Language Sciences 69 (2018) 125–135

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

The similar rates of occurrence of consonants across the world's languages: A quantitative analysis of phonetically transcribed word lists

Caleb Everett

University of Miami, Anthropology Department, P.O. Box 248106, Coral Gables, FL, 33124-2005, USA

A R T I C L E   I N F O

Article history:
Received 30 November 2017
Received in revised form 3 May 2018
Accepted 8 July 2018
Available online 23 July 2018

Keywords:
Phonetics
Phonology
Zipf
Typology
Ease of articulation

A B S T R A C T

Via an analysis of nearly 7000 phonetically transcribed word lists representing every major linguistic family, this study examines the frequency with which languages use particular consonants. The results suggest that there is an underlying global similarity in the frequency of occurrence of these consonants. Some consonants are uncommon across the word lists, and do not occur frequently even in those languages in which they are used. In contrast, a few other consonants represent a large portion of sounds in the word lists, across regions and linguistic families. The analysis quantifies the pervasiveness of such sounds that are known to be, impressionistically, common in speech. A new method captures the overall similarity of consonants' rates of occurrence across languages and families. This method offers one means of evaluating the extent to which individual languages or families deviate from, or adhere to, the typical patterns of consonant usage. It is suggested that the crosslinguistic similarity of consonant usage, in terms of rates of occurrence in word lists, likely owes itself to previously documented factors like the relative ease of articulation of some sounds. This new evidence suggests that the role of ease of articulation in shaping speech may be more influential than generally assumed, though other explanatory factors are also likely at work and these patterns require further exploration with more robust intralinguistic phonetic corpora.

1. Introduction

The world's linguistic diversity includes an impressive array of sounds used to make speech happen (Maddieson, 1984; Evans and Levinson, 2009). Languages have from 11 to 141 phonemes, judging from surveys of the extant phoneme databases, and the kinds of sounds used as phonemes are well known to vary substantially (Gordon, 2016). Additionally, languages vary with respect to permissible syllable structures and sundry other articulatory phenomena. The sounds used in speech are, simply, remarkably diverse. In principle, languages could also vary dramatically with respect to the frequency with which they use the individual sounds at their disposal. Yet it is unclear just how diverse languages are in terms of how much they rely on particular sounds. Most work on the typology of sound systems focuses on the composition of phoneme inventories, not on the relative frequency in speech of individual phonemes and their variants. These inventories have played a critical role in offering insights into phonology (e.g., Clements, 2003) and have also served as the basis for studies on putative extra-linguistic associations with speech (Hay and Bauer, 2007; Atkinson, 2011). Studies of the latter sort have claimed, *inter alia*, that a language's number of phonemes correlates positively with speaker population size and negatively with its distance

E-mail address: caleb@miami.edu.

https://doi.org/10.1016/j.langsci.2018.07.003

from Africa, though such claims have been argued against compellingly (Moran et al., 2012; Maddieson et al., 2011). Despite such extensive and cross-disciplinary attention paid to phoneme inventories, though, relatively little attention has been paid to the rates at which sound types are utilized in language. This is understandable since large corpora of transcribed texts are available for only a fraction of the world's languages. Some research has investigated the rates of occurrence of consonants within that fraction of languages (Yegerlehner and Voegelin, 1957; Gusein-Zade, 1987). Yet no study has documented, on a global scale, the rates of occurrence of sound types within languages, as opposed to their prevalence across phoneme inventories. This study addresses that lacuna in the literature by analyzing a database of phonetically transcribed word lists representing thousands of languages across every major geographic region and linguistic phylum. Patterns in this database suggest that, despite their well-known phonetic and phonological diversity, languages are often quite similar vis-à-vis the rates at which certain phones occur in their word lists.

The data analyzed here offer additional support for previous studies that have presented physiological motivations for the pervasiveness of some sounds in speech (These include Locke, 1983, Ohala, 1983, and Lindblom and Maddieson, 1988, *inter alia.*). They do so by offering the advantage of examining phones in such a large sample of languages. Phoneticians and phonologists have long been aware that such large samples are beneficial, given that sufficiently representative samples are critical to investigating universal biases in articulation and perception. As Ohala (1980:182) notes, "Another general issue concerns the problem of how to obtain a truly representative sample of sound patterns from a variety of languages such that the sample is not biased by including too many or too few languages having certain genetic, typological, or geographical linkages." Here that problem is addressed. Nevertheless, the limitations of the database used in this study should also be borne in mind from the outset. While the data used here offer some clear advantages given the number of language varieties represented, some conclusions are tempered because each language is only represented with a small phonetic sample. Still, the data are suggestive of strong tendencies in the usage of consonants, across the world's language families. The particular tendencies focused upon in this study are generally consistent with previous work in the literature that is based on very different methodologies and databases. The distributional tendencies in question are used to quantify the similarity of consonant usage across languages, and to derive a way of isolating languages and linguistic phyla that are typologically anomalous with respect to their usage of consonants. The patterns uncovered here, and the method introduced, merit further exploration with larger corpora for individual languages.

2. Data and general analysis

The database relied on for this study is the Automated Similarity Judgment Program (ASJP), which contains transcriptions of words that are resistant to borrowing in thousands of languages (For the list of 40 words found across all lists, see https://en.wikipedia.org/wiki/Automated_Similarity_Judgment_Program). Most lists in the database represent 40 concepts that are essentially a subset of the 100-word Swadesh list, though some lists have all 100 words (Swadesh, 2006). This database was developed to investigate the relatedness of languages, though it has been used for other purposes also (Wichmann et al., 2016; Brown et al., 2013; Blasi et al., 2016). There are presently over 7000 word lists in the database, though this study focuses on 6901 word lists for distinct dialects and languages. It excludes, for example, word lists based on constructed languages. A map depicting the locations of the represented language varieties is provided in Fig. 1.

