

Review



Cite this article: Everett C. 2021 The sounds of prehistoric speech. *Phil. Trans. R. Soc. B* **376**: 20200195.
<https://doi.org/10.1098/rstb.2020.0195>

Accepted: 16 October 2020

One contribution of 17 to a theme issue ‘Reconstructing prehistoric languages’.

Subject Areas:

evolution, behaviour

Keywords:

phonemes, prehistoric speech, Upper Palaeolithic, vowels, consonants

Author for correspondence:

Caleb Everett
e-mail: caleb@miami.edu

Electronic supplementary material is available online at rs.figshare.com.

The sounds of prehistoric speech

Caleb Everett

Department of Anthropology, University of Miami, Miami, FL, USA

CE, 0000-0002-5889-3935

Evidence is reviewed for widespread phonological and phonetic tendencies in contemporary languages. The evidence is based largely on the frequency of sound types in word lists and in phoneme inventories across the world’s languages. The data reviewed point to likely tendencies in the languages of the Upper Palaeolithic. These tendencies include the reliance on specific nasal and voiceless stop consonants, the relative dispreference for posterior voiced consonants and the use of peripheral vowels. More tenuous hypotheses related to prehistoric languages are also reviewed. These include the propositions that such languages lacked labiodental consonants and relied more heavily on vowels, when contrasted to many contemporary languages. Such hypotheses suggest speech has adapted to subtle pressures that may in some cases vary across populations.

This article is part of the theme issue ‘Reconstructing prehistoric languages’.

1. Introduction

The tools of historical linguistics, so critical to the reconstruction of the phonological characteristics of extinct languages, are generally considered incapable of reconstructing words or phoneme inventories of languages spoken prior to the Neolithic [1]. While some recent approaches have attempted to stretch this temporal window, it is still debated whether such methods allow for the discovery of anything specific about particular ancestral languages spoken during the Upper Palaeolithic [1,2]. Nevertheless, contemporary phonological and phonetic data do allow us to generate reasonable general hypotheses about the sound systems of the languages spoken during that epoch. After all, while the phonologies of the world’s languages are incredibly diverse, they also exhibit some well-known commonalities [3]. Alongside such established typological tendencies, there also exist more subtle commonalities that are coming into view via inspections of new sources of data. Most of these widespread tendencies, both familiar and recently uncovered, have likely existed since the development of sapiens’ anatomically modern vocal apparatus, bearing in mind the extant divergent views on when the dawn of speech occurred [4,5]. This review surveys evidence for several of these tendencies. An ancillary goal of this piece is to underscore that, like many other facets of human behaviour, sound systems are adaptive, evolving in accordance with pressures associated with the ease of articulation and perception. While the human vocal apparatus is largely uniform across populations, the biomechanical ease of producing some sound types may in some instances vary subtly across populations or environments. This particular point is disputed, however, and so throughout this piece, I will draw attention to the varying degrees of support for the relevant hypotheses mentioned. Some degree of conjecture is unavoidable when deriving hypotheses related to languages spoken so many millennia ago.

Most typologically oriented work on sounds focuses on patterns in phoneme inventories. This review relies largely on an alternate approach, emphasizing recent work on the frequency of sounds in transcribed words in many languages. Other contributions in this volume take a more phoneme-based approach (e.g. [6]). Such an approach is essential, but phonemes relate only obliquely to the frequency of particular articulations in speech, and that frequency is also

critical to understanding patterns that are most representative of speech. For instance, while the voiced postalveolar fricative in English (e.g. the last sound in 'rouge') is phonemic and contrastive, it is exceedingly infrequent in English word types or tokens. By contrast, the alveolar nasal (e.g. the first and last sound in 'none') is much more frequent, by an order of magnitude in fact [3]. Much of the work surveyed here relies on frequency-based data derived from transcriptions of common words in a large sample of the world's languages. Various scholars have recently used related approaches to address a host of non-physiological phenomena (e.g. [7,8]). This includes the phenomenon of systematic sound-meaning associations, as common associations between sound types and particular meanings have been uncovered in the world's languages. These systematic correspondences include the use of the nasal [n] sound in words for 'nose', and the [m] sound in words for 'breast' [7]. Given their prevalence in the world's languages, it seems likely that some correspondences like these also existed in the languages of the Upper Palaeolithic.

Of course, patterns in the world's phoneme inventories are also critical to deciphering patterns in prehistoric languages. Considering the data in PHOIBLE, the most extensive database of the world's phoneme inventories, we can see some clear cross-linguistic tendencies [9]. Out of the 2186 languages represented in that database, the median number of phonemes is 33 (mean = 34.9, s.d. = 13.4). The median number of distinctive vowel types is 9 (mean = 10.3, s.d. = 6.25), including vowels that are contrasted via factors like duration and nasality. The median number of consonant phonemes in a language is 22 (mean = 23.9, s.d. = 10.3). Extreme variability exists, with the size of phoneme inventories ranging from 11 to 161. While it is possible that contemporary languages are somehow a poor proxy for prehistoric languages vis-à-vis phoneme inventory sizes, there is only weak support for this notion. Some work has suggested that languages spoken by larger populations tend to have bigger phoneme inventories [10]. Were this association robust, then we might expect that some languages spoken today should have larger phoneme inventories than those typical in the pre-Neolithic, since extremely large populations are a byproduct of the agricultural revolution. Atkinson [11] suggested the possibility of a phoneme serial founder effect, whereby languages tended to lose phonemes over the course of humans' global circumambulation. Counterintuitively, perhaps, this suggestion was consistent with the putative correlation between population and phoneme inventory size, given that human populations in places like the Americas tend to be smaller than those in Africa. Yet, the association between phoneme inventory size and population did not hold once language relatedness was controlled for, at least in one follow-up study [12]. Additionally, scholars have offered other criticisms of Atkinson's hypothesis [13,14]. Some other work has suggested that a weak correlation between sound-inventory size and population may in fact exist [15]. Assuming this is a meaningful correlation, one possible motivation that has been presented for it is as follows: larger populations could place simplifying pressures on morphologies, more commonly yielding shorter words with modest or no affixation [16,17]. The reliance on shorter stems could generate pressures for larger phoneme inventories, in order to more effectively disambiguate words earlier during their production [16, p. 113; 18]. Recent work, based on a newly constructed database of reconstructed sound inventories for proto languages, does observe that some ancient inventories tend towards fewer consonants than their

descendant languages [19]. Yet, it is unclear how robust this tendency is across the world's language families. Overall, there is no clear evidence that the sizes of phoneme inventories of prehistoric languages differed in pronounced ways from those observed today. Given the above-mentioned facts about current phoneme inventories, it seems likely that prehistoric languages had an average of roughly 30 total phonemes, about two-thirds of which were consonants. The number of phonemes likely varied substantially across languages during the Upper Palaeolithic, judging from contemporary languages.

