LITERATURE REVIEW

THE SOUND **SYSTEMS OF LANGUAGES ADAPT**, BUT TO WHAT EXTENT? CONSIDERATIONS OF TYPOLOGICAL, DIACHRONIC AND MERCURIAL DATA

Caleb EVERETT (iD) ✉

University of Miami (UM)

ABSTRACT

The sound systems of the world's languages adapt to biomechanical, aerodynamic and cognitive pressures associated with sound production and discrimination. Such pressures help to yield the greater frequency of some sound types and the reduced frequency of others. In this paper I explore such adaptation, pointing out that sound systems not only adapt to such pressures in ways that are clear from a typological perspective, but that they adapt in more subtle ways that are only now becoming apparent. Furthermore, I survey a host of recent studies suggesting that sound systems likely adapt to some pressures that vary across populations and environments. While the extent of adaptation to such variable pressures is very debatable, the existence of phonetic/phonological adaptation to pressures that differ across populations is increasingly well supported. The evidence in support of such adaptation ranges from large-scale quantitative data, to biomechanical modeling, to the speech of Freddie Mercury. It is stressed that potentially adaptive processes do not require alternative diachronic mechanisms to well established ones. Instead, commonly acknowledged diachronic trajectories may simply be triggered differentially across populations in some instances.

RESUMO

Os sistemas de som das línguas do mundo se adaptam às pressões biomecânicas, aerodinâmicas e cognitivas associadas à produção e discriminação de som. Essas pressões ajudam a produzir a maior frequência de alguns tipos de som e a frequência reduzida de outros. Neste artigo, exploro essa adaptação, frisando que os sistemas de som não apenas se adaptam a essas pressões de maneiras que são claras de uma perspectiva tipológica, mas que se adaptam de maneiras mais sutis que só agora estão se revelando. Além disso, abordo uma série de estudos recentes sugerindo que os sistemas de som se adaptam a pressões que variam entre populações e ambientes. Embora a extensão da adaptação a tais pressões variáveis seja discutível, a mera existência de adaptação fonética / fonológica a pressões que diferem entre populações é cada vez mais bem sustentada. As evidências que apóiam essa adaptação incluem dados quantitativos em grande escala, dados derivados de modelagem biomecânica, até à fala de Freddie Mercury. Sublinho que tal evidência não implica que quaisquer processos adaptativos representem mecanismos alternativos a fenômenos bem conhecidos na fonologia diacrônica. Em vez disso, os mecanismos diacrônicos comumente reconhecidos podem simplesmente ser acionados diferencialmente entre as populações em alguns casos.

KEYWORDS
Phonetics; Phonology; Cultural Evolution; Adaptation.

PALAVRAS-CHAVE
Fonética; Fonologia; Evolução Cultural; Adaptação.

INTRO AND BACKGROUND

Humans are uniquely adaptive creatures. Through processes of cultural evolution, many of which are non-conscious, people have developed all sorts of behaviors that allow for their success. These behaviors, for instance the consumption of cod liver oil that helps prevent osteomalacia in some Nordic populations, or the elaborate cooking methods that allow Amazonians to consume manioc without its characteristically high rates of hydrocyanic acid, or limitless other examples, may be critical in some ecologies but not others. (Henrich 2015) While humans have also adapted genetically to particular environments, these genetic adaptations are comparably minor in scope as our species exhibits a remarkably low degree of cross-population genetic diversity when contrasted to other great apes. (Bowden et al. 2012) Instead, cultural evolution has taken over, guiding much of our behavior. Furthermore, cultural factors more than strictly environmental ones now place strong pressures on the selection of particular genes. (Richerson & Boyd 2005) These cultural factors may interact with environmental pressures in ways that are often not immediately apparent. Two recent, highly publicized examples underscore this point: First, human cultures exhibit parallel foraging and reproductive behavioral characteristics when contrasted to nonhuman mammal and bird species in similar physical environments. (Barsbai et al. 2021) Second, human sleep behavior is synchronized with the lunar cycle in ways that are only discernible with recently developed methods. (Casiraghi et al. 2021) Such examples illustrate how new sources of data are allowing us to detect adaptive elements in human behavior that would have been considered unlikely even recently. This theme carries through to linguistic behavior.

On some level it is clear that language, perhaps the most unique tool in the human cultural toolkit, is also adaptive. It is adaptive on the macro scale, having enabled or at least greatly facilitated humans' global circum-ambulation and success at the expense of competitor species. It is also adaptive at micro scales, since individual linguistic features present tremendous advantages in some contexts and also owe themselves to processes of cultural evolution. (see, e.g., Evans 2003) As such linguistic features develop they are selected for and elaborated as populations stumble upon their advantages. To cite one of many examples, number words gradually accumulate in given cultures and are selected for as they enable all sorts of interrelated behaviors like counting and precise measurement that may present advantages in some contexts. (Everett 2017a) Languages are adaptive in even less obvious ways as well. For instance kinship terms, for all their well-known variation, appear to evolve in ways that yield heightened communicative and cognitive efficiency. (Kemp et al. 2018) Unrelated aspects of language like syntax also exhibit adaptation for efficient communication. (Gibson et al. 2019)

The notion that languages are adaptive systems has become more commonplace in the literature over the last few years. With respect to color terms, Gibson et al. (2017) offered compelling evidence that basic color terms have evolved to fit the needs of speakers in environments with pervasive brightly colored foregrounded objects. Such salient objects have become more common since the advent of industrialization, due to the increasing use of dyed materials. Along those lines, other sensory terms are also adaptive, for example the words for smells used by tested hunter-gatherer populations seem to adapt to their greater needs for distinguishing some odors. (Majid et al. 2018) Less surprisingly, such adaptation is also evident in the terms for environmental features like ice and snow that are found in some ecologies and not others. (Regier et al. 2016) At the morphological level, languages appear to adapt to the size of their speakers' populations, with less complex morphological characteristics being selected for probabilistically in the languages of larger groups. (Lupyan & Dale 2010, Raviv et al. 2019) The latter association may be due to the greater simplicity required for morphological acquisition when a language has a high ratio of second-language learners, a point returned to below since it relates indirectly to the size of phoneme inventories.

While languages, like other forms of culturally acquired behaviors, are certainly adaptive at the level of words and likely at the level of morphosyntax, it is unclear how much such adaptation applies to sound systems. Do the phonological and phonetic characteristics of languages adapt to the needs or environments of their speakers? In one way, the answer to this question is an incontrovertible "yes". Yet it is hotly contested just how much adaptation is evident in the sounds used in the world's languages. As one example of the ways in which linguistic sounds adapt subtly to cognitive factors, Blasi et al. (2016) demonstrate that there are pervasive non-arbitrary sound-meaning correspondences across the world's languages, correspondences that have been overlooked in some cases until the advent of computational analyses of large typological datasets. Blasi et al. (2016) discovered, for example, that the word for 'nose' in the world's languages is inordinately likely to include a nasal sound, even after controlling for a host of confounds. The fact that humans use nasal sounds to refer to their nasal cavity has a likely proprioceptive motivation, since the air resonates in our nasal cavities when we produce nasal sounds. So these sounds are, in a sense, selected for in particular words since they have some inherent advantages for naturally conveying the meaning associated with the words in question.