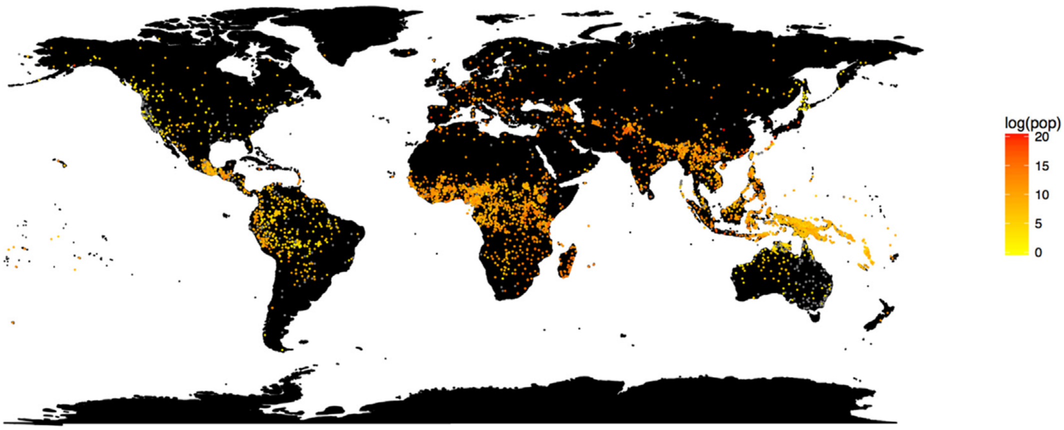

Fig. 1. Locations of the language varieties examined for this study, with log of speaker population. Extinct varieties are represented with gray dots.

There are 4571 distinct ISO codes represented in the data, so the majority of the world's approximately 7000 extant languages are represented. (*ISO* refers to the international standard of language codes, so the presence of 4571 distinct codes suggests that there are that many mutually unintelligible languages in the data.) Many languages are represented with word lists for more than one dialect. For that reason it is particularly critical to control for relatedness. These languages were grouped into families by relying on one commonly utilized language taxonomy, that of the World Atlas of Language Structures

(Dryer and Haspelmath, 2013). The transcribed words in the database tend to be frequent in speech, making them decent indicators of sounds' rates of occurrence within wider segments of speech (Calude and Pagel, 2011). The transcriptions are based on analyses of fieldworkers' transcriptions, written dictionaries, and other sources analyzed by the linguists who compile and curate the database. There are inevitable limitations to such a database. For example, function words are lacking, and function words are frequent in speech. In the case of English the voiced interdental fricative and the voiceless labiodental fricatives are not common to many words, but they are each found in a few highly frequent function words, like "the" and "of". It is worth noting, then, that the English data from the ASJP database are still fairly representative of the rates at which consonants occur. In another study (Everett, in press), the rates of occurrence of the consonants in the ASJP English list were contrasted with the rates of occurrence in a much larger corpus of naturalistic speech, and the association between the rates in the datasets was satisfactorily high ($R^2 = 0.82$) Nevertheless, the usage of this database admittedly sacrifices intralinguistic depth for the sake of crosslinguistic breadth, and there are limitations to any database representing so many languages (Vaux, 2013). Yet the database used here is the most comprehensive source available to analyze consonant usage across and within the majority of the world's languages. Other excellent global databases of sound types exist, but these focus on phoneme inventories and do not allow for the analysis of rates of occurrence (Moran et al., 2017; Maddieson, 2016; Maddieson and Precoda, 1990).

There are 34 basic consonant types represented in the ASJP database, as is evident in Table 1. Seven vowels are also represented in the database, but each of these vowels subsumes several vocalic configurations and the vowel symbols do not denote factors like tone, nasality, or length. Due to such factors, this study focuses on consonants.[1] The rate of occurrence for each of the 34 consonants encoded was obtained for each of the 6901 word lists. The rate of occurrence is the sum of the

Table 1
Phylogenetically controlled rates of occurrence for the 34 consonant types in the data. Rates of occurrence represent the ratio of all sound tokens, in a word list, represented by the sound type in question.

Main IPA sound (variants subsumed)	Rank, phylogenetically controlled	Average usage rate across families
n (n)	1	0.0588
k (k)	2	0.0546
m (m)	3	0.0449
t (t)	4	0.0445
r (r, ɽ, ɾ, ʀ)	5	0.0298
l (l)	6	0.0266
h (h, ɦ)	7	0.0256
p (p, ɸ)	8	0.0255
s (s)	9	0.0250
w (w)	10	0.0249
y (j)	11	0.0223
b (b, β)	12	0.0198
d (d)	13	0.0191
g (g)	14	0.0149
ʔ (ʔ)	15	0.0139
tʃ (tʃ)	16	0.0110
ŋ (ŋ)	17	0.0110
ʃ (ʃ)	18	0.0091
ts (ts, dz)	19	0.0081
x (x, ɣ)	20	0.0080
f (f)	21	0.0045
q (q)	22	0.0043
L (L, ɭ, ʎ)	23	0.0041
v (v)	24	0.0033
j (dʒ)	25	0.0031
! (!, ǀ, ʘ, ǂ, ǁ)	26	0.0026
z (z)	27	0.0024
c (c, ɟ)	28	0.0018
χ (χ, ʁ, ħ, ʕ)	29	0.0016
θ (θ, ð)	30	0.0012
ɲ (ɲ)	31	0.0012
ʒ (ʒ)	32	0.0009
ɢ (ɢ)	33	0.0002
n̪ (n̪)	34	0.00004

[1] Some fine-grained distinctions between consonants are also missed in the database, but it does consistently encode the 34 basic consonant types listed in Table 1. Some of these types include multiple related consonants. For example, the click symbol in the database represents all clicks, including lateral and bilabial clicks, and the rhotic symbol represents multiple rhotics. However, the main sound types focused on here exhibit a one-to-one relationship between sound type and database symbol.

instances of a sound, divided by the sum of the instances of all sounds in a list. Each member of a pair of coarticulated sounds, for instance the nasal and the stop in a "prenasalized" stop like [mb], was counted separately.[2] Diacritic marks for secondary articulatory factors, for instance the aspiration of a stop, were not included in the calculation of a sound's rate of occurrence. Brackets are used here when referring to specific sound types, since the transcriptions are phonetic. However the level of phonetic detail varies across word lists. Phonemic judgements and allophonic relationships are opaque in the database, and syllable structures are not encoded. For such reasons, this analysis focuses exclusively on the frequency of occurrence of basic consonant types. Word boundaries are evident in the database but, given the lack of phonemic information, speculations about the effect of word boundaries are avoided. Patterns of word-final devoicing, for example, are not explorable without greater phonological detail. (See, e.g., Locke, 1983:114–118.)

