PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Moran S, Lester NA, Grossman E. 2021 Inferring recent evolutionary changes in speech sounds. *Phil. Trans. R. Soc. B* **376**: 20200198.
https://doi.org/10.1098/rstb.2020.0198

Accepted: 22 December 2020

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

Subject Areas:
evolution, cognition

Keywords:
language evolution, phonology, language contact

Author for correspondence:
Steven Moran
e-mail: steven.moran@unine.ch

Inferring recent evolutionary changes in speech sounds

Steven Moran[1], Nicholas A. Lester[2] and Eitan Grossman[3]

[1]Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, G B35, 2000 Neuchâtel, Switzerland
[2]Department of Comparative Language Science, University of Zurich, Thurgauerstrasse 30/32, 8050 Zurich, Switzerland
[3]Department of Linguistics, Hebrew University of Jerusalem, 91905 Jerusalem, Israel

SM, 0000-0002-3969-6549; EG, 0000-0001-7479-7831

In this paper, we investigate evolutionarily recent changes in the distributions of speech sounds in the world's languages. In particular, we explore the impact of language contact in the past two millennia on today's distributions. Based on three extensive databases of phonological inventories, we analyse the discrepancies between the distribution of speech sounds of ancient and reconstructed languages, on the one hand, and those in present-day languages, on the other. Furthermore, we analyse the degree to which the diffusion of speech sounds via language contact played a role in these discrepancies. We find evidence for substantive differences between ancient and present-day distributions, as well as for the important role of language contact in shaping these distributions over time. Moreover, our findings suggest that the distributions of speech sounds across geographic macro-areas were homogenized to an observable extent in recent millennia. Our findings suggest that what we call the Implicit Uniformitarian Hypothesis, at least with respect to the composition of phonological inventories, cannot be held uncritically. Linguists who would like to draw inferences about human language based on present-day cross-linguistic distributions must consider their theories in light of even short-term language evolution.

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

1. Introduction

Research on language evolution tends to focus on evolutionary events and processes that occurred in the distant past. For example, Dediu *et al.* [1] focus on the evolution of the human vocal tract and the reconstruction of speech sounds in the prehistoric past, while Everett [2] explores the more recent prehistory of biases in sound systems. However, for most of the world's languages, 'prehistory' includes the relatively recent past; the vast majority of the 7000 or so languages spoken today have no premodern documentation. As such, earlier stages of these languages are prehistoric. In this paper, we investigate the impact of short-term evolutionary changes in the worldwide distribution of speech sounds and highlight the potentially large impact of language contact, particularly through linguistic colonial expansions, on today's languages.

Each spoken language has a set of linguistically contrastive speech sounds (so-called *phonemes*) that through phonological analysis are segmented from the speech stream (hence also referred to as *segments*), which are posited as the language's phonological inventory (or phonological 'repertoire'). We hypothesize that wide-scale language contact has made phonological inventories more similar in the past few hundred to few thousand years, which thereby substantially altered the worldwide distribution of speech sounds that we observe today. As a corollary, we hypothesize that earlier stages of human languages were less uniform with respect to their phonological inventories and they showed greater area-specific profiles.

The existence of phonological areas is of course not a novel suggestion, and the areal-linguistic literature is replete with observations about the positive

THE ROYAL SOCIETY PUBLISHING

Table 1. Cross-linguistically frequent voiced obstruents in South America versus worldwide coverage (in PHOIBLE versus PHOIBLE without South America).

sound	frequency in SA	frequency in PHOIBLE	Δ	frequency in PHOIBLE (without SA)	Δ
b	0.44	0.63	−0.19	0.67	−0.23
d	0.37	0.46	−0.09	0.47	−0.11
g	0.29	0.57	−0.28	0.62	−0.33
β	0.17	0.10	+0.07	0.09	−0.08
v	0.03	0.27	−0.24	0.31	−0.28
ð	0.02	0.05	−0.03	0.06	−0.04
z	0.06	0.30	−0.24	0.34	−0.26
ʒ	0.06	0.16	−0.10	0.18	−0.12
ɤ	0.06	0.14	−0.10	0.16	−0.10

features of particular areas. A well-known example is the prevalence of retroflex stops in South Asia [3] or high central vowels in the languages of Amazonia [4]. However, negative features, i.e. the lack of particular speech sounds (or their phonological features), are not as widely reported. But they are nonetheless found. Perhaps most famously, Australian languages generally lack contrastive fricatives and voicing contrasts in stop series [5–7]. And while labiodental sounds are frequent in present-day Europe and its surrounding areas, the lack of labiodentals carves out a particular area in eastern Europe and the northern Caucasus ([8], p. 136). Such negative features may persist over long periods of time.

We suggest that the large-scale borrowing of phonologically contrastive speech sounds (also known as 'phonological segments'), through lexical borrowing, may have made some speech sounds more prevalent in the world's languages, thereby levelling to some extent earlier areal-specific profiles. An interesting example is that of voiced obstruents in South America. Voiced stops are less common in South America than in other geographic macro-areas, occurring in fewer than 45% (at most) of the languages in a sample of South American languages, and voiced fricatives in the sample are even rarer: except for the voiced bilabial fricative /β/, the attestation of voiced fricatives in South America is much lower than expected based on the global distribution (table 1). The relative rarity of voiced obstruents in South America is even more salient when South America is compared to the rest of the world; notably, /β/ no longer shows a positive delta (difference in percentages) but rather a weak negative delta, showing that it is slightly under-represented in South America relative to the rest of the world. It is plausible that the paucity of voiced obstruents reflects a relatively old negative feature of South America. However, many South American languages, such as Mojeño (Arawakan, Bolivia) and Jauja Wanca Quechua (Quechuan, Bolivia) have recently acquired voiced obstruents as a result of contact with Spanish and Portuguese, thereby making this macro-area closer to the worldwide distribution of speech sounds.

Similarly, as early as Houis [9] and Maddieson [6], it has been noted as characteristic of Africa that voiceless bilabial stops are missing in many African languages, an observation replicated and nuanced by Clements & Rialland ([10], pp. 65–67). It is therefore interesting to note that in a sample of recent phonological segment borrowings in African languages, the most frequently borrowed speech sound is /p/ [11], and reports that /p/ is found only in loanwords such as Tigrinya (Semitic, Ethiopia), Tem (Central Gur, Togo), or !Xóõ (Tuu, Botswana and Namibia) are common ([10], pp. 65–67). Similar observations about cross-linguistically common speech sounds being missing from large macro-areas are not uncommon—even though such segments are reported to be frequent in the world's languages today.

However, it is difficult to estimate areal specificity based on qualitative studies that are themselves based on possibly dissimilar methodologies. Tables 5 and 6 in appendix A provide a brief quantitative overview of different geographical macro-areas in terms of the consonants that most differ from global distributions.[1] Table 5 shows the segments per area that are overrepresented with respect to global distributions, while table 6 shows the segments that are underrepresented with respect to global distributions.

The relationship between cross-linguistic frequency distributions, both at the global level and at the level of smaller geographical areas, and borrowability (i.e. the likelihood that a particular linguistic property will be borrowed), is crucial for this study. Earlier studies of language contact have typically assumed that the empirical frequency of borrowing is a measure of the probability that a property will be borrowed. For example, it is assumed that if nouns are more frequently borrowed than prepositions, nouns are more inherently borrowable than prepositions. We adopt another perspective on the matter, proposing that borrowability is best understood as a function of cross-linguistic frequencies. This is because in order to be borrowed from one language to another, a linguistic property has to be present in some languages and absent in others. For example, pharyngeal fricatives are unlikely to borrowed very frequently, because they are absent from the languages of most geographic regions in the world.[2] Similarly, the bilabial nasal /m/ is unlikely to be frequently borrowed, because most of the world's languages already have this contrastive speech sound. The interpretation of borrowability as a function of frequency makes a prediction, namely, that the most borrowable speech sounds are in the mid-range of their cross-linguistic frequency distribution. In other words, it is the sounds that are neither very rare nor very common that will be the most likely to be borrowed as the result of language contact [11,12].

In order to investigate the possibility of substantial but recent changes in the composition and structure of

phonological inventories in the past several millennia, we present quantitative methods for comparing ancient versus modern speech sound frequency distributions on an empirical basis. The data come from three phonological databases. The first, an expanded version of BDPROTO [13], is a database of more than 250 ancient and reconstructed phonological inventories from all of the world's major macro-areas. This allows us to approximate the distribution of phonological segments in the distant past, at least to the limits of the comparative-historical method. The second, PHOIBLE [14], is a comprehensive database of present-day phonological inventories. The third, SegBo [11], is a database of borrowed phonological segments. We leverage these resources to analyse the discrepancies between the cross-linguistic frequencies of sound classes across ancient and reconstructed languages versus present-day languages, and to investigate the extent to which phonological segment borrowing contributed to these discrepancies.

Finally, we raise a hitherto under-discussed challenge in language evolution research, which we call *temporal bias* [13]. That is, in addition to the usual biases involved in comparative linguistic research, including bibliographical, genealogical and geographical biases [15], we identify a temporal bias that stems from data sparsity at different time depths compounded by data sparsity at different phylogenetic distances from the present. Temporal bias presents a strong obstacle to the ability of linguists to evaluate the empirical validity of the Uniformitarian Hypothesis on the basis of databases (see §2). However, acknowledging the problems that temporal bias presents may spur thinking about ways to address them. In this article, we present a method for mitigating temporal bias, at least to an extent, for investigating evolutionary changes in the distribution of speech sounds in the world's languages.

The structure of this paper is as follows. In §2, we outline the motivation for our analysis. In §3, we describe our data sources. In §4, we discuss our statistical approaches and present our results. And lastly in §5, we discuss our results and their relevance for the empirical study of language evolution in the short-to-long term.

2. Motivation

Linguists often draw inferences about human language on the basis of a sample of languages. A well-known example is Hawkins' processing theory of word order, which proposes that harmonic word orders (e.g. verb-final order and postpositions) are beneficial for cognitive processing. Ideally, the universal probability of a linguistic type could be inferred from the empirical frequency of that type [16] or from other aspects of cross-linguistic distributions of properties.

However, the possibility of drawing valid inferences of this sort depends to a large extent on some version of the Uniformitarian Hypothesis (cf. Walkden [17]). The Uniformitarian Hypothesis, which in essence stresses the time-independent unity of human languages, has been interpreted in a variety of ways, e.g. as a constraint on language change or as a constraint on synchronic distributions [18]. One version of the Uniformitarian Hypothesis, formulated explicitly as '[H]uman languages have always been pretty much the same in terms of the typological distribution of the units that compose them' ([19], p. 360), is often assumed by theoretical linguists in order to infer universal and time-independent properties of human language directly from present-day distributions. We call this version the 'Implicit Uniformitarian Hypothesis', which we formulate as follows:

> **The Implicit Uniformitarian Hypothesis:** Throughout the history of what linguists call 'human language', cross-linguistic distributions of linguistic properties, whether simple or complex, have always been more or less the same. In particular, linguistic properties that are currently rare have always been rare, and linguistic properties that are currently frequent have always been frequent. That is, cross-linguistic distributions are time-independent.