Regardless of the number of phonemes in a given language, the frequency of occurrence of those phonemes follows a power-law distribution whereby a few phonemes and their variants are incredibly frequent with respect to the others [20–22]. In languages tested to date, a phoneme's intralinguistic frequency, F , is predictable by its intralinguistic frequency-based rank, r , when factoring in the total number of phonemes in the language, n [20–22]. Several equations have been proposed to account for this predictable relationship, for instance the following relatively straightforward equation that has been shown to be accurate across many languages [20]

$$F = \frac{\log(n+1) - \log r}{n}.$$

While debate exists regarding the best way to describe the power-law distribution of sounds in speech, and slightly more accurate formulae are derivable, there is no debate that sounds generally adhere to a power-law distribution, both in word lists and in transcriptions of discourse [20–22]. There is no motivation for suspecting that pre-Neolithic languages differed from contemporary languages according to this parameter either. This point underscores the need for a frequency-based approach to complement more traditional approaches: if we aim to characterize the world's sound systems and the pressures that have shaped them in preceding millennia, it is helpful to describe the very commonly used sounds in individual languages, and to better illuminate the pressures that lead to their inordinate frequency. Section 2 addresses some of these pressures associated with consonants, and §3 addresses vowels. Understanding these pressures allows us to generate some hypotheses about the sounds used in prehistoric speech. Each of the next two sections begins with the most well-grounded hypotheses and proceeds to the more speculative.

2. Consonants

A key factor influencing the frequency of the consonants used in speech is the biomechanical ease associated with their articulation [23]. It is unclear how much of a role the ease of articulation plays, particularly given that other factors are also at work in shaping phonetic and phonological tendencies, including needs for perceptual discriminability and efficiency. Nevertheless, the most common consonants in the world's languages tend to be easy to produce. Some of these sounds are alveolar consonants like [n] and [t], which require the relatively miniscule tip of the tongue to touch the alveolar ridge via a small movement, in contrast with harder-to-produce sounds requiring, for instance, movement of the larger tongue root towards the back of the pharyngeal wall. The former sounds are easy to make and frequently occur in syllables babbled by children across populations, while the latter sounds do not [24]. Other consonants that

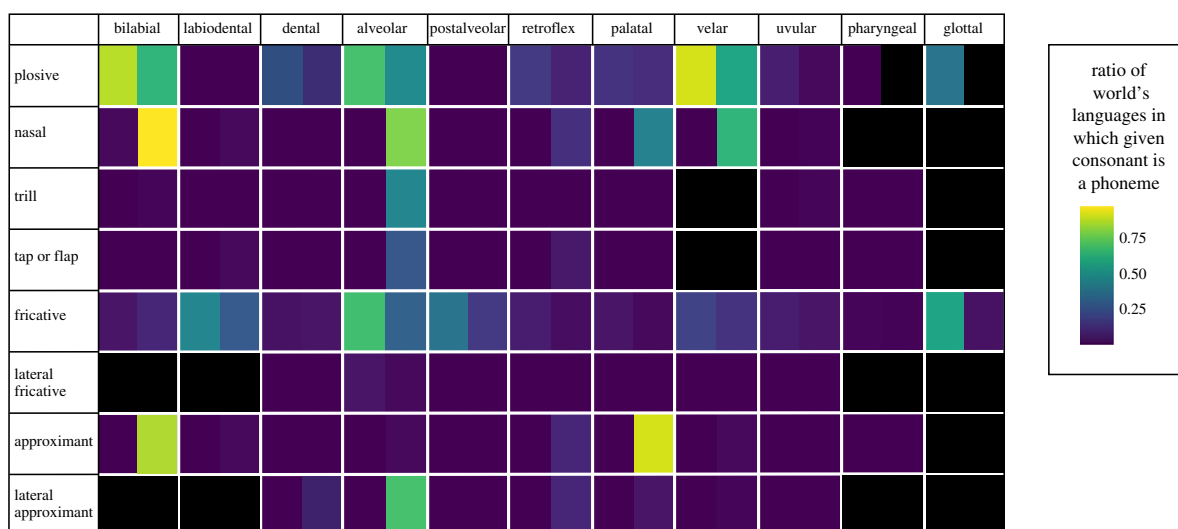


Figure 1. This chart represents the pulmonic consonants of the International Phonetic Alphabet. Each cell represents one consonant on the IPA chart. Cell brightness corresponds to a consonant's frequency as a phoneme across the world's languages, based on the data in PHOIBLE, a database with sound-inventory information for 2186 languages. Black cells correspond to sounds that do not occur because of physiological limitations.

are generally considered easy to produce include bilabials like [m] and [p], which babbling infants also commonly produce [24]. (Nasal sounds like [m] and [n] also require the lowering of the velum to create their characteristic resonance, but this is the default position of the velum during breathing.) The same could be said of the velar consonant [k], which requires little movement of the tongue to reach the nearby velum, especially among infants as they have a lower soft palate [24]. Given such factors, it is perhaps unsurprising that /m/, /k/, /p/, /n/ and /t/ are (in that order) the most well-represented consonants in the world's phoneme inventories, excepting vowel-like approximants. Based on the 2186 languages in PHOIBLE, /m/ is found in 96% of languages, /k/ in 90%, /p/ in 86%, /n/ in 78% and /t/ in 68% [9]. Despite such prevalent sounds, though, note that none are universal. The commonality of basic consonants across languages is depicted in figure 1. The figure displays pulmonic consonants with one manner of articulation and no secondary modulation, so some consonant types are excluded, for instance, affricates and ejectives.

