RESEARCH REPORT

The global dispreference for posterior voiced obstruents: A quantitative assessment of word-list data

Caleb Everett

University of Miami

This work explores the effect of ease of articulation on speech by examining the rates at which various consonants occur in word lists representing thousands of languages. The data reveal that obstruents produced with oral obstruction closer to the glottis are less likely to be voiced when contrasted with their counterparts produced in the anterior region of the vocal tract. While this finding is explainable via previously documented aerodynamic factors, these new data suggest that such factors may have a more powerful influence on speech than typically assumed. The pattern in question is evident even after controlling for the relatedness and areal proximity of language varieties. This study isolates and quantifies the decrease in consonant voicing associated with the reduction in size of the supralaryngeal cavity.*

Keywords: phonetics, phonology, ease of articulation, voicing, stops, fricatives, typology

1. Introduction. One way in which ease of articulation may impact the development of phonologies is through the differential preference for voiced stops observed at distinct places of articulation. It has previously been noted that phoneme inventories are more likely to exclude voiced velar stops than voiced bilabial or voiced alveolar stops (Maddieson 1984, 2013a, Ménard 2013, Napoli et al. 2014, inter alia). Here I offer evidence that, in addition to being less common in phoneme inventories, voiced velar stops are relatively uncommon in phonetically transcribed word lists. More broadly and interestingly, perhaps, the evidence suggests that the likelihood of stops and fricatives being voiced decreases in accordance with their place of articulation. Posterior voiced obstruents are relatively infrequent across thousands of word lists, even after controlling for the relatedness and geographic proximity of languages. Furthermore, the data suggest that the voiced variety of a posterior obstruent is comparatively uncommon even in languages in which both voiceless and voiced varieties of the obstruent are used. For instance, if a language employs both a voiceless and voiced velar stop, the latter is typically less frequent in a word list representing the language.

The ease-based explanation of such distributional tendencies relates to aerodynamic factors: voicing is theoretically more effortful to maintain as oral occlusion moves closer to the glottis. This increased effort owes itself, essentially, to air-pressure differentials. Vocal-fold vibration requires transglottal airflow and, therefore, a lower air pressure above the glottis than below. Supraglottal air pressure rapidly increases during the production of velar stops because they entail a relatively small supralaryngeal cavity. This rapid increase works against the airflow requisite for vocal-fold vibration (Ohala 1983, 1997, Keating 1984). Furthermore, enlargement of the oral cavity during the production of voiced bilabial stops may further facilitate their articulation (Ohala 1983). These aerodynamic factors, it has been argued, explain the relevant tendency previously described for phoneme inventories, namely the relative scarcity of voiced velar stops. It also potentially explains other patterns. For instance, the greater likelihood of devoicing high vowels, when contrasted with low vowels that have less oral constriction, may owe itself to the same phenomenon (Ohala 1983, Tsujimura 2007).

* I am grateful to the editor of *Language* and to two anonymous referees for very insightful comments.

Interrelatedly, the comparable rarity of voiced fricatives, when contrasted with voiceless ones, may be due to the fact that voiced fricatives require sufficient transglottal airflow to maintain both voicing and frication at an oral point of constriction (Ménard 2013). Judging from the results offered below, the bias against voiced fricatives is particularly strong for posterior places of articulation. There is some previously established intralinguistic support for such effects. For instance, in some languages voiced velar stops seem to be voiced for a shorter duration than their bilabial and alveolar counterparts (Klatt 1975, Smith & Westbury 1975). Such intralinguistic evidence lends support to the ease-based account of the relevant typological patterns in phoneme inventories.

Yet the question of how much these aerodynamic factors impact languages is an open one, judging from the literature on this topic. While presumably the aerodynamic effects are operative in the vocal tracts of all populations, the extent of their influence in phonological and phonetic patterns is unclear, though careful attention has been paid to the aforementioned relative scarcity of voiced velar stops across the world's phoneme inventories. For example, in one study it was found that 175 of the surveyed phoneme inventories had /g/, while 283 had /k/ (Shariatmadari 2006). However, it is unclear the extent to which such a disparity is due to the particular ease-based factor in question, since voiced stops are generally less frequent than voiceless stops. Judging from the phoneme-inventory data alone, the influence of the air-pressure-based physiological effect may be minor. For example, inspection of the data in Gordon 2016 (table 3.1), based on Maddieson 1984, suggests that voiced velar stops exist in 55% of the phoneme inventories surveyed, while voiced alveolar stops are found in 62%. In the same study it is observed that about 97% of phoneme inventories surveyed have /t/, while about 89% have /k/. Voiceless alveolar stops are 1.56 times (97% vs. 62%) more common than their voiced counterparts according to such data, whereas voiceless velar stops are 1.62 times (89% vs. 55%) more common than their voiced counterparts. Consistent with the ease-based effect on voicing, the disparity between voiceless and voiced bilabial stops is relatively low, with /p/ being 1.32 times (82% vs. 62%) as common as /b/. This relatively minor disparity, when contrasted with stops closer to the glottis, does suggest that the aerodynamic effect in question helps shape phonologies. Still, the differences between the rates of occurrence in phoneme inventories of voiceless and voiced stops, across places of articulation, are apparently modest, and some of these differences could be due in part to genealogical or areal confounds.