Or consider the subtle ways in which sounds are selected for in order to disambiguate meaning, some of which have gone unnoticed until recently. It has long been known that less probable words tend to be longer, given Zipf's law of abbreviation. (Zipf 1949) This implies that less probable words tend to have more sounds with which to disambiguate themselves with respect to other words. In a recent study exploring this

topic of disambiguation, King & Wedel (2020) demonstrate that less probable words also have sounds that convey more disambiguating information on average when compared to more probable words of the same length. This suggests that the sounds in such words are structured in ways that gradually adapt to efficiency needs. In King & Wedel's (2020:9) words, "lexicons are optimized" for efficient communication. King & Wedel (2020) also demonstrate that this pattern tends to be strongest towards the beginning of words. Early sounds within words contribute more to the disambiguation of those words from so-called lexical competitors, and this is particularly true for less probable words. They observe this pattern over twenty languages from several linguistic phyla. For example, less probable words hit the sound representing their "uniqueness point" earlier than more probable words of the same length. The uniqueness point is the point in a word at which that word can no longer be confused with lexical competitors. To illustrate: /kændl/ ('candle') and /kændi/ ('candy') are only separated by the fifth segment in each word, so their uniqueness point occurs relatively late in the word. The key finding, in the present context, is that sounds are sequenced within words in ways that make the disambiguation of words more straightforward. Critically, language users are unaware of these patterns that make the transfer of meaning, from speakers to listeners, more efficient. We can think of this as a subtle form of phonological adaptation to cognitive pressures. More obvious cases of adaptation to cognitive pressures for disambiguation would include, for example, the pressures towards maximizing the formant spaces in vowel phoneme inventories (e.g. Gordon 2016).

This all may seem like a bit of a misdirect, however. Sound patterns adapt to cognitive pressures, sure, but is this really relevant to the suggestion that sound systems are adaptive like other adaptations of cultural evolution? After all, subtle sound patterns in speech presumably do not impact the overall reproductive success of a culture. Furthermore, it is unclear whether these pressures on sound systems vary across populations, like the pressures placed on many other forms of behavior. Yet many cultural adaptations are similarly subtle, non-conscious, and simply make life a bit easier without a clear connection to reproductive success. Furthermore, it is unclear whether all pressures on sound systems are uniform across populations. Results like those in Blasi et al. (2016) and King & Wedel (2020) underscore that a) there are subtle indications of adaptations of sound use to previously unnoticed pressures and b) these adaptations may only be uncovered with large-scale quantitative tests of typological data. Next we will consider if there are analogously subtle but physiologically motivated adaptations evident in sound systems. Then we will turn to the more controversial question of whether there are physiologically motivated adaptations that vary across human populations and across environments, as seen in many non-linguistic cases of cultural evolution.

1. PHYSIOLOGICALLY MOTIVATED ADAPTATIONS SHARED BY LANGUAGES

In addition to their common phonetic and phonological tendencies that are driven by perceptual phenomena, such as the maximization of vowel formant space, languages also share common (though non-universal) tendencies driven by articulatory ease. (Napoli 2016, Gordon 2016) Some of these tendencies are motivated by straightforward physiological factors. For example, languages favor bilabial, alveolar, and velar stops that require relatively little articulatory effort and precision to produce while also allowing for perceptual discriminability due to these stops' distinctive effects on the formant values of adjacent vowels. The articulatory ease of these stops helps to explain their worldwide commonality in phoneme inventories, and is evident in their frequency in infant babbling. (Locke 1983) More recently, it has been observed that some consonants, including these stops and a few others, are exceedingly common in phonetically transcribed word lists across the world's languages, even after controlling for language relatedness. (Everett 2018a). Those consonants that tend to be very frequent in word lists across languages are sounds that are known to be easy to articulate and frequent in the speech of babbling infants.

Such commonalities in consonant use across the world's languages suggest that languages adapt to pressures of articulatory ease. While any given language may have sounds and sound patterns that are less straightforward in terms of articulation there is, so to speak, an underlying current pulling phonetic and phonological patterns towards a valley of articulatory ease. While linguists have long been aware of this current, its strength has perhaps been underestimated since so many diverse sounds are observed across the world's phoneme inventories. These inventories may obscure the "current" in some cases since they do not contain frequency information that reveal just how strong the pressures are to produce easy-to-articulate sounds. As an example of this, consider a more subtle way in which phonetic patterns have recently been shown to be adaptive. For decades, it has been suggested that posterior voiced obstruents may be slightly more difficult to articulate, when contrasted to their homorganic voiceless obstruents, because of aerodynamic factors. (Ohala 1983) For example, [g] is a bit more difficult to articulate than [k], according to this hypothesis, because [g] requires vocal cord vibration concomitant with oral cavity obstruction relatively close to the glottis. This obstruction near the glottis yields a small supralaryngeal cavity into which air enters from the lungs. Once the air pressure in the supralaryngeal cavity equals subglottal air pressure, voicing becomes impossible since vocal cord vibration requires a pressure differential across the glottis. This basic aerodynamic factor is well known, but the question is whether it has a material effect on the sounds used in speech. After all, voiced posterior obstruents like [g] are very common in the world's languages. Furthermore, speakers have all sorts of tricks at their disposal to

increase the size of the supralaryngeal cavity during the production of a voiced stop and reduce the effect of this factor, for instance they can move the tongue root forward. Typological analyses of phoneme inventories only weakly support the suggestion that this factor has any material effect on consonant use across languages. (Maddieson 2013)

However, frequency-based data gleaned from transcribed word lists paint a very different picture. In Everett (2018b), I offer relevant data from thousands of such lists, each representing 40-100 words in the ASJP database. (Wichmann et al. 2016) Those data suggest that [g] is typically much less common than [k] even within languages in which both sounds occur. Additionally, there is a more pervasive pattern that is evident in the word lists: As the place of articulation of obstruents gets closer to the glottis, the disparity between the frequency of voiceless obstruents and their homorganic voiced counterparts increases. For instance, [p] and [b] are equi-frequent across the world's languages, while [t] is more common than [d]. The relative disparity between voiceless and voiced stops is even greater for [k] and [g], and greater still at the uvular place of articulation. This main result is evident even after controlling for the potential influences of language contact and linguistic lineage via linear mixed modeling and random sampling.

