Language is not isolated from its wider environment: Vocal tract influences on the evolution of speech and language

CrossMark

Dan Dediu [a,b,*], Rick Janssen [a], Scott R. Moisik [a,c]

[a] Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, Nijmegen, The Netherlands
[b] Donders Institute for Brain, Cognition and Behaviour, Kapittelweg 29, Nijmegen, The Netherlands
[c] Division of Linguistics and Multilingual Studies, Nanyang Technological University, 50 Nanyang Avenue, Singapore, Singapore

ARTICLE INFO

ABSTRACT

Language is not a purely cultural phenomenon somehow isolated from its wider environment, and we may only understand its origins and evolution by seriously considering its embedding in this environment as well as its multimodal nature. By environment here we understand other aspects of culture (such as communication technology, attitudes towards language contact, etc.), of the physical environment (ultraviolet light incidence, air humidity, etc.), and of the biological infrastructure for language and speech. We are specifically concerned in this paper with the latter, in the form of the biases, constraints and affordances that the anatomy and physiology of the vocal tract create on speech and language. In a nutshell, our argument is that (a) there is an under-appreciated amount of inter-individual variation in vocal tract (VT) anatomy and physiology, (b) variation that is non-randomly distributed across populations, and that (c) results in systematic differences in phonetics and phonology between languages. Relevant differences in VT anatomy include the overall shape of the hard palate, the shape of the alveolar ridge, the relationship between the lower and upper jaw, to mention just a few, and our data offer a new way to systematically explore such differences and their potential impact on speech. These differences generate very small biases that nevertheless can be amplified by the repeated use and transmission of language, affecting language diachrony and resulting in cross-linguistic synchronic differences. Moreover, the same type of biases and processes might have played an essential role in the emergence and evolution of language, and might allow us a glimpse into the speech and language of extinct humans by, for example, reconstructing the anatomy of parts of their vocal tract from the fossil record and extrapolating the biases we find in present-day humans.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

[1]There is no doubt that language is a cultural phenomenon continuously evolving[2] (Dediu et al., 2013), a complex process influenced by a multitude of factors, most internal to language itself (for example, morpho-syntactic changes driven by reinterpretation (Bybee, 2001), phonetic by-products becoming important due to the erosion of other sounds, etc.; Campbell,

---

* Corresponding author. Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, Nijmegen, The Netherlands.
 E-mail address: Dan.Dediu@mpi.nl (D. Dediu).
 [1] Abbreviations: VT = vocal tract.
 [2] There is an intriguing ambiguity in "language evolution" – is it cultural? is it biological? or is it both? – to which we return in Section 4.

2004), but some having to do with the *wider context* of language use. While the first category is extensively studied and numerous general principles and fascinating particular cases have been discovered (Campbell, 2004; Hopper and Traugott, 2003), the latter is much broader and much less studied, but promises more articulated causal explanations better integrated with other scientific disciplines. In this paper we will briefly review some of the proposals currently available, with a particular focus on the anatomy and physiology of the vocal tract and their influence on phonetics and phonology, concluding with some suggestions about the origins and evolution of language in the auditory-vocal modality.

2. Language in its wider context

More and more, language is seen as a complex entity that has cultural, social and biological components in intimate interactions (Dediu et al., 2013), leading to a fascinating interplay between cultural and biological evolution. Also, the increasing realization that language is intrinsically multimodal (see, for example, the theme issue "Language as a multimodal phenomenon" of Phil. Trans Royal. Soc. B from 14 September, 2014, vol. 369, issue 1651; Wacewicz et al., 2016), involving simultaneously multiple channels (voice, face, hands), plays a major role in the proper understanding of language and its evolution.

There are several well-known examples of *gene-culture co-evolution* (Richerson et al., 2010) and *(cultural) niche construction* (Laland et al., 2010) concerning non-linguistic phenomena, such as the co-evolution of lactose tolerance and farming (Gerbault et al., 2011).[3] However, when it comes to language probably the most convincing cases of such a full co-evolutionary cycle are sign languages and possibly the descent of the larynx (Lieberman, 2012; Fitch, 2010). In this section we will briefly review some recent proposals linking language to its wider environment, starting with congenital hearing loss, moving to the way the physical and social environments experienced by the speakers affect their language(s), and ending with the possible influence of neuro-cognitive, perceptual and even genetic characteristics of the speakers themselves.

2.1. The cultural evolution of language

Culture is a full-blown evolutionary system in its own (Richerson and Christiansen, 2013) and language in particular is increasingly viewed from such a perspective (Tamariz and Kirby, 2016; Dediu et al., 2013; Pagel, 2009; Atkinson and Gray, 2005; Croft, 2000). Thus, many of the properties of language, including its universal tendencies and its amazing diversity (Evans and Levinson, 2009), ultimately emerge from its repeated use, learning and processing (within and across generations) in a range of social contexts, but of particular importance here is the capacity of cultural evolution to *amplify very weak biases*. At face value, it is easy to believe that such defining properties as the duality of patterning and compositionality must certainly reflect strong innate biases, but several lines of ongoing research, ranging from computer simulations (Kirby et al., 2007; Dediu, 2008; Smith and Kirby, 2008; Thompson et al., 2016), to animal models (Fehr et al., 2009) and human participants (Kirby et al., 2008; Caldwell and Smith, 2012; Smith and Wonnacott, 2010) clearly show that this is not the case.

The repeated use and transmission of culture (and, in particular, language) in a population of agents that have only very weak biases can result, under the right circumstances, in the amplification of these biases out of proportion, creating a strong emergent patterning of the resulting communicating system. A striking consequence of this is that the nature and strength of the biases cannot be directly inferred from the observed properties of the emergent communicative system (Kirby et al., 2008): a strong "universal" property of language does not immediately imply a strong universal bias in humans. A second important consequence concerns the locus and nature of these biases: while they might be part of the biological infrastructure of language (but even this apparently simple statement hides incredible complexities; Fisher and Vernes, 2015; Fisher, 2006; Dediu, 2015), this is not necessarily the case and they may result from properties of the wider environment in which language is used, and, in the following sections, we will encounter different types of such biases. Third, if the bias is universally present it might play a role in explaining the "design features" and universal tendencies of language, but if different human populations show different such biases (due to the different environments they inhabit or due to genetic differences), or with a different strength and/or direction, then this might help explain, through cultural evolution, patterns of cross-linguistic diversity (Dediu, 2011).

2.2. Sign languages and their co-evolution with genes

There are several communities around the world where the high incidence of recessive mutations resulting in congenital non-syndromic deafness coupled with inclusive attitudes towards the deaf members of the community have resulted in the *de novo* emergence of sign languages (famous examples are *Kata Kolok* on the island of Bali, Indonesia; De Vos, 2012 and *Al-Sayyid Bedouin Sign Language*, or ABSL, in Israel; Sandler et al., 2005; see also Meir et al., 2010). While the particular genetic mutations and socio-cultural practices involved are different, there seem to be certain commonalities (Dediu, 2015; Gialluisi et al., 2013) including the recessiveness of the mutation (making it very hard to predict which parents will have deaf children), its localized and early effects (hearing loss only present at birth), the (more or less complete) social acceptance of the deaf persons within the community, and social arrangements that increase the frequency of the mutation in the community (tendency to marry close relatives or people from within the village). When such conditions are maintained for several generations, the initially simple spontaneous *homesigns* of the first generation(s) of deaf children become gradually more

---

[3] For more examples see Dediu (2015, chapter 9) and Jobling et al. (2013).

complex, structured and communicatively effective across generations (Sandler et al., 2005). This is due to the repeated use of the emerging language in multiple contexts by multiple individuals (both deaf – native language users – but also hearing – second-language users) from the same or different generations (Levinson and Dediu, 2013; Dediu, 2015).

These *emergent* (or *village*) sign languages have fascinating characteristics and temporal dynamics, showing minimal influence from the community's spoken language(s) or the national sign languages that might be available (Meir et al., 2010; De Vos, 2012), and the gradual development of a phonology (Sandler et al., 2005). However, a very important aspect of the emergence of such languages is that, on the one hand, the higher frequency of the recessive mutation is a causal factor in the emergence of the sign language, while the sign language, in turn, facilitates the normal socialization and integration of the deaf members in the community, shielding them from the otherwise negative impact (in terms of a lower biological fitness due to discrimination and social exclusion) of hearing loss (see, for example, Gialluisi et al., 2013; Aoki and Feldman, 1991; Levinson and Dediu, 2013).