Further evidence in favor of this air-pressure-related effect is the following: some phoneme inventories have /p/, /b/, /t/, /d/, and /k/ but lack /g/. This sort of stop inventory has been found in forty-eight languages in the UCLA Phonological Segment Inventory Database (UPSID), while inventories that lack /b/ or /d/ (of the six stops in question) are represented by only eight languages each in the same database (Maddieson & Precoda 1990, Napoli et al. 2014). One recent study concluded that the low incidence of /g/ in phoneme inventories is not explainable via areal effects (Maddieson 2013a). For that study, 567 phoneme inventories were surveyed, and it was observed that /g/ is 'gapped', or missing, in thirty-four of these. In such cases a language has /p/, /t/, /k/, /b/, and /d/ phonemes, but not /g/. Maddieson (2013a) also makes the interesting observation that, while thirty-three of the languages in his sample have a 'gapped' /p/, these languages exhibit a very clustered geographic distribution, primarily in Africa. In contrast, languages that are missing the voiced velar stop exhibit a widespread global distribution, suggesting that the relative scarcity of this sound is not due simply to language contact or phylogeny. Maddieson's interpretation is further supported by the data offered below. However, the data presented below allow for greater confidence in the phenomenon, given that they are based on rates of occurrence that allow for intralinguistic tests, as well as crosslinguistic tests over a larger sample than those evident in previous work on phoneme inventories. Other phoneme-based evidence that has been offered on this topic also points in the direction of the proposed place-of-articulation effect, but with even smaller sample sizes. For example, in the UPSID database there are seven instances of /tː/, but only three of /dː/. Similarly, there are nine instances of /kː/, but only four of /gː/, perhaps since length is particularly difficult to achieve for these voiced stops (Shariatmadari 2006).

Some phonologists have pointed out that, if a universal ease-based influence on speech exists, it is perhaps surprising that languages do not converge more regularly on similar phoneme inventories (Kaye 1989, Ploch 2004). Others have suggested that this latter perspective caricaturizes the view that ease of articulation influences phonology (Shariatmadari 2006). According to the latter position, ease of articulation has a systematic but nondeterministic effect on speech as it competes with myriad other factors (see e.g. Lindblom 1983). While recent work points unambiguously to other ease-based effects on, for instance, signed language (Napoli et al. 2014), there is apparently still room for debate as to the extent of the influence that ease of articulation has in shaping speech. The results presented below suggest that differential ease of consonant voicing does lead to a kind of universal convergence in speech, since it is evident in robust patterns in word production across all languages and regions. These patterns in word production are consistent with the evidence previously gathered from phoneme inventories. This consistency is suggestive of a diachronic link between the high proportions of certain consonants, in phonetically transcribed lists of basic words, and the occurrence of those consonants as phonemes. The results in this work could serve as useful baselines for further explorations of the relationship between a sound's proportion in phonetic transcriptions and its prevalence in phoneme-inventory databases (e.g. Maddieson 2013b, Moran et al. 2017). The approach outlined below allows us to carefully control for the effects of phylogeny and areal proximity in order to isolate quantitatively the extent of the relevant articulatory bias on speech. Even if one does not doubt the plausibility of the pertinent aerodynamically based influence on articulation, this new method helps to elucidate the extent of that influence.

2. Employing proportion data to explore ease of articulation. One way to examine this issue is to focus on the rates at which sound types occur in transcriptions. The present study relies on a database of word lists for thousands of language varieties, the Automated Similarity Judgment Program (ASJP) (Wichmann et al. 2016). This database contains IPA-based phonetic transcriptions, transformed into basic typescript, of forty words (typically) in each of over 7,000 language varieties. (Some lists have more words, while others have slightly fewer than forty.) The ASJP database has been used for a variety of functions, including the establishment of common sound correspondences between languages and as a means for the computational classification of languages (Brown et al. 2008, Brown et al. 2013). More recently it has been used for distinct purposes like the exploration of statistical signals of iconicity in speech (Blasi et al. 2016), as well as the consideration of the potential effects of ambient aridity on phonation (Everett 2017).

The words in the ASJP database, which overlap substantially with the Swadesh list, are resistant to borrowing (Swadesh 1952, Haspelmath & Tadmor 2009). They are also typically frequent in speech, so they are decent, though of course not perfect, indicators of sound patterns in a given language (Calude & Pagel 2011). (Bearing in mind that larger digitized phonetic corpora are unavailable for most languages.) The ASJP transcriptions vary in terms of how narrow they are. Still, the transcription conventions used in the database denote the basic voicing distinction that is essential for the present purposes. Of course, voice onset time can vary in fine-grained ways, both intralinguistically and crosslinguistically, and such granular variation is not captured in transcriptions or in phoneme inventories. Still, the basic voiceless/voiced distinction is critical to speech, and it is one to which linguists and other transcribers are generally sensitive.

