chao Tian, Phoebe A. White, Lesley M. Butler, Gabriel Silva, Rick Kittles, Marta E. Alarcon-Riquelme, Peter K. Gregersen, John W. Belmont, Francisco M. De La Vega & Michael F. Seldin. 2009. An ancestry informative marker set for determining continental origin: validation and extension using human genome diversity panels. *BMC Genetics* 10. 39. DOI: <https://doi.org/10.1186/1471-2156-10-39>
- Newman, Mark. 2010. *Networks: An introduction*. Oxford, UK: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199206650.001.0001>
- Novembre, John, Toby Johnson, Katarzyna Bryc, Zoltán Kutalik, Adam R. Boyko, Adam Auton, Amit Indap, Karen S. King, Sven Bergmann, Matthew R. Nelson, Matthew Stephens & Carlos D. Bustamante. 2008. Genes mirror geography within Europe. *Nature* 456. 98–101. DOI: <https://doi.org/10.1038/nature07331>
- Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In Leiv E. Breivik & Ernst Håkon Jahr (eds.), *Language change: Contributions to the study of its causes*, 173–198. Berlin: Mouton de Gruyter. http://www.linguistics.berkeley.edu/~ohala/papers/pool_of_synchronic_var.pdf.
- Ounap, Katrin. 2016. Silver-Russell syndrome and Beckwith-Wiedemann syndrome: Opposite phenotypes with heterogeneous molecular etiology. *Molecular Syndromology* 7(3). 110–121. DOI: <https://doi.org/10.1159/000447413>
- Paschou, Peristera, Jamey Lewis, Asif Javed & Petros Drineas. 2010. Ancestry informative markers for fine-scale individual assignment to worldwide populations. *Journal of Medical Genetics* 47. 835–847. DOI: <https://doi.org/10.1136/jmg.2010.078212>
- Pickrell, Joseph K., Nick Patterson, Chiara Barbieri, Falko Berthold, Linda Gerlach, Tom Güldemann, Blesswell Kure, Sununguko Wata Mpoloka, Hiroshi Nakagawa, Christfried Naumann, Mark Lipson, Po-Ru Loh, Joseph Lachance, Joanna Mountain, Carlos D.

- Bustamante, Bonnie Berger, Sarah A. Tishkoff, Brenna M. Henn, Mark Stoneking, David Reich & Brigitte Pakendorf. 2012. The genetic prehistory of southern Africa. *Nature Communications* 3. 1143. DOI: <https://doi.org/10.1038/ncomms2140>
- Praveen, B. N., Amrutesh Sunitha, Pal Sumona, A. R. Shubhasini & Vaseemuddin Syed. 2011. Various shapes of soft palate: A lateral cephalometric study. *World Journal of Dentistry* 2. 207–210. DOI: <https://doi.org/10.5005/jp-journals-10015-1084>
- Rattansi, Ali. 2007. *Racism: A very short introduction*. Oxford, UK: Oxford University Press.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>.
- Reich, David. 2018. *Who we are and how we got here: Ancient DNA and the new science of the human past*. Oxford, UK: Oxford University Press. DOI: <https://doi.org/10.1093/actrade/9780192805904.001.0001>
- Reynoso, M. C., Albert Hernández, L. A. Lizcano-Gil, Aurelio Sarralde, Manuel C. Abreu, Zamira Nazará & Ruben Fragoso. 1994. Autosomal dominant congenital macroglossia: further delineation of the syndrome. *Genetic counseling* 5(2). 151–154.
- Ridgeway, Cecilia L. 2011. *Framed by gender: How gender inequality persists in the modern world*. Oxford, UK: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199755776.001.0001>
- Riquelme, Armindo & Larry J. Green. 1970. Palatal width, height, and length in human twins. *The Angle orthodontist* 40(2). 71–79.
- Safdar, Saba & Natasza Kosakowska-Berezecka. 2015. *Psychology of gender through the lens of culture: Theories and applications*. Switzerland: Springer. DOI: <https://doi.org/10.1007/978-3-319-14005-6>
- Sánchez, Linnett, D. Turner, Karen Sparrow, Sandra Buxton, Sarosh Kapadia, Sheridan Flint, Brenna Eckert, A. Howard, Quenten Iskov, N. Loades & Claire Loades. 2010. Prevalence of ear health and hearing problems in remote and indigenous school-age children. In *Audiology Australia XIX national conference*. Sydney, Australia.
- Schneider, David J. 2005. *The psychology of stereotyping*. New York: Guilford Press.
- Scobbie, James M., K. Sebregts & Jane Stuart-Smith. 2006. From subtle to gross variation: an ultrasound tongue imaging study of Dutch and Scottish English /r/. In *Labphon conference 10*. Paris, France. <https://slideplayer.com/slide/8240517/>.
- Setó-Salvia, Núria & Philip Stanier. 2014. Genetics of cleft lip and/or cleft palate: Association with other common anomalies. *European Journal of Medical Genetics* 57(8). 381–393. DOI: <https://doi.org/10.1016/j.ejmg.2014.04.003>
- Solé, Maria-Josep & Daniel Recasens i Vives. 2012. *The initiation of sound change: Perception, production, and social factors*. Amsterdam & Philadelphia: John Benjamins Publishing. DOI: <https://doi.org/10.1075/cilt.323>
- Stavness, Ian, Mohammad Ali Nazari, Pascal Perrier, Didier Demolin & Yohan Payan. 2013. A biomechanical modeling study of the effects of the orbicularis oris muscle and jaw posture on lip shape. *Journal of Speech, Language, and Hearing Research* 56(3). 878–890. DOI: [https://doi.org/10.1044/1092-4388\(2012/12-0200\)](https://doi.org/10.1044/1092-4388(2012/12-0200))
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, Mass: MIT Press.
- Stevens, Mary & Jonathan Harrington. 2014. The individual and the actuation of sound change. *Loquens* 1(1). 3. DOI: <https://doi.org/10.3989/loquens.2014.003>
- Stoakes, Hywel, Andrew Butcher, Janet Fletcher & Marija Tabain. 2011. Long term average speech spectra in Yolngu Matha and Pitjantjatjara speaking females and males. In *INTERSPEECH 2011*. Florence, Italy. <https://pdfs.semanticscholar.org/00b0/1139fcd636f3dcd2abcb415460bb419b1f6b.pdf>.