However, the possibility of inferring causes directly from present-day distributions has been called into question. For example, Dryer [20] and Nichols [21] have argued that present-day distributions may be the result of language contact, which can scale up to continent-sized macro-areas. Moreover, Maslova [22] notes that cross-linguistic properties have not reached a stationary distribution, such that the present-day distribution of linguistic properties in the world's languages is independent of an initial state; in other words, present-day distributions are likely still dependent on still-unknown initial states. Furthermore, Piantadosi & Gibson [23] have shown that the number of present-day language families and known areas does not provide enough independent data points to infer universal properties of language. These observations, taken together, point to the possibility that present-day distributions are at least to some extent artefacts of inheritance, on the one hand, and of language-external historical events, such as language contact, on the other.

It has also become increasingly clear that non-linguistic factors may shape language structure in ways that influence cross-linguistic distributions. Some proposed factors include different aspects of the social structure of linguistic communities [24–27], such as population size and complexity [28–36]; genes and aspects of anatomy [37–40], geography and other environmental factors [41–44]; technology, in particular food production [45]; and, purportedly, sexual mores [46]. Since these factors themselves may be subject to change, it cannot be ruled out that their influence on cross-linguistic distributions is dynamic. A striking case is that of labiodental speech sounds, such as /f/ and /v/, which have been shown to be relative latecomers to speech, even though they are among the best-documented speech sounds in present-day languages [45]; this lateness has been attributed to changes in food technology production, which in turn led to changes in articulatorily relevant anatomy. Furthermore, research on phonological segment borrowing has pointed to the large influence of relatively recent processes of colonialism and globalization on the sound systems of languages throughout the world [11]. In the light of all of these factors, it is likely that present-day distributions may conceal substantive evolutionary changes in human languages, even in the relatively recent past. Therefore, we think it is worthwhile to treat the Uniformitarian Hypothesis as an open question to be investigated empirically. Next we present our data sources and then we turn to the methods we use to investigate recent evolutionary changes in speech sounds.

3. Data

(a) Overview

Phonological inventories of thousands of languages have been compiled into several large-scale databases, which we

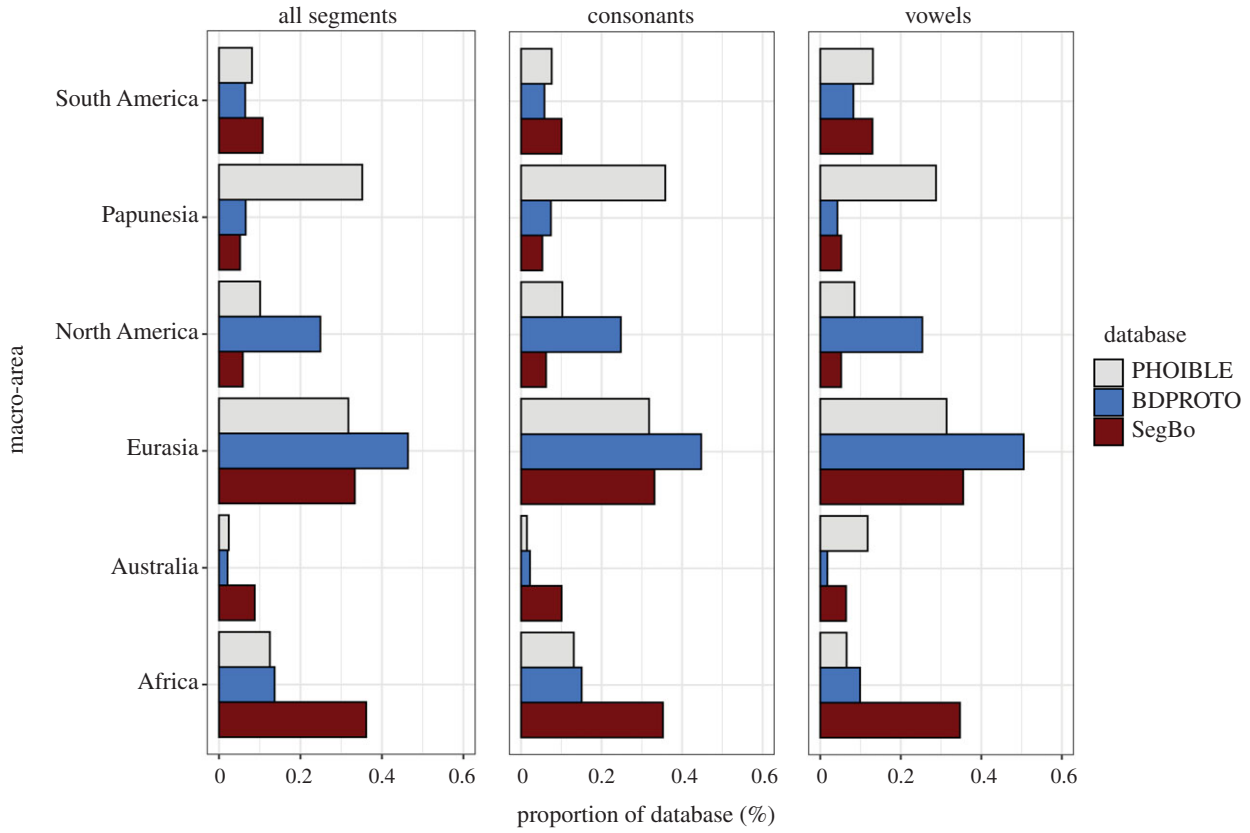

Figure 1. Comparison of sample sizes by macro-area and segment type. (Online version in colour.)

Table 2. Overview of the contents of each database.

database	phonological inventories	language varieties	language families	segment types
BDPROTO	257	192	75	637
PHOIBLE	3020	2186	174	3183
SegBo	531	514	113	219

use as input to our statistical models in this paper. Table 2 and figure 1 give an overview of the contents of these three databases. We first briefly describe each database, and then we discuss the issues that the contents of these databases raise for our models and analyses.

(b) BDPROTO

BDPROTO is a database of 257 phonological inventories from ancient and reconstructed languages that were extracted from historical linguistic reconstructions and then interpreted by experts [13,47].[3] These inventories come from publications by historical linguists, who applied the historical-comparative method to synchronic datasets (e.g. word lists, phonological descriptions, grammars) and reconstructed the contrastive sound systems of proto-languages.[4] The BDPROTO data build on the model of PHOIBLE and include phonological inventory data and metadata about the languages represented in the sample. In this study, we exclude several data points from BDPROTO (as noted in table 2) owing to the questionable nature of the proposed language families, e.g. Nostratic, in the original BDPROTO sample [48].

(c) PHOIBLE

PHOIBLE is a repository of cross-linguistic phonological inventory data, which have been extracted from source documents and tertiary databases and compiled into a single convenience sample [49].[5] PHOIBLE version 2.0 includes 3020 inventories that contain 3183 segment types found in 2186 distinct languages [14]. It is currently the most comprehensive cross-linguistic database on phonological inventories, which is openly available. PHOIBLE includes not only phonological inventories, but also a detailed phonological feature set for each sound, and accompanying metadata for each language.

(d) SegBo

SegBo is the first large-scale cross-linguistic database of borrowed phonological segments and it contains information on over 1600 borrowing events in 531 language varieties [11].[6] Phonological segment borrowing is a process in which a language acquires a new speech sound as the result of borrowing new words from another language. Contact-induced phonological change through language contact is rampant in languages and SegBo can be used to shed light on

borrowing events in human history, e.g. [11,50]. Data points in SegBo come from reports, such as grammars and phonological descriptions, that include detailed information on segments that have been borrowed from one language to another. Individual data points were coded for, inter alia, geographical area of the borrowing language, the source language (when known), the distribution of the segment in the borrowing language's lexicon, and whether the segment introduces new phonological distinctions into the language. Other metadata related to both borrowing and source languages is available via Glottocodes [51]. Importantly, most of the borrowing events documented in SegBo are recent, and in fact most probably date to the past approximately 500 years, and many of the donor languages are those that spread throughout the world through historical processes of colonialism and globalization, such as English, Spanish, Arabic, Russian and Indonesian [11].

(e) Issues related to the data and analyses

Table 2 shows that PHOIBLE is much larger than the other two databases. Figure 1 highlights the discrepancies in the size and macro-area coverage of the three databases. It provides an overview of the database sample sizes by geographical macro-area, divided into the overall number of segments, consonants and vowels. This plot shows us that each database favours one or more macro-areas more than the other two.[7] Furthermore, except for SegBo, which reports more borrowed vowels than consonants in Australia, the overall proportions (i.e. all segment types considered) more closely resemble the distributional biases we see for consonants across macro-areas than for vowels.

Clearly, BDPROTO, PHOIBLE and SegBo do not cover the same language families nor segment types. This first point hinders interpretability of any results on a worldwide scale because they could be driven by the behaviour of specific families (i.e. not global trends). The second point is more challenging. What does it mean for a database not to contain a single instance of a segment? Given the sizes of the databases, it is unlikely to always be the result of sampling error. For example, speech sounds that do not appear in SegBo can be absent because they are simply not borrowed (a true zero frequency). More importantly, even for those segments for which sampling error is to blame (a false zero), their rate of occurrence would still be expected to be close to zero. Other factors include how the languages are documented (e.g. are tones treated phonemically?) and how segments are treated (e.g. when does an individual researcher decide that a speech sound has been borrowed?). A final factor to consider is the areal, i.e. geographical, makeup of the databases. PHOIBLE heavily favours African languages; BDPROTO favours Eurasia and North America; and SegBo favours Papunesia. Therefore, we expect that the differences in geographical coverage among the databases will impact similarity estimates.

We take the following steps to mitigate these issues. First, we only consider languages from the families that appear at least once in all of the databases (see the electronic supplementary material for additional analyses without this constraint). Second, we define the set of possible segments as the union of all segment types that appear across the three databases. When a segment was not observed in a given database, its frequency was set to 0 (see the electronic supplementary material for a more detailed discussion of these factors and their impact on our results). Third, we attempt to account for sampling issues by employing several different resampling procedures (see §4).

Before moving on, we also raise a data issue regarding BDPROTO. As earlier studies have noted, the historical-comparative method may have some inherent biases (see [13] and references therein). The first relates to lack of accurate phonetic data for ancient and reconstructed languages. This cannot be easily mitigated by the curators of BDPROTO, and we note it as an inherent issue in the reconstruction of speech sounds more generally. A second and related issue concerns the potential over-normalization of segments in reconstructed phonological inventories. While the most frequent realization of a phoneme in a given language might bear some additional phonetic characteristics (e.g. aspiration of voiceless stops, as in English), such subtleties might be invisible to reconstructions, which might in turn inflate the representation of 'plain' segments at the cost of more complex ones. A third issue is that proto-segments are typically established on the basis of correspondence sets, which might lead to an over-estimation of the number of segments in a reconstructed inventory. Practitioners of the historical-comparative method are very aware of this issue, and much attention is devoted to identifying potentially overlapping distributions in order to mitigate the inflation of proto-segments. Finally, a reviewer raised the concern that 'exotic' sounds might be under-reported in BDPROTO. We think that this is unlikely, given the wide geographical and genealogical span of BDPROTO, taking into consideration the fact that 'exotic' sounds like ejectives, implosives, preaspirated nasals, prenasalized stops and a variety of segments with additional articulations are often reconstructed and therefore reported in phonological inventories in BDPROTO.