To get a sense of how aerodynamic factors might inordinately influence some voiced obstruents, the frequency of each consonant's occurrence was obtained for each word list. There are a total of thirty-four basic consonant types in the ASJP database. Five of these consonants actually subsume voiced and voiceless varieties of a given sound, due to the transcription conventions used in the database. This precludes their analysis here, but they fortunately are not critical to the analysis. The proportion of consonants, for each consonant, was calculated. 'Proportion of consonants' refers to the ratio of all consonant tokens, in a given list, represented by a particular consonant type. Proportions for each consonant type were calculated for each of 6,830 word lists. (Word lists for some languages, mainly constructed languages, were excluded.) For the principal analysis a carefully categorized subset of these word lists was utilized. The word lists represent over 4,500 separate ISO codes, with some languages represented by multiple lists for differing dialects. Proportions of consonants were calculated by summing all instances of a given sound type in a word list, and dividing that sum by the total number of consonants in the list. Secondary diacritics were ignored in this calculation, with the exception of the diacritic for glottalization. While glottalized stops are not common in the word lists, they were counted separately from voiceless and voiced stops to more precisely test the hypothesis. Given their scarcity, this choice had little impact on the results. This kind of study inevitably overlooks certain phonetic details for the sake of exploring global patterns; some phonetic minutiae are not encoded in the database, while some that are encoded are ignored in this analysis. While these word lists do not constitute large corpora, each typically includes hundreds of transcribed sounds. Combined with the large number of languages they represent and the associated opportunities for phylogenetic and areal controls, they offer a statistically robust means of investigating the issue at hand. Also, while these lists represent only relatively small corpora, there may be diminishing rates of return for larger corpora, at least in the case of some languages. For instance, in the case of English, the obstruents' proportions of consonants in the ASJP list are very similar to those obtained for a larger previously analyzed corpus (Martin 2007, analyzed in Gordon 2016).The correlation between the English ASJP proportions and the English proportions in that larger corpus is strong for obstruents (Spearman's rank correlation = 0.91, $R^2 = 0.82$).

The consonants with the highest proportions are [n], [k], [s], and [t], in that order. Previous work has pointed to the typically high intralinguistic frequency of these same four consonants (Gordon 2016). The purpose of the present study is not, however, to explore the relative frequencies of sounds in this general sense. Here I focus on the disparity between voiceless and voiced obstruents, in accordance with place of articulation. (Raw data and code are available in the supplemental material.[1])

I focus on the following stops and fricatives, for which the aerodynamics of voicing are potentially relevant: [p], [b], [t], [d], [k], [g], [q], [ɢ], [f], [v], [s], [z], [ʃ], and [ʒ]. The mean proportions of consonants for each of these sounds, across all word lists and

[1] All supplemental materials can be accessed at http://muse.jhu.edu/resolve/XX.

	PROPORTIONS			
	VOICELESS	VOICED	VOICELESS : VOICED RATIO	ORDER
STOPS				
Bilabial	0.038	0.049	0.768	1
Alveolar	0.081	0.039	2.073	2
Velar	0.086	0.033	2.629	3
Uvular	0.006	0.0003	24.43	4
FRICATIVES				
Labial	0.011	0.009	1.211	1
Alveolar	0.046	0.008	5.573	2
Postalveolar	0.012	0.002	5.778	3

TABLE 1. Ranking of ratios of voiceless : voiced proportions of obstruents, from lowest to highest, by place of articulation. These ratios are based on the mean proportions across all lists in the data set, without any controls.

prior to applying any controls, are provided in Table 1.[2] Inspection of the table reveals the pattern we might expect if ease of voicing impacts the usage of sounds. (For a graphical depiction of this pattern, see the heatmap visualization in supplemental Figure A.) Voiceless consonants are generally more common than their homorganic voiced counterparts, and this voiceless-voiced disparity is much more apparent for stops and fricatives as one moves closer to the glottis. In fact, voiced bilabial stops are actually slightly more common than their voiceless counterparts in these basic frequency data. The data in Table 1 offer preliminary findings that are consistent with the suggested heightened difficulty of voicing for sounds with constrictions closer to the glottis. Consider that, according to these data, voiced velar stops are just over one third as common as their voiceless counterparts, while voiced alveolar stops are just under half as common as their voiceless counterparts. These patterns make sense if consonant production is significantly affected by the aerodynamic bias in question. Still, we must apply more refined analyses to isolate this potential bias. To do so, this study relies on the AUTOTYP database, which categorizes languages into 312 distinct linguistic stocks and, critically, twenty-four well-motivated geographic regions (Bickel & Nichols 2017). The languages in the ASJP database were cross-referenced with those of the AUTOTYP database. This yielded 3,341 language varieties, representing 2,132 distinct languages as judged by ISO codes, whose obstruents' proportions were then analyzed. (Supplemental Figure B presents the locations of the 3,341 language varieties, as well as the twenty-four AUTOTYP regions.) As we will see below, this sample is more than sufficient for high degrees of confidence in the results obtained.