Several recent studies have relied on the frequency of sounds within the word lists in the ASJP database [25–28]. These lists consist of phonetic and phonological transcriptions of common concepts across a large cross-section of the world's languages. While such lists do not necessarily include all phonemes in a given language and are coarsely transcribed in some cases, they offer a means of detecting the relative frequency of common consonants in basic words while sampling the bulk of the world's language taxa. (Bearing in mind the limitations of this approach, it is worth underscoring that phoneme inventories offer no information on the relative frequency of particular sounds or associated articulatory gestures.) Even in such word-list data, the power-law distribution of sound types is evident. A few consonants are extremely common within individual word lists, when contrasted to some other sounds in the same list [27]. The frequency of each basic consonant, as a proportion of all the phonetic segments in each word list, was ascertained [27]. These ratios were then averaged within each language family, and the family means were then averaged to generate a global mean 'usage rate'. Based

on that approach, [n] was found to have a mean usage rate of 0.059. That is, [n] represents 5.89% of all transcribed sounds across families. The usage rate of [k] was 5.46%, while those of [m] and [t] were 4.49% and 4.45%, respectively. These four consonants were the only ones with global mean usage rates over 3%. While the articulation of these consonants is not identical across languages or individuals, their commonality underscores how languages rely heavily on consonants made at the lips, the alveolar ridge and at the velum. Even languages with well-known differences in phoneme inventory sizes and phonotactic characteristics often exhibit similar frequency of use of these basic sounds. For example, while Hawaiian has a very small phoneme inventory and English a fairly average-sized inventory, both languages rely on the alveolar nasal to about the same degree in the word lists, and it is the most common consonant in each [27]. Given the robust reliance on these sound types within languages of different families and regions, it is very plausible that many prehistoric languages not only used [m], [k], [p], [n] and [t], or some subset thereof, but that they did so frequently.

Note that the most common stop consonants in speech are voiceless. Voiceless stops have marked effects on the formants of adjacent vowels, giving them clear perceptual correlates despite their characteristic absence of both vocal cord vibration and oral aperture [29]. Voiced stops, while not uncommon sounds, are overall less common than their voiceless counterparts despite requiring the same manipulation of the oral articulators. This is apparently due to a subtle ease-based phenomenon that impacts speech, namely that it is more difficult to maintain vocal cord vibration alongside oral occlusion [30]. Vocal cord vibration requires transglottal airflow that is contingent on supralaryngeal air pressure being lower than subglottal air pressure. During the production of stops, air pressure across the glottis equalizes, making vocal cord vibration more difficult to achieve. This factor has been noted for decades, but it has remained unclear how much it actually affects speech since voiced stops are still quite common within and across languages. It has long been suggested that voiced stops produced closer to the glottis are less common typologically than their anterior counterparts,

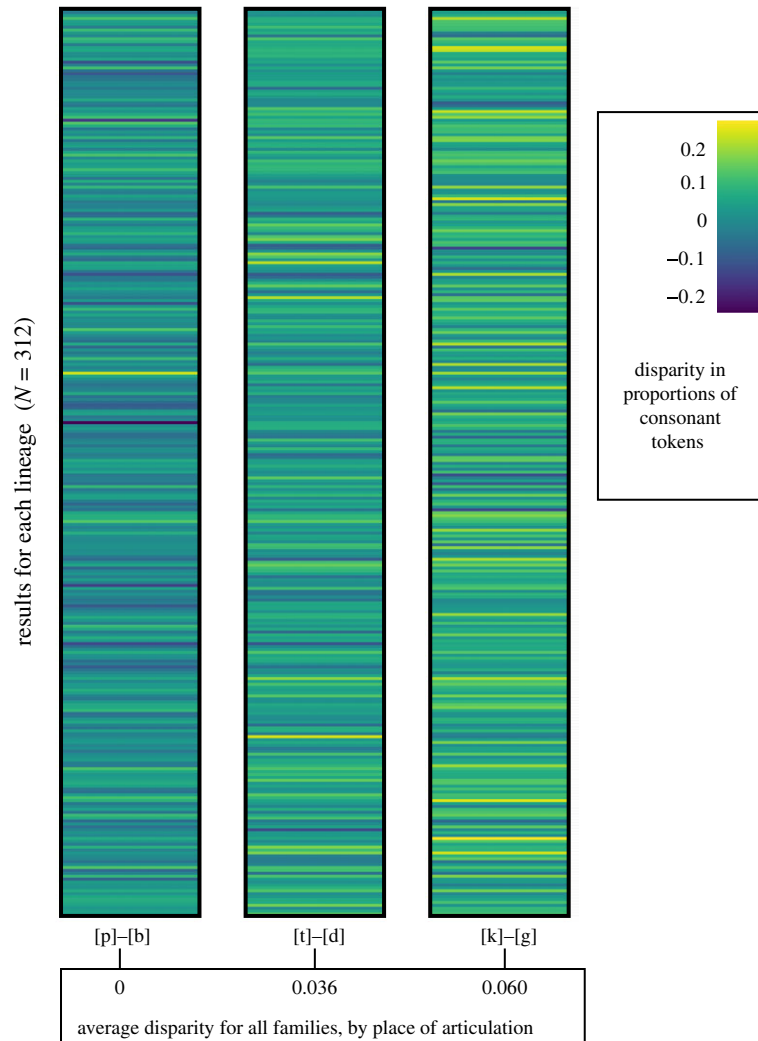


Figure 2. Average disparity in usage rates of plosive consonants at the same place of articulation, for 312 language lineages. Based on Everett [26]. The lineages are organized alphabetically (see electronic supplementary material for list). Voiced plosives become relatively less common as place of articulation approaches the glottis. The mean difference in the frequencies of [k] and [g], calculated as proportions of all transcribed consonants, is 0.06. This is the average disparity obtained across the 312 lineages represented by 3341 documented language varieties or doculects. [k] is used about three times as much as [g] across all lineages, [t] is about two times more common than [d], while [p] and [b] are equiprecquent. (See [26, p. e6].)

but evidence from phoneme inventories is equivocal [30]. A recent study presents frequency data from word lists demonstrating that, as the place of articulation of obstruents moves farther back in the mouth, voiced varieties of those obstruents become much less common in speech [26]. This pattern holds robustly across 3341 word lists even after controlling for language contact and relatedness via linear mixed modelling. In figure 2, the pattern is visualized in a new way, using the data for stop consonants in Everett [26] for three key places of articulation. The pattern also appears to hold for places of articulation that are less well represented in speech [26]. The rows in the figure represent the average disparities in the usage rates of voiceless and voiced stops across 312 language lineages, i.e. families or language isolates, based on the AUTOTYP database [31]. The global dispreference for posterior voiced obstruents suggests that minor aerodynamic effects can impact languages in subtle yet pervasive ways. Given that the size of the supralaryngeal cavity has not changed substantially since the African exodus, this tendency likely existed in the languages of the Upper Palaeolithic.