4. Methods and results

The aims of this study are (i) to see if we can detect distributional changes between BDPROTO and PHOIBLE and (ii) to ask whether, and if so how, the borrowing of speech sounds played a role, by comparing their segment distributions against SegBo. In order to do so, we conducted several series of analyses, described in the sections below. For each, we describe the methods, followed immediately by the results. We begin by testing the overall similarity of each of the databases and then we correct against sampling biases in the databases by using a leave-one-out resampling approach.

(a) Overall comparisons of the databases

We begin by testing the overall similarity of each of the databases to each other by comparing the frequency distributions of segments in the three databases. We also explore in what ways these similarity estimates might be affected by individual segments, geographical macro-areas or segment class (consonants or vowels).

We operationalize the difference between any two frequency distributions as the Jensen–Shannon divergence (JSD; [54]). JSD is a measure of how much information on average one would need to transform one distribution into another, i.e. information regarding how probabilities per outcome must be shifted so that the distributions match.[8]

Suppose we have two probability distributions P and Q taken over the set of possible segments. The JSD between them is defined as:

$$\mathrm{JSD}(P\|Q) = 0.5*(\mathrm{KLD}(P\|M) + \mathrm{KLD}(Q\|M)) \tag{4.1}$$

where

$$\mathrm{KLD}(A\|B) = \sum A(x)*\log_2\left(\frac{A(x)}{B(x)}\right) \tag{4.2}$$

and

$$M = 0.5*(P + Q). \tag{4.3}$$

In this case, the discrete random variable x is the distribution of segments in each database, where $P(x)$ and $Q(x)$ each sum up to one. For calculating JSD, we use the *philentropy* package [55] in the *R* programming language [56]. The resulting JSD value is given in terms of bits, i.e. in terms of $\log_2$.

For this analysis, we compute JSD for the segment frequency distributions for each of our comparisons of interest (BDPROTO > PHOIBLE, BDPROTO > SegBo and PHOIBLE > SegBo), using all inventories that conform to the stipulations mentioned above. The values for JSD are scaled from 0 to 1, with 0 indicating identity and 1 total independence.

Family-controlled and total-sample JSD estimates are nearly identical, so we only report the former (being that it is the more conservative of the two; all estimates can be found in the electronic supplementary material). The frequency distributions of segments in BDPROTO and PHOIBLE are more similar (JSD = 0.14) than BDPROTO versus SegBo (JSD = 0.34) or PHOIBLE versus Segbo (JSD = 0.31). This is to be expected because the frequency distribution of segments in both BDPROTO and PHOIBLE represent full phonological inventories, whereas SegBo contains observations of individually borrowed speech sounds across languages. It is difficult to interpret the magnitudes of the differences in JSD in real-world terms, as there are no theory-based predictions for what different values or effect sizes should reflect. However, what is important here is that these values differ significantly from random baselines. We recompute JSD based on a leave-one-out (LOO) sampling method in the next section. This baseline means that we can also explore in what ways these similarity estimates might be affected by specific speech sounds, geographical macro-areas and language families.

(b) Leave one out resampling

Next, we examine whether these JSD estimates are driven heavily by any specific speech sounds using a LOO resampling technique. LOO works by removing one sound at a time—leaving all other sounds present—from each of the two frequency distributions being compared until each sound has been removed once. On each iteration, we compute a new JSD. This method allows us to associate any changes in JSD with a single segment. Taking this step provides us with two pieces of information. First, we can measure how stable our estimates of similarity are between databases. If the presence or the absence of specific sounds leads to drastic shifts in our similarity estimates, then we cannot rely on any aggregate analysis to reveal something meaningful about how these databases are related. If, on the other hand, the similarity estimates remain largely stable across iterations, then we have reason to believe that the aggregate estimates of similarity are reasonable. Moreover, this approach allows us to see how much of a difference any individual sound makes when comparing the databases. The extremes of these estimates (sounds that either radically increase or decrease similarity estimates) tell us specifically how the databases disagree. Second, the spread of values we observe informs us about possible sampling biases. We know that these databases offer different coverage of geographical areas, language families, as well as a range of other variables. These differences could produce systematic, sample-based biases in our estimates of similarity (sounds being either over- or under-represented in one or more of the databases simply because of the types of languages or locations that were sampled (cf. tables 1, 5 and 6). If these occur, we would also like to know the size of the effect and what sounds or groups of sounds are responsible.

The main result of this analysis is that the variability falls in a very narrow range, indicating that (a) there is a high degree of similarity between the frequency distributions in the PHOIBLE and BDPROTO databases, and (b) only a relatively small number of segments play any role in distinguishing between the two databases. Detailed analyses can be found in the electronic supplementary material.

(c) Poisson model

We dig deeper into the relationships between individual segment frequencies across databases. We perform a quasi-Poisson regression to model SegBo segment frequency as a function of PHOIBLE and BDPROTO segment frequencies (scaled to account for differences in sample size). Rather than the traditional Poisson model, we apply the quasi-Poisson technique to account for significant overdispersion of the SegBo frequencies. The null hypothesis is that there is no relation between these segment frequencies and those found in SegBo. The alternative hypothesis is that the most frequently borrowed segments will correspond to those segments that distinguish BDPROTO and PHOIBLE.

The dependent variable is SegBo frequency. Predictors are the segment frequencies from PHOIBLE and BDPROTO (scaled from 0 to 10), the source of the frequencies (BDPROTO, PHOIBLE) and the interaction between the two. An exploratory generalized additive model with separate data-driven smooths for PHOIBLE and BDPROTO frequencies suggested a strongly nonlinear relationship between both and SegBo frequency. Therefore, we include orthogonal polynomial terms for our frequency predictor up to the seventh degree. We prefer the generalized linear model in this case because it allows us to directly test the difference between BDPROTO and PHOIBLE frequencies via the interaction term. See the electronic supplementary material for the code and detailed results.

Furthermore, since the relationships between frequency distributions may differ by geographical area, we expand the modelling dataset by adding frequencies per macro-area per database. We explore several model structures. First, we leave the model structure from the full model intact, but split the data up by macro-area. We then run two generalized linear mixed-effect Poisson regressions with macro-area and family as random effects (crossed and nested, respectively). For reasons of space, we relegate the bulk of these analyses to the electronic supplementary material. We then take a

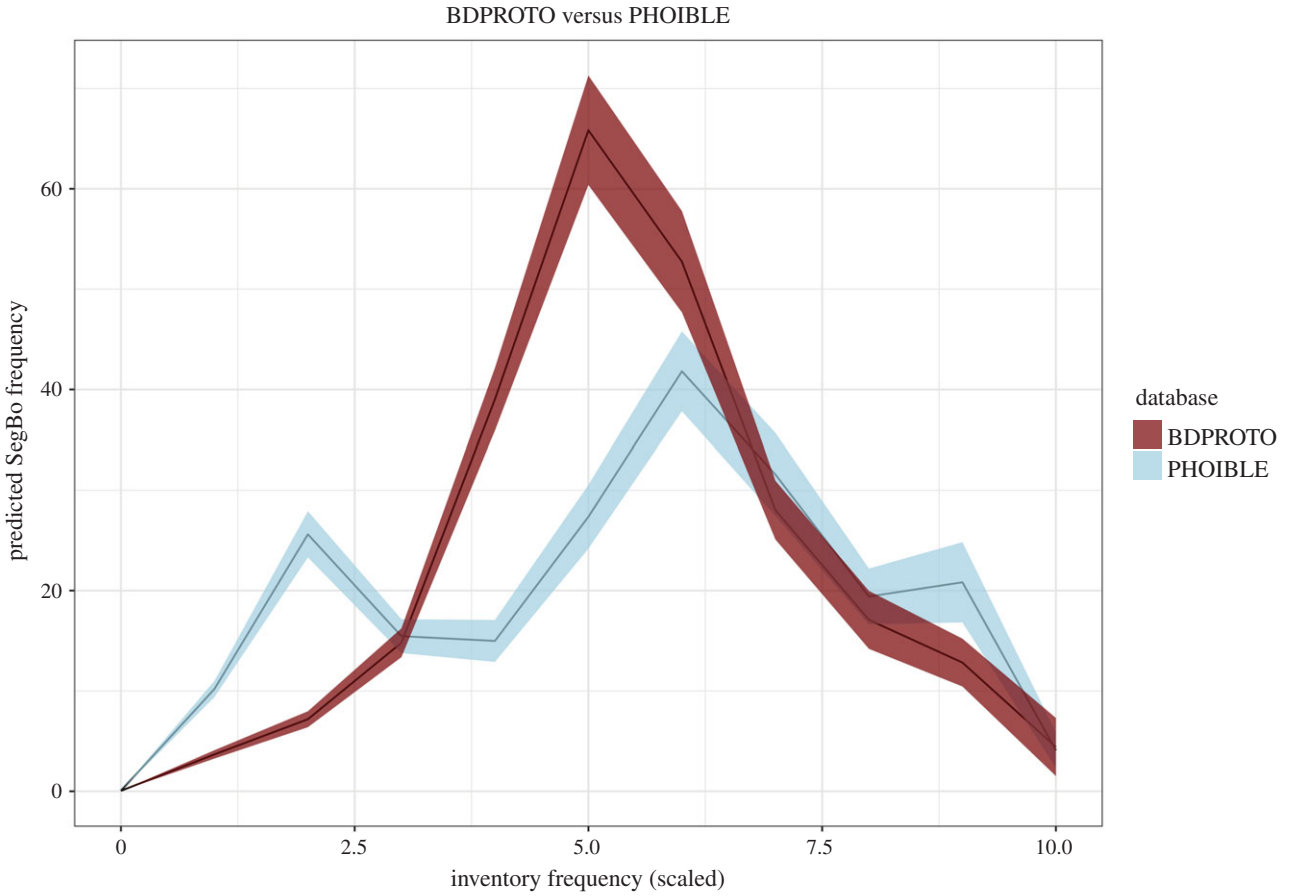

Figure 2. Results of the quasi-Poisson model. (Online version in colour.)

closer look at Africa, the area for which we have the most data, particularly in PHOIBLE (which allows us to take a more fine-grained perspective on areal factors). We divide the African macro-area by the Saharan desert region, with different types of languages appearing on either side (and indeed, with differing contact environments). We make this split in two ways. First, we split according to latitude (languages that fall above latitude 23.806078). Second, we split according to family, given that Afro-Asiatic or Nilo-Saharan languages are generally associated with the Northern part of the continent, and Niger-Congo and others with the Subsaharan region. Here we only report the results of the latitude-based split. The other analysis can be found in the supplementary material. The results are similar for the Subsaharan region, but BDPROTO behaves differently in the family-split analysis. In both cases, however, the overall relationships for the frequency bands of interest are similar. This small case study is intended to determine to what extent the method we have applied broadly to the global sample reveals some kernel of truth about the behaviour of well-defined subparts of the data.

