

# *Estimating historical probabilities of natural and unnatural processes\**

**Gašper Beguš**

University of California, Berkeley

---

This paper presents a technique for estimating the influences of channel bias on phonological typology. The technique, based on statistical bootstrapping, enables the estimation of historical probability, the probability that a synchronic alternation arises based on two diachronic factors: the number of sound changes required for an alternation to arise and their respective probabilities. I estimate historical probabilities of six attested and unattested alternations targeting the feature [voice], compare historical probabilities of these alternations, perform inferential statistics on the comparison and, to evaluate the performance of the channel bias approach, compare outputs of the diachronic model against the independently observed synchronic typology. The technique also identifies mismatches between the typological predictions of the analytic bias and channel bias approaches. By comparing these mismatches with the observed typology, this paper attempts to quantitatively evaluate the distinct contributions of the two influences on typology in a set of alternations targeting the feature [voice].

---

## 1 Introduction

Typological literature in phonology has long revolved around the question of which factors influence the observed typology. Two major lines of thought have emerged in this discussion: the ANALYTIC BIAS approach and the CHANNEL BIAS approach (Moreton 2008, Yu 2013b).<sup>1</sup> The analytic bias approach argues that the observed typology results primarily from differences in the learnability of phonological processes; the channel bias approach argues that the inherent directionality of sound changes based

\* E-mail: [BEGUS@BERKELEY.EDU](mailto:BEGUS@BERKELEY.EDU).

I would like to thank Kevin Ryan, Jay Jasanoff, Adam Albright, Donca Steriade, Edward Flemming, Patrick Mair, Morgan Sonderegger, the editors and three anonymous reviewers at *Phonology*, and the audiences at NELS 48, AMP 2016, 2017, WCCFL 35 and CLS 53 for useful comments on earlier versions of this paper. All remaining mistakes are my own.

<sup>1</sup> Other names have been used for the two approaches, such as Evolutionary Phonology *vs.* Amphichronic Phonology (Blevins 2004, Kiparsky 2006, 2008).

on phonetic precursors (articulatory and perceptual) shapes typology (for further discussion, see Hyman 1975, 2001, Greenberg 1978, Ohala 1981, 1983, 1993, Kiparsky 1995, 2006, 2008, Blevins 2004, Moreton 2008, 2012, Moreton & Pater 2012a, b, de Lacy & Kingston 2013, Garrett & Johnson 2013, Cathcart 2015, among others).

Empirical evidence often supports both approaches equally well. Typologically rare processes have in many cases been shown to be more difficult to learn, which supports the analytic bias approach (Kiparsky 1995, 2006, 2008, Wilson 2006, Becker *et al.* 2011, de Lacy & Kingston 2013, White 2017; for an overview of the experimental analytic bias literature, see Moreton & Pater 2012a, b). On the other hand, typologically frequent processes are often shown to directly result from the phonologisation of underlying articulatory or perceptual phonetic precursors (e.g. sound change in progress that results in a typologically common pattern), whereas rare or unattested processes lack such precursors, lending support to the channel bias approach (cf. Hyman 1975, Greenberg 1978, Ohala 1981, 1983, 1993, Lindblom 1986, Barnes 2002, Blevins 2004, 2006, 2007, 2008a, b; see Garrett & Johnson 2013 for an overview of the literature). This ambiguity of evidence poses the primary challenge in typological research.

The stance of this paper is that the typology is influenced by both factors (as has been argued in a growing body of recent research; see Hyman 2001, Myers 2002, Moreton 2008, Moreton & Pater 2012a, b, de Lacy & Kingston 2013). However, the role of phonological research is to quantitatively evaluate which aspects of typology are more likely to result from one factor or the other. The question that this paper addresses is whether some observed typological distributions (e.g. those targeting the feature [voice]) are primarily influenced by different degrees of the learnability of different processes or by the different diachronic trajectories underlying different processes. In particular, this paper proposes a technique based on which channel bias influences on typology can be estimated.