Results like those in Everett (2018a, 2018b) suggest that, alongside more obvious ease-of-articulation influences on the sounds used in speech worldwide, there are less obvious typological patterns as well. The frequency with which sounds are used across the world's languages point to undercurrents that pull sound systems in similar directions, undercurrents that are barely visible on the surface, some almost imperceptible in phoneme inventories. Sound systems adapt to subtle and minor aerodynamic pressures that are shared across human populations. Note, however, that this does not imply that the pressures are identical across human populations. For instance, the size of the pharyngeal and oral cavities, and the ratio between their lengths, can vary in minor ways across populations. (Xue et al. 2006) This variance in size hints that the aforementioned aerodynamic factors, while operative across all populations, may differ in minor ways across some populations. This fact requires further exploration.

2. POTENTIAL ADAPTATIONS OWING TO ANATOMICAL DIFFERENCES ACROSS POPULATIONS

Compelling evidence already exists that some sound patterns are more likely to occur in certain populations as sound systems gradually adapt to particular anatomical characteristics of their speakers. Researchers have now used a variety of approaches like biomechanical modeling to demonstrate that the articulatory effort required to produce some sounds can vary across populations and individuals. This work suggests that

individual and population-level differences in vocal tract characteristics may impact the use of rhotics and clicks, for instance. In the case of rhotics, Dediu and Moisik (2019) demonstrate that individual anatomical variation affects the production strategies used in the articulation of the American English /r/. Their work also suggests that the unusual typological distribution of phonemic clicks may be due to the anatomical characteristics of some populations. (Dediu & Moisik 2017) Certain Sub-Saharan populations are known to have smooth anterior palates, without the pronounced alveolar ridge that characterizes the mouths of most populations. Biomechanical modeling suggests that this characteristic makes the production of most clicks slightly less effortful, and this fact could help explain why clicks are so prevalent as phonemes in languages in southern Africa. Counter-arguments exist, however: Clicks are common as paralinguistic gestures worldwide in many populations without the relevant oral characteristics. Also, this account does not explain why bilabial clicks are also so common in southern Africa, though the latter distribution could be the result of clicks migrating to other places of articulation to fill perceptual space. Still, the account is suggestive of a way in which phoneme inventories may adapt to the physical characteristics of specific populations.

Other suggestions along these lines have been made, most prominently in a paper in *Science* in 2019. Dámian Blasi, Steve Moran, and other colleagues (including Dediu and Moisik) demonstrate in that paper that the biomechanical effort required to produce labiodental consonants is reduced in mouths with overbite and overjet, i.e. in the mouths in which the top incisors overlap vertically and horizontally with the bottom incisors. (Blasi et al. 2019) These dental characteristics have become more prevalent in the last few millennia as softer diets have become more common worldwide due to the adoption of intensive agriculture. Soft diets are more likely to yield overbite and overjet in the mouths of adults, for reasons discussed in Blasi et al. (2019). Additionally, eating utensils have contributed to less wear and tear on the teeth of agriculturalists and those in industrialized societies, further contributing to the prevalence of overbite and overjet. In contrast, in those contemporary societies with diets that are more similar to those of humans prior to the widespread use of agriculture, i.e. in hunter-gatherer societies, adults tend to have edge-to-edge bite types. For mouths with such bite types, the bottom lip does not rest immediately under the top incisors, meaning that labiodentals require slightly more effort to produce. Charles Hockett was actually the first scholar to suggest that these cross-population differences in bite types, owed to differences in diet, impact the sounds used in languages. (Hockett 1985) More specifically, Hockett suggested that labiodental fricatives have become much more pervasive since the advent of softer diets, and offered some general typological evidence in favor of this hypothesis. Blasi, Moran, and colleagues offered extensive typological evidence for the account based on phoneme inventories. They also

presented evidence from the history of Indo-European supporting the hypothesis, in addition to their evidence from biomechanical modeling.
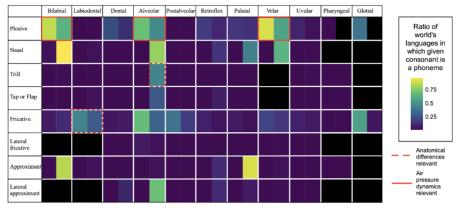
In recent work (Everett & Chen 2021), Sihan Chen and I present new sorts of evidence that cross-population variations in bite type impact not just the likelihood that a given population's language has labiodental phonemes but actually impact the frequency with which labiodental sounds are used. Across 2,729 phonetically transcribed word lists (again using the ASJP data), we found that the languages of hunter gatherers use labiodentals at extremely low rates, when contrasted to the languages of agriculturalists with softer diets. We also observe the same pattern in transcriptions of short texts for a smaller set of dozens of languages. This pattern holds globally even after controlling for language contact and lineage via frequentist and Bayesian linear mixed modeling. Only one language family is an apparent exception to the pattern, the Arawá language family of Amazonia. In Arawá languages, spoken by cultures without agriculture, labiodental consonants are nevertheless prevalent in the word lists. As we note in the paper, however, labiodentals are a very recent development in Arawá languages. Worldwide, labiodentals appear to have become much more prevalent across the last few hundred years, during the period of colonization. (Moran et al. 2020) They are amongst the most frequently borrowed sounds in the world's languages and, of course, have been borrowed by hunter-gatherers in many cases in which those groups have encountered European languages in which labiodentals are found. This appears to be the case in Arawá languages. The typological data are remarkably consistent, then: languages of hunter-gatherers do not traditionally rely on labiodentals as phonemes and even today such sounds are very infrequent in the words of such languages, judging from the transcriptions so far tested.

Such results suggest that cross-population differences in bite-type impact speech or, framed differently, that sound systems adapt to the specific anatomical characteristics of their speakers. Is such adaptation evident at the level of individuals? To answer this question, we examined the speech of ten famous English speakers with differing bite types. (Everett & Chen 2021) Transcribing their speech, we observed that the speakers with overbite used labiodentals with much greater frequency than the speakers with edge-to-edge bites. For example, the speaker with the most pronounced overbite in our data set, *Queen* lead singer Freddie Mercury, exhibited the highest rate of labiodental production, even replacing many putative bilabial consonants with labiodentals. In contrast, the speakers with no overbite/overjet, for instance Michael Phelps, produced labiodentals at the lowest rate. In fact, Phelps produced some supposed labiodental phonemes as bilabials--the reverse pattern of that evident in Mercury's speech. Such patterns were observed for nine of the ten speakers examined, across thousands of transcribed consonants. In short, bite type seems to affect the production of labiodentals both at the level of individuals and at the level of populations. At the population level, such bite type characteristics are

impacted by cultural factors like subsistence and diet. Given that the heavy reliance on agricultural technologies is influenced, in part, by environmental factors, this effect of diet on sound use is indicative of a very indirect effect of environments on language. In the next section we will consider whether environmental factors might influence the sounds of speech in more direct ways.