The overall model revealed significant main effects of frequency (analysis of deviance: $\chi^2(7,\ N = 5554) = 17014.9$, $p < 0.001$) and the source of the frequency ($\chi^2(1,\ N = 5554) = 32.5$, $p < 0.001$). Furthermore, we find a significant interaction between the source and frequency ($\chi^2(7,\ N = 5554) = 355.1$, $p < 0.001$).

Figure 2 illustrates the model differences between PHOIBLE and BDPROTO as they relate to SegBo frequencies. Digging into the results, the segments that are drawn to the median frequency band across time, based on the comparison of BDPROTO and PHOIBLE, are mostly fricatives and affricates, including the labiodentals /f, v/, the alveolars and postalveolars /s, z, ʃ, ʒ/, the voiced velar /ɤ/, the voiced pharyngeal /ʕ/ and the glottal /h/, as well as the postalveolar affricates /t͡ʃ d͡ʒ/. Additionally, the low-frequency rhotic /ɾ/ is drawn into the mid-frequency range. Broadly, these are sounds that are significantly underrepresented in one or more of the world's macro-areas. They are also among the most frequently borrowed sounds in SegBo.

For some of these sounds, we have independent evidence that they have become more frequent in the past millennia. For example, in present-day languages, labiodental sounds including /f/ and /v/ are among the top 40 most frequent in the world's languages, which means that they are not extremely common: /f/ and /v/ are found in 44% and 27% of the world's languages, respectively [14]. However, since most of the segments documented in PHOIBLE are found in $<0.01\%$ of the world's languages, these are clearly among the most frequently-attested sounds known. Importantly, Blasi *et al.* [45] demonstrate that labiodental fricatives became frequent in human languages only after the development and diffusion of agriculture, and in particular milling, and hence soft diets. These sounds are overwhelmingly absent from ancient and reconstructed languages [13,57], and first diffused in areas where agriculture developed. Finally, and most strikingly, these are among the most frequently borrowed sounds in SegBo, which documents relatively recent borrowing events. Taken together, these findings point to a later, second wave of diffusion of labiodental segments in the world's languages, plausibly resulting from the spread of

Table 3. Top-10 borrowed segments per macro-area.

	Africa	Australia	Eurasia	North America	Papunesia	South America
1	p	o	f	g	d͡ʒ	g
2	h	t̪	ʒ	b	t͡ʃ	b
3	z	s	z	d	h	f
4	ʃ	f	d͡ʒ	f	r	d
5	f	v	x	ɾ	g	r
6	v	n	v	l	z	o
7	d͡ʒ	eː	ts	p	v	e
8	g	e	t͡ʃ	z	l	s
9	x	ʎ	ʃ	t͡ʃ	b	l
10	r	ʃ	ʕ	o	p	ɾ

large colonial languages. And in the present context, the diffusion of labiodentals in recent times has had a homogenizing effect on the phonological inventories of the world's languages. However, this diffusion has not had the same effect everywhere: the indigenous languages of Australia, North America and South America still show a lower-than-average presence of labiodentals relative to the global average. The end result is that these sounds have come to occupy a position of flux; they are neither so rare nor so frequently available as to preclude borrowing. Instead, they appear in the middling range frequency ranks, a position which indicates that these formerly infrequent sounds have been drawn into roughly average prominence. We further note that this change in frequency suggests that these sounds are optimally positioned for further spread; they are both rare enough to fill gaps in inventories and accessible enough to be successful in filling such gaps.

It is plausible that a similar logic underlies other sounds that drove the shift from earlier to present-day distributions. While we have not undertaken in-depth research on each segment or segment class (cf. the study on labiodentals through time by Blasi *et al.* [45]), we can raise some hypotheses. At a global level, we hypothesize that the correlation between the segments that best distinguish between BDPROTO and PHOIBLE, on the one hand, and SegBo, on the other, resulted from large-scale segment borrowing events in the recent past, which had an homogenizing effect on today's worldwide distribution of speech sounds.

At an areal level, however, the data are trickier to interpret. We investigate this issue from several perspectives. First, we turn to the quantitative estimation of the over- or under-representation of sounds with respect to global averages (represented as Δ) in particular macro-areas. Essentially, overrepresented sounds (+Δ) are positive areal signals or features, while underrepresented ones (−Δ) are negative areal signals or features, as discussed in §1.[9] Sounds with strong positive or negative deltas (taking more than 15% as an arbitrary cutoff) in particular areas, like the strong positive delta for retroflex consonants in Australia or aspirated stops in Eurasia, or the negative delta for voiced stops in Australia or palatal and velar nasals in North America, may be the result of older phylogenetic and areal effects. Strong positive deltas might also facilitate the further diffusion of sounds within a given macro-area, by making these sounds highly available for languages that lack them. For example, even though Africa is well above the global average for some voiced stops (e.g. /g, b/) and some voiced fricatives (e.g. /v, z/), these are among the segments most frequently borrowed in relatively recent times. Similarly, Eurasia has a small positive delta for the fricatives /x/ and /ʒ/, which means that these sounds are generally frequent enough to be borrowed but not so frequent that no language could borrow them; these are among the more frequently borrowed sounds in this macro-area. On the other hand, if a particular area is so saturated in terms of particular segments that there are few languages that could possibly borrow them, then a strong positive delta should not correlate with the most frequently borrowed segments. This is the case in Australia, in which retroflex consonants are found in nearly all languages, and indeed, do not show up as frequently borrowed sounds. A similar case is the prevalence of /k/ in South America, which is found in 98% of the languages in the area; it is therefore unsurprising that /k/ has almost no chances of being borrowed.

Strong negative deltas for a given segment in a particular area are rarely associated with a high frequency of borrowing, again highlighting the important role of cross-linguistic frequency for borrowing. However, there are some cases in which a speech sound underrepresented in a particular area was nonetheless frequently borrowed in that area. This is the case for voiced stops and fricatives like /f, z/ in the Americas, and for some fricatives and affricates in Papunesia. In such cases, it may be that a historical lack of these sounds in a particular area has been reduced by recent contact with languages originating from outside the area. Such a scenario is supported by SegBo data, which shows that the main donor languages in these areas originated outside the macro-area (e.g. Ibero-Romance in the case of South America and parts of North America; English, Arabic and Portuguese in the case of Papunesia), or were local languages that spread well beyond their homeland, like Indonesian in Papunesia.

However, weak deltas in either direction, on their own, are unlikely to reveal the effects of borrowing. In order to see more clearly the homogenizing effect of the diffusion of individual speech sounds across macro-areas, we present

Table 4. Comparing segment frequencies in the lower and middle bands of PHOIBLE/BDPROTO.

low-frequency				mid-frequency			
BDPROTO		PHOIBLE		BDPROTO		PHOIBLE	
f	1.9 (10)	ʒ	1.7 (2.2)	g	5.8 (5.1)	f	5.1 (10)
d͡ʒ	0.7 (5.1)	x	1.9 (1.8)	h	5.4 (4.3)	ð	3.1 (5.1)
t͡ʃ	1.8 (4.8)	ts	2.4 (1.5)	o	5.9 (2)	t͡ʃ	4.3 (4.8)
v	1.2 (3.7)	ɤ	1.6 (1.2)	e	5.9 (1.7)	h	6.0 (4.3)
ʒ	0.6 (2.2)	ʕ	0.2 (0.9)	ts	3.2 (1.5)	z	3.4 (4)
ɾ	0.5 (1.9)	q	0.8 (0.9)	ŋ	5.3 (1.5)	v	3.3 (3.7)
ɤ	1.6 (1.2)	ə	2.5 (0.7)	ʔ	4.7 (1.4)	r	4.6 (3.5)
ʕ	0.7 (0.9)	dz	1.1 (0.7)	ɲ	3.4 (1.2)	d	5.0 (3.3)
q	2.0 (0.9)	ʎ	0.5 (0.7)	eː	3.1 (0.3)	ʃ	3.8 (2.9)
ɛ	1.5 (0.7)	ɸ	0.4 (0.7)	aː	3.4 (0.2)	ʔ	3.7 (1.4)

table 3, based on Grossman *et al.* [11] and Eisen [12]. We clearly see the prevalence of the fricatives and affricates across macro-areas, as well as several more area-specific borrowing profiles. Importantly, most of these speech sounds figure in table 4 as sounds driving the difference between BDPROTO and PHOIBLE. Furthermore, a comment about segments that have low frequencies in PHOIBLE is in order. All of these segments, despite having relatively low frequencies in PHOIBLE, are still among those speech sounds that are the most frequent in the world's languages, keeping in mind that PHOIBLE has a strong positive skew and many segments in PHOIBLE are documented in a very small percentage of the world's languages. None of the segments in table 4 are part of this long tail. We raise this issue because even speech sounds that reach these cross-linguistic frequencies may have undergone an increase in frequency relative to BDPROTO.

In each of the macro-areas, the segments frequently borrowed rose in frequency per area in relation to an earlier state of affairs. We can be confident about the relative lateness owing to the outsized influence of large expansionist languages on SegBo [11,12]. Many of these segments are, again, precisely those that best distinguish the frequency distributions of speech sounds in BDPROTO and PHOIBLE. However, some of these distinctive segments are not among the most borrowed, such as /ɸ/, and with respect to these speech sounds it is likely that an alternative explanation should be sought.

We now raise another empirical issue of interest. For the most part, voiced stops are not picked out by the analyses as drivers of the difference between BDPROTO and PHOIBLE. However, these are among the most borrowed sounds in several areas, particularly in the Americas and Papunesia. We suggest that these speech sounds are so frequent in both BDPROTO and PHOIBLE that they do not distinguish the two distributions at a global scale, but they did bring the languages of the Americas and Papunesia closer to today's global distributions, raising deltas that would otherwise be lower.

Next we turn to the Poisson regression analysis of each macro-area, plotted in figure 3. For the simple by-area models (which most closely resemble the overall model in structure), we see largely similar shapes for the behaviour of both BDPROTO and PHOIBLE frequency distributions. The major exceptions are spikes in the high-frequency PHOIBLE range for Africa and Australia. Further, Australia shows the most distinctive pattern, having a much higher proportion of highly borrowed sounds.

For the random-effect models, the most stable features are (i) the greater amount of highly borrowed sounds in the mid-frequency range of PHOIBLE relative to BDPROTO and (ii) the rapid decreases on either end of the frequency spectrum. We also see a much more pronounced effect of BDPROTO in the high-frequency range (e.g. 7.5 in scaled frequency). Finally, the random-effect models are more sensitive to the high-band spikes for PHOIBLE that we observe for Africa, Eurasia, and Australia.

We now briefly explore the sounds that distinguish BDPROTO from PHOIBLE in the higher frequency band for each database in the African area, in order to understand the source of the secondary high-band spike. Examining the specific sounds that drive the difference between BDPROTO and PHOIBLE for Africa overall, we find that the high-frequency sounds in each database that were also found to be frequently borrowed include the voiced stops /b, d/ in BDPROTO and the relatively heterogeneous but simple sounds /f, d, p, l, e, t/ in PHOIBLE. As noted above, we divided Africa into a northern and a southern region in two different ways. The results are plotted in figure 4.