If we examine the presence of the relevant sounds in each stock, we can consider the data in a way that is analogous to traditional binary phonemic status. Each stock was tested to see whether a given obstruent is represented in at least one of its word lists. These tests suggested patterns similar to those previously described for phoneme inventories. For instance, the bilabial, alveolar, and velar stops are found in most families. Consider the percentages of the language stocks in which each of the relevant stops is found to occur: [p]: 86.9%, [b]: 83.1%, [t]: 97.4%, [d]: 81.5%, [k]: 98.1%, and [g]: 73.8%. The disparity in the cross-family presence of voiceless and voiced stops increases from the bilabial to alveolar place of articulation, and from the alveolar to velar. Yet all of the stops are quite common, as is evident in previous work on phoneme in-

[2] Note that the ratios in Tables 1 and 2 were calculated from the unrounded raw data rather than from the mean proportions reported there.

ventories (Gordon 2016, Maddieson 1984). The disparities in the actual proportions of the sounds are more pronounced than the disparities evident in such 'stock representation' findings. For that reason I focus next on the proportions within individual word lists, within stocks and across stocks.

Table 2 presents the mean proportions of the pertinent consonants, across AUTOTYP stocks, in the narrower sample of 3,341 languages. To arrive at these figures, the proportions of consonants for each sound were averaged within stocks, and then those means were averaged. These figures are generally similar to the uncontrolled proportions in Table 1, and the main predicted findings following from an ease-of-articulation account are also evident in Table 2. Voiced stops and fricatives are less common, at each place of articulation, when contrasted with their voiceless homorganic counterparts. However, the disparity in usage between voiceless and voiced sounds increases substantially as place of articulation moves further back in the vocal tract.

	PROPORTIONS			
	VOICELESS	VOICED	VOICELESS : VOICED RATIO	ORDER
STOPS				
Bilabial	0.043	0.043	0.999	1
Alveolar	0.076	0.040	1.90	2
Velar	0.095	0.035	2.75	3
Uvular	0.008	0.0003	25.4	4
FRICATIVES				
Labial	0.009	0.006	1.36	1
Postalveolar	0.015	0.002	8.95	2
Alveolar	0.044	0.004	10.6	3

TABLE 2. Ranking of phylogenetically controlled ratios of voiceless : voiced proportions of obstruents, from lowest to highest, by place of articulation. Proportions are the means of the average proportions observed within each of the stocks.

As evident in Table 2, the ratios between the proportions of voiceless and voiced stops are quite consistent with the account. For bilabial stops that ratio is about 1 : 1, for alveolar stops it is 1.9 : 1, and for velar stops it is 2.75 : 1. Table 2 also includes data for uvular stops and postalveolar fricatives, which are not focused on in the two principal analyses below. Some comments on these sounds are in order. First, for postalveolar fricatives: the lower-frequency peaks in the spectra of these fricatives are indicative, like palatography data, of less narrow constriction when contrasted with alveolar fricatives (Stevens 2000). So, while postalveolar fricatives are produced slightly closer to the glottis than alveolar fricatives are, their less narrow constriction means that there is no clear prediction for their voiceless-voiced disparity, at least when compared with that of alveolar fricatives. There is, however, a prediction that postalveolar fricatives should have a larger voicing disparity than that evident for labiodental fricatives. As seen in Tables 1 and 2, this prediction is met. Nevertheless, I set aside postalveolar fricatives for the main portion of the analyses below, given the turbid predictions for them. Similarly, I do not consider uvular stops in the main analysis because of their very low rates of occurrence when contrasted with other stops. However, it is worth observing that, as evident in Table 2, voiceless uvular stops are about twenty-five times more common than voiced uvular stops. This is actually the most pronounced voiceless-voiced disparity evident for stops at one place of articulation, as we might predict since uvular stops are located closer to the glottis than any of the other consonants tested.

In some cases languages lack voiced velar stops entirely, and this fact could contribute to the disparity between the phylogenetically controlled proportions of voiceless and voiced velar stops. Voiced alveolar stops could also be missing in some word lists. Interestingly, though, even if we focus only on those language varieties that have all six of the relevant stops represented somewhere in their word list, the pattern is still evident. In other words, it is not simply that posterior voiced stops are less likely to exist in a given language. The pattern in question is also due to the fact that, in languages with homorganic voiced and voiceless posterior stops, the voiceless varieties are generally more common in the word lists. A total of 2,686 word lists have at least one token of each of [p], [b], [t], [d], [k], and [g]. The mean proportions, across these 2,686 word lists, are as follows: [p] = 0.041, [b] = 0.059, [t] = 0.079, [d] = 0.051, [k] = 0.084, and [g] = 0.046. The voicing disparity is once again greatest for velar stops (ratio = 1.83) and less pronounced for alveolar stops (ratio = 1.55). In the case of bilabial stops, voiced variants are actually more common than their voiceless counterparts (ratio = 0.69). The same findings hold when stock affiliation is factored in. Cross-referencing the 2,686 word lists with AUTOTYP stocks yields 1,294 word lists (in 167 stocks) that have all six of the relevant stops. For this even narrower selection of lists, that pattern is still observed. The mean proportion of each stop was found for each stock, and then these means were averaged. The mean proportions obtained via this method were as follows: [p] = 0.047, [b] = 0.057, [t] = 0.070, [d] = 0.055, [k] = 0.091, and [g] = 0.051 ([p] : [b] ratio = 0.82, [t] : [d] ratio = 1.27, [k] : [g] ratio = 1.78). In short, even if we restrict our attention to languages that have all of the relevant stops represented in transcriptions, the pattern in question still surfaces. This smaller set of languages is returned to below.