In Figure 1, I adapt a usage-based IPA chart I developed in other work (Everett 2021) to highlight some of the sounds discussed in these sections. In the chart the frequency of pulmonic consonants across 2,186 phoneme inventories is visualized. These data are based on the PHOIBLE database. (Moran & McCloy 2019) Cell brightness corresponds to greater prevalence across the world's phoneme inventories. The cells highlighted with red outlines represent some of the key sounds discussed in this section and in the preceding section. Note that the highlighted sounds are amongst the most common consonants in the world's phoneme inventories. In other words, the adaptations of sound systems to aerodynamic and anatomical pressures actually relate to some of the most prevalent sounds in phoneme inventories.

The assumption that sound systems change because of pressures that do not vary across populations, both today and in the past, is maintained by many scholars in linguistics and is sometimes referred to as the "uniformitarian" hypothesis. To a growing number of language researchers including myself, however, this assumption is neither well motivated nor well supported. In fact, to some of us it would be surprising to find that articulatory pressures are exactly the same across populations. After all, the vocal tract of *sapiens* varies in minor ways across contemporary populations and not just in terms of dental or alveolar characteristics, for instance cross-population variation in nasal cavities and laryngeal cavities have been well documented. (e.g. Noback et al. 2011, Xue et al. 2006) Additionally, the vocal tract has certainly changed extensively over long periods and language may have evolved alongside such changes. (e.g. Boë et al. 2019, D. Everett 2017) So it is unclear at which time scale the uniformitarian assumption would hold. There may be an understandable reluctance by some scholars to acknowledge that any cross-population disparities in physiology impact the sounds used in speech, given the history of simplistic ideas on such topics, but the research in physical anthropology and in other disciplines suggests that cross-population differences in vocal tract characteristics are a reality. There are, of course, no overall advantages to any of these characteristics, but this does not mean they are irrelevant to minor factors of human behavior like, e.g., voice onset time. Finally, with respect to the uniformitarian hypothesis, recent analyses of reconstructed phoneme inventories suggest that inventories have shifted dramatically over the last few centuries and millennia, in ways that are also problematic for the hypothesis. (Moran et al. 2021) In short, the uniformitarian hypothesis, while no doubt a useful heuristic in studies of specific kinds of sound change, is not well grounded empirically.

Figure 1. IPA chart of pulmonic consonants in which cell brightness corresponds to the frequency of a given sound across the world's phoneme inventories. Based on approximately 2,186 phoneme inventories in PHOIBLE. (Moran & McCloy 2019) Adapted from Everett (2021). The cells highlighted with red boundaries relate to phenomena discussed in this and the preceding sections. The cells highlighted with dashed red rectangles highlight sounds that appear to vary in accordance with individual and/or population-level anatomical differences.

3. POTENTIAL ADAPTATIONS OWING TO ENVIRONMENTAL VARIATION ACROSS POPULATIONS

It is unsurprising that sound systems adapt to the general constraints of the human vocal tract, to constraints of the human auditory capacities, and to cognitive constraints shared across cultures. What is perhaps more surprising, from the perspective of many linguists anyhow, is that sound systems apparently adapt to factors that vary across populations. In the case of the increased prevalence of labiodentals associated with softer diets, this adaptation is ultimately due to cultural factors and indirectly due, at least in part, to environmental factors. Sound systems may also adapt to variable social environmental factors since larger populations tend to have larger phoneme inventories. (Hay & Bauer 2007, Wichmann et al. 2011, though see Moran et al. 2012) Relatedly, Lupyan & Dale (2010) suggest that languages spoken by small populations tend to develop more complex morphological systems, systems that include robust sets of affixes when compared to the languages of larger populations. One interpretation of the pattern, evident across the roughly 2,236 languages sampled by Lupyan & Dale (2010), is that widely spoken languages adapt to accommodate the needs of adult learners. Since adults tend to learn second languages poorly when compared to children learning their native language, it has been argued that languages with high ratios of non-native speakers should tend towards reduced morphological complexity to aid in second-language acquisition. Critically in the present context, this comparably reduced affixation may yield shorter words on average in larger populations. Such shorter words may require that a larger number of phonemes be used in order to make the words easier to decipher. (See relevant discussion in Bentz 2018.)

This suggestion is consistent with the findings in King & Wedel (2020), discussed above. Just as phoneme inventories may gradually adapt to anatomical constraints, then, they may also adapt to social environments due to such proximate cognitive factors.

Is it possible that sound systems adapt to nonsocial environmental factors, much like other phenomena on which cultural evolution operates? It is unclear at present, but some research suggests that such direct environmental adaptation may exist. The relevant work suggests that variation in ambient air pressure and aridity may impact the operation of the vocal cords, to the extent that sound systems are subtly and probabilistically affected over long time periods. The work in question remains primarily correlational in nature, though it is supported by indirect experimental evidence in some cases. While correlational approaches face clear limitations, such approaches are critical to run preliminary tests and to see which hypotheses merit further scrutiny. In one paper I suggested that ejectives might be slightly easier to articulate in high-elevation regions, and offered correlational support for the idea that this modeled ease-based factor impacts the likelihood that languages use ejective phonemes. This hypothesis is not dissimilar, at its core, to the better-supported hypothesis that posterior voiced obstruents are slightly more difficult to produce than their voiceless homorganic counterparts due to subtle aerodynamic factors. In Everett (2013) I showed that languages with ejectives tend to be located in five of the six major regions of high elevation across the world's continents. This high-altitude distribution was also recently supported with a much larger dataset, though the study in question also found support for a weaker association between uvular phonemes and high elevation. (Urban & Moran 2021) The latter association hints that the ejective-altitude correlation is likely spurious. In other work I test the frequency of ejective consonants in phonetically transcribed word lists from 1,991 dialects, once again using the ASJP database, along with numerous other consonant types. (Everett, in progress) I observe that bilabial, alveolar, and velar ejectives are each positively correlated with elevation, even after controlling for language relatedness and contact via several approaches. The data also suggest that voiceless uvular stops associate with elevation, though more weakly, while voiced uvular stops and uvular fricatives do not associate with elevation. Ejective consonants are the only basic consonant type (not voicing/place/manner combination) that are more likely to be used in high elevation regions. In Figure 2, the 176 word lists with any ejectives, of the 1,991 tested in Everett (in progress), are plotted on a world map. Brighter dots correspond to word lists with higher ratios of ejectives as a proportion of all the consonants in the lists. As seen in the map, the dots tend to cluster in the five high-altitude regions in question, i.e. the North American Cordillera, the Andes, the Southern African Plateau, the highlands of East Africa, and the Caucasus. The association between ejectives and high elevation exists both in terms of phoneme inventories and in terms of the frequency of ejectives used in word lists. Associations of this type can be used as starting points to further inquiry, though in the case

of ejectives there are reasons to doubt that the association is due to the motivation suggested in Everett (2013). (Urban & Moran 2021)
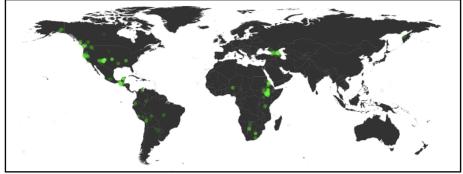

Figure 2. Locations of 176 dialects whose Swadesh-type word lists contain ejective consonants, out of 1,991 tested in Everett (in progress). Brightness corresponds to the ratio of all consonants in a word list that are ejectives. See also Urban & Moran (2021).