A consonant's rate of occurrence was calculated as follows: The total number of tokens of a given consonant in a word list was found, and that number was then divided by the total number of consonant and vowel tokens. For example, in one word list for Hawaiian there are 34 instances of [l] and 396 total consonants and vowels, meaning the rate of occurrence of [l] in Hawaiian is 34/306 or 0.111. In contrast, there are 12 instances of [p] in the word list, meaning the rate of occurrence for [p] is 12/306 or 0.039. Calculations like these were done for all 34 basic consonant types in each of 6901 word lists, so a total of 234,634 rates of occurrence (6901×34) were obtained (The code used for the analysis is available upon request; Raw data for all lists is available in the supplemental material.) Such findings are admittedly somewhat coarse, and do not offer fine-grained detail on, for instance, the range of features employed in a given phonological system (For discussions of the use of such features in the structuring of phonologies, see Clements (2003)). Still, they can be used to assess the relative frequency of major consonant types. In the case of Hawaiian, for example, we cannot use the database to assess any details of allophonic variation of/l/ and /p/, respectively, but we can claim that [l] is more common than [p] in the hundreds of phone tokens evident. Furthermore, the fact that many of the 34 transcribed consonants are lacking in the Hawaiian word lists tells us something about the language, which is known to have a small phoneme inventory. So the presence or absence of consonants in a language's word list can be used in a manner that is analogous to, though of course not identical to, the presence or absence of a sound in a language's phoneme inventory. In short, the word lists can help uncover useful information about the relative frequencies of consonants across and within languages.

As one visual representation of the findings on the relative frequency of consonants, consider Fig. 2 in which a heat map is used to represent the rates of occurrence of the obstruents and nasals. It is clear in the figure that some consonants are used much more frequently than others. For example, the columns representing [k], [n], [s], [t], and [m] are very bright, while those representing [n̪], [ɢ], and [ʒ] are particularly dark. The former sounds are known to be common in the world's phoneme inventories (Gordon, 2016).

To control for relatedness, the average rates of occurrence within each of the 261 phylogenetic groupings was obtained for each of the 34 consonants. These family averages were then averaged to obtain global averages for each consonant type. The resultant phylogenetically controlled global rates of occurrence of each consonant are listed in Table 1. The consonants are ranked from most to least common. Some consonants are particularly ubiquitous in the word lists, even after controlling for relatedness. There is a power-law association between consonant ranking and frequency, similar though not identical to the Zipfian association well-known to surface in other facets of speech like word usage (Zipf, 1949; Piantadosi, 2015) (The association is technically closer to a Yule-ian association–see Tambovtsev and Martindale (2007)).

A few consonant types play a fairly outsized role within and across the world's language taxa. This is particularly true of the alveolar nasal, the voiceless velar stop, the bilabial nasal, and the voiceless alveolar stop. The inordinate frequency of these sounds is also evident in the data without family controls, as seen in Fig. 2. As is also apparent in both Table 1 and Fig. 2, amongst fricatives the voiceless alveolar variety is particularly common. The five most common consonants, [n], [m], [k], [t], and [s], are all represented by individual graphemes in the ASJP database, without subsuming other sounds. So the fact that they are so frequent in usage, across linguistic phyla, is not explained by some feature of the transcription method utilized in the database. Which is not to say that there are not minor distinctions, missed in this analysis, with respect to how these sounds are articulated in different languages. Tellingly, though, [n], [m], [k], and [t] are also the most common consonants in the world's phoneme inventories. Ranked according to prevalence in those inventories, the top four consonants are as follows: 1. [n], 2. [t], 3. [m], 4. [k] (Gordon, 2016). In terms of frequency in the word lists, the ordering is similar: 1. [n], 2. [k], 3. [m], 4. [t]. Our consonant ranking is also similar to that found in previous work on intralinguistic sound frequency in 34 languages (Gordon, 2016).

The similarity of the usage ranking in Table 1, when contrasted with the more common phoneme-prevalence rankings, betrays a disparity between the rate of occurrence of a given consonant and its likelihood of serving as a phoneme: /n/ occurs in about 99% of the world's phoneme inventories, judging from previous work, while /t/ occurs in about 97%, /m/ in 93%, and /k/ in 89% (Maddieson, 1984). Many other consonants occur in over 50% of inventories, as there is a gradual decline in the cross-linguistic frequency of phonemes. For instance, /g/ is found in about 55% of phoneme inventories, over half as many as /n/ , while /s/ is found in about 84% (Gordon, 2016). In contrast, there is a clearer decline evident in the frequency of consonants in word lists. Sticking with the same examples, the rate of occurrence of [g], across families, is about one-fourth that

[2] However, the nasal segments in prenasalized stops represent only a small fraction of nasals in the word lists. The rates of occurrence of prenasalized stops, across all lists, are as follows: [mb] = 0.0016, [nd] = 0.0017, and [ŋg] = 0.0010. This methodological tack therefore does not have a major effect on the findings discussed here.