There are only 421 word lists in which all four of the most relevant fricatives, [f], [v], [s], and [z], occur. (And only 184 of these also have AUTOTYP classifications, precluding more detailed analysis for the others.) For this small subset of lists, the pattern is still observed. The mean proportions of the consonants are as follows: [f] = 0.030, [v] = 0.033, [s] = 0.051, and [z] = 0.036. The voiceless : voiced ratio for the labiodental fricatives is 0.91, while that for the alveolar fricatives is 1.41. On average, then, voicing disparity is more pronounced for the more posterior fricatives, even in languages that have at least one instance of all four of the pertinent fricatives.

To better isolate the extent to which place of articulation impacts the rate at which languages utilize voiced and voiceless sounds, two approaches were adopted to control simultaneously for the confounds of phylogenetic and areal relatedness. There are various approaches to controlling for such factors (see e.g. Bickel 2008). Two of these are based on random sampling and mixed modeling, and both are adopted here since they offer complementary ways to isolate the size of the effect in question while controlling for region and relatedness with a set of continuous data. (In contrast, some other approaches require that languages be categorized into discrete groups, for example, by whether they have a voiced velar stop phoneme.) Both the random-sampling portion of the analysis and the mixed-modeling portion focus on the differing rates of occurrence for each member within the following sound pairs: (i) [k] and [g], (ii) [t] and [d], (iii) [p] and [b], (iv) [s] and [z], and (v) [f] and [v]. The predictions of the account are clearest for the typologically common labial, alveolar, and velar consonants. Since the voicing distinction is collapsed for velar fricatives in the ASJP database, they are not considered, but labial and alveolar fricatives are examined. The results below suggest that alveolar fricatives are less likely to be voiced, even after controlling for language genealogy and contact, than labiodental fricatives. This fact, like the aforementioned greater likelihood of high-vowel devoicing, is potentially explicable via the same aerodynamic factors that appear to explain the discrepancies in the proportions of stops. After all, even the alveolar obstruction associated with fricatives should theoretically increase supralaryngeal air pressure. While this increase to supralaryngeal air pressure may not be as pronounced as that associated with complete oral occlusion, what is most relevant is whether voiced alveolar fricatives require greater articulatory effort when contrasted with voiceless alveolar fricatives, not when contrasted with stops. The fricative data considered below suggest that they do in fact require greater articulatory effort. (Though perceptual factors may also motivate the higher proportions of some fricative types.)

Here is how the random-sampling method proceeded: for the subset of 3,341 language varieties represented in both the ASJP and the AUTOTYP database, one variety was randomly selected from each of the 312 AUTOTYP stocks. From this random sample of 312 word lists, one language was then randomly selected from each of the twenty-four AUTOTYP regions. Each iteration of this sampling process therefore generated twenty-four word lists, each of which represented a unique language family and a unique region. Next, for this controlled sample of twenty-four lists the proportions of consonants for each of two stops or fricatives at a specific place of articulation were tabulated. For each list, the proportion of the voiced consonant was then subtracted from the proportion of its homorganic voiceless counterpart, and this 'voicing disparity' was then averaged across the twenty-four lists for the controlled sample. Each iteration of the sampling technique therefore yielded a mean controlled disparity in proportion, for twenty-four word lists, between the pertinent voiceless and voiced consonants. One thousand iterations of the procedure were run for each of the five sound pairs. The mean disparities for each of these 5,000 iterations are visualized in Figure 1. As can be seen in the figure, the mean disparities for the controlled samples follow the predicted trajectory: as consonant constriction moves closer to the glottis, the disparity between voiceless and voiced stops, and between voiceless and voiced fricatives, increases. This is not altogether surprising given the results in Table 2, but the results in Fig. 1 allow for more confidence regarding the strength of the association.

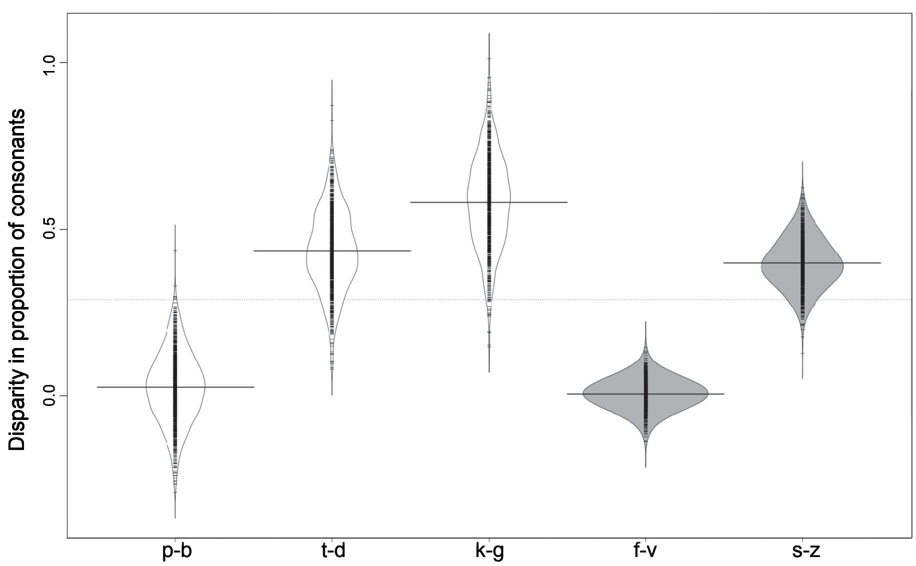

FIGURE 1. Beanplots of the values obtained via the 5,000 iterations of the sampling technique. Bean width corresponds to density of the sampled values for the proportion disparities. The value of each sample is represented with a very short solid line in the middle of a bean. Long horizontal solid lines represent the means for each place of articulation. Disparities for stops are represented with white beans, disparities for fricatives with gray beans. The thin horizontal line across the entire plot represents the mean disparity for all places of articulation.