More substantively, my research on this topic has suggested that extreme ambient aridity may impact the sounds used in speech. This suggestion is more substantive because it is supported, at least indirectly, by extensive experimental evidence suggesting that very dry ambient air affects the operation of the vocal cords. Such aridity leads to greater phonatory effort and the perception that vocal cord vibration is more tiring. (e.g. Leydon et al. 2009, Sundarrajan et al. 2017) It also leads to greater rates of jitter and shimmer, suggesting that precision of vocal cord vibration is somewhat harder to maintain. (Sivasankar & Erickson-Levendoski 2012) (Such facts are consistent with anecdotal evidence offered by some singers who take effort not to spend time in very dry places prior to performances or recordings.) The latter fact led myself and two colleagues to test whether languages with complex tonality are less likely to develop in very dry regions. We found evidence in support of a negative association between complex tonality and ambient air desiccation. (Everett et al. 2015) Nevertheless, this correlational finding may be due to coincidental patterns of borrowing across languages. (Collins 2016) To further complicate matters, associations between tonality and population-level genetic factors exist. (Dediu & Ladd 2007, Wong et al. 2020) Such varied and complex associations point to the difficulty of disentangling causal mechanisms in such typological data, particularly since direct experimental evidence is difficult to derive for such hypotheses based on probabilistic mechanisms operating over millennia. Nevertheless, it is worth stressing that indirect experimental evidence is consistent with the hypothesis.
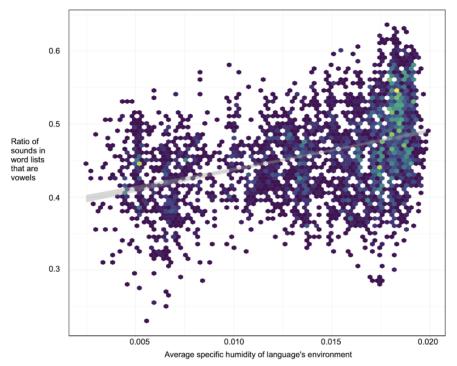
The same could be said of another hypothesis, one that is interrelated to the tone-aridity hypothesis. The hypothesis was motivated by this question: Given that experimental evidence suggests that vocal cord vibration is slightly more effortful in very dry contexts, and given that languages exhibit such a clear tendency for reduced articulatory effort, do

they exhibit a tendency for reduced vocal cord vibration in very dry ecologies? This question leads to a simple idea: Vowels, the sounds that require vocal cord vibration in nearly all cases, often at high amplitude, could be relied on a bit less in very dry environments. Note that this hypothesis too is probabilistic. Given sufficient time and given sufficient languages, languages in very dry contexts might adapt in subtle and non-conscious ways by relying on vowels a little less. Note that this hypothesis does not relate to the number of vowels in phoneme inventories. After all, some languages with small vowel inventories have characteristically simple syllable structures and therefore high proportions of vowels in the speech stream. The hypothesis relates to the ratio of sounds that are vowels in actual words. Once again, the underlying causal mechanism is one that is itself uncontroversial: Ease of articulation effects the use of particular sound types.

In Everett (2017b) I examined the ratio of vowels, as a proportion of all transcribed sounds, in word lists for 4,012 dialects represented in the ASJP database. As noted above, this database contains transcribed words for 40-100 basic concepts for each doculect. This database has the global coverage necessary to run preliminary tests on such a hypothesis but there are of course limitations as well given the shallow depth of representation for each represented doculect. For instance, Easterday (personal communication) observes that "There are some languages for which all or the majority of consonant clusters occur in contexts of inflection and other morphological processes", hinting at one of the limitations of the ASJP data for tests of this sort. Nevertheless, it should also be noted that the data correspond well with external sources. For example, the same regions that Easterday (2019) describes as having characteristically complex syllable structures, clustered due to language contact, are found to be regions with low "vowel ratios" in Everett (2017b). (These are the Pacific Northwest, the Caucasus region, the Atlas Mountains region, Patagonia, the Northeastern US, the Sonoran Desert, Northern New Guinea, and Northeast Asia.) Also, as noted in Everett (2021), the families with the lowest vowel ratios are families like Salishan that are known to allow many complex syllable types and frequent consonant clusters, while those with the highest vowel ratios are families that are known to rely primarily on simple syllable structures.

To test more specifically the correspondence between the ASJP-based vowel ratios and our general understanding of syllable structures and consonant clusters, I examined the languages in Easterday's genealogically and geographically balanced analysis of syllable structure types. (Easterday 2019) Of this sample, 24 languages are categorized into the "simple" syllable structure type, and 25 are classified as having "highly complex" syllable types. For these 49 languages, 40 had vowel ratio data available based on Everett (2017b). For the languages in Easterday's "highly complex" category, the vowel ratio was found to be 0.393 on average. (N=21, s.d. 0.062) For the languages in the "simple" category, the mean vowel ratio was 0.535. (N=19, s.d. 0.049) This is an extremely marked difference, given the

inherent limitations of vowel-ratio variation, and reveals that the vowel ratios are generally indicative of broader patterns in syllable structures: Languages with complex syllables have fewer vowels, as a ratio of all sounds, as we would expect. This difference between ASJP vowel ratios based on known syllable-structure types is robust. (p<0.0001, two-tailed Mann-Whitney) In addition to these points vis-à-vis vowel ratios, in the case of other phenomena that have been explored with the ASJP data and tested with larger datasets, namely the prevalence of [k] and [g] and the prevalence of labiodentals, there is a clear correspondence between the ASJP data and other sources. Still, the vowel ratios are merely indicative of patterns that need to be tested with much larger datasets containing transcriptions of naturalistic discourse. Unfortunately, at present we do not have such datasets with sufficient global coverage to run preliminary tests that factor in ecological variables while controlling for Galton's problem adequately.