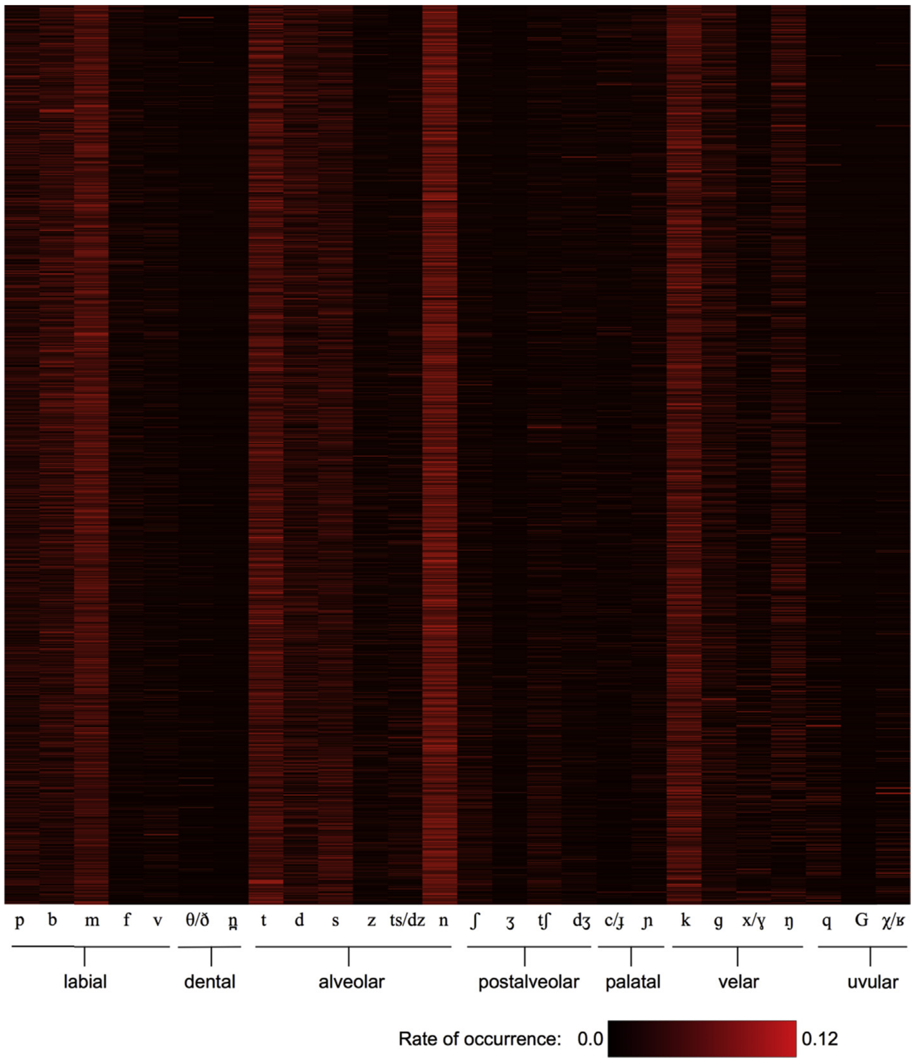

Fig. 2. Heat map of the rates of occurrence of the 26 obstruents and nasals in the data. Each row represents the consonants' rates of occurrence for one of 6901 word lists. Given the number of lists, similarly colored cells blend together. Each column represents one of the 26 consonants in question. (See Table 1 for information on all 34 consonants.) The brightness/darkness contrast between certain columns reflects the clear disparities in the relative frequencies of some consonants in the database. For example, [n] is clearly relied upon heavily across the bulk of the 6901 word lists, while [ɢ] is hardly used.

of [n], while that of [s] is less than half that of [n]. (See Table 1). In short, while a variety of consonants are quite common across the world's phoneme inventories, a few of these consonants are particularly preponderant in speech judging from these data. Consonant usage in these word lists is much more homogeneous than we might expect, revolving as it does around a few basic sounds, unlike the more heterogeneous phoneme inventories. (Bearing in mind, again, that some minor variation within sound types is not captured in these lists.)

Consider this illustration of the pattern in question: At least some dialects of English have 24 consonant phonemes (Jensen, 1993). In contrast, Xhosan languages are well known for their extremely large consonant inventories. Xhosan itself has 64 consonant phonemes, which do not include all English consonants (Jessen, 2002). Many of these, like clicks, are subsumed by one or a few symbols in ASJP data. Yet, in terms of usage, [n], [l], and [t] are three of the top four consonants in both languages. In fact, in Xhosan [n] is the most frequent and accounts for 12.9% of transcribed consonants, while in English it accounts for 12.8%. These three consonants, [n], [l], and [t], account for 33.3% of Xhosan consonants in the transcriptions, and 33.9% of those in English. Xhosan serves as an interesting contrast with English since the former language is famous for its enormous consonant inventory. But if one examines a language that occupies the other end of the phoneme diversity spectrum, consonant usage may still be similar. Hawaiian, like other Polynesian languages, is known for having few consonant phonemes, only eight. Yet in Hawaiian [n] and [l] are again two of the four most common consonants. Combined, [n] and [l] represent 32.2% of the consonants in the main Hawaiian word list. In other words, while Xhosan, English, and Hawaiian represent diverse points on the spectrum of phoneme inventory types, they exhibit some similarity in terms of the rates at which they use certain consonants. While the languages only share a few consonant phonemes out of the many they have (a combined 97), a few of these shared consonants carry very large loads in speech in all three languages. This is not to suggest that there are not important differences in the ways that sounds like [n] and [l] are utilized in these three languages. But we can nevertheless conclude that alveolar nasals and alveolar laterals represent a large portion of the consonants in word lists in all three languages.[3] Furthermore, this sort of similarity of usage is evident, to varying degrees, across all language families.

3. Similarity of usage across language families

The ranking of sound types in Table 1 also describes with some accuracy the ranking in individual languages and language families. In other words, we do not need to have any prior knowledge about a language to predict, upon being given the consonants that exist in the language, roughly how much each is used in a given word list.

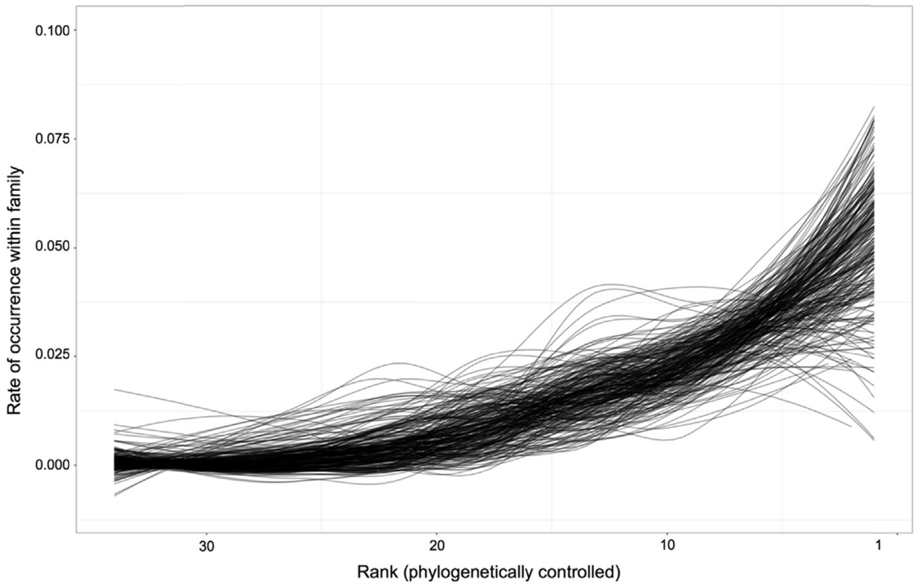

Fig. 3. Similarity in consonant usage across the 261 linguistic phyla in the data. Each of the 261 lines is a locally weighted polynomial regression (LOESS) contour depicting the associations between the 34 consonants' rates of occurrence, within a family, and their global ranking. The most frequent consonant worldwide, [n], is represented with a 1 on the x-axis. See Table 1 for the overall ranking of consonants.