As is apparent in Fig. 1, there is little difference in the controlled rates at which voiceless and voiced labial stops are used. The mean disparity between the sounds, using this approach, is only 0.005. In contrast, the mean disparity between the proportions of voiceless and voiced alveolar stops is about nine times greater, or 0.044, and the mean disparity between voiceless and voiced velar stops is about twelve times greater, or 0.059. In the case of fricatives, voiceless labiodental varieties are about as common as their voiced counterparts, with the mean disparity in proportions hovering around 0 (0.0009). However, the mean disparity between voiceless and voiced alveolar fricatives is comparatively pronounced, at 0.041. In short, the sampling-based approach reveals a clear association between likelihood of consonant voicing and place of articulation.

While subtracting proportions is useful, examining the ratios of proportions is also elucidative. To do this, however, we must restrict our attention to language varieties in which the relevant sounds have a proportion greater than zero. As noted above, 2,686 word lists have proportions greater than zero for each of the six crucial stops. However, only 1,294 of these language varieties are also represented in the AUTOTYP database. The sampling approach was used to test the ratios between the proportions of homorganic stops for these word lists. The same sampling procedure from above was followed but, instead of subtracting a voiced-stop proportion from the homorganic voiceless-stop proportion, the latter proportion was divided by the former. The resultant ratio was found for the twenty-four lists in each sample, and these twenty-four ratios were averaged. Again, 1,000 iterations of the sampling technique were applied to each place of articulation for the stops. The mean ratios between voiceless and voiced stops, across all random samples, were as follows: [p] : [b] = 1.91, [t] : [d] = 2.60, and [k] : [g] = 3.53. These phylogenetically and areally controlled figures demonstrate that, even in languages that have a 'complete' set of voiceless and voiced stops, posterior voiced stops are relatively dispreferred.

Random sampling is not the only potential approach to such data. It offers the advantage of giving us a fairly clear quantitative assessment of the relative disparity between rates of voicing at each place of articulation, even after controlling for region and phylogeny. Another useful and complementary approach is a linear mixed-effects model (Bates et al. 2015). The advantage of the latter approach is that it allows us to simply test the significance of the effect of place of articulation on the likelihood of voicing. Also, utilizing two distinct approaches allows for greater confidence that the results are not unduly biased by the underlying assumptions of any one method. (For further discussion of the advantages of mixed-effect models for such studies, see Jaeger et al. 2011.)

First, the mixed-model approach was used for stops. The disparity between voiceless and voiced bilabial, alveolar, and velar plosives was calculated across all 3,341 word lists, for a total of 10,023 calculated disparities (3,341 × 3). This 'voicing disparity' was treated as the dependent variable. A mixed-intercepts model was created in which AUTOTYP stock and AUTOTYP region were treated as random variables, while place of articulation was treated as an ordered fixed categorical variable:

$$\text{VoicingDisparity}_i = \beta_{POA} + \alpha_{STOCK} + \alpha_{REGION} + \varepsilon.$$

This model suggested that place of articulation was a clear contributing factor to voicing disparity, with velar stops being associated with an increase in voicing disparity when compared to alveolars. The increase in disparity was 0.015 ± 0.001 (standard error). Alveolars were associated with an increase in voicing disparity when contrasted with bilabials. The difference was 0.045 ± 0.001 (standard error). For a null model, AUTOTYP stock and region were treated as random variables without the inclusion of a fixed variable. A likelihood ratio test revealed a highly significant disparity between the fixed and null models ($\chi^2(2) = 1927$, $p < 0.00001$).

The same approach was applied to the 1,294 languages (167 stocks) that are classifiable via AUTOTYP and that also have a 'complete' set of stops. This yielded 3,882 calculated disparities (1,294 × 3). Once again, place of articulation was found to be a clear contributing factor to voicing disparity. The increase in disparity was 0.012 ± 0.001 (standard error) for velars when contrasted to alveolars. Alveolars were again associated with an increase in voicing disparity when contrasted with bilabials, with a difference of 0.040 ± 0.001 (standard error). For a null model, AUTOTYP stock and region were treated as random variables without the inclusion of a fixed variable. A likelihood ratio test revealed a highly significant disparity between the fixed and null models ($\chi^2(2) = 804$, $p < 0.00001$).