This feedback loop between the genetics of hearing loss and the emergent sign language is probably one of the best cases of gene-culture co-evolution currently known for language, but these dynamics are also at work (at least partially) in less extreme situations as well. The celebrated *Nicaraguan Sign Language* (Senghas, 1995) shows that language emergence is a general phenomenon when the right socio-cultural conditions are present, no matter what the biological cause of hearing loss is, but more generally, the attitudes towards deafness in the larger society can be shown to have an impact on the population genetics of hearing loss. For example, the introduction of the *American Sign Language* (ASL) and the various institutions/ schools for the deaf in the early 19th century US have resulted, through the better social integration, opportunities, assortative mating (more frequent inter-marriage between deaf individuals) and linguistic homogamy (marriage between signers even if one is a hearing child of deaf parents), in the relaxation of selection against recessive mutations causing hearing loss. This, in a feedback cycle, allowed the frequency of these mutations and of ASL use to increase (Nance et al., 2000; Arnos et al., 2008).

2.3. Physical environment and phonetics–phonology

It seems almost obvious that features of the physical environment in which speech is produced, transmitted and perceived must play a role in influencing the sounds of language, just as they do when it comes to non-human animal communication (Everett et al., 2016), but there are currently very few concrete proposals and even fewer can be considered as being convincing. Everett (2013) has recently suggested that the altitude at which a language is spoken is related to the presence or absence of ejective consonants, the main causal factor being atmospheric pressure (thus, higher altitude results in a lower atmospheric pressure, which eases the production of ejectives); while very interesting, this proposal is widely open to criticism as the statistical association adduced in its support is based on debatable classifications of languages into altitude groups and it is not entirely clear how a lower atmospheric pressure might actually favor the production of ejectives.

A more recent proposal involving the same author (Everett et al., 2015) suggests that air humidity influences the presence or absence of linguistic tone, with the causal link being that air dryness results in reduced control of the vocal folds, causing reduced control of pitch production, and thus an absence of (complex) tone. While this proposal seems supported by a much better database, better statistical analyses and a more fleshed-out mechanistic explanation, not everybody is currently convinced, with issues concerning the classification of tone, the statistical control for language history and contact, the relationship between tone and intonation, and the validity of the proposed physiological mechanism (see the debate in the 1st issue of the *Journal of Language Evolution*, January 2016).

Finally, developing the observation that the environment impacts non-human animal communication systems in predictable ways, Maddieson and Coupé (2015) propose that in environments that do not transmit high frequency sounds well (e.g., dense vegetation or high temperature), languages will tend to be "more sonorous", affecting the ratio of sonorant to obstruent segments. This proposal is currently being tested using extensive databases and advanced statistical methods, and the effect seems to be robust (Christophe Coupé, p.c. march 2016), supporting the causal reasoning.

2.4. Physical environment and the lexicon

It is not just the phonetics and phonology that can be influenced by environmental factors, but also a language's lexicon: Lindsey and Brown (2002) and Brown and Lindsey (2004) have proposed that the amount of ultraviolet light (UV-B) impacting upon the eyes can negatively affect the perception of the blue/green distinction (phototoxicity). Given that the incidence of UV-B strongly correlates with latitude, being much higher at the equator than at the poles, this mechanism would (partly) explain why languages at lower latitudes tend to lack a word for "blue" from their color lexicon more often than languages at higher latitudes. Normal colour perception in humans relies on three types of cone cells in the retina, sensitive to short wavelength ("blue") light, to medium wavelength ("green") light, and to long wavelength ("red") light, respectively (Deeb, 2006). In an interesting twist, the more frequent loss of the blue/green distinction closer to the equator generates stronger selective pressures against also losing the red/green distinction in these regions compared to the higher latitudes, as not having both distinctions simultaneously results in a massive impact on visual perception. This proposal might explain the increased incidence of anomalous colour perception involving red and green with increasing distance from the equator, and is supported by the statistical analysis of a relatively large dataset[4] but also by a clear physiological mechanism as well as some experimental evidence.

---

[4] We are currently trying to extend this database and apply better controls for language contact and genetic relationships.

2.5. Social structure and language

Features of the language users' societies will affect the lexicon (see, for example, the elaborated kinship terms in many Australian languages; Dousset, 2011) but also other aspects of language. Extending Wray and Grace (2007)'s dichotomy *exoteric* (large speaker base, including non-natives) – *esoteric* (closed communities mainly of native speakers), Lupyan and Dale (2010) and Dale and Lupyan (2012) introduce the *Linguistic Niche Hypothesis* (namely the proposal that languages tend to reflect the social environments in which they are learned and used by adapting to their learners' constraints and biases; Dale and Lupyan, 2012) with a particular emphasis on the relationship between population size and morpho-syntactic complexity. More precisely, the statistical analysis of a large typological database (Lupyan and Dale, 2010) coupled with agent-based computer models (Dale and Lupyan, 2012) strongly support this hypothesis, also suggesting that the causal mechanism is represented by the difficulty of non-native adult learners to acquire a complex foreign morpho-syntax. In turn, this results in the tendency of languages with a high proportion of non-native speakers to have a less complex morpho-syntax.

There are various proposals that there is also a correlation between the phonological inventory of a language and its speaker base (Hay and Bauer, 2007; Atkinson, 2011). For instance, Atkinson (2011) suggests that when a smaller daughter population splits off a source population, some of the phonemic diversity is lost in the process, but this explanation (paralleling the founder effect in population genetics) does not seem to work (Cysouw et al., 2012). Moreover, when using a larger database and language relatedness is properly controlled for, the reported correlations between population size and phonological complexity seem to vanish (Moran et al., 2012; Cysouw et al., 2012).

A recent proposal by Bromham et al. (2015) connects population size and the rate of gain and loss of basic vocabulary cognates in a set of Polynesian languages, and shows that larger populations tend to gain new words (i.e., introduce a new lexeme for a semantic category that is not cognate with any other member of the family) while smaller ones tend to lose existing ones (i.e., having lost an ancestral cognate present in other members), suggesting that at work are mechanisms parallel to the population genetic processes (Dediu, 2015) of positive selection (more effective in large populations) and random drift (more effective in small populations). While extremely interesting, it remains to be seen if these results generalize outside Polynesia, to a larger part of the lexicon, and even to other aspects of language.

2.6. Neuro-cognitive processing and language

Our neuro-cognition clearly provides constraints and affordances that shape language, as argued in a landmark paper by Christiansen and Chater (2008), that in many respects language must be seen as being "shaped by the brain". This proposal is very general, but there are also some fascinating case studies. An example is the *Neuronal Recycling Hypothesis*, which suggests that features shared by writing systems across the globe (such as lines meeting at junctions that do not depend on the viewing angle) result from our brain's pre-existing functional specialization for visual processing (Szwed et al., 2009, 2011). More precisely, the so-called *visual word form area* (VWFA) in the fusiform gyrus is being ontogenetically re-purposed for the processing of written language during the acquisition of literacy, but there has not been enough time for biological evolution to have specifically tuned it for this particular task (writing was invented in the last 6000 years or so, but its widespread use is much more recent), meaning that its pre-existing properties impose constraints and biases on its new functions related to processing written language.

However, probably the best example to date is the recent proposal that neuro-cognitive processing generates biases against the *ergative* case marking (Bickel et al., 2015). Briefly, languages such as English mark the subject (S) of an intransitive verb ("he" in "he slept") and the agent (A) of a transitive verb ("he" in "he sold a book") in the same way, while other languages such as Hindi mark the transitive verb agent with a special ergative marker. Interestingly, a statistical analysis of diachronic tendencies across 617 languages shows that there is a bias acting against ergative marking (i.e., it is unlikely for a language without ergative marking to gain it, but likely for a language with ergative marking to lose it) on the timescale required for language families to form (hundreds to thousands of years). This large-scale statistical tendency is further supported by experimental evidence showing that even the speakers of a language with ergative case marking (Hindi) show a bias towards assuming initially that the agent comes first in a sentence, requiring a later re-interpretation/ disambiguation (generating a biphasic N400 + P600 effect in an Event-Related Potentials paradigm). Such type of research, featuring a strong linguistically-motivated hypothesis, advanced statistical analyses on large databases and experimental support, should become the norm as it allows a better articulation of the causal mechanisms involved (Ladd et al., 2015).