During the course of childhood and adolescence, the human bite type naturally transitions from an overbite/

overjet configuration to an edge-to-edge bite type wherein the top incisors and the bottom incisors form a flat plane when the mouth is closed. This ontogenetic shift to an edge-to-edge bite is due to three processes associated with heavy wear of the teeth: lingual tipping, dental eruption and mesial drift [32]. Since heavy wear and chewing have decreased alongside the introduction of softer diets during the Neolithic, the bites of adult agriculturalists are now more likely to exhibit overbite and overjet when compared with adult hunter-gatherers. This fact led Hockett to suggest in 1985 that labiodental consonants like [f] and [v], which require the movement of the bottom lip to the top teeth, may have become common only after the advent of agriculture [33]. Hockett's hypothesis was largely ignored until Blasi *et al.* [32] offered evidence in its favour. This evidence included biomechanical modelling revealing that the articulatory effort required to produce labiodentals is significantly reduced in individuals with overbite/overjet. Most critically, perhaps, it was demonstrated that labiodentals are notably lacking in the phoneme inventories of contemporary hunter-gatherer populations. In follow-up work, the hypothesis was tested with frequency-based data [34]. It was found that labiodental consonants represent

about 2% of all consonants in the word lists of agricultural populations but are nearly absent in the word lists of hunter–gatherers. This latter study was based on data for 2756 populations for which subsistence data were also available. It was also observed that individuals with different bite types rely on labiodental consonants to differing degrees. The bite type of English speakers was found to be a strong predictor of the speakers' use of labiodental sounds [34]. All these points suggest that labiodental consonants have become more common during the Neolithic, in association with cultural changes that led to less heavy dental wear. This finding implies that the sound systems of languages are adaptive not just to physiological pressures owing to the vocal tract characteristics shared by all human populations, as in the case of pressures on obstruent voicing, but that some pressures may differ slightly across populations. This sort of adaptation is the norm for many kinds of socially transmitted human behaviour, as cultures adapt in non-conscious ways to biological and environmental factors [35]. Intriguingly, while labiodental consonants may have been less common in the Upper Palaeolithic, they are very common across the world's languages today in part because they have been transmitted extensively across languages during the last few centuries. In fact, /f/ has been the most commonly borrowed phoneme during that time [36]. This is due principally to the fact that many languages of colonizers have labiodentals [36].

Other recent work also suggests that consonant use may vary cross-linguistically due in part to pressures that differ across populations. For instance, the prevalence of clicks in sub-Saharan Africa may owe itself partially to the absence of pronounced alveolar ridges in key populations [37]. Biomechanical modelling is consistent with this suggestion, since oral cavities with less pronounced alveolar ridges do allow for click production with reduced muscular effort [37]. Two potentially problematic points for this hypothesis are that bilabial clicks are also absent as phonemes outside of southern Africa (the ease of production of these sounds is not related to the shape of the alveolar ridge), and that many clicks are common as paralinguistic gestures worldwide [38]. With respect to the former point, though, it is possible that bilabial clicks became prevalent in some languages after their speakers used other click types. Assuming for the moment that this hypothesis is on the right track, it suggests that Upper Palaeolithic languages spoken in southern Africa would have been more likely to have click phonemes as well, as there is no evidence that the relevant oral features are recent innovations.

Finally, another controversial hypothesis suggests that ejective consonants are more common at higher elevations since the compression of the pharyngeal cavity that is required for their articulation could be somewhat easier to achieve in regions of reduced air pressure [39]. The hypothesis is supported by the fact that ejective phonemes occur in languages at or near regions of high elevation [39]. The hypothesis has met with skepticism, though it should be noted that the general phenomenon at its core, as in the case of the click and labiodental hypotheses, is itself fairly unremarkable: the ease of articulation impacts the sounds used in speech. The unresolved question is whether ambient air characteristics do impact the ease of production of ejectives somehow but, unlike the case of the click hypothesis, no biomechanical modelling evidence exists. If there is such an interaction, we might infer that ejectives have likely been

a feature of languages of the southern African plateau, where they exist today, for many millennia. Additionally, ejectives might have been selected for, probabilistically over many millennia, in different high-altitude regions that humans entered prior to the Neolithic. As in the case of the click-related hypothesis, however, this suggestion remains speculative. Recent work replicates the key finding in Everett [39] with a much larger sample of languages, finding a significant worldwide correlation between ejectives and altitudes [40]. However, that work also questions the interpretation of the distribution offered in Everett [39], while acknowledging that room for such an interpretation remains [40].

3. Vowels and prosody

It has long been posited that peripheral vowels are a critical feature of human speech, and that other related animals and hominids lack(ed) the capacity to produce vowels like [i], [a] and [u] [41]. This latter point has been called into question since other species are capable of producing peripheral vowels even if not with the degree of ease exhibited by humans [42–44]. Regardless of their role in the evolution of language, peripheral vowels certainly play a critical role in contemporary speech. While the median number of vowels in the phoneme inventories in PHOIBLE is 9, this figure includes vowels that are distinguished by nasality and/or length. Most languages have fewer distinctive vowel positions and [i], [a] and [u] are the most common of these. Each is found in over 86% of PHOIBLE inventories [9]. Given that peripheral vowels help to maximize formant space and allow for ease of perceptual discriminability, we can be confident that they were ubiquitous in speech during the pre-Neolithic and likely prior to the African exodus. After all, the perceptual factors motivating peripheral-vowel discriminability would have held for many tens of millennia, given that human auditory capacities do not vary substantially across populations.