These plots reveal that the southern languages are the ones driving the secondary spike, a result that is not affected by the particular method of dividing Africa into northern and southern regions. In other words, northern Africa, like the global sample and like most macro-areas, shows the recurrent pattern of mid-frequency speech sounds being the more frequently borrowed, while in southern Africa, some sounds that are highly frequent in PHOIBLE are often borrowed. Importantly, we observe, again, the remarkable consistency, between these smaller-than-continent-sized areas and the plots we have observed elsewhere: an asymmetry at the low-end of the frequency spectrum (BDPROTO sounds tend to be more borrowable), as well as higher frequency of borrowing

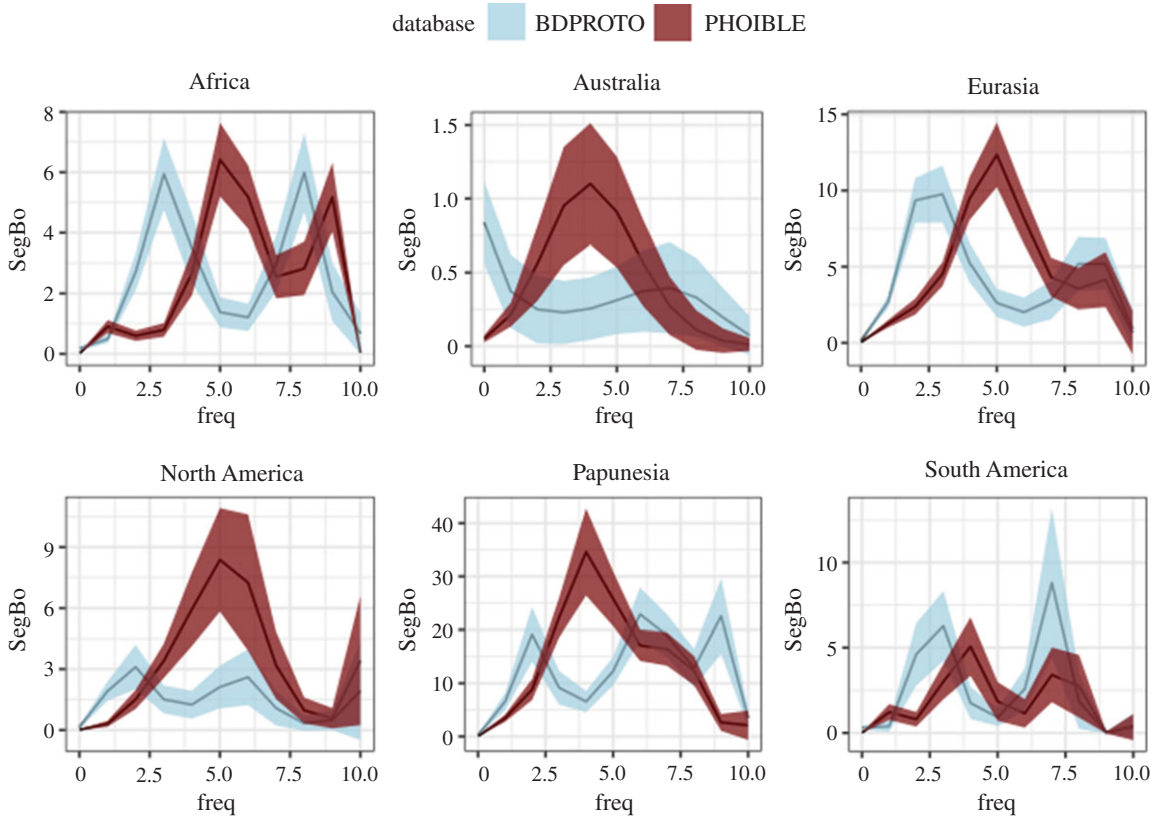

Figure 3. Results of Poisson regression analysis by area. (Online version in colour.)

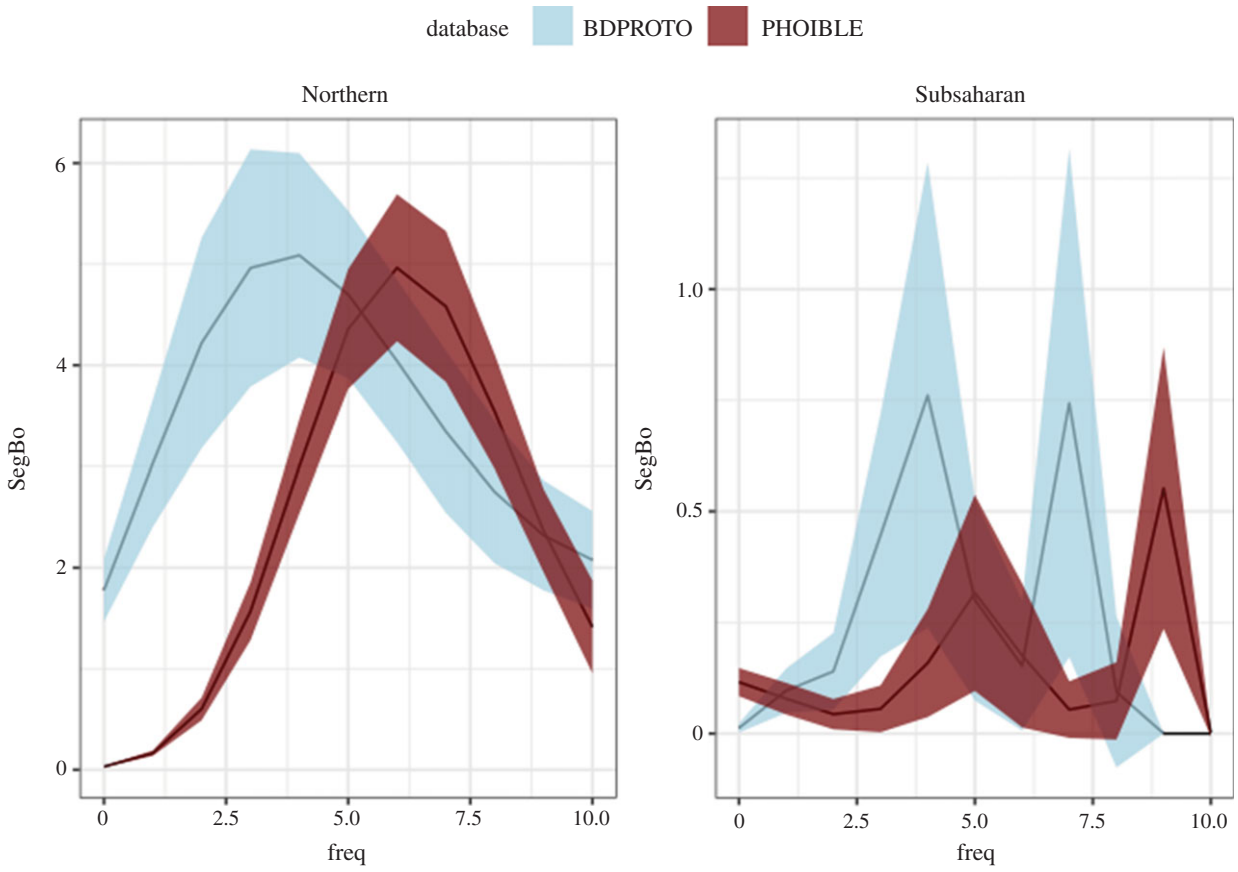

Figure 4. Northern versus Subsaharan Africa. (Online version in colour.)

for PHOIBLE vis-à-vis BDPROTO in the middle range (sample sizes have been reduced, hence the standard error increases; nevertheless, the centroid shift is apparent).

Summarizing the results of this analysis, over time, the segments that are drawn towards the median frequency band across languages are those that are reportedly most frequently borrowed. Conversely, the least frequently borrowed sounds are drawn to the extremes of the frequency range. This fits with the idea that both typologically dispreferred sounds (e.g. sounds that are rare because they are difficult

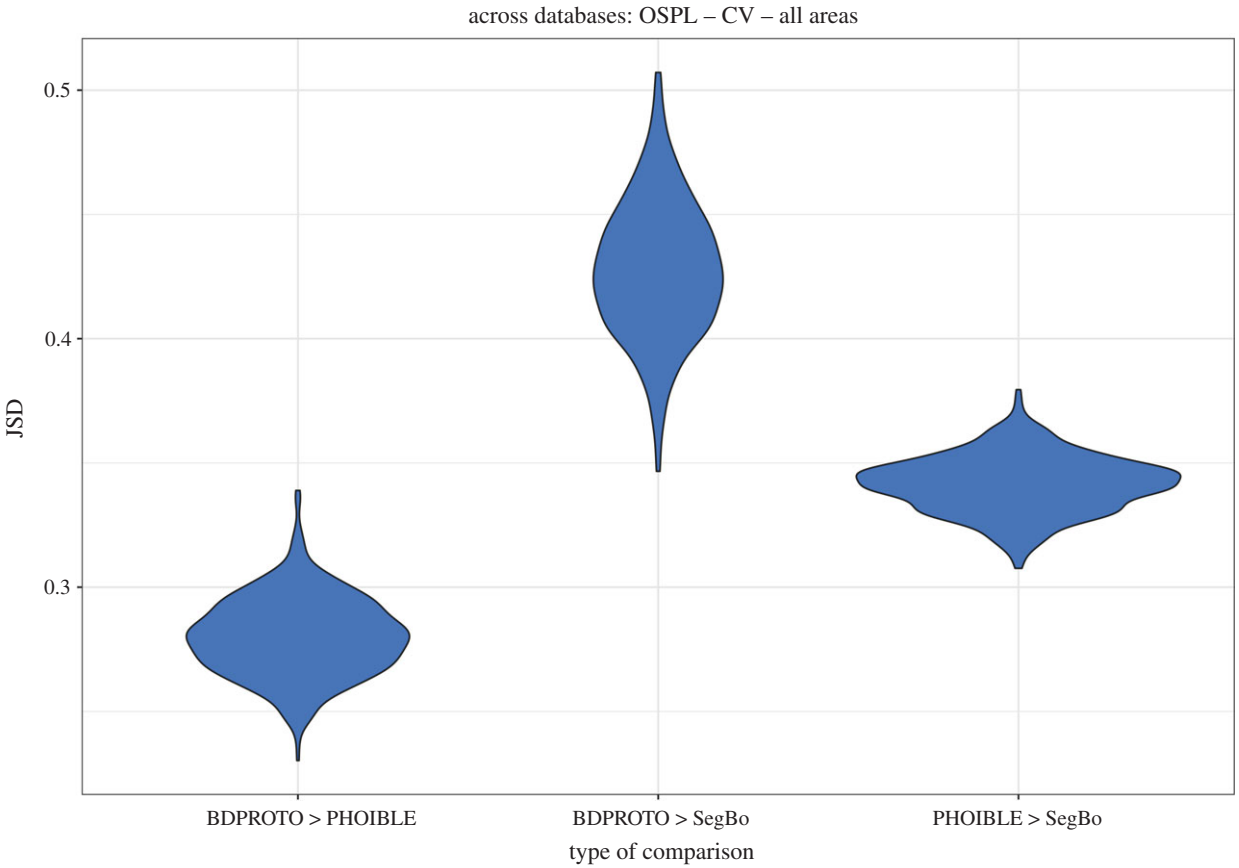

Figure 5. JSD analysis (500 iterations). CV, consonants and vowels. (Online version in colour.)

to articulate or perceive) and sounds that are typologically preferred (i.e. the sounds that languages tend to already have) are the least likely to be borrowed in contact scenarios [12]; whereas those in the mid-range of the curve are the most likely to be borrowed. Furthermore, we emphasize that the largest discrepancy between BDPROTO and PHOIBLE arises for those sounds that are the most frequent in SegBo.