2.7. Auditory perception and language

While total hearing loss can sometimes engender fascinating examples of (emergent) sign languages (Section 2.2), more subtle changes in auditory perception that affect a sizeable proportion of a speech community for a long time can also influence phonetics and phonology. A convincing example is due to Butcher (2006), who suggests that *Chronic Otitis Media* (COM), a severe chronic infection of the middle ear usually resulting in conductive hearing loss with decreased sensitivity to low frequency ($\leq$500 Hz) but also sometimes to high frequency ($\geq$4000 Hz) sounds, might play a role in explaining intriguing features of the phonetics and phonology of Australian Aboriginal languages. More precisely, COM seems to affect a large

number of Australian Aboriginal children, resulting in populations with a significant proportion of speakers with altered hearing sensitivity, a situation apparently persisting for an extremely long historical time. In turn, this forced Australian languages to adapt to this new "auditory fitness landscape" by avoiding vowels with formant frequencies below approximately 500 Hz, voicing contrasts and fricatives, while exploiting the intact frequency range, with, for example, many place-of-articulation distinctions. This hypothesis is extremely intriguing but requires more experimental work, an investigation of other populations with a high incidence of COM, and a better understanding of the temporal depth and population patterns of COM incidence in Australia.

2.8. Do genes affect language?

An essential aspect of the wider environment of language is represented by its speakers' biology, comprising their brains, their ears, their lungs and larynges, their hands and eyes, among others; importantly, all these have ultimately a *genetic component* (Dediu, 2015). All we currently know from modern genetics, evolutionary and developmental biology, the neurocognitive sciences and linguistics, points away from any sort of biological determinism and towards a complex and fascinating interplay between genetic information and the multi-faceted environment (Dediu, 2015; Dediu and Christiansen, 2016). In this view, apparently intuitive and simple concepts such as "innate" and "learned" reveal unexpected subtleties (Mameli and Bateson, 2006), and the causal connections between the molecular and the linguistic levels are not by any means obvious nor direct (Fisher and Vernes, 2015). Nevertheless, our genome influences language in profound ways, as revealed by the outcome of various attempts at training non-human animals to use language[5] on one hand, and by the multitude of various pathologies with a genetic component affecting language (see for example OMIM, 2016; Dediu, 2015), on the other.

But how about the vast amounts of inter-individual *normal genetic variation*? Do these affect inter-individual variation in language and speech, and, if so, how exactly? Currently there are few conclusive studies, not least because of the practical difficulties in amassing the large samples required, in devising and applying appropriate measurement instruments, and because of the intrinsic complexity of the genetic bases of normally-functioning speech and language (see Dediu, 2015; Fisher and Vernes, 2015). However, those that do exist, coupled with insights from other cognitive domains such as intelligence (Plomin et al., 2013) and even "simple" but easily measurable phenotypes such as height (Visscher et al., 2010), suggest that at least some aspects of the variation between normal individuals in speech and language have a genetic component.

The next step is based on the species-wide patterning of genetic variation in present-day humans: due to our relatively recent origins and expansion from Africa (Barbujani and Colonna, 2010) with limited admixture with other (probably language-using) varieties of humans (Dediu and Levinson, 2013), there are relatively few genetic differences between various human populations. Nevertheless, such differences do exist, forming continuous multi-variate patterns across multiple genetic loci, with genetic diversity decreasing with distance from Africa; interestingly, these differences are mostly distributed among individuals from the same population, with a very small proportion attributable to variation between continents (Barbujani and Colonna, 2010; Jobling et al., 2013). The existence of (small) genetic differences between populations raises the possibility that they might also play a role in explaining aspects of linguistic diversity.

Dediu and Ladd (2007) make the case that the distribution of *tone languages* is partly due to genetic differences between populations at two loci involved in brain growth and development (*ASPM* and *MCPH1*). The authors bring statistical support for this proposal showing that in a relatively large database of populations widely distributed across the world, the population frequency of specific alleles at these two loci predicts the presence of linguistic tone better than expected by chance when compared to a large set of genetic loci and linguistic structural features. The suggested mechanism (Dediu and Ladd, 2007; Ladd et al., 2008) is based on the *amplification of biases* that are very weak at the individual level but that are intensified by means of language use and transmission across generations in populations of similarly-biased individuals, resulting in population-wide effects (Dediu, 2011). This process of *genetically-biased cultural transmission of language* is the causal mechanism through which weak biases may result in large-scale observable patterns, either in terms of cross-linguistic differences (when populations differ in the strength and/or the direction of the bias) or in universal tendencies (where all human populations share the same bias).

While intriguing, this particular proposal relating linguistic tone and *ASPM* and *MCPH1* is currently far from having been established or definitely rejected. On the positive side, the inter-population statistical correlation in Dediu and Ladd (2007) seems robust (but more data is required, covering more populations, especially in the Americas and Australia), the mechanism of cultural amplification of weak biases in certain conditions is supported both by computer models (Kirby et al., 2007; Dediu, 2008) and experimental approaches (Kirby et al., 2008; Smith and Wonnacott, 2010), and Wong et al. (2012) found an association between *ASPM* and pitch processing in normal individuals. On the negative side, Wong et al. (2012) is clearly underpowered and the effect is apparently in the opposite direction from the predicted one,[6] the distribution of tone

---

[5] It seems a human genome is required at some level, as even after intensive training non-human primates have a low command of language (see, for example, Koko the gorilla Patterson and Cohn, 1990; Perlman et al., 2012).

[6] However, a different interpretation of the tasks as actually measuring intonation and not tone, suggests that the effect is actually in the predicted direction; see Caldwell-Harris et al. (2015) for more details.

languages might be due to non-genetic climatic factors (Everett et al., 2015), but probably the most important issue, in our view, is that – despite a series of experiments exploring various candidates such as the perception of the missing fundamental[7] (Ladd et al., 2013) and a segmentation task where tone distinguishes words from non-words (Caldwell-Harris et al., 2015) – it is not clear what precisely the postulated bias is, how it might affect tone and what are the molecular mechanisms linking *ASPM* and *MCPH1* to it.

In conclusion, while the jury is still out on this particular proposal, the general idea of genetic biases affecting language and speech seems worth pursuing (Dediu, 2011), but a more fruitful approach would be to focus on aspects of speech and language that are mechanistically simpler, easier to quantify and to measure reliably, such as articulatory phonetics.

2.9. Summing up

As this very brief review of some recent proposals shows, the idea that non-linguistic factors, ranging from genetics to perception to neuro-cognition to climate and social structure might affect language change, resulting in observed linguistic properties, is an active area of current research. However, most of these proposals are in their early stages and need to be replicated, extended and supported with experimental evidence in order to not only have confidence in their reality but also to have a complete understanding of the complex and fascinating causal mechanisms behind them (Ladd et al., 2015). Nevertheless, firmly anchoring language in its wider environment is essential for a proper science (or a set of sciences) of language that fits seamlessly in the larger scientific landscape.

3. Genes, the vocal tract, and linguistic diversity

The production of speech is an extremely complex process that crucially involves the precise coordination of several organs such as the lungs, the larynx, the tongue, the soft palate, the lower jaw and the lips (Gick et al., 2013). Leaving aside for a moment the neuro-cognitive component and focusing on the articulators themselves, these are usually idealized, being conceived as essentially uniform across our species (at least in what concerns their properties relevant for speech). However, like anything else biological, there is a massive amount of inter-individual and even inter-population differences in almost any aspect of the vocal tract one cares to look at, but, in general, this variation is not considered relevant for "normal" speech and language, with attention being devoted only to "pathological" variation such as cleft palate or too big or too small a tongue (macro- and micro-glossia, respectively). So, how much variation is there in vocal tract structures and how is it geographically patterned? Does it have a genetic basis? And does it have any effects on speech production at all?