The ratio of sounds that are vowels was obtained for 6901 lists of transcribed words in the world's languages [28]. The ratios ranged from 0.23 to 0.65, with a mean of 0.46. **Figure 3** depicts in a new manner the ratios of sounds in speech that are vowels, across major language families. The contemporary language families that have the lowest vowel ratios are known to rely heavily on complex syllable structures with multiple consonants for each vowel nucleus. This is particularly true of Salishan languages, which have the lowest vowel ratios across these families [45]. Those families with characteristically high vowel ratios are conversely not known to rely on complex syllable structures [45]. The ranges of vowel ratios in languages today offer some indication as to how much languages of the Upper Palaeolithic used vowels, since there is no motivation to believe those languages were outliers in this respect.

There is a clear correlation between languages' vowel ratios and the characteristic ambient specific humidity of the regions in which the languages developed [28] (**figure 4**). Based on experimental work in laryngology demonstrating how dry air impacts the vocal cords (e.g. [46–48]), it has been suggested that this association may be due to the small increases in effort required to vibrate the vocal cords in very dry/cold climates. After all, vowels require vocal

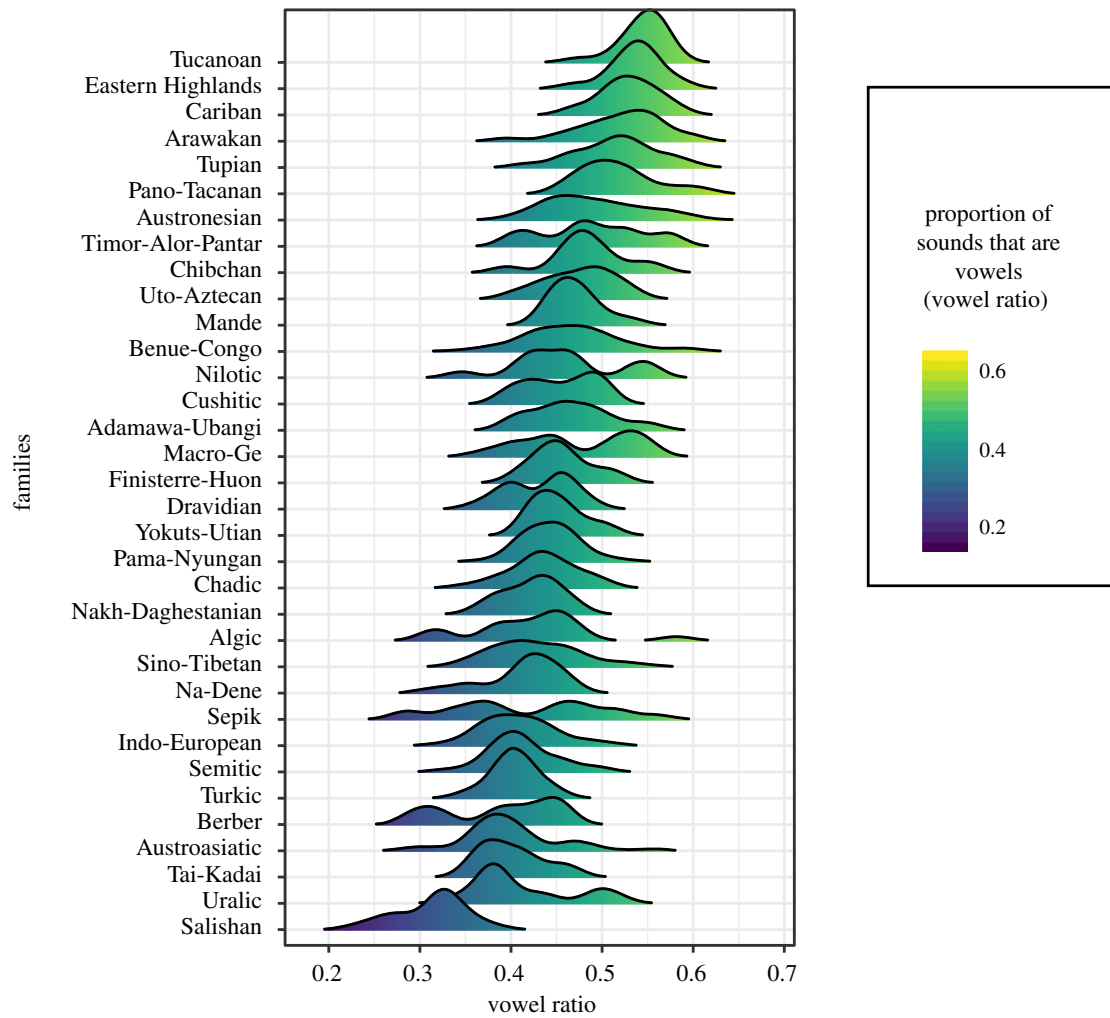


Figure 3. Ridgeline plot representing the proportion of sounds that are vowels, i.e. ‘vowel ratios’, across the 35 language families with at least 10 representative samples in the dataset used in Everett [28]. Each ‘hill’ represents a density distribution of the vowel proportions across the word lists for a given family. (Gaussian density plot used.) Families are ordered from highest mean vowel ratio (top) to lowest.

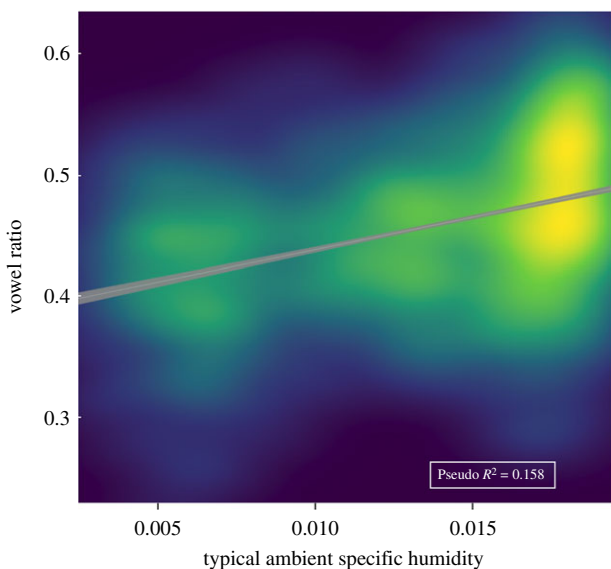


Figure 4. Heatmap of the global association between vowel ratios and ambient humidity, based on data for 4012 word lists from 2632 distinct languages [28]. This association may owe itself in part to the minor increase in effort required of vocal cord vibration in dry contexts. Pseudo- R^2 value is based on the β -regression analysis in Everett [28].

cord vibration, often at high amplitudes and relatively long durations. This is another potential case in which minor ease-based factors associated with voicing could impact

speech in subtle ways, as in obstruent voicing patterns. The association between ambient humidity and vowel ratios was found to be robust to the confounds of language relatedness and contact based on various statistical techniques [28,49]. (Though it should be noted that some of the tests in [49] relate to phenomena like the number of vowel phonemes and acoustic vowel space, neither of which are directly relevant to the hypothesis in [28].) Other scholars have offered acoustically based hypotheses to explain interrelated tendencies in the world’s languages [50,51].