Importantly, these results hold both at a global level and for each macro-area tested. As such, we conclude that our results support the alternative hypothesis: the most frequently borrowed segments in the recent past largely correspond to those segments that distinguish BDPROTO and PHOIBLE. We take this as support for the plausible role of recent language contact in shaping the distribution of speech sounds in the languages of the world today.

(d) Testing the JSD model with random resampling

The frequency estimates have so far demonstrated a difference in the distribution of frequently borrowed sounds within BDPROTO and PHOIBLE. However, we still lack a clear picture of the global estimates of similarity (i.e. JSD). That is, we know how the frequency profiles stack up, but not how similar they are on a sound-by-sound basis (i.e. which sounds show what frequencies across the databases) when accounting for the various biases we have outlined above. It is therefore important to determine whether our overall JSD estimates of similarity are skewed and/or variable across macro-areas or families. Moreover, we have not yet examined whether the differences we observe in general significantly distinguish BDPROTO and PHOIBLE in terms of their similarity to SegBo. Doing so requires two steps. First, we test how robust the JSD estimates are by applying a one-sound-per-language (OSPL) random resampling technique. On each iteration, we randomly sample a single sound per inventory and compute the new frequency distribution. We then take the JSD between the new distributions. Because of differences in areal coverage of the databases, we constrain the set of inventories from which we sample by macro-area. This process is repeated 500 times per macro-area to produce a distribution of JSDs. This process is designed to guard against biases in the frequency distributions of the databases, e.g. owing to the potential overrepresentation of particular language families or geographical areas, with otherwise low-frequency segments. Second, these JSDs are modelled as a function of comparison type (BDPROTO > SegBo versus PHOIBLE > SegBo). We further include an interaction term to test whether the effect differs across macro-areas. In this way, we aim to (a) account for the geographical biases inherent to each database and (b) see whether differences in JSD associated with BDPROTO or PHOIBLE generalize across geographical regions.

All consonants and vowels were included, and tones were excluded. Predictors were comparison type (BDPROTO > SegBo, PHOIBLE > Segbo) and macro-area (Africa, Australia, Eurasia, North America, Papunesia, South America), along with the interaction between the two. Because SegBo is partially derived from PHOIBLE (i.e. a portion of the inventories used in SegBo were taken directly from PHOIBLE), we performed a second regression analysis on only those languages that are not shared by these two databases.

This model produced results almost identical to the one based on the full set of inventories above. We therefore report the model based on the more complete sample here, presented in figure 5 (all analyses are available in the electronic supplementary material).

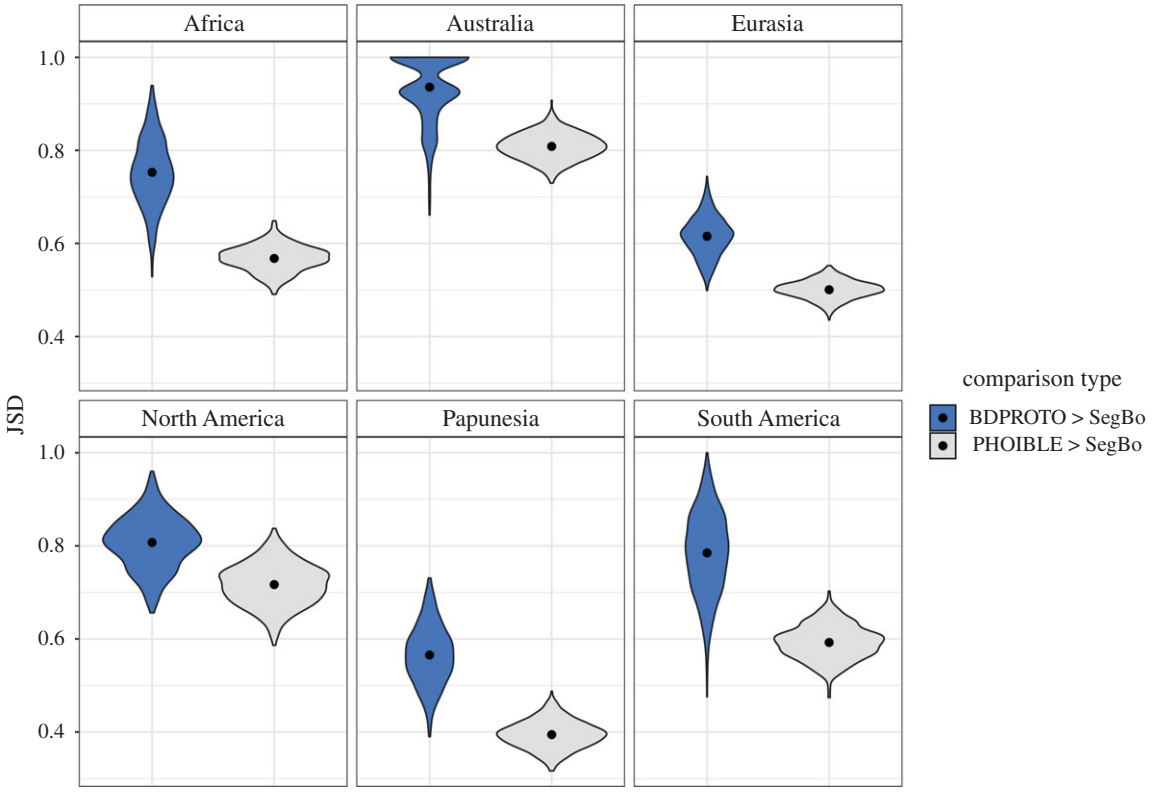

Figure 6. JSD for database pairs. Points are predicted values. Shaded areas represent the density (width) and range (height) of partial residuals. (Online version in colour.)

The model revealed that SegBo was significantly more similar to PHOIBLE than to BDPROTO ($F_{1,5988} = 12025.77$, $p < 0.001$) and that the macro-areas differed in average JSD ($F_{5,5988} = 7300.15$, $p < 0.001$). Finally, the effect of comparison type varied across macro-areas ($F_{5,5988} = 159.82$, $p < 0.001$; see figure 6). Crucially, the difference was in the same direction for all macro-areas. The closest matches between databases occur in Papunesia and Eurasia; the worst matches were found for Australia; the Americas were in the middle. Figure 6 also reveals that the variability in JSD is generally greater for BDPROTO than for PHOIBLE. Thus, the relationship between PHOIBLE and SegBo is not only closer on average, but also more stable than the relationship between BDPROTO and SegBo. We note finally that the scale of the shift across macro-areas is fairly consistent. Unfortunately, we cannot offer an interpretation of the effect size (i.e. we cannot state *a priori* how similar the distributions should be, or what would constitute a large shift). Nevertheless, the effect appears to be of moderate size; the means are generally shifted by about 0.2, or 20%. In other words, PHOIBLE is roughly 20% more similar to SegBo than BDPROTO, irrespective of macro-area.

We consider this evidence for the hypothesis that language contact, specifically large-scale borrowing events, has driven languages in all macro-areas of the world to a more homogeneous distribution of speech sounds in the relatively recent past.

5. Discussion

We have compared the cross-linguistic frequency distributions of speech sounds from ancient and reconstructed languages against present-day languages and we have measured how they differ from each other, and how these differences are related to relatively recent contact-induced changes. We report our two main findings. First, we find substantial differences between ancient and reconstructed languages, on the one hand, and present-day languages, on the other. Specifically, the distributions of specific speech sounds, in particular fricatives and affricates, have increased over time.

Second, and crucially, the greatest disparities between earlier and present-day distributions turn out to be largely owing to those sounds that have spread in the world's languages owing to relatively recent borrowing events. The result of these mass borrowing events in recent times is that languages have become more homogeneous, with correspondingly reduced areal-specific profiles, in terms of their phonological inventories. In this respect, the period in the past approximately 500 years or so seems to have been a watershed in the evolution of phonological systems.

This raises the issue of the relationship between the so-called 'functional factors' and historically-contingent 'event-based factors' [58]. The former refer to factors 'grounded in the biological/cognitive or social/communicative conditions of language, such as specific processing preferences … or specific sociolinguistic constellations … that systematically bias the way linguistic structures evolve' ([58], 42). Important examples of this in the present context are inherent biases related to the ease of production and perception of particular sounds. 'Event-based factors' are historical events whose main effect is to bring speakers of different languages into and out of contact with speakers of other languages. While we do not directly test for the role of event-based triggers in this study, previous research [11] makes it plausible that the large-scale expansion of a small handful of colonial languages (and other globalization processes) augmented functional factors related to the relative ease of production and perception of particular sounds, as well as other

historically contingent events, such as changes in food-technologies and resulting changes in aspects of human physiology relevant to speech production [45]. Such a scenario is reminiscent of proposals found in Bickel [59] and related work, according to which the spread of particular grammatical properties in Eurasia were the result of interacting functional and event-based factors. Whatever the ultimate causes of the homogenization of phonological inventories in relatively recent times, the present study shows that there is a strong correlation between contact-induced language change and the present-day cross-linguistic distribution of speech sounds.

In brief, the analyses presented here do not support the null hypothesis of time-independence of cross-linguistic distributions. Rather, they support the alternative hypothesis explored, which is that in terms of the distribution of their speech sounds, present-day languages are substantially different from ancient and reconstructed languages.

Moreover, we suggest that the problems posed by temporal bias [13], which is essentially the problem of autocorrelation compounded by increasing data sparsity as we look deeper into the past, may be mitigated by the appropriate use of resampling techniques. The analyses presented here provide a measure of confidence that the results are not artefacts of standard sampling techniques.

Our findings suggest that the Implicit Uniformitarian Hypothesis, at least with respect to the composition of phonological inventories, cannot be held uncritically. Insofar as cross-linguistic distributions may have differed substantially in earlier times, theories about, for example, markedness, may have to recognize that present-day distributions may be contingent on the historical events that led to new language contact situations to a larger extent than previously supposed. In other words, generalizations about present-day sound systems of spoken languages may not be time-independent, and linguists who would like to draw inferences about human language based on cross-linguistic distributions may have to consider their theories in light of the pressures contributing to language evolution, even in the short term.

Data accessibility. Data and code required to reproduce all results presented in this paper are available as electronic supplementary material at: https://github.com/bambooforest/inferring-paper.

Competing interests. We declare we have no competing interests.

Funding. Steven Moran was funded by the Swiss National Science Foundation (Grant No. PCEFP1_186841).

Acknowledgements. We would like to thank Haim Dubossarsky, Elad Eisen, Caleb Everett, Dmitry Nikolaev, Yves Tillé, and two anonymous reviewers for their suggestions and feedback. The usual disclaimers apply.