## 1.1 Analytic bias

If typologically infrequent processes are experimentally shown to be more difficult to learn than typologically frequent processes (for an overview, see Moreton & Pater 2012a, b), a reasonable conclusion would be that typological differences result precisely from these differences in learnability. A challenge that the analytic bias approach faces is that artificial grammar learning experiments frequently show no learnability differences between typologically rare or non-existent unnatural processes and typologically frequent natural processes when the structural complexity of the tested alternations is controlled for. The influences of analytic bias can be subdivided into SUBSTANTIVE BIAS and COMPLEXITY BIAS (Wilson 2006, Moreton 2008, Moreton & Pater 2012a, b). For substantive bias, phonetically motivated processes are easier to learn than unmotivated (or unnatural) ones; for complexity bias, alternations involving more

conditioning features are more difficult to learn than simpler alternations (Moreton 2008). A survey of experimental literature on analytic bias in Moreton & Pater (2012a, b) shows that there exist consistent differences in experimental results testing the two biases. While complexity bias was consistently confirmed by the majority of studies surveyed, experimental outcomes of the substantive bias were mixed. Several studies that tested the learning of unnatural alternations as defined in §2 found no effect of substantive bias (Pycha *et al.* 2003, Seidl *et al.* 2007, Kuo 2009, Skoruppa & Peperkamp 2011, Moreton & Pater 2012a, b, Do *et al.* 2016, Glewwe 2017, Glewwe *et al.* 2018). A smaller subset of studies, however, do report positive results (Carpenter 2006, 2010, Wilson 2006).

L1 acquisition and L2 acquisition of word-final stops by speakers of L1s that ban obstruent codas are the only areas where differences between the natural and unnatural pair of alternations are observed. Learners acquire word-final voiceless stops earlier than voiced stops, and devoice voiced stops more frequently than they voice voiceless stops word-finally (Clark & Bowerman 1986, Kong *et al.* 2012 and references therein; for an overview, see Broselow 2018). It is likely, however, that this type of experiment tests differences in the learning of more complex *vs.* less complex articulations (Kong *et al.* 2012), and not the abstract phonological learning that is observed, for example, in artificial grammar learning experiments (e.g. where complex alternations are more difficult to learn than simple alternations, which is independent of articulatory factors; Moreton & Pater 2012a, b). Articulation of segments that require more articulatory effort in a given position is expected to be learned less successfully: ‘cross-language differences in the age of children’s mastery of adult-like voiced stops are typically explained in terms of the relative difficulty of the laryngeal gestures for the language’s voice onset time distributions’ (Kong *et al.* 2012: 725). The identical mechanism is in fact responsible for final voicing within the channel bias approach: even adult L1 speakers with full contrast devoice final voiced stops gradiently and passively, due to their greater articulatory complexity, which can result in a typologically common sound change that operates in an adult population (cf. Labov 1994). These L1 and L2 learning differences thus likely reflect differences in articulatory effort that should be modelled as a channel bias influence. It is in fact not trivial to show how differences in L1 articulatory learning would result in phonological typology (cf. Rafferty *et al.* 2013), given that, after some developmental stage, children reproduce their input with a high degree of faithfulness (e.g. in the acquisition of the voicing contrast; see Kong *et al.* 2012).

## 1.2 Channel bias

One of the objections to the channel bias approach to typology is that it fails to explain why some processes are unattested (Kiparsky 2006, 2008, de Lacy & Kingston 2013). Kiparsky (2006), for example, lists several diachronic trajectories that would lead to final voicing, yet final voicing is arguably not attested as a productive synchronic process. More generally,

combinations of sound changes could conspire to yield a number of processes that are never attested as productive synchronic alternations. In the absence of a diachronic explanation, Kiparsky invokes grammatical constraints and learnability to explain these typological gaps.

Most of the current models of typology within the channel bias approach are indeed insufficient for explaining such typological gaps, because they do not quantify the probability of the occurrence of sound changes or combinations of sound changes. The default explanation within the channel bias approach has long been a qualitative observation that common processes are frequent because they are produced by frequent sound changes or because they require fewer sound changes (Blevins 2015: 485; also Greenberg 1978: 75–76). Such reasoning does not provide sufficient outputs for a quantitative comparison of different influences on phonological typology.