In Everett (2017b) I tested whether "vowel ratio" positively associates with ambient specific humidity. (Note that the relevant factor is specific humidity, not relative humidity, for reasons discussed in that paper.) In Figure 3 I present the data for the 4,012 doculects, adapted from Everett (2017b). A series of tests suggested that the association in Figure 3 is not due to clear confounds like language relatedness or language contact. In fact, the association between vowel use and humidity actually becomes much more robust, not less so, once such factors are included in an analysis via mixed effect modeling or random sampling. Of course, the association could still be coincidental. Nevertheless, there is indirect experimental evidence consistent with a motivated relationship between the variables. The tendency for languages in humid regions to rely on simple syllable structures and the tendency for languages in very dry, cold regions to rely on complex syllable structures may be motivated in part by ecological factors. To some this may seem implausible given our understanding of more proximate diachronic factors at work but, once again, the question is whether those factors are triggered differentially in particular contexts. For example, vowel deletion processes may be particularly favored over long periods in very dry, cold regions, and lexical variants with elided vowels may gradually become more prevalent in such contexts. This suggestion does not seem radical in the light of, for instance, the very minor aerodynamic effects on the production of voiced posterior obstruents, effects that nevertheless have a pervasive impact on the use of obstruents worldwide. Subtle ease-based articulatory effects clearly can impact the sounds used in languages in probabilistic but pervasive ways, suggesting that the distribution in Figure 3 may not be coincidental. Future tests and non-correlational data are required to explore this issue and the hypotheses discussed in this section can be considered exploratory.

Figure 3. Hexbin plot demonstrating correlation between vowel use and ambient water vapor. Brighter bins correspond to greater language density. Line represents locally weighted regression with confidence interval. (Beta regression Pseudo R^2 = 0.158.) Based on 4,012 word lists, adapted from Everett (2017b). See data and analysis in that study for further discussion.

4. HOW "EXTERNAL" FACTORS ARE COMPATIBLE WITH OUR UNDERSTANDING OF DIACHRONIC PHONOLOGY

One point worth underscoring is that potential "external" influences on sound systems do not require the postulation of novel diachronic mechanisms, when contrasted to those that are well established in historical linguistics. Nor are they incompatible with the fact that certain sound patterns in some regions are more common due to well-established cases of language contact. In this section I discuss how well-known diachronic trajectories could account for the distributions observed above if they are simply triggered at different rates over time across populations and environments. Given that sociolinguistic factors play a known role in sound change, such triggering could be mediated by such factors as influential speakers in particular networks have their speech affected by external effects. Differential environmental triggering could then, through a process of iteration across language learners and sociolinguistic networks, promote the sorts of distributions observed. This discussion is not meant to further buttress the hypothesized ecological influences that, as noted above, vary markedly in terms of extant support. The purpose of

the discussion is merely to illustrate how our present understanding of relevant diachronic mechanisms is not incompatible with the suggested hypotheses. Let us consider, in turn, some relevant diachronic accounts related to labiodentals, tone, ejectives and vowels.

i. Labiodentals

In response to Blasi et al. (2019), some suggested that well-established lenition processes, principally the lenition of /p/ to /f/, are sufficient to account for the global distribution of labiodentals observed in that study. While we know that lenition is common cross-linguistically, however, a lenition-based account alone could not in principle explain the typological distributions noted by Blasi et al. (2019) or Everett & Chen (2021). A key question is whether certain bite types cause such lenition to be more prevalent in populations with overbite/overjet. The only study that has actually looked at the effect of bite type on the production of labiodentals across speakers shows that bite type is a clear synchronic predictor of whether some bilabials are produced as labiodentals. (Everett & Chen 2021) We need data from many individuals within populations to test such hypotheses, data that are coded for factors like bite type in addition to the factors typically considered in research on sociolinguistically mediated sound change (e.g. gender, age, socioeconomic class).

The labiodental/bite-type interaction may prove relevant to a host of other well-known diachronic processes. To choose one of many, consider the shift of place of articulation in nasals in prefixation, as in the *in-*, *im-* prefixes in English that reflect well-known diachronic trajectories. The production of such prefixes actually varies across individuals with distinct bite types, at least in some cases. Some of the individuals we observed in Everett & Chen (2021) do not actually produce [im-] with a bilabial nasal but with a labiodental nasal. Now let us imagine that in a given population the amount of people that produce bilabial nasals as labiodentals is 2%, but in another it is 20% given clear disparities in overbite/overjet prevalence across populations. Such variation is not trivial and the way such phonetic tendencies might impact sound change trajectories is uncertain. It would not be describable without extensive quantitative data across speakers and populations. Critically, we lack such data. Some of this is due to the limitations of fieldwork, which often entails recording a few speakers only, or even just one. Factors like bite-type are not considered in descriptions of phonetic and phonological patterns. Instead such descriptions are often filled with "rules" that may betray the complexity of the synchronic data and, *ipso facto*, the diachronic picture.

ii. Tonogenesis

Voiced consonants can lead to lowered pitch on following vowels. If the voicing contrast of the consonant is subsequently lost via merger, the pitch distinction of the following vowel may become meaningful assuming it is maintained. (Kingston 2011) Perhaps in some

ecological contexts speakers are less likely to retain the pitch of the following vowel with sufficient fidelity, even if they use pitch for non-phonemic purposes, over the countless productions of relevant phonetic tokens. Disentangling this possibility would require plenty of quantitative data across individuals, variants, and environments. Relatedly and perhaps more critically, we know that tonality is a highly regional feature and that lexical borrowing plays a substantive role in its regional spread. This has been invoked as evidence against an ecological effect on language, but it need not be: What if speakers in some environments are subtly favored to copy complex tones with higher fidelity over many iterations of "tone-copying", across generations of speakers? This seems reasonable in the face of the laryngology data and the very regions in which we would expect to lack complex tones are the ones that lack them. While this could certainly be a coincidence, the key point here is that the traditional diachronic mechanisms associated with tonogenesis and the introduction of tones via contact do not represent incompatible hypotheses to an ecologically driven account: the question is simply whether such mechanisms are favored or disfavored, in probabilistic ways evident over many generations of language use, in particular environments.

iii. Ejectives

Ejectives may arise from the fusion of a glottal stop and an adjacent stop consonant, in some cases after metathesis that yielded that adjacency. Variants of this fusional process have been described for a number of language families. (see e.g. Fallon 1998) Yet such explanations fail to explain why there is such a marked regional clustering to the fusional processes, particularly given that the conditioning factors leading to ejectives are again not uncommon worldwide. (Of course, the distribution of ejectives could still be a historical coincidence, see e.g. Urban & Moran (2021)). The question, in this context, is whether certain effects on the vocal cords cause glottalization to be slightly more likely to be maintained when a glottal stop and an oral stop are initially fused in the speech of individuals. Current studies on the topic do not offer the sorts of data that would be needed to eliminate this possibility. What we have instead are interesting observations like Fallon's that "Yapese, to my knowledge, is one of only two (of over 900) Austronesian languages to have ejectives". (1998:441) Given that glottal stops are pervasive in some Austronesian languages, one wonders why the fusional process has not been triggered more commonly in them.