In this section we will briefly overview what is known about these questions focusing on a few examples that our team is currently investigating and that might ultimately help generate better models of the origins and evolution of language.

3.1. Patterns of variation in the vocal tract

Speech is produced using structures of the *vocal tract* (VT), such as the larynx, the pharynx, the oral and nasal cavities (Gick et al., 2013). There has been a lot of interest in pathological variation, its causes (including genetics), its effects on speech and language development, and possible therapeutic approaches, especially in the medical literature. For example, *cleft lip and/or palate* (CLP) has been extensively investigated both as part of larger syndromes and as isolated (non-syndromic) phenomena, with a special accent on their (genetic) causes (e.g., Dixon et al., 2011; OMIM, 2016), so that we have currently a good understanding of the genetic networks involved in the development of the palate and lips; a complementary view is provided by evo-devo studies bringing a larger perspective on these genetic networks and an interest in normal development as well (Cai et al., 2005). By contrast, there is comparatively little work looking at the patterns of normal variation in VT structures, such as, for example, in the soft palate (Kumar and Gopal, 2011; You et al., 2008), the hard palate both within and between populations (Lammert et al., 2013; Byers et al., 1997), the tongue (Harshman et al., 1977), and the longitudinal development of various vocal tract parameters (Fitch and Giedd, 1999; Vorperian et al., 2005). Nevertheless, even this rather limited body of systematic work coupled with clinical observations and the experience of practising phoneticians[8] suggests that there is a lot of under-appreciated variation in the normal population concerning practically all parameters of the vocal tract. An intriguing example of such inter-population variation (to which we will return later) is given by the shape of the *alveolar ridge* in speakers of the click languages of South Africa (Moisik and Dediu, 2015).

3.2. The causes of VT variation

Besides strong findings from the medical and evo-devo literatures of a genetic basis for vocal tract development (Dellavia et al., 2007; Dixon et al., 2011; Reynoso et al., 1994; Anthwal et al., 2013), there are also indications that the inter-individual

---

[7] The phenomenon whereby a pair of sounds, each composed only of a set of higher harmonics but not their fundamental frequency, in such as way that the (missing) fundamental frequency and the (present) higher harmonics are ordered differently, are perceived differently by different individuals or by the same individual in different conditions (see Ladd et al., 2013 for a clear explanation with actual examples).

[8] For example, dental casts acquired during dentistry procedures or palatographic experiments.

patterns of variation in VT structures also have a genetic component (Patel et al., 2008; Bowden and Goose, 1968; Schwab et al., 2006; Townsend et al., 1990), suggesting that a fruitful approach might be to conduct large genetic association studies in the normal populations either testing candidate genes involved in VT pathological variation and development, or doing unbiased genome-wide scans (Lewis and Knight, 2012). However, we should never fall into the trap of genetic determinism and ignore the important role of the environment, with work such as von Cramon-Taubadel (2011)'s being powerful reminders of the role cultural practices (such as the consistency of food) play in shaping structures of the vocal tract.

3.3. Does VT variation affect speech?

Even if there is an impressive amount of inter-individual and inter-population variation in vocal tract anatomy, does it actually affect speech? Or are we perfectly capable of completely over-riding these anatomical differences (through *compensation*) and producing the same, anatomy-independent phonetic output? If the answer were "yes, compensation is perfect", then the phoneticians' traditional lack of interest in anatomical variation would be justified. However, it is becoming clearer and clearer that, while being able to compensate (to a certain extent) even for major anatomical differences such as missing the tongue (Gerdeman and Fujimura, 1990) or for the artificial manipulation of hard palate shape (Brunner et al., 2006), even small anatomical differences may result in different articulatory strategies and affect the acoustic output (Brunner et al., 2009; Hiki and Itoh, 1986; Lammert et al., 2011; Tiede et al., 2004; Weirich, 2010; Weirich and Fuchs, 2011; Zhou et al., 2007).

Thus, taken together, these suggest the possibility that inter-individual differences in the anatomy of the vocal tract, with a genetic component, might be able to influence the articulatory behavior and even the acoustic output, contributing to inter-individual phonetic variation. Even more, given the apparent existence of stable patterns of variation in vocal tract anatomy between geographic populations, it is possible that such individual phonetic differences become amplified and "fixed" by language use and transmission, influencing not only the phonetics but also the phonology of the world's languages. In the following we will focus on a few cases that we are currently exploring.

An important issue in this context is, on one hand, the relationship between *speech* and *language*, and, on the other, between *phonetics* and *phonology*. Clearly language is different from speech (as shown by sign languages and by the multimodality of spoken language) but nevertheless, at least now and presumably for a very long time in our prehistory (Dediu and Levinson, 2013), speech is (and was) the main medium of language, and speech and language are (and were) continuously shaping each other (Dor, 2015). Likewise, the relationship between phonetics and phonology is far from clear but they are not identical and there is certainly scientific utility in keeping both perspectives in mind when approaching speech sound systems (Archangeli and Pulleyblank, 1994; Scobbie, 2005; Ohala, 2005), and it could be argued that while the properties of the vocal tract might affect phonetics, they are irrelevant to phonology, it being concerned with cognitive, abstract, meaning-carrying contrasts. However, phonological contrasts are implemented in phonetics and, importantly, they emerge, disappear and change with changing phonetic details, as shown by many examples of sound change driven by phonetic factors (Campbell, 2004) such as the emergence of tone (Yip, 2002; Hyslop, 2009). This means that processes that affect phonetics (such as constraints and affordances due to vocal tract anatomy and physiology) have the potential to create the preconditions for sound change resulting in changes to the phonological system through repeated cycles of language use and transmission (Levinson and Dediu, 2013).

3.4. The G[3]bils project

The "Genetic biases in language and speech" (G[3]bils) project aims to quantitatively investigate patterns of inter-individual and inter-population variation in components of the vocal tract that may play a role in cross-linguistic phonetic and phonological diversity. The main idea behind the project is that stable anatomical variation in vocal tract structures (mostly quantitative in nature) between geographic populations (due, for example, to genetic differences, different cultural practices or environmental factors) might result in very weak biases at the individual level that can, nevertheless, be amplified by language use and transmission in populations of similarly-biased speakers (Dediu, 2011; Janssen and Dediu, in press). To this end, we are pursuing a multi-pronged and multi-disciplinary approach, with the main directions being:

(a) the quantitative modeling of anatomical variation in vocal tract structures,
(b) the use of computer models to systematically study the effects of anatomical variation and population structure on phonetics and phonology,
(c) the acquisition of primary structural, articulatory and acoustic data from a large multi-ethnic sample, and
(d) the statistical analysis of correlations between patterns of anatomical, genetic and linguistic variation in large databases.

As an example of quantitative modeling of anatomical variation (a), we (Janssen et al., 2015) have developed a model of the midsagittal shape of the human hard palate based on Bézier curves. We used published and original midsagittal scans of hard palates to fit a Bézier model and extensive tests showed that not only our model fits the data at least as well as the classic Principal Component Analysis approach (e.g., Lammert et al., 2013) with as few as two or three easy-to-interpret free parameters, but that our model can generate novel realistic hard palate shapes much better than alternative models. Therefore,

we can summarize an individual's hard palate shape with a small set of meaningful variables, making the statistical analysis of large-scale databases of inter-individual variation possible.

Moreover, we have implemented this Bézier model of the hard palate in the widely-used VocalTractLab (a geometric computer model of the vocal tract that can generate acoustic output given the position of the articulators; http://www.vocaltractlab. de; Birkholz, 2005), with the aim of studying the influence of hard palate shape on the acoustic behavior of computer agents (Janssen et al., 2016). We are currently using this Bézier-extended-VTL to implement populations of agents with known anatomy and experimenting with language use and transmission across generations, in order to systematically explore the nature and strength of the influence of anatomical biases (here, hard palate anatomy) on the phonetics and phonology of the resulting artificial languages (b). We have also integrated a preliminary extension of the Bézier model to three dimensions (using a parabolic model of the coronal shape of the hard palate) in the biomechanical modeling system ArtiSynth (www.artisynth.org; Lloyd et al., 2012), with the aim of exploring the influence of hard palate shape on articulatory biomechanics.