As observed in Fig. 3, the language families exhibit similar patterns of consonant usage in the lists, in that the same few consonants dominate usage across linguistic phyla. A strong association between language-specific sound rankings and language-specific usage, for any given language, is not unexpected. What is notable about these data is that the one global ranking (in Table 1) has a fair amount of descriptive power vis-à-vis the frequency of consonants within each of the world's language families. How surprising this consistency is may depend somewhat on one's theoretical orientation. Yet, even if one is not particularly surprised by the relative consistency across lists, this analysis offers a means of more precisely quantifying that crosslinguistic similarity, which is arguably obscured somewhat by the phonemic heterogeneity of languages. For 261/261 families, a positive association was observed between the global rankings of the consonants and the average rates of occurrence of the consonants within each family's word lists. Once again, the limitations of the data should be kept in mind, since there are many granular distinctions between consonants that are missed in the lists. Still, the data do give us a sense of the dominance of some basic consonants in the word lists across the bulk of languages, and the comparable rarity of other consonants.

Tests were run to examine the relationships between the global average rates of occurrence and each family's average rates of occurrence. Since we are dealing with ratio data, a straightforward regression between the two rates of occurrence is not entirely appropriate. So the global averages were logit-transformed. Separate regressions were then run, for each of the families, between the average rates of occurrence of the sounds (within a family's word lists) and the logit-transformed global rates of occurrence of the sounds. The mean R^2 values for these 261 regressions was 0.33, and the R^2 values ranged from 0.02 to 0.55. In other words, some language families are much more consistent with the global averages than others. Notably, the

[3] At least in English, this is true for much larger word lists as well. /l/ and /n/are the first and third most common consonants, respectively, judging from thousands of lexical items. See Roberts (1965) and Locke (1983).

two language families that are the least consistent with the global averages are K'xa and Tu, two language families of sub-Saharan Africa that have many clicks in their lists. Since clicks are lacking in the vast majority of word lists, the exceptionality of these two language families is perhaps not surprising but also indicates that the regressions do offer a means of quickly highlighting language taxa that are typologically remarkable with respect to consonant usage.

Separate regressions were also run between the consonants' rates of occurrence in each of the 6901 lists, and the logit-transformed global rates of occurrence. These regressions revealed positive associations for every single list in the data set. This degree of consistency is perhaps surprising. Again, though, there was a fair amount of crosslinguistic variation, with R^2 values that ranged from 0.0001 to 0.61. The mean R^2 value across all lists was 0.28. Fig. 4 offers a clearer representation of the ranges of R^2 values obtained for these regressions.

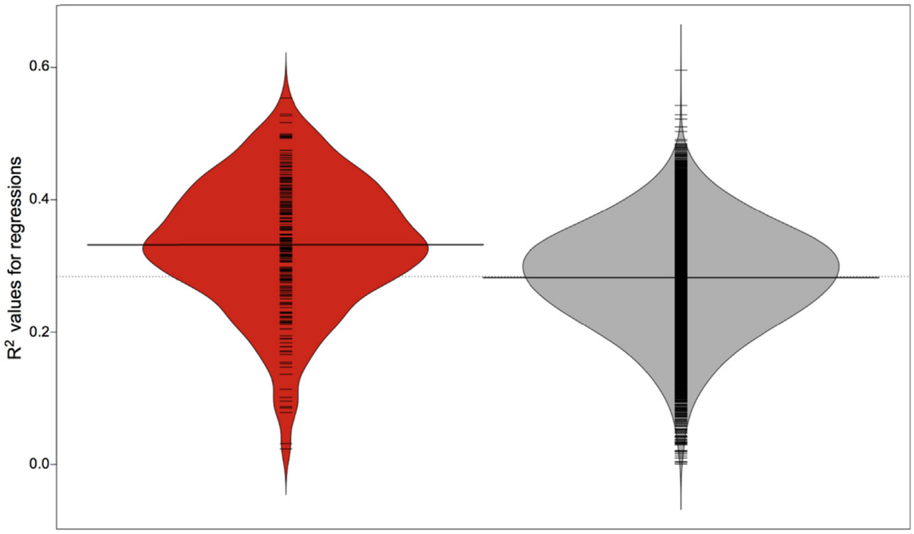

Fig. 4. Beanplots of R^2 values for the positively sloped regressions between rates of occurrence and the logit-transformed global average rates of occurrence (phylogenetically controlled). Beanplot A depicts the range of R^2 values obtained for each of 261 language families. Beanplot B depicts the range of R^2 values for the 6901 regressions based on individual word lists. For beanplots, the width of the bean corresponds to the density of the values obtained along the y-axis. The long dark horizontal lines represent the averages for each of the two beans. The small black horizontal lines in the center of each bean represent the R^2 values of the individual regressions (So there are 261 small black lines in beanplot A, and 6901 in beanplot B.)

Particularly strong or weak instances of the association should be approached with caution, given the limitations of any one list. The predictability of consonants' rates of occurrence is more apparent at the level of language families, which are represented with multiple word lists. Given that the words in the lists of related languages represent cognates in some cases, some cognates may contribute in pronounced ways to the association evident for particular families. However, any cognates that rely heavily on uncommonly used sounds would also work against the association for a given language family. Despite the latter possibility, however, we see that the associations are positive across all language families. They are also positive across all specific language varieties.

4. The role of articulatory ease

These results, like many previous results on phoneme inventories, suggest that languages are biased towards the use of certain consonants. It is worth recalling that there should be an explanation for such biases. As Lindblom and Maddieson (1988:64) note in a discussion of phoneme inventories, "we have no *a priori*, formal reason to expect any particular relationship between the relative frequencies of these classes of sound." Explanations of relative frequencies are likely to be found in articulatory and perceptual factors, and I - should stress that the results offered here are consistent with previous suggestions of some articulatory biases. Here again, the conclusions of Lindblom and Maddieson (1988:76) are relevant: "… if we know the number of consonants a language uses, we can make pretty good estimates of: (a) the number of obstruents and sonorants it will have; and (b) what their relative degree of articulatory complexity will be". The conclusions here are distinct but clearly related: If we know what consonant phones a language uses (judging from phonetically transcribed word lists), we can predict to some degree the relative frequency of these consonants, since the relative frequencies are often similar crosslinguistically. We can predict, for example, that [n], [k], [s], and [t] will be utilized quite frequently in languages that have these sounds. In other words, these data give us a new way of analyzing the "magnetic" pull (Lindblom and Maddieson, 1988:69) that certain articulatory configurations have on consonant production, and on speech more generally.