Despite these objections, mechanisms exist within the channel bias approach to derive typology that go further than the simple statement that rare sound changes produce rare alternations. Based on a typological study of an unnatural process, postnasal devoicing, Beguš (2019) argues that unnatural processes require at least three sound changes (as opposed to at least two for unmotivated processes and at least one for natural processes; the MINIMAL SOUND CHANGE REQUIREMENT), which explains the relative rarity of processes with different degrees of naturalness. The idea that unmotivated processes are rare because they require a complex history is certainly not new (Bell 1970, 1971, Greenberg 1978: 75–76, Cathcart 2015, Morley 2015), but the Minimal Sound Change Requirement explains why unnatural processes are the least frequent (compared to natural or unmotivated processes; see §2). The Minimal Sound Change Requirement on its own, however, does not explain why some unnatural processes are attested, while others are not. To quantify the channel bias influences on typology further, the concept of the Minimal Sound Change Requirement has to be combined with the estimation of probabilities of individual sound changes that are required for each synchronic alternation to arise.

Two models have thus far attempted to quantify probabilities of the occurrence of various primarily static phonotactic processes and explain the relative rarity of some processes based on diachronic factors. Bell (1970, 1971) and Greenberg (1978) propose a ‘state-process model’. Their model operates with typological states (phonological, morphological and syntactic) that can arise from other states, depending on the number of previous states, transitional probabilities from one state to another and the rest probabilities of each state. This approach is most suitable for modelling the probabilities of various phonotactic restrictions. Modelling the probabilities of transitions (processes) in the version of the model in Bell (1971) involves relative probabilities that only tangentially reflect the frequencies of the processes in the samples. The main ideas behind Bell’s and Greenberg’s models are similar to what will be proposed in this paper, but their proposal lacks inferential statistical tests. Crucially, by estimating

uncertainty behind the distributions with bootstrapping, we can compare the historical probabilities of alternations, and perform hypothesis testing on the comparisons. The proposal here also uses substantially more elaborate historical samples.

A different model for calculating the probabilities of the combination of sound changes is offered by Cathcart (2015), who computes permutations of sound changes that lead to a certain process (in this case, final voicing), and compares that to permutations of all sound changes in a given survey to gain an estimate of the probability of certain processes. Due to its design, however, Cathcart's (2015) model relies on the representativeness of diachronic surveys for all sound changes, not only for those that are estimated (see also §3.2), and is computationally demanding, making it difficult to implement. The models in Greenberg (1978) and Cathcart (2015) also do not take into consideration the crucial distinctions made in Beguš (2019: 744): 'the subdivision of unusual rules into unnatural versus unmotivated rules, paired with the proof that the latter require at least three sound changes to arise'.<sup>2</sup> The model proposed in this paper has the disadvantage that the trajectories of sound changes that lead to a certain alternation need to be identified manually (similar to the Bell and Greenberg models), but this also means that samples of sound changes need be representative only for the sound changes being estimated.

### 1.3 Objectives

The goal of this paper is to propose a quantitative method for estimating the influences of channel bias on phonological typology using a statistical method called *BOOTSTRAPPING* (Efron 1979, Efron & Tibshirani 1994). The technique estimates the *HISTORICAL PROBABILITY*, the probability of an alternation arising based on two diachronic factors: the number of sound changes required for an alternation to arise (the Minimal Sound Change Requirement; §2), and their respective probabilities, estimated from surveys of sound changes. Using the proposed technique, we can (i) estimate the historical probability of any alternation (§4), (ii) compare two alternations, attested or unattested, and perform statistical inferences on the comparison (§5.2) and (iii) compare outputs of the historical model with independently observed typology to evaluate the performance of the channel bias approach (§5.3). The assumptions of the model are discussed in §3.3. The paper also identifies mismatches between typological predictions of the analytic and channel bias approaches (§5.3 and §6). By testing these mismatched predictions against the observed typology, we can at least partially control for one factor when testing the other, and

<sup>2</sup> The model of automated reconstruction in Bouchard-Côté *et al.* (2013) estimates the probabilities of individual sound changes, but does not deal with combinations of sound changes. Other quantitative approaches to sound change (e.g. Kirby & Sonderegger 2013, 2015, Hruschka *et al.* 2015) do not directly deal with estimating the probabilities of sound changes that operate in combination, but model the initiation and propagation of single sound changes computationally.