Once again, then, the processes that have been established in diachronic phonology are not incompatible with potential ecological influences. In the case of ejectives the association with altitude may be a coincidence and this is likely the weakest recent case that has been made that speech adapts to external pressures. However, the core factor motivating the hypothesis, namely that ease of articulation impacts sound use, is itself uncontroversial. No broader mechanism besides this ease of articulation is required, so any

diachronic processes that relate to ease-of-articulation could in principle be triggered differently by such external factors.

iv. Vowel reduction

There are many well-established processes that could affect the "vowel ratio" in a given language. As I ask in a previous paper: "Are vowel elision and vowel epenthesis slightly more and less likely, respectively, to occur in very dry regions? Are innovators of sociolinguistic change more likely (however, slightly) to rely on easier-to-articulate lexical variants with reduced vowels, in dry regions? Are elderly speakers more likely to produce easier-to-articulate variants, given that they seem particularly prone to the effects of humidity on phonation?" (Everett 2017b)

In their typologically based analysis of vowel reduction processes, Kapatsinski, Easterday, and Bybee (2020:20-21) make two observations that are relevant in the current context:

i)   "Detailed phonetic knowledge developed by individuals is transmitted from generation to generation as part of the community norms. Thus phonetic variation in online production may not always just attend to the immediate needs of the participants but also be influenced by subtle aspects of past usage."

ii)   "Reduction-favoring contexts include …. phonological environments that make the target sounds difficult to articulate."

Questions arise. With respect to i), do subtle aspects of past usage include minor variations in vowel reduction rates, e.g. duration of vocal cord vibration, in the lexical tokens produced by individuals in very dry climates? With respect to ii), if we know that articulation difficulty in some phonological environments impacts vowel reduction rates, perhaps effects of aridity on vowel articulation would be more likely to surface precisely in such phonological environments. This may not be the case--as I have openly acknowledged the hypotheses discussed in this section are exploratory. Nevertheless, such factors cannot be ruled out simply because there are well-known diachronic mechanisms at play. The subtle environmental influences could manifest themselves precisely through such diachronic mechanisms. (Relatedly, see Bickel's (2017) discussion of "functional triggers".)

5. CONCLUSION

The sounds people produce during speech, like other kinds of human behavior, exhibit signs of adaptation. Sound systems adapt to pressures to encode information efficiently, as well as to pressures to reduce the effort of production and perception. Some of these pressures appear to apply equitably in all cultures, but in other cases the pressures appear to vary across populations. Variable pressures may relate to minor anatomical differences and to differences in speakers' social and physical environments. In some sense, such adaptation should be unsurprising since it is ubiquitous in other forms of socially transmitted human behavior as evidenced by the growing literature on cultural evolution. From the perspective of some researchers like myself, it is problematic to assume that phonological systems are immune to such forms of adaptation.

Nevertheless, it is also clear from this survey that work on the latter topic, viz. the potential adaptation of sound systems to population-specific factors, is in its nascent stages. The possibilities discussed here remain candidates, suggestive of such adaptation and not yet conclusive. Some cases of potential adaptation, for instance the prevalence of labiodental consonants in populations with overbite/overjet, are relatively well supported. Others, for instance the case of ejectives at high elevation, may turn out to represent "false starts" in this broad line of inquiry. Ultimately new experimental methods will need to be developed to test some of the more specific hypotheses. Innovative methods have recently been developed to test other kinds of linguistic adaptation, so perhaps other approaches will be developed to test adaptation of the sort discussed here. (Raviv et al. 2019)

Finally, it is worth noting that many features of human behavior are maladaptive, and no claim is being made here that sound systems always change in strictly adaptive ways. Adaptive pressures can be operative and the resultant linguistic characteristics may still exhibit maladaptive characteristics of a completely different sort. For instance, a recent strand of experimental research has demonstrated that certain sound types are more likely to produce aerosol particles, particles that are so critical to the transmission of SARS-CoV-2 and other airborne viruses. (Asadi et al. 2019) Early work on this topic suggests the high-front vowel [i] is associated with particularly high quantities of aerosol particles. (Asadi et al. 2020) Of course, the [i] vowel also presents many advantages in terms of perceptual discrimination, helping to maximize formant space. The point here is that certain sounds could be maladaptive in one sense and adaptive in another. They are not, however, necessarily neutral in terms of their adaptive characteristics. This too is unsurprising from the perspective of human behavior more broadly, since many types of human behavior exhibit both adaptive and maladaptive features in addition to neutral ones that are due simply to "drift". In this paper I have simply discussed how some of the adaptive characteristics of languages may help shape the phoneme inventories of the world's

languages, along with the frequency of sounds in words within and across languages. The extent of phonetic/phonological adaptation in the world's languages requires further exploration in the coming years.

REFERENCES

Asadi, S. *et al.* 2019. Aerosol emission and superemission during human speech increase with voice loudness. *Scientific Reports* 9, 2348

Asadi, S. *et al.* 2020. Effect of voicing and articulation manner on aerosol particle emission during human speech. *PLoS ONE* 15, e0227699

Barsbai, T, Lukas D, Pondorfer, A. 2021. Local convergence of behavior across species. *Science* 371.292-295.

Bentz, C. 2018. *Adaptive Languages: An Information-Theoretic Account of Linguistic Diversity*. De Gruyter: Mouton.

Bickel, B. 2017. Areas and universals. In *The Cambridge Handbook of Areal Linguistics*, ed. R. Hickey, 40-54. https://doi.org/10.1017/9781107279872.004

Blasi D, Wichmann S, Hammarstrom H, Stadler P, Christiansen M. 2016. Sound-meaning association biases evidenced across thousands of languages. *Proc. Natl Acad. Sci. USA* 113 (39) 10818-10823

Blasi, D. et al. 2019. Human sound systems are shaped by post-neolithic changes in bite configuration. *Science* 363.eaav3218.

Boë LJ, Sawallis T, Fagot J, Badin P, et al. 2019. Which way to the dawn of speech? *Science Advances* 5.12 eaaw3916

Bowden R, MacFie TS, Myers S, Hellenthal G, Nerrienet E, et al. Genomic Tools for Evolution and Conservation in the Chimpanzee: Pan troglodytes ellioti Is a Genetically Distinct Population. *PLoS Genet*, 2012 DOI: 10.1371/journal.pgen.1002504

Casiraghi L, et al. 2021. Moonstruck sleep: Synchronization of human sleep with the moon cycle under field conditions. *Science Advances* DOI: 10.1126/sciadv.abe0465

Cohen Priva U. 2017. Informativity and the actuation of lenition. *Language* 93. 565-597.

Collins J. 2016. Commentary: The role of language contact in creating correlations between humidity and tone. *Journal of Language Evolution*, 1.46–52.

Dediu D & Ladd R. 2007. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proc. Natl Acad. Sci. USA* 104.10944-10949.

Dediu, D. & S. Moisik. 2017. Anatomical biasing and clicks: Evidence form biomechanical modeling. *Journal of Language Evolution* 2.37-51.