As noted above, a sound's frequency within and across languages is interrelated to, but still distinct from, the odds that the sound is phonemic in a given language. For instance, while it has previously been shown that we can predict with confidence that a randomly chosen language *x* is a bit more likely to have a voiceless velar stop phoneme than a voiced alternate, these data suggest that, in any language *x* with both of these sounds, [k] will also be used more frequently than [g]. It has been suggested that [k] is more likely to serve as a phoneme, when compared to [g], because [k] is easier to articulate (Ohala, 1983; Keating, 1984). The latter velar sound theoretically requires less effort than its voiced counterpart, since vibrating the vocal cords demands significant transglottal airflow. This transglottal airflow requires supralaryngeal air pressure to be lower than sublaryngeal air pressure. Supralaryngeal air pressure increases during vocal cord vibration, as air passes through the glottis. Oral constriction at the velum creates a relatively small pocket of air above the glottis, meaning that supralaryngeal pressure rises quickly and voicing is made more difficult for velar stops than for, say, bilabial stops. Phoneticians have long been aware of this phenomenon, and intralinguistic data suggest that duration of voicing differs in accordance with place of articulation (Ohala and Riordan, 1979; Napoli, 2014).[4]

Nevertheless it remains unclear just how much aerodynamic factors like this actually influence language in terms of their rate of reliance on particular stops. Additionally, it is worth noting that some linguists have claimed that ease of articulation does not significantly impact the evolution of phonologies, since languages do not converge on one set of easy-to-articulate phonemes, and supposedly do not converge phonologically in any pervasive manner (Kaye, 1989; Ploch, 2004). Even with respect to a phonetically well-grounded phenomenon like the crosslinguistic disparity between [k] phonemes and [g] phonemes, there is indeed room for a cautious approach. As mentioned above, [g] occurs in about 55% of surveyed phoneme inventories, while [k] occurs in about 89%. This is a major disparity, to be sure, but the phoneme surveys have not demonstrated how robust the disparity is across language families and regions, and similar disparities are observed within- other homorganic pairs of voiceless/voiced stops. In these data 99% of language families have [k] (260/261), while 74% have [g] (196/261). Where our data may advance the understanding of this issue, however, is in the examination of those 196 language families that have both [k] and [g] represented in their word lists. For one of these families, [k] and [g] have the same average rate of occurrence. For about 10% (19/196), the average rate of occurrence of [g] exceeds that of [k], by a margin of 0.017. Yet for about 90% (176/196) of these language families the average rate of occurrence of [k] exceeds that of [g], and by a much more pronounced figure of 0.038. Overall, then, [k] has a higher rate of occurrence than [g] in 244 out of 261 language families, and often by a very noticeable margin. This disparity is significant according to a Mann-Whitney-Wilcoxon test ($W = 6556.5$, $P = 0.000$) Framed differently: Phoneme-inventory data, traditionally relied upon in discussions of ease of articulation, may underestimate the worldwide disparity between [k] and [g] in speech. Once again, though, it should be stressed that these word list data are not meant to tell the entire story. For instance, the higher rates of [k], when contrasted to [g], do not tell us how many of the [k] tokens in word lists are allophones of [g] that have devoiced in particular contexts. They simply tell us that [k] is much more frequent than [g] in phonetically transcribed word lists, a fact that nevertheless warrants attention. Framed differently: The average rate of occurrence of [k], across families, is 0.055, while that of [g] is 0.015. The disparity between [k] and [g] can be expressed as a ratio of 3.67:1. In contrast, consider that the average rate of occurrence of [p] is 0.026, while that of [b] is 0.02. The ratio for this disparity is only 1.3:1. For alveolar stops, the average rate of occurrence of [t] is 0.045, while that of [d] is 0.19, for a ratio of 2.37:1. Note that the ratios for voiceless:voiced stops increase from the bilabial place of articulation (1.3:1), to the alveolar (2.37:1), and are greatest for the velar (3.67:1). This is what we might expect given the aerodynamic factors at work, and this fact mirrors some previous observations, with much smaller crosslinguistic samples, for phoneme data (See e.g. Ohala, 1983:202, For a fuller treatment of this particular topic, see Everett, in press).

In addition to the potential influence of place of articulation on voicing rates, it has also long been noted that voiceless stops are generally easier to articulate than their homorganic voiced counterparts. This point recurs in the literature. For instance, the work of Ohala (1983) suggests that, due to the physiological effort of maintaining transglottal airflow despite oral closure, voiced stops are relatively difficult to produce. There is extensive crosslinguistic phonemic evidence for this as well. For instance, based on a survey of 706 languages in Ruhlen (1975), Ohala offers data on the occurrence of voiced stops in phoneme inventories. According to those data, 702 of the 706 languages have voiceless stops, while only 540 have voiced stops. Voiced fricatives are also difficult to articulate, when contrasted to voiceless fricatives, since they require sufficient airflow for both vocal cord vibration and frication and the point of constrictions (Ohala, 1983:202). Inspection of the same data in Ohala (1983) reveals that 662 of the 706 languages have voiceless fricatives, while only 398 languages have voiced fricatives. A natural question that arises, and that is testable with the basic consonant data under scrutiny here, is whether these patterns are also evident in the rates at which sounds actually occur in words. The results here suggest that the relative difficulty of producing voiced stops and fricatives, when contrasted to their homorganic voiceless counterparts, is even more starkly evident in the word-list data. The combined rate of occurrence of [p], [t], and [k] in the data, based on the cross-family averages of each sound type, is 0.125. In contrast, the combined rate for [b], [d], and [g] is 0.054, less than half that of the voiceless stops. The usage-based disparity of fricatives is even more apparent: The combined rate of occurrence of [f], [s], and [ʃ] in the data, based on the cross-family averages of each sound type, is 0.039. In contrast, the combined rate for [v], [z], and [ʒ] is 0.0066. In other words, there are about six times more voiceless fricatives in the database, across the same three places

[4] It should be noted, however, that the relative facility of voicing anterior stops, when contrasted with velar stops, is not simply due to the passive differences in supralaryngeal air cavity sizes across places of articulation. As Ohala (1983:200) notes, active increasing of the oral cavity size, at least for bilabial stops, appears to contribute to the disparate ease of voicing across these two places of articulation.

of articulation, than voiced fricatives. These data support previous claims on the comparable difficulty of production of voiced consonants. However, they do so with phonetic transcriptions that represent the bulk of the world's languages and language families. Overall, the data suggest that the disparities between voiceless and voiced obstruents, in terms of phonetic occurrence, is even more pronounced than that evident in phoneme inventories.