In the *ArtiVarK* subproject (c), we collected data from a large sample of 96 participants (38 female, age range 18–61) coming from four widely-defined ethnic groups ("Chinese", "European", "North Indian", and "South Indian"), including detailed questionnaires about the languages they speak, their ethnic origins and orthodontic history, intra-oral high resolution 3D optical scans, and structural and real-time MRI scans during the production of various speech sounds of interest, including *click consonants*, to which we turn now.

The use of clicks as phonemes is famously circumscribed geographically to languages spoken mostly in southern Africa, and they have very interesting phonetic and articulatory characteristics (not to mention having been wrongly claimed to be remnants of primordial languages; Knight et al., 2003 but see Güldemann, 2007). An old observation, going back to Tony Traill (Traill, 1985), suggests that typical click language speakers tend to lack an alveolar ridge and this feature might, in turn, help with the production of certain classes of clicks. We decided to formally test this conjecture[9] using a three-pronged approach. First, we conducted a review of the relevant anthropological literature (Moisik and Dediu, 2015) and we confirmed, with much larger sample sizes than in Traill (1985), that the absence of alveolar ridge occurs at much higher frequencies in the populations speaking click languages than in other populations (again, this highlights the continuous nature of inter-population variation). Second, we used ArtiSynth (Lloyd et al., 2012) to implement models with various degrees of alveolar ridge prominence and to simulate the release of a generic alveolar click (Moisik and Dediu, 2015); the results of this simulation, while clearly in need of refinement, do support the conjecture that a less prominent alveolar ridge simultaneously reduces the muscular effort required and favors a faster release, presumably resulting in easier to produce and more acoustically salient clicks. Third, we are currently analyzing the statistical relationship between the shape of various components of the vocal tract and the production of clicks in our ArtiVarK sample and, while very preliminary, our results seem to indicate a significant link between the anatomy of the alveolar and post-alveolar areas and the easiness of click production (Dediu and Moisik, 2016). Taken together, these findings lend support to the idea that very small, quantitative differences between populations in the anatomy of the vocal tract (here, the alveolar ridge) can influence the articulatory and acoustic properties of a class of sounds (here, post-dental/alveolar clicks) facilitating their inclusion in, and their further elaboration as part of, the phonological inventory, resulting in cross-linguistic variation.

The final direction in our project (d) concerns the investigation of statistical associations between genetics, anatomy and linguistic diversity in large databases. While correlation famously is not causation, such studies are essential both as exploratory instruments and as components of hypothesis-testing approaches (Ladd et al., 2015) and we will pick here only one illustrative example currently being investigated. As discussed above, the genetic bases of various pathologies affecting the anatomy of the vocal tract are an active area of research, but the bases of normal variation are largely neglected. Therefore, we are trying to address this gap in our knowledge by conducting a large-scale genetic association study for normal anatomical variation. More precisely, we first placed a number of standardized landmarks (using a computer-assisted approached developed in house) on a large set of MRI structural scans from approximately 3000 healthy Dutch participants from the Cognomics Consortium (http://www.cognomics.nl), then we quantify their anatomy using classic and geometric morphometric (Zelditch et al., 2012) approaches, finally conducting candidate-gene (considering normal variants of genes involved in pathologies of the vocal tract collected from the primary literature and databases such as OMIM, http://www.omim.org, and FaceBase, https://www.facebase.org) and genome-wide genetic association studies. While the statistical power afforded by such sample sizes is not overwhelming, this study will allow a better estimation of the heritability of various vocal tract structures and the proposal of candidate genes for future replication studies and functional investigations in the wet lab.

4. Environment, genes, anatomy, and language evolution

Scientifically studying the origins and evolution of language is notoriously difficult, but we are now in an age where new data, methods and concepts make this more feasible then ever (Dediu and de Boer, 2015). It is interesting to note, in this context, the intrinsic ambiguity of "language evolution"[10]: it can be taken to mean the biological (phylogenetic) evolution of the capacity for language, or it can be interpreted to mean the cultural evolution of languages (language change), but probably

---

[9] We became aware of it thanks to Didier Demolin.
[10] See for example the recent post by Martin Haspelmath, https://dlc.hypotheses.org/894.

a much better way forward is to embrace the fact that both aspects are needed and that they evolve in concert (Dediu et al., 2013).

What can we learn about language evolution from the type of evidence and inquiry reviewed in this article? First and foremost is the fact that such an *embedded* view of language allows us to use varied cues coming from multiple disciplines to make inferences about the origins and evolution of speech and language: archeology, paleoanthropology, paleoecology, paleoclimatology, etc. are all relevant as they give us an understanding of the environment, biology and culture of past humans; evolutionary theory gives us the broad framework for asking such questions; genetics (modern and ancient) is relevant as we can infer something about the extinct humans' biology and neurocognition; animal models allow us to dissect components of the capacity for language; finally, the languages sciences are needed to understand not only the production, perception, processing, acquisition and transmission of modern languages by modern humans, but also the processes that generate the observed patterns of linguistic diversity (*contra* Hauser et al., 2014). Such an encompassing view of language, using a sort of Bayesian view of science where all the evidence available is rationally weighted against explicit prior as-sumptions resulting in probabilistic conclusions, allows us to consider the possibility that language and speech are very old, that they evolved gradually through a continuous co-evolution between the biological underpinnings and language as a cultural entity (and most probably continue to evolve to this day), and that other forms of humanity such as the Neandertals and Denisovans also probably had recognizably modern (but of course not identical to our own) speech and language (Dediu and Levinson, 2013).[11]

Trying to apply our understanding of the ways in which the anatomy and physiology of the present-day human vocal tract influence phonetics and phonology to questions about language evolution has a long history (it suffices to mention here the classic inquiry into the position of the larynx in Neandertals and its putative influence on vowel production; Lieberman and Crelin, 1971), but the type of data, methods and inferences allowed by the renewed interest in anatomical biasing will hopefully allow much finer-grained, reliable and valid inferences about the languages of past humans. For example, if indeed the shape of the anterior part of the hard palate does facilitate the inclusion of clicks in the phonological inventory, or if the overall shape of the hard palate impacts the production of retroflex consonants, then paleoanthropological data might allow us to infer subtle details about the phonetics and phonology of long-gone languages. Of course, the fossil record conserves at best the hard structures of the vocal tract (hard palate, lower jaw, dentition, hyoid bone, cervical vertebrae) – and even here the findings tend to be rare and fragmentary – but we can, at least in principle, extrapolate back in time, using our under-standing of modern humans and non-human primates, to generate informed probabilistic distributions of the soft tissues (for example, the thicknesses of the alveolar ridge) in these fossil findings. As an example of a potential first step, we are currently investigating with collaborators in the department of Human Osteology and Funeral Archeology at the University of Leiden, The Netherlands, the possibility of making inferences about the hard palate and lower jaw (Bosman, 2016) of historical populations from carefully selected and well-studied graves from post-medieval Netherlands.

5. Conclusions

Language is not an ethereal entity isolated from its wider environment, but it actively interacts with multiple factors in this environment including the climate and vegetation, the biology of the speakers, and the socio-cultural system, its evolution being influenced by them and influencing some of them in turn. While currently we do not yet have definitive answers on the precise influence of such factors, we do have some proposals that have generated a healthy discussion and a renewed interest in quantitative methods and large databases of linguistic diversity, and there are several projects and teams around the world exploring such ideas, developing methods and collecting data. Our own contribution to these efforts focuses on the anatomy of the vocal tract and its potential influence on phonetics and phonology, the collection of relevant primary and secondary data, and the development of statistical and computational methods that will hopefully prove useful to more general questions as well. Moreover, considering all the channels/modalities used in linguistic interactions should allow us to better identify the loci of such extra-linguistic biases; while our own current work focuses almost exclusively on the verbal channel (the vocal tract and its influence on speech), it is absolutely conceivable that similar types of constraints and affordances also shape co-speech gestures (Kelly et al., 2010), the division of labor between the verbal and visual channels, and the properties of sign languages (Sandler, 2006), and might even help in the ongoing debate between the gestural and verbal origins of language (Wacewicz et al., 2016; Arbib et al., 2008; Sterelny, 2012).