The mixed-models approach was also used to analyze the two sets of fricatives for which there are the clearest predictions: labiodental and alveolar fricatives. The voicing disparities at these two places of articulation were calculated across all lists with stock information, for a total of 6,682 disparities (3,341 × 2). Voicing disparity was once again treated as the dependent variable, stock and region were treated as random intercept variables, and fricative place of articulation was considered a fixed categorical variable. This model also revealed that place of articulation was a clear contributor to voicing disparity, with voicing disparity increasing for alveolar fricatives when contrasted with labiodentals. The increase was 0.037 ± 0.001 (standard error). For the null model, stock and region were again treated as random variables, but place of articulation was not included in the model. A likelihood ratio test again revealed a clear disparity between the fixed and null models ($\chi^2(1) = 1771$, $p < 0.00001$).

In short, both the random-sampling and mixed-model approaches point to a clear influence of place of articulation on obstruent voicing rates, even after controlling for relatedness and region. This influence surfaces across a global sample of languages. It surfaces even when we restrict our attention to those languages that have at least one instance of each of the six most relevant stops, or at least one token of each of the four most relevant fricatives.

3. Discussion and conclusion.

There are various factors that help motivate the prevalence of some sounds in speech at the expense of others. For instance, perceptually robust sounds are more prevalent both across the phoneme inventories of languages and within individual languages (Wedel & Winter 2016). The results presented here are suggestive of an articulatory bias that also helps motivate the prevalence of some sounds. Ease of articulation, and more specifically ease of voicing, has for some time been posited as a potential shaper of speech. Yet debate remains as to the extent of its influence, and some scholars have questioned why it is that, if ease of articulation is such an important factor in speech, phonologies do not exhibit more crosslinguistic convergence. This study has offered a new approach to the question of just how much one particular ease-based factor, namely transglottal air-pressure differential, impacts how much languages rely on particular sounds. The data I have offered suggest a clear effect of this factor on speech, even after controlling for the confounds of linguistic genealogy and contact. These results are consistent with previous work suggesting that the factor in question impacts the phoneme inventories of the world's languages. There is evidence in the literature that the high functional load of some sounds leads to them being maintained as phonemes (Wedel et al. 2013). The relatively high word-list proportions (and likely high functional loads) of sounds like /k/ may help explain their relative prevalence in phoneme inventories when contrasted with their voiced homorganic counterparts.

The mechanisms through which the less effortful obstruents come to be more frequent in word lists requires exploration. One potential factor is the rate at which stops are voiced word-finally. After all, it is well known from diachronic studies that obstruent devoicing is more likely to occur word-finally (see Crowley & Bowern 2010 for one of many discussions of this phenomenon). It is possible that word-final devoicing applies to posterior stops at a higher rate, when contrasted with more anterior stops, or that posterior stops are more likely to be voiceless at syllable boundaries. While the ASJP data do not include syllable boundaries, they are amenable to analysis of word-final position: across all word lists, word-final voiceless velar stops are 3.9 times more common than their voiced counterparts. Recall from Table 1 that voiceless velar stops are generally only about 2.6 times more common than voiced velar stops across all lists, so these word-final disparities for velar stops are particularly pronounced. For alveolar stops, however, voiceless variants are only 1.9 times more common in word-final position—about the same ratio as that exhibited overall (see Table 1). Furthermore, word-final voiceless bilabial stops are 1.7 times more common than their word-final voiced counterparts, which is somewhat surprising given that voiceless bilabials are actually less common overall in the uncontrolled data described in Table 1. Still, these figures, particularly those related to velar stops, suggest that the interaction of the phenomenon with word boundaries, and potentially syllable boundaries, requires further exploration. Ultimately, the diachronic mechanisms motivating the pattern described here require substantive investigation with more robust intralinguistic data.

Even if one already accepts the notion that the aforementioned aerodynamic factors yield universal biases on phonetic and phonological patterns, the present work offers strong additional evidence for this acceptance. These results suggest that differential ease of voicing has a pervasive effect on how languages use their stops and fricatives. It is not simply that certain stops, such as the voiced velar stop, are slightly less common in phoneme inventories. These data do suggest that voiced velar stops are less likely to exist in a given language family, when compared to voiceless velar stops. More pervasively, though, there is a decline in the likelihood that stops and fricatives are voiced as place of articulation progresses toward the glottis. This decline is apparent even after controlling for the areal proximity and relatedness of languages. The decline is explainable if the aforementioned aerodynamic pressures do in fact yield biases on sound usage. The uncovered patterns in languages' relative reliance on obstruents are consistent with previously observed, though less pronounced, patterns in phoneme inventories.

Finally, it is hoped that this study serves as an illustration of the ways that an analysis of typologically diverse word lists can be used as a useful complement to other better-established approaches in the exploration of sound patterns in the world's languages.

REFERENCES

Bates, Douglas; Martin Mächler; Ben Bolker; and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1–48. DOI: 10.18637/jss.v067.i01.

Bickel, Balthasar. 2008. A refined sampling procedure for genealogical control. *STUF—Language Typology and Universals* 61.221–33. DOI: 10.1524/stuf.2008.0022.

Bickel, Balthasar, and Johanna Nichols. 2017. *The AUTOTYP genealogy and geography database*. Online: http://www.autotyp.uzh.ch/.

Blasi, Damián E.; Søren Wichmann; Harald Hammarström; Peter F. Stadler; and Morten H. Christiansen. 2016. Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences of the United States of America* 113.10818–23. DOI: 10.1073/pnas.1605782113.