It has been suggested that very dry ambient air might affect not just the ease of vocal cord vibration but the ease with which precise tones can be produced with the vocal cords [52]. Experimental evidence from laryngology offers support for this hypothesis, though the support is indirect. The hypothesis is supported primarily by correlational data from the world’s tonal languages, but the putative effect of dry ambient air on tonality is contested for numerous reasons (see [49,53,54]). It is worth noting as well that genetic factors may help explain the distribution of languages with phonemic tones, though the genetic and environmental hypotheses are not mutually exclusive [55,56]. The various hypotheses surrounding such distributions underscore the challenges of interpreting newly uncovered global correlations, challenges that will likely persist given the difficulty of testing such hypotheses experimentally. Among these challenges is the fact that there is no widespread

consensus on the best approaches to investigating large-scale synchronic patterns in speech with a view towards illuminating long-term diachronic trends. A common approach is to rely on Bayesian inferences to produce the most likely tree of a given language family, allowing scholars to examine the development of some feature(s) in that family (e.g. [57,58]). This approach is not without its detractors, however, who point out that tree-based approaches improperly de-emphasize the role of areal diffusion effects, producing tree models of language families at the exclusion of wave and linkage models that may also account for the data [59,60]. Some recent approaches rely less on synchronic patterns and emphasize the diachronic transitions within particular families, relying on the incorporation of statistical methods and the traditional comparative method [60]. Such approaches may benefit from new databases such as BDPROTO, a collection of reconstructed phoneme inventories from proto languages [19]. These sorts of data can help scholars to reconstruct the probabilities of certain phonological changes over time. Such thorny methodological issues aside, and assuming for the moment that the effect of aridity on vocal cord vibration did probabilistically impact how languages came to rely on vowels and tonality, we can infer that languages tended towards heavier reliance on vowel use and tonality prior to sapiens' incursions into very cold/dry regions.

4. Discussion and conclusion

This survey has focused largely on tendencies motivated by biomechanical ease of articulation, but there are of course other higher order factors that help shape languages, some of which relate indirectly to commonalities in sound systems. For instance, recent work suggests that the information rate of speech is similar across languages, regardless of syllable types [61]. Other recent work points to a variety of ways in which languages adapt to cognitive constraints [62–64] and ecological factors [65–68], so uncovering some adaptivity in phonetic and phonological patterns is perhaps expected, given the trend towards detecting adaptivity in speech.

The data and hypotheses surveyed here allow us to reasonably derive several hypotheses regarding the sounds of prehistoric languages. Unsurprisingly given the time depth involved, the hypotheses vary in terms of how well supported they are. Yet, the core idea motivating each of them is itself uncontroversial, namely that the ease of articulation and discrimination affected the sounds that languages relied upon during the Upper Palaeolithic, as those factors still do. It should be stressed that all the hypotheses discussed here are not deterministic but probabilistic and are not necessarily incompatible with our understanding of near-term sound changes [69,70]. As noted in recent work, specific forms of language change can be motivated by 'functional triggers' that may be grounded on biological or cognitive factors. While such factors affect the transition probabilities between particular language states in systematic ways, this does not mean that they hold uniformly across all populations. As noted by Bickel, 'a functional trigger may be tied to a biological or social condition that has itself a limited distribution in the world' [60, p. 3]. This would be true in the case of the potential effects of extreme aridity or certain bite

types on the sounds used in speech. Varied phenomena that subtly promote changes to the sounds used in basic words, and to phoneme inventories, may interact in complex ways. (These phenomena could also include the so-called 'event-based triggers' of language change, which are due to the copying of linguistic features for social reasons that are idiosyncratic and not preferred across linguistic lineages in some motivated or systematic manner [60].) In the case of vowel use, for example, there are obviously well-known diachronic phenomena like vowel epenthesis and elision that are relevant to the global pattern evident in figure 4. The question is whether such processes are triggered at different rates in particular ecologies due to very subtle variations in the ease of production, given sufficient time and bearing in mind that such factors are mediated by sociolinguistic phenomena. Answering such a question with greater confidence will help us to better elucidate the characteristics of the sound systems of prehistoric languages. It should also be mentioned that a question like this, as with a number of the other questions raised in this piece, is unlikely to be settled by language researchers alone. Understanding the nature of prehistoric speech is a notoriously complex issue requiring the engagement of linguists with fields such as genetics, physical anthropology, archaeology, laryngology and statistics. Such engagement and associated collaboration is fortunately becoming more prevalent.

Some inferences about the sounds of prehistoric speech can be made with confidence: most languages in the Upper Palaeolithic likely had about two to three dozen phonemes, roughly two-thirds of which were consonants. In each language, some phonemes played a pronounced role, given the way intralinguistic sound occurrence can be described via a power-law function. It is likely that [m], [n], [k], [p] and [t] were used frequently in prehistoric speech, alongside some of the other consonants that are prevalent in today's speech (figure 1), while other consonants played a smaller role. Most if not all prehistoric languages likely relied on the peripheral vowels [i], [a] and [u] and their variants. Vowels probably represented about half of the phonetic segments as they do today in most language families, due to the prevalence of simple syllable structures. It is possible vowels were relied on a bit more heavily prior to humans' incursion into cold/dry climates, but this point is more tenuous. Prehistoric languages likely relied less on voiced posterior obstruents, when contrasted to voiceless alternatives, given basic aerodynamic factors associated with vocal cord vibration. It seems unlikely that languages in the Upper Palaeolithic used labiodental consonants much. These latter suggestions point to adaptation of sound systems to minor physiological pressures, pressures that in some cases may vary subtly across populations and eras.