Easterday S. 2019. *Highly complex syllable structure: A typological and diachronic study*. Language Science Press.

Evans, N. 2003. Context, culture, and structuration in the languages of Australia. *Annual Review of Anthropology* 2003.13-40.

Everett, C. In progress. Into thin air once more: Ejectives, uvulars, and high elevation.

Everett, C. 2013. Evidence for direct geographic influences on linguistic sounds: the case of ejectives. *PLoS ONE* 8:e65275. doi: 10.1371/journal.pone.0065275

Everett, C. 2017a. *Numbers and the Making of Us*. Cambridge: Harvard University Press.

Everett, C. 2017b. Languages in drier climates use fewer vowels. *Front. in Psych.* 8.1285. *doi:10.3389/fpsyg.2017.01285*

Everett, C. 2018a. The similar rates of occurrence of consonants in the world's languages. *Language Sciences* 69.125-135.

Everett, C. 2018b. The global dispreference for posterior voiced obstruents: A quantitative assessment of word list data. *Language* 94.e311-e323.

Everett C. 2021. The sounds of prehistoric speech. *Philosophical Transactions of the Royal Society B*. DOI: 10.1098/rstb.2020.0195

Everett C & Chen S. 2021. Language adapts to differences in dentition within and across human populations. *Scientific Reports* 11.1066. https://doi.org/10.1038/s41598-020-80190-8

Everett, C., D. Blasi & S. Roberts. 2015. Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proc. Natl Acad. Sci. USA* 112.1322-1327 +SI. doi:10.1073/pnas.1417413112.

Everett, D. 2017. *How Language Began*. New York: W.W. Norton.

Fallon, Paul. 2002. *The Synchronic and Diachronic Phonology of Ejectives*. Abingdon, England: Routledge Press.

Gibson E, Futrell R, Jara-Ettinger J, Mahowald K, Bergen L et al. 2017. Color naming across languages reflects color use. *Proc. Natl Acad. Sci. USA* 114.10785-10790.

Gibson E, Futrell R, Piantadosi S, Dautriche I et al. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences* 1894.1-19.

Gordon M. 2016. *Phonological Typology*. Oxford University Press, Oxford.

Hay J, Bauer L. 2007. Phoneme inventory size and population size. *Language* 83.388-400

Henrich J. 2015. *The Secret of Our Success*. Princeton: Princeton University Press.

Hockett C. 1985. Distinguished lecture: F. *Am. Anthropol.* 87:263-28.

Kapatsinski V, Easterday S, and Bybee J. 2020. Vowel reduction: A usage-based perspective. *Italian Journal of Linguistics* 32.19-44.

Kemp C, Xu Y, Regier T. 2018. Semantic typology and efficient communication. *Annual Review of Linguistics* 4.109-128.

King A & Wedel A. 2020. Greater early disambiguating information for less-probable words: The lexicon is shaped by incremental processing. *Open Mind* 4.1–12. https://doi. org/10.1162/opmi_a_00030

Kingston, John. 2011. Tonogenesis. In Marc van Oostendorp; Colin J Ewen; Elizabeth Hume; Keren Rice (eds.). *The Blackwell companion to phonology, Volume 4: Phonological interfaces*. Chichester: Wiley-Blackwell. pp. 2304–2333.

Leydon C, Sivasankar M, Falciglia D, Atkins C, Fisher K. 2009. Vocal fold surface hydration: a review. *J. Voice* 23, 658–665. doi: 10.1016/j.jvoice.2008.03.010

Locke J. 1983. *Phonological acquisition and change*. New York: Academic Press.

Lupyan G, Dale R. 2010. Language structure is partly determined by social structure. PLOS ONE. https://doi.org/10.1371/journal.pone.0008559

Maddieson I. 2013. Voicing and gaps in plosive systems. In: Dryer M, Haspelmath M,

Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Majid, A et al. 2018. Differential coding of perception in the world's languages. *Proc. Natl Acad. Sci. USA* 115.11369-11376.

Moran S, Grossman E, Verkerk A. 2020. Investigating diachronic trends in phonological inventories using BDPROTO. *Language Resources and Evaluation*. https://doi.org/10.1007/s10579-019-09483-3

Moran S, Lester, N. & Grossman E. 2021. Inferring recent evolutionary changes in phonological inventories. *Philosophical Transactions of the Royal Society B*.

Moran S, McCloy D. 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. (Available online at http://phoible.org, Accessed on 2020-08-06.)

Moran, S, McCloy D, Wright R. 2012. Revisiting population size vs. phoneme inventory size. *Language* 88.877-893

Napoli, DJ. 2014. On the linguistic effects of articulatory ease, with a focus on sign languages. *Language* 90.424–456.

Noback M., Harvati K. & Spoor F. 2011. Climate-related variation of the human nasal cavity. *American Journal of Physical Anthropology* 145.599-614.

Ohala, John. 1983. The origin of sound patterns in vocal tract constraints. T*he production of speech*, ed. by Peter MacNeilage, 189–216. New York: Springer

Raviv L, Meyer A, Lev-Ari S. 2019. Larger communities create more systematic languages. *Proceedings of the Royal Society B* 286 (1907) https://doi.org/10.1098/rspb.2019.1262

Regier T., Carstensen A., Kemp C. 2016. Words for snow revisited. *PLOS ONE* https://doi.org/10.1371/journal.pone.0151138

Richerson, P. J., & Boyd, R. 2005. *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.

Sivasankar M, Erickson-Levendoski E. 2012. Influence of obligatory mouth breathing, during realistic activities, on voice measures. *J. Voice* 26, e9–e13. doi: 10.1016/j.jvoice.2012.03.007

Sundarrajan A, Brinton Fujiki R, Loerch S, Venkatraman A, Sivasankar P. 2017. Vocal loading and environmental humidity effects in older adults. *J. Voice* doi: 10.1016/j.jvoice.2017.02.002

Urban M, Moran S. 2021. Altitude and the distributional typology of language structure: Ejectives and beyond. *PLoS ONE* doi:10.1371/journal.pone.0245522.

Wichmann S., Rama, T. & E. Holman 2011. Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguistic Typology* 15.177-197.

Wichmann S., Holman E.W. & Brown, C. 2016. The ASJP database (version 17). http:/asjp.clld.org/.

Wong, P et al. 2020.*ASPM*-lexical tone association in speakers of a tone language: Direct evidence for the genetic-biasing hypothesis of language evolution. *Science Advances* DOI: 10.1126/sciadv.aba5090

Xue S, Hao G. & Mayo R. 2006. Volumetric measurements of vocal tracts for male speakers from different races. *Clinical Linguistics & Phonetics* 20(9): 691–702

Zipf, G. 1949. *Human Behavior and the Principle of Least Effort*. Cambridge, Mass.: Addison-Wesley Press, Inc.