We believe that this rebirth of "embedded language" will not only give a better understanding of its form today, but will also allow principled, valid and detailed inferences about the past, about the speech and language of long-gone humans.

---

[11] It is probably worth noting that in the intervening years since its publication, the type of reasoning and suggestions in this paper have been mostly strengthened by evidence from genetics, archeology and paleoanthropology (Johansson, 2015).

References

Anthwal, N., Joshi, L., Tucker, A.S., 2013. Evolution of the mammalian middle ear and jaw: adaptations and novel structures. J. Anat. 222 (1), 147–160.
Aoki, K., Feldman, M.W., 1991. Recessive hereditary deafness, assortative mating, and persistence of a sign language. Theor. Popul. Biol. 39 (3), 358–372.
Arbib, M.A., Liebal, K., Pika, S., 2008. Primate vocalization, gesture, and the evolution of human language. Curr. Anthropol. 49 (6), 1053–1063 discussion 1063–76.
Archangeli, D., Pulleyblank, D., 1994. Grounded Phonology. MIT Press.
Arnos, K.S., Welch, K.O., Tekin, M., Norris, V.W., Blanton, S.H., Pandya, A., Nance, W.E., 2008. A comparative analysis of the genetic epidemiology of deafness in the United States in two sets of pedigrees collected more than a century apart. Am. J. Hum. Genet. 83 (2), 200–207.
Atkinson, Q.D., 2011. Phonemic diversity supports a serial founder effect model of language expansion from Africa. Science 332 (6027), 346–349.
Atkinson, Q.D., Gray, R.D., 2005. Curious parallels, curious connections – phylogenetic thinking in biology and historical linguistics. Syst. Biol. 54 (4), 513–526.
Barbujani, G., Colonna, V., 2010. Human genome diversity: frequently asked questions. Trends Genet. 26 (7), 285–295.
Bickel, B., Witzlack-Makarevich, A., Choudhary, K.K., Schlesewsky, M., Bornkessel-Schlesewsky, I., 2015. The neurophysiology of language processing shapes the evolution of grammar: evidence from case marking. PLoS One 10 (8), e0132819.
Birkholz, P., 2005. 3D-Artikulatorische Sprachsynthese (Ph.D. thesis). Fakultät für Informatik und Elektrotechnik, Universität Rostock, Logos Verlag, Germany, Berlin. http://www.vocaltractlab.de/publications/birkholz-2005-dissertation.pdf.
Bosman, A., 2016. Talking Heads (MSc thesis). Human Osteology and Funerary Archaeology, University of Leiden, The Netherlands.
Bowden, D.E., Goose, D.H., 1968. The inheritance of palatal arch width in human families. Arch. Oral Biol. 13 (10), 1293–1295.
Bromham, L., Hua, X., Fitzpatrick, T.G., Greenhill, S.J., 2015. Rate of language evolution is affected by population size. Proc. Natl. Acad. Sci. 112 (7), 2097–2102.
Brown, A.M., Lindsey, D.T., 2004. Color and language: worldwide distribution of Daltonism and distinct words for "blue". Vis. Neurosci. 21 (3), 409–412.
Brunner, J., Fuchs, S., Perrier, P., 2009. On the relationship between palate shape and articulatory behavior. J. Acoust. Soc. Am. 125 (6), 3936–3949.
Brunner, J., Hoole, P., Perrier, P., Fuchs, S., 2006. Temporal development of compensation strategies for perturbed palate shape in German/sch/-production. In: Proceedings of the 7th International Seminar on Speech Production, pp. 247–254.
Butcher, A., 2006. Australian aboriginal languages. In: Harrington, J., Tabain, M. (Eds.), Speech Production: Models, Phonetic Processes, and Techniques. Macquarie Monographs in Cognitive Science. Psychology Press, NY, pp. 187–210.
Bybee, J.L., 2001. Phonology and Language Use. Cambridge University Press.
Byers, S.N., Churchill, S.E., Curran, B., 1997. Identification of Euro-Americans, Afro-Americans, and Amerindians from palatal dimensions. J. Forensic Sci. 42, 3–9.
Cai, J., Ash, D., Kotch, L.E., Jabs, E.W., Attie-Bitach, T., Auge, J., Mattei, G., Etchevers, H., Vekemans, M., Korshunova, Y., Tidwell, R., Messina, D.N., Winston, J.B., Lovett, M., 2005. Gene expression in pharyngeal arch 1 during human embryonic development. Hum. Mol. Genet. 14 (7), 903–912.
Caldwell, C.A., Smith, K., 2012. Cultural evolution and perpetuation of arbitrary communicative conventions in experimental microsocieties. PloS One 7 (8), e43807.
Caldwell-Harris, C.L., Lancaster, A., Ladd, D.R., Dediu, D., Christiansen, M.H., 2015. Factors influencing sensitivity to lexical tone in an artificial language. Stud. Second Lang. Acquis. 37 (02), 335–357.
Campbell, L., 2004. Historical Linguistics: An Introduction. Edinburgh University Press, Edinburgh.
Christiansen, M.H., Chater, N., 2008. Language as shaped by the brain. Behav. Brain Sci. 31 (5), 489–508 discussion 509–58.
Croft, W., 2000. Explaining Language Change: An Evolutionary Approach. Pearson Eduction Limited, Harlow, England.
Cysouw, M., Dediu, D., Moran, S., 2012. Comment on phonemic diversity supports a serial founder effect model of language expansion from Africa. Science 335 (6069), 657.
Dale, R., Lupyan, G., 2012. Understanding the origins of morphological diversity: the linguistic niche hypothesis. Adv. Complex Syst. 15, 1150017.
De Vos, C., 2012. Sign-spatiality in Kata Kolok: How a Village Sign Language of Bali Inscribes its Signing Space (Ph.D. thesis). Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.
Dediu, D., 2008. The role of genetic biases in shaping language-genes correlations. J. Theor. Biol. 254, 400–407.
Dediu, D., 2011. Are languages really independent from genes? If not, what would a genetic bias affecting language diversity look like? Hum. Biol. 83 (2), 279–296.
Dediu, D., 2015. An Introduction to Genetics for Language Scientists: Current Concepts, Methods, and Findings. Cambridge University Press.
Dediu, D., Christiansen, M.H., 2016. Language evolution: constraints and opportunities from modern genetics. Top. Cogn. Sci. 8, 361–370. http://dx.doi.org/10.1111/tops.12195.
Dediu, D., Cysouw, M., Levinson, S., Baronchelli, A., Christiansen, M., Croft, W., Evans, N., Garrod, S., Gray, R., Kandler, A., Lieven, E., 2013. Cultural evolution of language. In: Richerson, P., Christiansen, M. (Eds.), Cultural Evolution: Society, Technology, Language, and Religion. Vol. 12 of Strüngmann Forum Reports. MIT Press, Cambridge, Mass, pp. 303–332.
Dediu, D., de Boer, B., 2015. Language evolution needs its own journal. J. Lang. Evol. lzv001.
Dediu, D., Ladd, D.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 (26), 10944–10949.
Dediu, D., Levinson, S.C., 2013. On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. Front. Lang. Sci. 4, 397.
Dediu, D., Moisik, S., 2016. Anatomical biasing of click learning and production: an MRI and 3d palate imaging study. In: Roberts, S.G., Cuskley, C., McCrohon, L., Barcel-Coblijn, L., Fehr, O., Verhoef, T. (Eds.), The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11). Online at: http://evolang.org/neworleans/papers/57.html.
Deeb, S.S., 2006. Genetics of variation in human color vision and the retinal cone mosaic. Curr. Opin. Genet. Dev. 16 (3), 301–307.
Dellavia, C., Sforza, C., Orlando, F., Ottolina, P., Pregliasco, F., Ferrario, V.F., 2007. Three-dimensional hard tissue palatal size and shape in down syndrome subjects. Eur. J. Orthod. 29 (4), 417–422.
Dixon, M.J., Marazita, M.L., Beaty, T.H., Murray, J.C., 2011. Cleft lip and palate: synthesizing genetic and environmental influences. Nature reviews. Genetics 12 (3), 167–178.
Dor, D., 2015. The Instruction of Imagination: Language as a Social Communication Technology. Oxford University Press.
Dousset, L., 2011. Australian Aboriginal Kinship: An Introductory Handbook with Particular Emphasis on the Western Desert. pacific-credo Publications. https://hal.archives-ouvertes.fr/hal-00626243.
Evans, N., Levinson, S.C., 2009. The myth of language universals: language diversity and its importance for cognitive science. Behav. Brain Sci. 32, 429–492.
Everett, C., 2013. Evidence for direct geographic influences on linguistic sounds: the case of ejectives. PLoS One 8 (6), e65275.
Everett, C., Blasí, D.E., Roberts, S.G., 2015. Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. Proc. Natl. Acad. Sci. 201417413.
Everett, C., Blasí, D.E., Roberts, S.G., 2016. Language evolution and climate: the case of desiccation and tone. J. Lang. Evol. 1 (1), 33–46.