Brown, Cecil H.; Eric W. Holman; and Søren Wichmann. 2013. Sound correspondences in the world's languages. *Language* 89.4–29. DOI: 10.1353/lan.2013.0009.

Brown, Cecil H.; Eric W. Holman; Søren Wichmann; and Viveka Velupillai. 2008. Automated classification of the world's languages: A description of the method and preliminary results. *STUF—Language Typology and Universals* 61.285–308. DOI: 10.1524/stuf.2008.0026.

Calude, Andreea S., and Mark Pagel. 2011. How do we use language? Shared patterns in the frequency of word use across 17 world languages. *Philosophical Transactions of the Royal Society B* 366.1101–7. DOI: 10.1098/rstb.2010.0315.

Crowley, Terry, and Claire Bowern. 2010. *An introduction to historical linguistics*. Oxford: Oxford University Press.

Everett, Caleb. 2017. Languages in drier climates use fewer vowels. *Frontiers in Psychology* 8:1285. DOI: 10.3389/fpsyg.2017.01285.

Gordon, Matthew. 2016. *Phonological typology*. Oxford: Oxford University Press.

Haspelmath, Martin, and Uri Tadmor. 2009. *Loanwords in the world's languages: A comparative handbook*. Berlin: De Gruyter Mouton.

Jaeger, T. Florian; Peter Graff; William Croft; and Daniel Pontillo. 2011. Mixed effect models for genetic and areal dependencies in linguistic typology. *Linguistic Typology* 15.281–319. DOI: 10.1515/lity.2011.021.

Kaye, Jonathan. 1989. *Phonology: A cognitive view*. Mahwah, NJ: Lawrence Erlbaum.

Keating, Patricia. 1984. Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics* 59.29–34. Online: https://escholarship.org/uc/item/2497n8jq.

Klatt, Dennis H. 1975. Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18.687–703. DOI: 10.1044/jshr.1804.686.

Lindblom, Björn. 1983. Economy of speech gestures. *The production of speech*, ed. by Peter MacNeilage, 217–45. New York: Springer.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian. 2013a. Voicing and gaps in plosive systems. *The world atlas of language structures online*, ed. by Matthew Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: http://wals.info/chapter/5.

Maddieson, Ian. 2013b. Consonant inventories. *The world atlas of language structures online*, ed. by Matthew Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: http://wals.info/chapter/1.

Maddieson, Ian, and Kristin Precoda. 1990. Updating UPSID. *UCLA Working Papers in Phonetics* 74.104–11. Online: https://escholarship.org/uc/item/71s1701m.

Martin, Andy. 2007. *The evolving lexicon*. Los Angeles: University of California, Los Angeles dissertation.

Ménard, Lucie. 2013. Sensorimotor constraints and the organization of sound patterns. *New perspectives on the origins of language*, ed. by Claire Lefebvre, Bernard Comrie, and Henri Cohen, 257–78. Amsterdam: John Benjamins.

Moran, Steven; Daniel McCloy; and Richard Wright. 2017. *PHOIBLE online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: http://phoible.org.

Napoli, Donna Jo; Nathan Sanders; and Rebecca Wright. 2014. On the linguistic effects of articulatory ease, with a focus on sign languages. *Language* 90.424–56. DOI: 10.1353/lan.2014.0026.

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

Ohala, John. 1997. The relation between phonetics and phonology. *The handbook of phonetic sciences*, ed. by William Hardcastle, John Laver, and Fiona Gibbon, 654–77. Oxford: Blackwell.

Ploch, Stefan. 2004. What is wrong with linguistics. *SOAS Working Papers in Linguistics and Phonetics* 13.363–82.

Shariatmadari, David. 2006. Sounds difficult? Why phonological theory needs 'ease of articulation'. *SOAS Working Papers in Linguistics* 14.207–26. Online: https://www.soas.ac.uk/linguistics/research/workingpapers/volume-14/file37823.pdf.

Smith, B. L., and J. R. Westbury. 1975. Temporal control of voicing during occlusion in plosives. Paper presented at the 89th meeting of the Acoustical Society of America.

Stevens, Kenneth. 2000. *Acoustic phonetics*. Cambridge, MA: MIT Press.
Swadesh, Morris. 1952. Lexico-statistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society* 96.452–63. Online: https://www.jstor.org/stable/3143802.
Tsujimura, Natsuko. 2007. *An introduction to Japanese linguistics*. Oxford: Blackwell.
Wedel, Andrew; Abby Kaplan; and Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128.179–86.
Wedel, Andrew, and Bodo Winter. 2016. Languages prefer robust phonemes. *The evolution of language: Proceedings of the 11th International Conference (EVOLANG 11)*, New Orleans. DOI: 10.17617/2.2248195.
Wichmann, Søren; Eric Holman; and Cecil Brown (eds.) 2016. *The ASJP database (version 17)*. Online: http://asjp.clld.org.

Department of Anthropology
University of Miami
Coral Gables, FL 33124-2005
[caleb@miami.edu]

[Received 13 December 2017;
revision invited 26 February 2018;
revision received 31 March 2018;
accepted with revisions 2 June 2018;
revision received 6 June 2018;
accepted 10 June 2018]