- Stone, Maureen, Susan Rizk, Jonghye Woo, Emi Z. Murano, Hegang Chen & Jerry L. Prince. 2012. Frequency of Apical and Laminal /s/ in Normal and Postglossectomy Patients. *Journal of Medical Speech-Language Pathology* 20(4). 19. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4492454/>.
- Swadesh, Morris. 1961. The sounds of language. An inquiry into the role of genetic factors in the development of sound systems. L. F. Brosnahan. Hefner, Cambridge, England, 1961. 250 pp. *Science* 134(3479). 609–609. DOI: <https://doi.org/10.1126/science.134.3479.609>
- The 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature* 526(7571). 68–74. DOI: <https://doi.org/10.1038/nature15393>
- Thompson, Bill, Simon Kirby & Kenny Smith. 2016. Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences* 113(16). 4530–4535. DOI: <https://doi.org/10.1073/pnas.1523631113>
- Tiede, Mark K., Suzanne E. Boyce, Carol Y. Espy-Wilson & Vincent L. Gracco. 2010. Variability of North American English /r/ production in response to palatal perturbation. In Ben Maassen & Pascal van Lieshout (eds.), *Speech motor control*, 53–68. Oxford, UK: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199235797.003.0004>
- Tiede, Mark K., Suzanne E. Boyce, Christy K. Holland & K. Ann Choe. 2004. A new taxonomy of American English /r/ using MRI and ultrasound. *The Journal of the Acoustical Society of America* 115(5). 2633–2634. DOI: <https://doi.org/10.1121/1.4784878>
- Topol, Eric J. 2014. Individualized medicine from prewomb to tomb. *Cell* 157(1). 241–253. DOI: <https://doi.org/10.1016/j.cell.2014.02.012>
- Torres-Durán, María, Juan Miguel Barros-Dios, Alberto Fernández-Villar & Alberto Ruano-Ravina. 2014. Residential radon and lung cancer in never smokers. A systematic review. *Cancer Letters* 345(1). 21–26. DOI: <https://doi.org/10.1016/j.canlet.2013.12.010>
- Townsend, Grant C., Lindsay Clem Richards, Mitsuo Sekikawa, Tasman Brown & Tadashi Ozaki. 1990. Variability of palatal dimensions in South Australian twins. *The Journal of Forensic Odonto-Stomatology* 8(2). 3–14.
- Tsujino, Keiichiro & Y. Machida. 1998. A longitudinal study of the growth and development of the dental arch width from childhood to adolescence in Japanese. *The Bulletin of Tokyo Dental College* 39(2). 75–89.
- Twist, Alina, Adam Baker, Jeff Mielke & Diana Archangeli. 2007. Are “covert” /ɹ/ allophones really indistinguishable? *University of Pennsylvania Working Papers in Linguistics* 13(2). 16. <https://repository.upenn.edu/pwpl/vol13/iss2/16/>.
- Ubbanowski, Ann & Joan Wilson. 1947. Tongue curling. *Journal of Heredity* 38(12). 365–366. DOI: <https://doi.org/10.1093/oxfordjournals.jhered.a105674>
- Uldall, Elisabeth. 1958. American “molar” R and “flapped” T. *Bulletin of Laboratório de Fonetica Experimental Da Faculdade de Letras Da Universidade de Coimbra* 4. 103–106.
- Vendryès, J. 1902. Some thoughts on sound laws. In A. R. Keiler (ed.), *A Reader in Historical and Comparative Linguistics*, 109–120. New York: Holt, Rinehart and Winston.
- Vorperian, Hourii K. Ray D. Kent, Mary J. Lindstrom, Cliff M. Kalina, Lindell R. Gentry & Brian S. Yandell. 2005. Development of vocal tract length during early childhood: a magnetic resonance imaging study. *The Journal of the Acoustical Society of America* 117(1). 338–350. DOI: <https://doi.org/10.1121/1.1835958>
- Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. *Empirical foundations for a theory of language change*. Austin, TX: University of Texas Press. http://mnytud.arts.unideb.hu/tananyag/szoclingv_alap/wlh.pdf.
- Weirich, Melanie & Susanne Fuchs. 2011. Vocal tract morphology can influence speaker specific realisations of phonemic contrasts. In *Proceedings of the ISSP*, 251–258. <http://>