Endnotes

[1]Owing to the paucity of tones and vowels in some of the databases analysed in this study, the tables in the appendix highlight consonant distributions. However, the full dataset is given in the Supplementary Information: https://github.com/bambooforest/inferring-paper.
[2]However, it is interesting to note that in regions where languages have such speech sounds, such as southwest Asia, they are readily borrowed.
[3]https://github.com/bdproto/bdproto.
[4]The historical-comparative method is considered the gold standard for demonstrating language relatedness and for reconstructing grammar and lexicon. However, the quality of individual reconstructions may vary with respect to the nature of the data to which the method is applied and with respect to the rigour of the application of the method. The inclusion or exclusion of individual datasets in BDPROTO relied primarily on the authors' estimation of the quality of the reconstruction, taking into account these two caveats.
[5]https://phoible.org/.
[6]https://github.com/segbo-db/segbo.
[7]The term 'Papunesia', coined by editors of the Glottolog [51], is a portmanteau of Papua New Guinea and Austronesia, the latter of which refers to the islands of Insular Southeast Asia and Oceania, but excludes Australia. Other terms that have been used are 'Multinesia' [52], and 'Greater New Guinea Area' [53]. We eschew the term 'Oceania', since it typically includes Australia. We thus conform here to Glottolog's usage in order to avoid multiplying terms further.
[8]JSD is a symmetrical variant of the Kullback–Leibler divergence (KLD).
[9]We refer the reader to tables 5 and 6.

Appendix A

Tables 5 and 6 are given on the following page.

royalsocietypublishing.org/journal/rstb Phil. Trans. R. Soc. B 376: 20200198

Table 5. Sounds overrepresented with respect to the global mean, by area.

macro-area	segment	freq	PHOIBLE freq	Δ	PHOIBLE freq (without macro-area)	Δ
Africa	f	0.84	0.44	0.40	0.27	0.57
Africa	d	0.80	0.46	0.35	0.31	0.49
Africa	g	0.88	0.57	0.31	0.44	0.44
Africa	gb	0.42	0.12	0.30	0.00	0.42
Africa	kp	0.42	0.12	0.30	0.00	0.42
Africa	ɲ	0.71	0.42	0.29	0.30	0.41
Africa	b	0.91	0.63	0.28	0.52	0.40
Africa	z	0.56	0.30	0.26	0.19	0.37
Africa	v	0.52	0.27	0.25	0.16	0.35
Africa	s	0.89	0.67	0.22	0.58	0.31
Australia	ɳ̥	0.87	0.13	0.74	0.00	0.87
Australia	ʈ̥	0.72	0.10	0.61	0.00	0.72
Australia	ɻ̥	0.66	0.10	0.56	0.01	0.65
Australia	ɳ̊	0.66	0.13	0.53	0.04	0.61
Australia	l̥	0.63	0.12	0.51	0.03	0.60
Australia	ʈ̥	0.57	0.16	0.41	0.09	0.48
Australia	n̥	0.58	0.18	0.40	0.11	0.47
Australia	ɭ̥	0.46	0.07	0.40	0.00	0.46
Australia	ŋ	1.00	0.63	0.37	0.57	0.43
Australia	t̪	0.50	0.23	0.26	0.19	0.30
Eurasia	pʰ	0.52	0.20	0.33	0.07	0.45
Eurasia	kʰ	0.52	0.20	0.32	0.08	0.44
Eurasia	d̪	0.36	0.14	0.21	0.07	0.29
Eurasia	t̪ʰ	0.31	0.13	0.18	0.07	0.24
Eurasia	x	0.36	0.19	0.17	0.13	0.23
Eurasia	b	0.80	0.63	0.17	0.57	0.23
Eurasia	ɖ̥	0.24	0.09	0.16	0.03	0.21
Eurasia	g	0.72	0.57	0.15	0.51	0.21
Eurasia	t̪ʰ	0.19	0.06	0.13	0.01	0.18
Eurasia	ʒ	0.29	0.16	0.13	0.11	0.18
North America	ʔ	0.86	0.37	0.48	0.35	0.51
North America	k’	0.42	0.08	0.34	0.05	0.36
North America	h	0.85	0.56	0.28	0.55	0.30
North America	ʃ	0.64	0.37	0.28	0.35	0.29
North America	t̪ʃ	0.67	0.40	0.27	0.39	0.28
North America	t̪ʃ’	0.32	0.06	0.25	0.04	0.27
North America	ts’	0.29	0.04	0.25	0.03	0.27
North America	t’	0.30	0.05	0.25	0.04	0.26
North America	kʷ	0.35	0.12	0.23	0.11	0.25
North America	p’	0.29	0.06	0.23	0.04	0.25
Papunesia	ʔ	0.52	0.37	0.14	0.37	0.15
Papunesia	s	0.76	0.67	0.09	0.67	0.09
Papunesia	β	0.17	0.10	0.07	0.10	0.07
Papunesia	d	0.52	0.46	0.06	0.45	0.07
Papunesia	g	0.62	0.57	0.06	0.56	0.06
Papunesia	ŋ	0.69	0.63	0.06	0.63	0.06

(*Continued*.)

Table 5. (*Continued.*)

macro-area	segment	freq	PHOIBLE freq	Δ	PHOIBLE freq (without macro-area)	Δ
Papunesia	n̪|n	0.11	0.05	0.05	0.05	0.06
Papunesia	ɸ	0.10	0.05	0.05	0.05	0.05
Papunesia	t̪|t	0.10	0.05	0.05	0.05	0.05
Papunesia	mb	0.15	0.10	0.04	0.10	0.05
South America	ɾ	0.72	0.26	0.47	0.17	0.56
South America	t	0.91	0.68	0.23	0.64	0.27
South America	t̪ʃ	0.63	0.40	0.22	0.36	0.27
South America	h	0.71	0.56	0.15	0.54	0.17
South America	ʔ	0.52	0.37	0.14	0.35	0.17
South America	ʃ	0.47	0.37	0.10	0.35	0.12
South America	ts	0.32	0.22	0.10	0.20	0.12
South America	p	0.95	0.86	0.09	0.84	0.10
South America	k	0.98	0.90	0.08	0.89	0.09
South America	β	0.17	0.10	0.07	0.09	0.08

Table 6. Sounds underrepresented with respect to the global mean, by area.

macro-area	segment	freq	PHOIBLE freq	Δ	PHOIBLE freq (without macro-area)	Δ
Africa	ʈ	0.03	0.16	−0.13	0.22	−0.19
Africa	p^h	0.06	0.20	−0.13	0.25	−0.18
Africa	k^h	0.07	0.20	−0.13	0.25	−0.18
Africa	ɳ̊	0.00	0.13	−0.13	0.19	−0.18
Africa	n̪	0.05	0.18	−0.12	0.23	−0.18
Africa	ɾ	0.14	0.26	−0.12	0.31	−0.17
Africa	ɭ̊	0.01	0.12	−0.11	0.17	−0.16
Africa	t̪	0.13	0.23	−0.10	0.28	−0.14
Africa	ɻ̊	0.00	0.10	−0.10	0.14	−0.14
Africa	t^h	0.06	0.13	−0.08	0.17	−0.11
Australia	s	0.00	0.67	−0.67	0.79	−0.78
Australia	b	0.02	0.63	−0.61	0.74	−0.71
Australia	h	0.01	0.56	−0.55	0.66	−0.65
Australia	g	0.02	0.57	−0.54	0.66	−0.64
Australia	d	0.02	0.46	−0.44	0.53	−0.51
Australia	f	0.00	0.44	−0.44	0.51	−0.51
Australia	t̪ʃ	0.00	0.40	−0.40	0.47	−0.47
Australia	ɲ	0.04	0.42	−0.38	0.48	−0.44
Australia	z	0.00	0.30	−0.29	0.35	−0.34
Australia	d̪ʒ	0.00	0.27	−0.27	0.32	−0.32
Eurasia	t	0.45	0.68	−0.23	0.77	−0.32
Eurasia	w	0.59	0.82	−0.23	0.90	−0.31
Eurasia	n	0.67	0.78	−0.10	0.82	−0.15
Eurasia	ɻ̊	0.01	0.10	−0.09	0.14	−0.13
Eurasia	k^w	0.03	0.12	−0.09	0.15	−0.12
Eurasia	ɓ	0.02	0.10	−0.07	0.13	−0.10

(*Continued.*)

Table 6. (*Continued.*)

macro-area	segment	freq	PHOIBLE freq	Δ	PHOIBLE freq (without macro-area)	Δ
Eurasia	mb	0.04	0.10	−0.07	0.13	−0.09
Eurasia	d	0.39	0.46	−0.06	0.48	−0.09
Eurasia	ŋg	0.03	0.10	−0.06	0.12	−0.09
Eurasia	nd	0.03	0.10	−0.06	0.12	−0.09
North America	ŋ	0.24	0.63	−0.38	0.66	−0.41
North America	ɲ	0.11	0.42	−0.30	0.44	−0.32
North America	f	0.20	0.44	−0.24	0.45	−0.25
North America	g	0.35	0.57	−0.21	0.58	−0.23
North America	r	0.24	0.44	−0.20	0.46	−0.22
North America	z	0.11	0.30	−0.18	0.31	−0.19
North America	v	0.09	0.27	−0.18	0.28	−0.19
North America	b	0.46	0.63	−0.17	0.64	−0.18
North America	d	0.29	0.46	−0.16	0.47	−0.17
North America	d͡ʒ	0.12	0.27	−0.15	0.28	−0.16
Papunesia	ʃ	0.07	0.37	−0.30	0.39	−0.32
Papunesia	z	0.07	0.30	−0.22	0.31	−0.24
Papunesia	j	0.71	0.90	−0.19	0.91	−0.21
Papunesia	v	0.10	0.27	−0.17	0.28	−0.18
Papunesia	t͡ʃ	0.24	0.40	−0.16	0.41	−0.17
Papunesia	ɲ	0.26	0.42	−0.16	0.43	−0.17
Papunesia	ts	0.06	0.22	−0.16	0.23	−0.17
Papunesia	ʒ	0.01	0.16	−0.14	0.17	−0.15
Papunesia	p^h	0.06	0.20	−0.14	0.20	−0.14
Papunesia	k^h	0.06	0.20	−0.14	0.21	−0.14
South America	ŋ	0.22	0.63	−0.41	0.71	−0.49
South America	r	0.06	0.44	−0.38	0.51	−0.45
South America	f	0.08	0.44	−0.37	0.51	−0.43
South America	l	0.32	0.68	−0.35	0.74	−0.42
South America	g	0.29	0.57	−0.27	0.62	−0.33
South America	v	0.03	0.27	−0.24	0.31	−0.28
South America	z	0.06	0.30	−0.23	0.34	−0.28
South America	b	0.44	0.63	−0.19	0.67	−0.23
South America	t̪	0.05	0.23	−0.19	0.27	−0.22
South America	n̪	0.02	0.18	−0.16	0.21	−0.19

References

1. Dediu D, Moisik SR, Baetsen WA, Bosman AM, Waters-Rist AL. 2021 The vocal tract as a time machine: inferences about past speech and language from the anatomy of the speech organs. *Phil. Trans. R. Soc. B* **376**, 20200192. (doi:10.1098/rstb.2020.0192)
2. Everett C. 2021 The sounds of prehistoric speech. *Phil. Trans. R. Soc. B* **376**, 20200195. (doi:10.1098/rstb.2020.0195)
3. Arsenault P. 2017 Retroflexion in South Asia: typological, genetic, and areal patterns. *J. South Asian Lang. Ling.* **4**, 1–53. (doi:10.1515/jsall-2017-0001)
4. Aikhenvald AY. 2012 *Languages of the Amazon*. Oxford, UK: Oxford University Press.
5. Gasser E, Bowern C. 2014 Revisiting phonotactic generalizations in Australian languages. In *Proc. Annual Meetings on Phonology* 1. (doi:10.3765/amp.v1i1.17)
6. Maddieson I. 1984 *Patterns of sounds*. Cambridge, UK: Cambridge University Press.
7. Maddieson I. 2013 Absence of common consonants. In *The world atlas of language structures online* (eds MS Dryer, M Haspelmath). Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/18.
8. Stolz T, Levkovych N. 2017 Convergence and divergence in the phonology of the languages of Europe. In *The Cambridge handboook of areal linguistics* (ed R Hickey), pp. 122–160. Cambridge, UK: Cambridge University Press. (doi:10.1017/9781107279872.007)
9. Houis M. 1974 A propos de /p/. *Afrique et langage* **1**, 35–38.