Beyond the rate of occurrence of various stops, ease of articulation may help explain other patterns in sound usage. Bilabial, alveolar, and velar consonants are particularly frequent in the data. This is unsurprising given the literature on the commonality of these places of articulation in phoneme inventories (Locke, 1983; Ohala, 1983). It has previously been observed that alveolar consonants are especially common in phoneme inventories, accounting for 15.3% of phonemes while labials account for 14.3% and velars 12.6% (Menard, 2013). Here again, though, the disparities in terms of rates of occurrence are more pronounced: The combined family-controlled rates of occurrence of all consonants made at the alveolar ridge is 0.179. That is, about 18% of all sounds people make are alveolar consonants, judging from these data anyhow. This is much greater than the rates observed for all consonants made at either the velum (9.2%) or the lips (9%). This alveolar dominance may be due partially to articulatory ease as well. In contrast to sounds made at other places of articulation, alveolar sounds require movement of the tongue tip as opposed to the heavier and less maneuverable back and root of the tongue. Some support for this idea comes from work with deaf infants, whose babbling at the age of 6–9 months is dominated by alveolar consonants.[5] (See discussion in Locke, 1983:42–46.) Such dominance in deaf language learners suggests that the crosslinguistic alveolar prevalence of adults' consonant productions is likely due to articulatory rather than perceptual factors, or at least that it is not simply explainable in terms of perceptual factors. The prevalence of some of the consonants in our data, particularly [n], [m], and [b], also seems consistent with evidence from child acquisition on the relative facility with which such sounds are produced. For example, amongst German-speaking and Japanese-speaking children, these consonants are generally produced with greater accuracy early in language acquisition. (See Locke, 1983:74–75.) Still, it would be inaccurate to say that all of the frequent sounds in our data are associated with low error rates during acquisition, or are evident early on in the babbling of deaf infants. Moreover, a variety of sensorimotor and neurophysiological constraints influence the overall ease of articulation of sounds, not just the overall muscle force required of a sound type. In fact, other kinds of sensorimotor effects associated with ease of both perception and articulation may be evident in sounds' rates of occurrence amongst adults, as may other factors like contextual biases favoring the use of some sounds (Martin, 2007).

The aforementioned patterns are also evident in the word lists even without appealing to the average ratios of sounds within and across families. Another way to look at the data is more similar to the binary method typically employed in the relevant literature on phoneme inventories, which considers simply whether or not a given sound is present in a language. In this case we are not considering whether a given sound is meaningfully contrastive in a given language, but whether or not it occurs in a language variety's word list. For example, out of all 6901 lists, [p] is found in 5452 lists–79.0%. In contrast, [k] and [t], the other two prevalent voiceless stops, are found in 96.0% and 96.4% of lists, respectively. Based on these results we can arrive at the simple implication that most languages are likely to utilize [t] *and* [k], and a substantive portion use [p], [t], and [k]. Voiced stops pattern quite differently, in terms of place of articulation: 81.2% of the 6901 lists have [b], 74.8% have [d], and 66.2% have [g]. In terms of this binary approach, we see that [b] > [d] > [g], as voicing becomes less likely with more posterior places of articulation. Note that voiced stops are generally less common in word lists, when contrasted to voiceless stops, but that these binary-based disparities are not as pronounced as those observed above for rates of occurrence. For voiceless fricatives, there is a clear disparity between the alveolar place of articulation and the other two major places represented: [s] occurs in 84.4% of lists, while [f] (34.1%) and [ʃ] (32.5%) each occur in about one-third of the lists. In the case of voiced fricatives, [v] (26.1%) and [z] (25.9%) each occur in about one-fourth of the lists, while [ʒ] occurs in only 8.9%. As with stops, the voiceless varieties of fricatives are more common than their voiced counterparts, even in terms of the number of lists in which they occur. Finally, with respect to nasals, [n] is found in nearly all lists (98.7%), and the same can be said of [m] (97.2%). Unsurprisingly, perhaps, [ŋ] is much less common, but it is still found in 60.9% of word lists.[6]

Such findings point to the same general patterns observed in the rate of occurrence data. Yet several of these sounds are found in a very large portion of lists, limiting the usefulness of the binary approach. Some, for instance [k], [n], [m], and [t], are found at near-ceiling levels, with each being represented in over 96% of the lists. As was observed above, however, there are clear disparities between the rates of occurrence of such sounds, even after controlling for relatedness. So while the binary approach is useful, it seems to obfuscate some interesting patterns in the data.

In short, these data offer potentially stronger support for previous claims on the relative ease of articulation of some sounds. This is not to suggest that all these patterns can be understood in terms of ease of articulation. There are also perceptual factors involved in the primacy of some sound types. For example, the distinct spectral characteristics of [s] (Stevens, 2000) may help to explain its commonality. Yet, whatever approach one takes to such global phonetic data, they seem to underscore previous claims on the role of ease of articulation in helping to shape speech. (See also Everett, in press.)

[5] Interestingly, prior to this age there is a velar dominance. However, this dominance may result, at least in part, from physiological necessity early on in life. In Locke's words, perhaps "the palate is so low that lingual contact is difficult to avoid" (Locke, 1983:46).

[6] It is interesting to note how these data are similar to those mentioned above, based on Gordon (2016), on the rates at which sounds occur in phoneme inventories. Here, again, are the rates at which the most common consonants appear in phoneme inventories, with the percentages of ASJP word lists in which they occur offered parenthetically: $n = 99\%$ (99%), $k = 89\%$ (96%), $m = 93\%$ (97%), $t = 97\%$ (96%), and $s = 84\%$(84%).

5. Conclusion

While it has long been recognized that some consonants are common in phoneme inventories, these results demonstrate that this phonemic commonality potentially undersells the prevalence of some consonants in speech. Certain consonants occur at high rates across the bulk of word lists that represent the majority of the world's languages. One reasonable account of the particularly high rate of occurrence of some consonant types, and of the particularly low rates of occurrence of others, relates to ease of articulation. While many linguists recognize that ease of articulation plays some role in shaping speech, the extent of its influence is debated and some linguists remain skeptical that it has any major effect on speech. According to this latter position, ease of articulation is an unlikely candidate as a significant shaper of language because, were it in fact operative, it would yield very similar sounds across the world's languages. This putative similarity is said to be contradicted by the diversity of phoneme inventories. The present study suggests that there is evidence, at a global scale, that languages do exhibit convergence in the rates at which they employ their consonants, despite the well known diversity in the spectrum of sounds they use. This result requires further exploration with more robust and detailed intralinguistic phonetic corpora. It should be stressed that these results are not meant to devalue work on phoneme inventories, but to serve as a useful complement to them. Phoneme inventory data offer advantages over the sorts of frequency data discussed here. For one, they are richer in terms of detail and allow us to see, for example, how adults come to rely on features that abstract away from some of the detail in the acoustic signal, while maximizing certain combinations of features.(See discussion in Clements, 2003.) They therefore allow for a more refined exploration of some of the perceptual and cognitive factors shaping phonologies. Yet the phonetic data discussed here, despite their limitations, also offer some clear benefits. Unlike inventory data, they provide some indication of just how much speakers of given languages rely on particular consonants and vocal tract configurations.