Fehr, O., Wang, H., Saar, S., Mitra, P.P., Tchernichovski, O., 2009. De novo establishment of wild-type song culture in the zebra finch. Nature 459 (7246), 564–568.

Fisher, S.E., 2006. Tangled webs: tracing the connections between genes and cognition. Cognition 101 (2), 270–297.

Fisher, S.E., Vernes, S.C., 2015. Genetics and the language sciences. Annu. Rev. Linguist. 1 (1), 289–310.

Fitch, W.T., 2010. The Evolution of Language. Cambridge University Press.

Fitch, W.T., Giedd, J., 1999. Morphology and development of the human vocal tract: a study using magnetic resonance imaging. J. Acoust. Soc. Am. 106 (3 Pt 1), 1511–1522.

Gerbault, P., Liebert, A., Itan, Y., Powell, A., Currat, M., Burger, J., Swallow, D.M., Thomas, M.G., 2011. Evolution of lactase persistence: an example of human niche construction. Philos. Trans. R. Soc. B Biol. Sci. 366 (1566), 863–877.

Gerdeman, B., Fujimura, O., 1990. Speaking without a tongue. J. Acoust. Soc. Am. 87 (S1). S89–S89.

Gialluisi, A., Dediu, D., Francks, C., Fisher, S.E., 2013. Persistence and transmission of recessive deafness and sign language: new insights from village sign languages. Eur. J. Hum. Genet. 21 (9), 894–896.

Gick, B., Wilson, I., Derrick, D., 2013. Articulatory Phonetics, first ed. Wiley-Blackwell, Malden, MA.

Güldemann, T., 2007. Clicks, Genetics, and "proto-world" from a Linguistic Perspective. Inst. für Afrikanistik.

Harshman, R., Ladefoged, P., Goldstein, L., Sep. 1977. Factor analysis of tongue shapes. J. Acoust. Soc. Am. 62 (3), 693–713.

Hauser, M.D., Yang, C., Berwick, R.C., Tattersall, I., Ryan, M., Watumull, J., Chomsky, N., Lewontin, R., 2014. The mystery of language evolution. Front. Psychol. 5, 401.

Hay, J., Bauer, L., 2007. Phoneme inventory size and population size. Language 83 (2), 388–400.

Hiki, S., Itoh, H., 1986. Influence of palate shape on lingual articulation. Speech Commun. 5, 141–158.

Hopper, P.J., Traugott, E.C., 2003. Grammaticalization. Cambridge University Press.

Hyslop, G., 2009. Kurtöp tone: a tonogenetic case study. Lingua 119 (6), 827–845.

Janssen, R., Dediu, D., in press. Genetic biases affecting language: what do computer models and experimental approaches suggest? In: Poibeau, T., Villavicencio, A. (Eds.), Language, Cognition and Computational Models. Cambridge University Press, Cambridge, UK.

Janssen, R., Dediu, D., Moisik, S., 2016. Simple agents are able to replicate speech sounds using 3d vocal tract model. In: Roberts, S.G., Cuskley, C., McCrohon, L., Barcel-Coblijn, L., Fehr, O., Verhoef, T. (Eds.), The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11). Online at: http://evolang.org/neworleans/papers/97.html.

Janssen, R., Dediu, D., Moisik, S., 2015. Bezier modelling and high accuracy curve fitting to capture hard palate variation. In: The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences. University of Glasgow, Glasgow, UK. http://www.icphs2015.info/pdfs/Papers/ICPHS0327.pdf.

Jobling, M.A., Hollox, E., Hurles, M., Kivisild, T., Tyler-Smith, C., 2013. Human Evolutionary Genetics. Garland Science, New York.

Johansson, S., 2015. Language abilities in neanderthals. Annu. Rev. Linguist. 1 (1), 311–332.

Kelly, S.D., Ozyürek, A., Maris, E., 2010. Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. Psychol. Sci. 21 (2), 260–267.

Kirby, S., Cornish, H., Smith, K., 2008. Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. Proc. Natl. Acad. Sci. 105 (31), 10681–10686.

Kirby, S., Dowman, M., Griffiths, T.L., 2007. Innateness and culture in the evolution of language. Proc. Natl. Acad. Sci. USA 104 (12), 5241–5245.

Knight, A., Underhill, P.A., Mortensen, H.M., Zhivotovsky, L.A., Lin, A.A., Henn, B.M., Louis, D., Ruhlen, M., Mountain, J.L., 2003. African Y chromosome and mtDNA divergence provides insight into the history of click languages. Curr. Biol. CB 13 (6), 464–473.

Kumar, D.K., Gopal, K.S., 2011. Morphological variants of soft palate in normal individuals: a digital cephalometric study. J. Clin. Diagnostic Res. 5 (6), 1310–1313.

Ladd, D., Ainscough, R., Assmann, C., Caldwell-Harris, C., Ganushchak, L., Swoboda, K., Woodfield, V., Dediu, D., 2013. Patterns of individual differences in the perception of missing-fundamental tones. J. Exp. Psychol. Hum. Percept. Perform. 39 (5), 1386–1397.

Ladd, D.R., Dediu, D., Kinsella, A.R., 2008. Languages and genes: reflections on biolinguistics and the nature-nurture question. Biolinguistics 2 (1), 114–126.

Ladd, D.R., Roberts, S.G., Dediu, D., 2015. Correlational studies in typological and historical linguistics. Annu. Rev. Linguist. 1 (1), 221–241.

Laland, K.N., Odling-Smee, J., Myles, S., 2010. How culture shaped the human genome: bringing genetics and the human sciences together. Nat. Rev. Genet. 11 (2), 137–148.

Lammert, A., Proctor, M., Katsamanis, A., Narayanan, S., 2011. Morphological variation in the adult vocal tract: a modeling study of its potential acoustic impact. In: Twelfth Annual Conference of the International Speech Communication Association. In: http://www.mproctor.net/docs/lammert11_IS2011_morphology.pdf.

Lammert, A., Proctor, M., Narayanan, S., 2013. Morphological variation in the adult hard palate and posterior pharyngeal wall. J. Speech Lang. Hearing Res. 56, 521–530.

Levinson, S.C., Dediu, D., 2013. The interplay of genetic and cultural factors in ongoing language evolution. In: Richerson, P., Christiansen, M. (Eds.), Cultural Evolution: Society, Technology, Language, and Religion. Vol. 12 of Strüngmann Forum Reports. MIT Press, Cambridge, Mass, pp. 219–232.

Lewis, C.M., Knight, J., 2012. Introduction to genetic association studies. Cold Spring Harb. Protoc. 2012 (3) pdb.top068163.

Lieberman, P., 2012. Vocal tract anatomy and the neural bases of talking. J. Phonetics 40 (4), 608–622.

Lieberman, P., Crelin, E., 1971. On the speech of Neanderthal man. Linguist. Inq. 2, 203–222.

Lindsey, D.T., Brown, A.M., 2002. Color naming and the phototoxic effects of sunlight on the eye. Psychol. Sci. 13 (6), 506–512.

Lloyd, J.E., Stavness, I., Fels, S., 2012. ArtiSynth: a fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In: Soft Tissue Biomechanical Modeling for Computer Assisted Surgery. Springer, pp. 355–394. http://link.springer.com/chapter/10.1007/8415_2012_126.

Lupyan, G., Dale, R., 2010. Language structure is partly determined by social structure. PLoS One 5 (1), e8559.