10. Clements GN, Ryalland A. 2008 Africa as a phonological area. In *A linguistic geography of Africa* (eds B Heine, D Nurse), pp. 36–85. Cambridge, UK: Cambridge University Press.
11. Grossman E, Eisen E, Nikolaev D, Moran S. 2020 SegBo: a database of borrowed sounds in the world's languages. In *Proc. 12th Int. Conf. on Language Resources and Evaluation*, Marseille, France, 11–16 May 2020, pp. 5316–5322. Marseille, France: European Language Resources Association. https://www.aclweb.org/anthology/2020.lrec-1.654.
12. Eisen E. 2019 *The typology of phonological segment borrowing*. MA thesis, Hebrew University of Jerusalem, Jerusalem, Israel.
13. Moran S, Grossman E, Verkerk A. 2020 Investigating diachronic trends in phonological inventories using BDPROTO. *Lang. Resour. Evaluat.* 1–25. (doi:10.1007/s10579-019-09483-3)
14. Moran S, McCloy D (eds). 2019 *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. See https://phoible.org/. (doi:10.5281/zenondo.2593234)
15. Bakker D. 2011 Language sampling. In *Handbook of linguistic typology* (ed. JJ Song), pp. 90–127. Oxford, UK: Oxford University Press.
16. Cysouw M. 2011 Understanding transition probabilities. *Linguistic Typol.* **15**, 415–431. (doi:10.1515/lity.2011.028)
17. Walkden G. 2019 The many faces of uniformitarianism in linguistics. *Glossa: J. Gen. Ling.* **4**, 52. (doi:10.5334/gjgl.888)
18. Janda R, Joseph B. 2003 On language, change, and language change – or, of history, linguistics, and historical linguistics. In *Handbook of historical linguistics* (eds B Joseph, R Janda), pp. 3–180. Malden, MA: Blackwell.
19. Newmeyer FJ. 2002 Uniformitarian assumptions and language evolution research. In *The transition to language* (ed. Alison Wray), pp. 359–375. Oxford, UK: Oxford University Press.
20. Dryer MS. 1989 Large linguistic areas and language sampling. *Stud. Lang.* **13**, 257–292. (doi:10.1075/sl.13.2.03dry)
21. Nichols J. 1992 *Linguistic diversity in space and time*. Chicago, IL: University of Chicago Press.
22. Maslova E. 2000 A dynamic approach to the verification of distributional universals. *Linguistic Typol.* **4**, 307–333. (doi:10.1515/lity.2000.4.3.307)
23. Piantadosi ST, Gibson E. 2014 Quantitative standards for absolute linguistic universals. *Cogn. Sci.* **38**, 736–756. (doi:10.1111/cogs.12088)
24. De Busser R, LaPolla RJ. 2015 *Language structure and environment: social, cultural, and natural factors*. Amsterdam, The Netherlands: John Benjamins Publishing Company.
25. Haudricourt AG. 1961 Richesse en phonèmes et richesse en locuteurs. *L'Homme* **1**, 5–10. (doi:10.3406/hom.1961.366337)
26. Nettle D. 1996 Language diversity in West Africa: an ecological approach. *J. Anthrpol. Archaeol.* **15**, 403–438. (doi:10.1006/jaar.1996.0015)
27. Trudgill P. 1989 Interlanguage, interdialect and typological change. In *Variation in second language acquisition: psycholinguistic issues* (ed. Susan Gass *et al*.), pp. 243–253. Clevedon, UK: Multilingual Matters.
28. Bentz C, Bodo W. 2014 Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change* **3**, 1–27. (doi:10.1163/22105832-13030105)
29. Donohue M, Nichols J. 2011 Does phoneme inventory size correlate with population size? *Linguistic Typol.* **15**, 161–170. (doi:10.1515/lity.2011.011)
30. Greenhill SJ. 2014 Demographic correlates of language diversity. In *The Routledge handbook of historical linguistics* (eds C Bowern, B Evans), pp. 557–578. London, UK: Routledge.
31. Hay J, Bauer L. 2007 Phoneme inventory size and population size. *Language* **83**, 388–400. (doi:10.1353/lan.2007.0071)
32. Lupyan G, Dale R. 2010 Language structure is partly determined by social structure. *PLoS ONE* **5**, 1–10. (doi:10.1371/annotation/0eabb45a-f9da-4c9d-8555-0efee6e777f8)
33. Moran S, McCloy D, Wright R. 2012 Revisiting population size vs. phoneme inventory size. *Language* **88**, 877–893. (doi:10.1353/lan.2012.0087)
34. Trudgill P. 1997 Typology and sociolinguistics: linguistic structure, social structure and explanatory comparative dialectology. *Folia Linguistica* **31**, 349–360. (doi:10.1515/flin.1997.31.3-4.349)
35. Trudgill P. 2011 *Sociolinguistic typology: social determinants of linguistic complexity*. Oxford, UK: Oxford University Press.
36. Widmer M, Jenny M, Behr W, Bickel B. 2020 Morphological structure can escape reduction effects from mass admixture of second language speakers: evidence from Sino-Tibetan. *Stud. Lang.* **36**. (doi:10.1075/sl.19059.wid)
37. Creanza N, Ruhlen M, Pemberton TJ, Rosenberg NA, Feldman MW, Ramachandran S. 2015 A comparison of worldwide phonemic and genetic variation in human populations. *Proc. Natl Acad. Sci. USA* **112**, 1265–1272. (doi:10.1073/pnas.1424033112)
38. Dediu D, Janssen R, Moisik SR. 2017 Language is not isolated from its wider environment: vocal tract influences on the evolution of speech and language. *Lang. Commun.* **54**, 9–20. (doi:10.1016/j.langcom.2016.10.002)
39. Dediu D, Ladd DR. 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)
40. Moisik SR, Dediu D. 2017 Anatomical biasing and clicks: evidence from biomechanical modeling. *J. Lang. Evol.* **2**, 37–51. (doi:10.1093/jole/lzx004)
41. 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)
42. Everett C, Blasi DE, Roberts SG. 2015 Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *Proc. Natl Acad. Sci. USA* **112**, 1322–1327. (doi:10.1073/pnas.1417413112)
43. Everett C, Blasi DE, Roberts SG. 2016 Language evolution and climate: the case of desiccation and tone. *J. Lang. Evol.* **1**, 33–46. (doi:10.1093/jole/lzv004)
44. Maddieson I, Coupé C. 2015 Human spoken language diversity and the acoustic adaptation hypothesis. *J. Acoust. Soc. Am.* **138**, 1838–1838. (doi:10.1121/1.4933848)
45. Blasi DE, Moran S, Moisik SR, Widmer P, Dediu D, Bickel B. 2019 Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* **363**, eaav3218. (doi:10.1126/science.aav3218)
46. Ember C, Ember M. 2007 Climate, econiche, and sexuality: influences on sonority in language. *Am. Anthropol.* **109**, 180–185. (doi:10.1525/aa.2007.109.1.180)
47. Marsico E, Flavier S, Verkerk A, Moran S. 2018 BDPROTO: A database of phonological inventories from ancient and reconstructed languages. In *Proc. 11th Int. Conf. on Language Resources and Evaluation*, pp. 1654–1658. Paris, France: European Language Resources Association. (See https://www.aclweb.org/anthology/L18-1262.pdf).
48. Marsico E. 1999 What can a database of proto-languages tell us about the last 10,000 years of sound changes? In *Proc. of the XIVth Int. Congress of Phonetic Sciences San Francisco, CA, USA, 1–7 August 1999* (eds JJ Ohala, Y Hasegawa, M Ohala, Daniel Granville, AC Bailey), pp. 353–356. Berkeley, CA: University of California.
49. Moran S. 2012 *Phonetics Information Base and Lexicon*: University of Washington dissertation. See https://digital.lib.washington.edu/researchworks/handle/1773/22452.
50. Grossman E, Eisen E, Nikolaev D, Moran S. 2020 Revisiting the Uniformitarian Hypothesis: can we detect recent changes in the typological frequencies of speech sounds? In *The evolution of language: Proc. of the 13th Int. Conf. (EvoLang13)* (eds A Ravignani, C Barbieri, M Martins, M Flaherty, Y Jadoul, E Lattenkamp, H Little, K Mudd, T Verhoef), doi:10.17617/2.3190925. http://brussels.evolang.org/proceedings/paper.html?nr=182.
51. Hammarström H, Forkel R, Haspelmath M, Bank SG. 2020 *Glottolog 4.2.1*. Jena: Max Planck Institute for the Science of Human History. https://glottolog.org/, accessed 25-09-2020..
52. Hammarström H, Donohue M. 2014 Some principles on the use of macro-areas in typological comparison. *Lang. Dyn. Change* **4**, 167–187. (doi:10.1163/22105832-00401001)
53. Hammarström H. 2016 Linguistic diversity and language evolution. *J. Lang. Evol.* **1**, 19–29. (doi:10.1093/jole/lzw002)
54. Lin J. 1991 Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **37**, 145–151. (doi:10.1109/18.61115)
55. Drost HG. 2018 Philentropy: information theory and distance quantification with R. *J. Open Source Softw.* **3**, 765. (doi:10.21105/joss.00765)

56. R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. (https://www.R-project.org/).
57. Hockett CF. 1985 Distinguished lecture: F. *Am. Anthropol.* **87**, 263–281. (doi:10.1525/aa.1985.87.2.02a00020)
58. Bickel B. 2017 Areas and universals. In *The Cambridge handbook of areal linguistics* (ed. R Hickey), pp. 40–54. Cambridge, UK: Cambridge University Press.
59. Bickel B. 2015 Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In *The Oxford handbook of linguistic analysis,* 2nd edn. (eds Bernd Heine & Heiko Narrog), pp. 901–923. Oxford, UK: Oxford University Press.