The world's sound systems are certainly quite diverse. Yet, while languages differ markedly with respect to phoneme inventories, syllable structures, prosody, and numerous other phonological and phonetic phenomena not even touched upon in this study, there is also a less visible but powerful undercurrent evident in the rates of occurrence of consonants in speech. Judging from the data tested in this study, anyhow, this undercurrent yields a clear and quantifiable similarity with respect to the manner in which the world's languages rely on their consonants.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.langsci.2018.07.003.

References

Atkinson, Q., 2011. Phonemic diversity supports a serial founder effect model of language expansion from Africa. Science 332, 346–349.
Blasi, D., Wichmann, S., Hammarström, H., Stadler, P., Christiansen, M., 2016. Sound-meaning association biases evidenced across thousands of languages. Proc. Natl. Acad. Sci. U.S.A. 113, 10818–10823.
Brown, C., Holman, E., Wichmann, S., 2013. Sound correspondences in the world's languages. Language 89, 4–29.
Calude, A.S., Pagel, M., 2011. How do we use language? Shared patterns in distribution of the frequency of word use across 17 world languages. Philos. Trans. R. Soc. B 366, 1101–1107. https://doi.org/10.1098/rstb.2010.0315.
Clements, N., 2003. Feature economy in sound systems. Phonology 20, 287–333.
Dryer, Matthew S., Haspelmath, Martin (Eds.), 2013. The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig.
Evans, N., Levinson, S., 2009. The myth of language universals: language diversity and its importance for cognitive science. Behav. Brain Sci. 32, 429–448.
Everett, Caleb, 2018. The global dispreference for posterior voiced obstruents: a quantitative assessment of word list data. Language in press.
Gordon, M., 2016. Phonological Typology. Oxford University Press.
Gusein-Zade, S.M., 1987. On the frequency of meeting of key words and on other ranked series. Sci. Tech. Inf. Ser. 2: Inf. Process. Syst. 1, 28–32.
Hay, J., Bauer, L., 2007. Phoneme inventory size and population size. Language 83, 388–400.
Jensen, J., 1993. English Phonology. John Benjamins, Amsterdam.
Jessen, Michael, 2002. An acoustic study of contrasting plosives and click accompaniments in Xhosa. Phonetica 59, 150–179.
Kaye, J., 1989. Phonology: a Cognitive View. Lawrence Erlbaum Associates, New Jersey.
Keating, P., 1984. Physiological effects on stop consonant voicing. UCLA Work. Pap. Phonetics 59, 29–34.
Lindblom, B., Maddieson, I., 1988. Phonetic universals in consonant systems. In: Li, C., Hyman, L.M. (Eds.), Language, Speech and Mind. Routledge, London, pp. 62–78.
Locke, J.L., 1983. Phonological Acquisition and Change. Academic Press, New York.
Maddieson, I., 2016. Consonant inventories. In: Dryer, M., Hasplemath, M. (Eds.), The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig.
Maddieson, I., 1984. Patterns of Sounds. Cambridge University Press.
Maddieson, I., Bhattacharya, T., Eric Smith, D., Croft, W., 2011. Geographical and phonological complexity. Linguist. Typol. 15, 267–279.
Maddieson, I., Precoda, K., 1990. Updating UPSID. UCLA Work. Pap. Phonetics 74, 104–111.
Martin, A., 2007. The Evolving Lexicon. PhD dissertation. UCLA.
Menard, L., 2013. Sensorimotor constraints and the organization of sound patterns. In: Lefebvre, C., Comrie, B., Cohen, H. (Eds.), New Perspectives on the Origins of Language. John Benjamins, Amsterdam, pp. 257–278.
Moran, S., McCloy, D., Wright, R., 2017. PHOIBLE Online. Max Planck Institute for Evolutionary Anthropology, Leipzig.
Moran, S., McCloy, D., Wright, R., 2012. Revisiting population size vs phoneme inventory size. Language 88, 877–893.
Napoli, D.J., 2014. On the linguistic effects of articulatory ease, with a focus on sign languages. Language 90, 424–456.
Ohala, J.J., 1983. The origin of sound patterns in vocal tract constraints. In: MacNeilage, P.F. (Ed.), The Production of Speech. Springer-Verlag, New York.
Ohala, J.J., 1980. Moderator's Summary of Symposium on 'Phonetic Universals in Phonological Systems and Their explanation.' Proceedings of the Ninth International Congress of Phonetic Sciences. Institute of Phonetics, Copenhagen, pp. 181–194.

Ohala, J.J., Riordan, C.J., 1979. Passive vocal tract enlargement during voiced stops. In: Wolf, J.J., Klatt, D.H. (Eds.), Speech Communication Papers. Acoustical Society of America, New York.
Piantadosi, S., 2015. Zipf's word frequency law in natural language: a critical review and future directions. Psychonomic Bull. Rev. 21, 1112–1130.
Ploch, S., 2004. What is wrong with linguistics. SOAS Work. Pap. Linguist. Phon. 13, 363–382.
Roberts, A.H., 1965. A Statistical Linguistic Analysis of American English. Mouton, The Hague.
Ruhlen, M., 1975. A Guide to the Languages of the World. Stanford University, Palo Alto.
Swadesh, M., 2006. The origin and diversification of language. Routledge Press, Abingdon.
Stevens, K., 2000. Acoustic Phonetics. MIT Press, Cambridge.
Tambovtsev, Y., Martindale, C., 2007. Phoneme frequencies follow a Yule distribution: the form of the phonemic distribution in world languages. SKASE J. Theor. Ling. 4.2.
Vaux, B., 2013. The role of features in a symbolic theory of phonology. In: Raimy, E., Cairns, C. (Eds.), Contemporary Views on Architecture and Representations in Phonology. MIT Press, Cambridge, MA, pp. 75–97.
Wichmann, S., Holman, E., Brown, C., 2016. The ASJP Database (Version 17).
Yegerlehner, J., Voegelin, F., 1957. Frequencies and inventories of phonemes from nine languages. Int. J. Am. Ling. 23, 85–93.
Zipf, G.K., 1949. Human Behavior and the Principle of Least Effort. Addison-Wesley, Reading, MA.