Maddieson, I., Coupé, C., 2015. Human spoken language diversity and the acoustic adaptation hypothesis. J. Acoust. Soc. Am. 138 (3), 1838–1838.

Mameli, M., Bateson, P., 2006. Innateness and the sciences. Biol. Philosophy 21, 155–188.

Meir, I., Sandler, W., Padden, C., Aronoff, M., 2010. Emerging sign languages. In: Marschark, M., Spencer, P.E. (Eds.), Oxford Handbook of Deaf Studies, Language, and Education, vol. 2. Oxford University Press, Oxford.

Moisik, S., Dediu, D., 2015. Anatomical biasing and clicks: preliminary biomechanical modelling. In: Hannah, L. (Ed.), The Evolution of Phonetic Capabilities: Causes Constraints, Consequences. Glasgow, UK, pp. 8–13. In: https://ai.vub.ac.be/sites/default/files/proceedingsfinal.pdf.

Moran, S., McCloy, D., Wright, R., 2012. Revisiting population size vs. phoneme inventory size. Language 88 (4), 877–893.

Nance, W.E., Liu, X.Z., Pandya, A., 2000. Relation between choice of partner and high frequency of connexin-26 deafness. Lancet 356 (9228), 500–501.

Ohala, J.J., 2005. The marriage of phonetics and phonology. Acoust. Sci. Technol. 26 (5), 418–422.

OMIM, April 2016. Online Mendelian Inheritance in Man, OMIM®. http://omim.org.

Pagel, M., 2009. Human language as a culturally transmitted replicator. Nat. Rev. Genet. 10, 405–415.

Patel, S.R., Frame, J.M., Larkin, E.K., Redline, S., 2008. Heritability of upper airway dimensions derived using acoustic pharyngometry. Eur. Respir. J. 32 (5), 1304–1308.

Patterson, F.G.P., Cohn, R.H., 1990. Language acquisition by a lowland gorilla: Koko's first ten years of vocabulary development. WORD 41 (2), 97–143.

Perlman, M., Patterson, F.G., Cohn, R.H., 2012. The human-fostered gorilla Koko shows breath control in play with wind instruments. Biolinguistics 6 (3), 433–444.

Plomin, R., Haworth, C.M.A., Meaburn, E.L., Price, T.S., Davis, O.S.P., 2013. Common DNA markers can account for more than half of the genetic influence on cognitive abilities. Psychol. Sci. 24 (4), 562–568.

Reynoso, M.C., Hernndez, A., Lizcano-Gil, L.A., Sarralde, A., Abreu, M.C., Nazar, Z., Fragoso, R., 1994. Autosomal dominant congenital macroglossia: further delineation of the syndrome. Genet. Counsling 5 (2), 151–154.

Richerson, P.J., Boyd, R., Henrich, J., 2010. Gene-culture coevolution in the age of genomics. Proc. Natl. Acad. Sci. USA 107 (Suppl 2), 8985–8992.

Richerson, P.J., Christiansen, M., 2013. Cultural evolution: Society, Technology, Language, and Religion.

Sandler, W., 2006. Sign language: an overview. In: Encyclopedia of Language & Linguistics, second ed. Elsevier, pp. 328–338.

Sandler, W., Meir, I., Padden, C., Aronoff, M., 2005. The emergence of grammar: systematic structure in a new language. Proc. Natl. Acad. Sci. USA 102 (7), 2661–2665.

Schwab, R.J., Pasirstein, M., Kaplan, L., Pierson, R., Mackley, A., Hachadoorian, R., Arens, R., Maislin, G., Pack, A.I., 2006. Family aggregation of upper airway soft tissue structures in normal subjects and patients with sleep apnea. Am. J. Respir. Crit. Care Med. 173 (4), 453–463.

Scobbie, J.M., 2005. The Phonetics Phonology Overlap. QMU Speech Science Research Centre Working Papers, WP-1. http://eresearch.qmu.ac.uk/138.

Senghas, A., 1995. Children's Contribution to the Birth of Nicaraguan Sign Language (Ph.D. thesis). Massachusetts Institute of Technology.

Smith, K., Kirby, S., 2008. Cultural evolution: implications for understanding the human language faculty and its evolution. Philos. Trans. R. Soc. B 363, 3591–3603.

Smith, K., Wonnacott, E., 2010. Eliminating unpredictable variation through iterated learning. Cognition 116 (3), 444–449.

Sterelny, K., 2012. Language, gesture, skill: the co-evolutionary foundations of language. Philos. Trans. R. Soc. B 367 (1599), 2141–2151.

Szwed, M., Cohen, L., Qiao, E., Dehaene, S., 2009. The role of invariant line junctions in object and visual word recognition. Vis. Res. 49 (7), 718–725.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., Cohen, L., 2011. Specialization for written words over objects in the visual cortex. Neuroimage 56 (1), 330–344.

Tamariz, M., Kirby, S., 2016. The cultural evolution of language. Curr. Opin. Psychol. 8, 37–43.

Thompson, B., Kirby, S., Smith, K., 2016. Culture shapes the evolution of cognition. Proc. Natl. Acad. Sci. 113 (16), 4530–4535.

Tiede, M.K., Boyce, S.E., Holland, C.K., Choe, K.A., 2004. A new taxonomy of American English /r/ using MRI and ultrasound. J. Acoust. Soc. Am. 115 (5), 2633–2634.

Townsend, G.C., Richards, L.C., Sekikawa, M., Brown, T., Ozaki, T., 1990. Variability of palatal dimensions in South Australian twins. J. Forensic Odontostomatol. 8 (2), 3–14.

Traill, A., 1985. Phonetic and Phonological Studies of !Xóõ Bushman. H. Buske.

Visscher, P.M., McEvoy, B., Yang, J., 2010. From Galton to GWAS: quantitative genetics of human height. Genet. Res. 92 (5–6), 371–379.

von Cramon-Taubadel, N., 2011. Global human mandibular variation reflects differences in agricultural and hunter-gatherer subsistence strategies. Proc. Natl. Acad. Sci. 108 (49), 19546–19551.

Vorperian, H.K., Kent, R.D., Lindstrom, M.J., Kalina, C.M., Gentry, L.R., Yandell, B.S., 2005. Development of vocal tract length during early childhood: a magnetic resonance imaging study. J. Acoust. Soc. Am. 117 (1), 338–350.

Wacewicz, S., Zywiczynski, P., Orzechowski, S., 2016. Visible movements of the orofacial area: evidence for gestural or multimodal theories of language evolution? Gesture 15 (2), 250–282.

Weirich, M., 2010. Articulatory and acoustic inter-speaker variability in the production of German vowels. ZAS Pap. Linguist. 52, 19–42.

Weirich, M., Fuchs, S., 2011. Vocal tract morphology can influence speaker specific realisations of phonemic contrasts. Proc. ISSP, 251–258. http://www.sprachwissenschaft.uni-jena.de/germsprach_multimedia/Downloads/weirich/Weirich_Fuchs_Vocal+tract+morphology.pdf.

Wong, P.C.M., Chandrasekaran, B., Zheng, J., 2012. The derived allele of ASPM is associated with lexical tone perception. PLoS One 7 (4), e34243.

Wray, A., Grace, G.W., 2007. The consequences of talking to strangers: evolutionary corollaries of socio-cultural influences on linguistic form. Lingua 117 (3), 543–578.

Yip, M., 2002. Tone. Cambridge Textbooks in Linguistics. Cambridge University Press, UK.

You, M., Li, X., Wang, H., Zhang, J., Wu, H., Liu, Y., Miao, J., Zhu, Z., 2008. Morphological variety of the soft palate in normal individuals: a digital cephalometric study. Dentomaxillofac. Radiol. 37 (6), 344–349.

Zelditch, M.L., Swiderski, D.L., Sheets, H.D., 2012. Geometric Morphometrics for Biologists: A Primer. Academic Press.

Zhou, X., Espy-Wilson, C.Y., Tiede, M., Boyce, S., 2007. An articulatory and acoustic study of "retroflex" and "bunched" American English rhotic sound based on MRI. In: INTERSPEECH, pp. 54–57. http://www.isr.umd.edu/Labs/SCL/publications/conference/t_zhou_etal_icslp_07.pdf.