Data accessibility. While no new data were collected for this review, it does include new visualizations of data available online in the databases and sources cited. Figure 1 is based on data from this site: <https://phoible.org/parameters>. Figure 2 is based on data in the electronic supplementary material, spreadsheet derived from the data in this paper: <https://muse.jhu.edu/article/712565>. Figures 3 and 4 are based on data available here: <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01285/full#supplementary-material>.

Competing interests. I declare I have no competing interests.

Funding. I received no funding for this study.

References

- Campbell L, Poser W. 2008 *Language classification: history and method*. Cambridge, UK: Cambridge University Press.
- Pagel M, Atkinson Q, Calude A, Meade A. 2013 Ultraconserved words point to deep language ancestry across Eurasia. *Proc. Natl Acad. Sci. USA* **110**, 8471–8476. (doi:10.1073/pnas.1218726110)
- Gordon M. 2016 *Phonological typology*. Oxford, UK: Oxford University Press.
- Dediu D, Levinson S. 2012 On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Front. Psychol.* **4**, 00397. (doi:10.3389/fpsyg.2013.00397)
- Everett D. 2017 *How language began: the story of humanity's greatest invention*. London, UK: Profile.
- Moran S, Lester N, Grossman E. 2021 Inferring recent evolutionary changes in speech sounds. *Phil. Trans. R. Soc. B* **376**, 20200198. (doi:10.1098/rstb.2020.0198)
- 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**, 10 818–10 823. (doi:10.1073/pnas.1605782113)
- Cohen PU. 2017 Informativity and the actuation of lenition. *Language* **93**, 569–597. (doi:10.1353/lan.2017.0037)
- Moran S, McCloy D. 2019 *PHOIBLE 2.0*. Jena, Germany: Max Planck Institute for the Science of Human History. See <http://phoible.org> (accessed on 6 August 2020).
- Hay J, Bauer L. 2007 Phoneme inventory size and population size. *Language* **83**, 388–400. (doi:10.1353/lan.2007.0071)
- Atkinson Q. 2011 Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science* **332**, 346–349. (doi:10.1126/science.1199295)
- Moran S, McCloy D, Wright R. 2012 Revisiting population size vs. phoneme inventory size. *Language* **88**, 877–893. (doi:10.1353/lan.2012.0087)
- Jaeger F, Graff P, Croft W, Pontillo D. 2011 Mixed effect models for genetic and areal dependencies in linguistic typology. *Linguist. Typol.* **15**, 281–319. (doi:10.1515/lity.2011.021)
- Bybee J. 2011 How plausible is the hypothesis that population size and dispersal are related to phoneme inventory size? Introducing and commenting on a debate. *Linguist. Typol.* **15**, 147–153. (doi:10.1515/lity.2011.009)
- Wichmann S. 2011 Phonological diversity, word length, and population sizes across languages: the ASJP evidence. *Linguist. Typol.* **15**, 177–197. (doi:10.1515/lity.2011.013)
- Bentz C. 2018 *Adaptive languages: an information-theoretic account of linguistic diversity*. Berlin, Germany: De Gruyter Mouton.
- Lupyan G, Dale R. 2010 Language structure is partly determined by social structure. *PLoS ONE* **5**, e8559. (doi:10.1371/journal.pone.0008559)
- 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. (doi:10.1162/opmi_a_00030)
- Moran S, Grossman E, Verkerk A. 2020 Investigating diachronic trends in phonological inventories using BDPROTO. *Lang. Resour. Eval.* (doi:10.1007/s10579-019-09483-3)
- Tambotsev Y, Martindale C. 2007 Phoneme frequencies follow a Yule distribution: the form of the phonemic distribution in world languages. *SKASE J. Theor. Linguist.* **4**, 2.
- Gusein-Zade SM. 1988 On the distribution of letters of the Russian language by frequencies. *Probl. Transm. Inf.* **23**, 102–107. [In Russian.]
- Macklin-Cordes J, Round E. 2020 Re-evaluating phoneme frequencies. *Front. Psychol.* **11**, 570895. (doi:10.3389/fpsyg.2020.570895)
- Napoli DJ. 2014 On the linguistic effects of articulatory ease, with a focus on sign languages. *Language* **90**, 424–456. (doi:10.1353/lan.2014.0026)
- Locke J. 1983 *Phonological acquisition and change*. New York, NY: Academic Press.
- Wichmann S, Brown C, Holman E. 2020 *The ASJP Database*.
- Everett C. 2018 The global dispreference for posterior voiced obstruents: a quantitative assessment of word list data. *Language* **94**, e311–e323. (doi:10.1353/lan.2018.0069)
- Everett C. 2018 The similar rates of occurrence of consonants in the world's languages. *Lang. Sci.* **69**, 125–135. (doi:10.1016/j.langsci.2018.07.003)
- Everett C. 2017 Languages in drier climates use fewer vowels. *Front. Psychol.* **8**, 1285. (doi:10.3389/fpsyg.2017.01285)
- Ladefoged P, Maddieson I. 1996 *The sounds of the world's languages*. Hoboken, NJ: Wiley Blackwell.
- Maddieson I. 2013 Voicing and gaps in plosive systems. In *The world atlas of language structures online* (eds M Dryer, M Haspelmath). Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology.
- Bickel B, Nichols J. 2017 *The AUTOTYP genealogy and geography database*. See <http://www.autotyp.uzh.ch/>.
- Blasi D et al. 2019 Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* **363**, eaav3218. (doi:10.1126/science.aav3218)
- Hockett C. 1985 Distinguished lecture: F. Am. *Anthropol.* **87**, 263–228. (doi:10.1525/aa.1985.87.2.02a00020)
- Everett C, Chen S. 2021 Speech adapts to differences in dentition within and across populations. *Sci. Rep.* **11**, 1066. (doi:10.1038/s41598-020-80190-8)
- Henrich J. 2015 *The secret of our success*. Princeton, NJ: Princeton University Press.
- Grossman E, Eisen E, Nikolaev D, Moran S. 2020 SegBo: a database of borrowed sounds in the world's languages. In *Proc. 12th Language Resources and Evaluation Conf., May, Marseille, France*, pp. 5316–5322. Paris, France: European Language Resources Association.
- Dediu D, Moisik S. 2017 Anatomical biasing and clicks: evidence from biomechanical modeling. *J. Lang. Evol.* **2**, 37–51. (doi:10.1093/jole/lzx004)
- Gil D. 2013 Para-linguistic usages of clicks. In *The world atlas of language structures online* (eds M Dryer, M Haspelmath). Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology.
- 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)
- Urban M, Moran S. 2021 Altitude and the distributional typology of language structure: ejectives and beyond. *PLoS ONE* **16**, e0245522. (doi:10.1371/journal.pone.0245522)
- Lieberman P. 1968 Primate vocalizations and human linguistic ability. *J. Acoust. Soc. Am.* **44**, 1574–1584. (doi:10.1121/1.1911299)
- Fitch T, de Boer B, Mathur N, Ghazanfar A. 2016 Monkey vocal tracts are speech ready. *Sci. Adv.* **2**, e1600723. (doi:10.1126/sciadv.1600723)
- Boë LJ et al. 2019 Which way to the dawn of speech? *Sci. Adv.* **5**, 12. eaaw3916. (doi:10.1126/sciadv.aaw3916)
- Everett C. 2017 Yawning at the dawn of speech: response to Fitch et al. (2016). *Sci. Adv.* eLetter. See <https://advances.sciencemag.org/content/2/12/e1600723/tab-e-letters>.
- Easterday S. 2019 *Highly complex syllable structure: a typological and diachronic study*. Berlin, Germany: Language Science Press.
- 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)
- Sundarrajan A, Brinton Fujiki R, Loerch S, Venkatraman A, Sivasankar P. 2017 Vocal loading and environmental humidity effects in older adults. *J. Voice* **31**, 407–413. (doi:10.1016/j.jvoice.2017.02.002)
- 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)
- Roberts S. 2018 Robust, causal, and incremental approaches to investigating linguistic adaptation. *Front. Psychol.* **9**, 166. (doi:10.3389/fpsyg.2018.00166)
- Maddieson I. 2017 Language adapts to environment: sonority and temperature. *Commun.* **3**, 28. (doi:10.3389/fcomm.2018.00028)
- Fought J, Munroe R, Fought C, Good E. 2004 Sonority and climate in a world sample of languages. *Cross-Cult. Res.* **38**, 27–51. (doi:10.1177/1069397103259439)