- www.sprachwissenschaft.unijena.de/germsprach_multimedia/Downloads/weirich/Weirich_Fuchs_Vocal + tract + morphology.pdf.
- Weirich, Melanie & Susanne Fuchs. 2013. Palatal morphology can influence speaker-specific realizations of phonemic contrasts. *Journal of Speech, Language, and Hearing Research* 56(6). 1894–1908. DOI: [https://doi.org/10.1044/1092-4388\(2013/12-0217\)](https://doi.org/10.1044/1092-4388(2013/12-0217))
- Wells, Jonathan C. K. 2009. Thrift: a guide to thrifty genes, thrifty phenotypes and thrifty norms. *International Journal of Obesity* 33(12). 1331–1338. DOI: <https://doi.org/10.1038/ijo.2009.175>
- World Health Organization. 2004. *Chronic suppurative otitis media: burden of illness and management options*. Geneva, Switzerland: World Health Organization. <http://www.who.int/iris/handle/10665/42941>.
- You, M., X. Li, H. Wang, J. Zhang, H. Wu, Y. Liu, J. Miao & Z. Zhu. 2008. Morphological variety of the soft palate in normal individuals: a digital cephalometric study. *Dento maxillo facial radiology* 37(6). 344–349. DOI: <https://doi.org/10.1259/dmfr/55898096>
- Yu, Alan C. L. 2013. *Origins of sound change: Approaches to phonologization*. Oxford, UK: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199573745.001.0001>
- Zelditch, Miriam Leah, Donald L. Swiderski & H. David Sheets. 2012. *Geometric morphometrics for biologists: A primer*. Amsterdam: Academic Press.
- Zemlin, Willard R. 1998. *Speech and hearing science: Anatomy and physiology*. London, UK: Allyn & Bacon 4th edn.
- Zhou, Xinhui, Carol Y. Espy-Wilson, Suzanne Boyce, Mark Tiede, Christy Holland & Ann Choe. 2008. A magnetic resonance imaging-based articulatory and acoustic study of “retroflex” and “bunched” American English /r/. *The Journal of the Acoustical Society of America* 123(6). 4466–4481. DOI: <https://doi.org/10.1121/1.2902168>
- Zipf, George Kingsley. 1949. *Human behaviour and the principle of least effort*. Oxford, England: Addison-Wesley Press.
- Zollo, Fabiana, Petra Kralj Novak, Michela Del Vicario, Alessandro Bessi, Igor Mozetič, Antonio Scala, Guido Caldarelli & Walter Quattrociochi. 2015. Emotional dynamics in the age of misinformation. *PLoS One* 10(9). e0138740. DOI: <https://doi.org/10.1371/journal.pone.0138740>

How to cite this article: Dediu, Dan and Scott R. Moisić. 2019. Pushes and pulls from below: Anatomical variation, articulation and sound change. *Glossa: a journal of general linguistics* 4(1): 7.1–33, DOI: <https://doi.org/10.5334/gjgl.646>

Submitted: 03 March 2018 **Accepted:** 01 November 2018 **Published:** 11 January 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