52. Everett C, Blasi D, Roberts S. 2015 Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *Proc. Natl Acad. Sci. USA* **112**, 1322–1327 +S1. (doi:10.1073/pnas.1417413112)
53. Everett C, Blasi D, Roberts S. 2016 Language evolution and climate: the case of desiccation and tone. *J. Lang. Evol.* **1**, 33–46. (doi:10.1093/jole/lzv004)
54. Collins J. 2016 Commentary: the role of language contact in creating correlations between humidity and tone. *J. Lang. Evol.* **1**, 46–52. (doi:10.1093/jole/lzv012)
55. 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**, 10 944–10 949. (doi:10.1073/pnas.0610848104)
56. Wong P, Kang X, Wong K, So H, Choy K, Geng X. 2020 ASPM-lexical tone association in speakers of a tone language: direct evidence for the genetic-biasing hypothesis of language evolution. *Sci. Adv.* **6**, eaba5090. (doi:10.1126/sciadv.aba5090)
57. Gray R, Drummond A, Greenhill S. 2009 Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479–483. (doi:10.1126/science.1166858)
58. Gray R, Jordan F. 2000 Language trees support the express-train sequence of Austronesian expansion. *Nature* **405**, 1052–1055. (doi:10.1038/35016575)
59. Francois A. 2014 Trees, waves and linkages: models of language diversification. In *The Routledge handbook of historical linguistics* (eds C Bower, B Evans), pp. 161–189. London, UK: Routledge.
60. Bickel B. 2017 Areas and universals. In *Cambridge handbook of areal linguistics* (ed. R Hickey), pp. 40–54. Cambridge, UK: Cambridge University Press.
61. Coupé C, Oh Y, Dediu D, Pellegrino F. 2019 Different languages, similar encoding efficiency: comparable information rates across the human communicative niche. *Sci. Adv.* **5**, eaaw2594. (doi:10.1126/sciadv.aaw2594)
62. Kemp C, Xu Y, Regier T. 2018 Semantic typology and efficient communication. *Annu. Rev. Linguist.* **4**, 109–128. (doi:10.1146/annurev-linguistics-011817-045406)
63. Gibson E *et al.* 2019 How efficiency shapes human language. *Trends Cogn. Sci.* **1894**, 1–19. (doi:10.31234/osf.io/w5m38)
64. Piantadosi S, Tily H, Gibson E. 2011 Word lengths are optimized for efficient communication. *Proc. Natl Acad. Sci. USA* **108**, 3526–3529. (doi:10.1073/pnas.1012551108)
65. Gibson E *et al.* 2017 Color naming across languages reflects color use. *Proc. Natl Acad. Sci. USA* **114**, 10 785–10 790. (doi:10.1073/pnas.1619666114)
66. Bentz C, Dediu D, Verkerk A, Jäger G. 2018 The evolution of language families is shaped by the environment beyond neutral drift. *Nat. Hum. Behav.* **2**, 816–821. (doi:10.1038/s41562-018-0457-6)
67. Nolle J, Staib M, Fusaroli R, Tylen K. 2018 The emergence of systematicity: how environmental and communicative factors shape a novel communication system. *Cognition* **181**, 93–104. (doi:10.1016/j.cognition.2018.08.014)
68. Raviv L, Meyer A, Lev-Ari S. 2019 Larger communities create more systematic languages. *Proc. R. Soc. B* **286**, 20191962. (doi:10.1098/rspb.2019.1262)
69. Labov W. 1994 *Principles of linguistic change: internal factors*. Hoboken, NJ: Wiley-Blackwell.
70. Crowley T, Bower C. 2010 *An introduction to historical linguistics*. Oxford, UK: Oxford University Press.