*vice versa*, which allows for quantitative evaluation of the distinct contributions of the analytic bias and channel bias factors to phonological typology (§6).

While the proposed method can be applied to any natural–unnatural pair of alternations, the paper focuses on a subset of typology: three natural–unnatural alternation pairs that target feature  $[\pm\text{voice}]$ . We estimate the historical probabilities of postnasal voicing (e.g.  $/p/ \rightarrow [b] / m \_$ ) and postnasal devoicing ( $/b/ \rightarrow [p] / m \_$ ), intervocalic voicing ( $/p/ \rightarrow [b] / V \_ V$ ) and intervocalic devoicing ( $/b/ \rightarrow [p] / V \_ V$ ), and final devoicing ( $/b/ \rightarrow [p] / \_ \#$ ) and final voicing ( $/p/ \rightarrow [b] / \_ \#$ ). The feature  $[\pm\text{voice}]$  is chosen because phonetic naturalness is probably best understood precisely for this feature (see Ohala 1983, 2011 and Westbury & Keating 1986 on the ‘Aerodynamic Voicing Constraint’, and for detailed argumentation Beguš 2019), and all alternations have been well researched typologically. The alternations also differ in their degree of synchronic attestedness (see Fig. 3 below), providing a good basis for a comparison of different approaches to phonological typology.

## 2 Background

This paper makes use of several diachronic concepts from Beguš (2019). First, it adopts the division of phonological processes into NATURAL, UNMOTIVATED and UNNATURAL. Natural processes, such as postnasal and intervocalic voicing and final devoicing, represent phonetically well-motivated universal phonetic tendencies. Unmotivated processes lack phonetic motivation, but do not operate against universal phonetic tendencies, which are defined by Beguš (2019: 691) as ‘phonetic pressures motivated by articulatory or perceptual mechanisms ... that passively operate in speech production cross-linguistically and result in typologically common phonological processes’. An example of an unmotivated process would be Eastern Ojibwe ‘palatalisation’ of  $/n/$  to  $[j]$  before front vowels (Buckley 2000). Unnatural processes not only lack phonetic motivation, but also operate against universal phonetic tendencies. Examples of unnatural alternations that operate against universal phonetic tendencies include postnasal devoicing, intervocalic devoicing and final voicing, which will be discussed in this paper.

We limit the modelling of sound change to a non-analogical phonetically driven sound change that targets a single feature value (Hyman 1976). We also assume that a single sound change is a change of a single feature or a deletion of a feature matrix (the ‘minimality principle’; Picard 1994). That a single sound change targets only one feature value is, at least in the great majority of cases, suggested by historical typology. Phonetic precursors also support the minimality principle: phonetic precursors that lead to sound change are usually articulatorily and perceptually minimal (Moreton 2008). For discussion, see Donegan & Stampe (1979), Picard (1994) and Beguš (2019).

This paper adopts two diachronic concepts for the derivation of typology within the channel bias approach that are proposed in Beguš (2019): the Blurring Process and the Minimal Sound Change Requirement. Typological surveys of unnatural processes targeting the feature [voice] conducted in Beguš (2018, 2019) and Beguš & Nazarov (2018) identify thirteen languages in which postnasal devoicing has been reported either as a productive synchronic alternation or as a sound change. Based on this typological survey, a hypothesis about how unnatural processes arise diachronically is proposed: the BLURRING PROCESS, which states that unnatural alternations arise through a combination of a specific set of three natural (phonetically motivated) sound changes: (i) a sound change that leads to complementary distribution, (ii) a sound change that targets changed or unchanged segments in complementary distribution and (iii) a sound change that blurs the original complementary distribution (Beguš 2019). For example, postnasal devoicing results from three sound changes: (i) frication of voiced stops except postnasally ([b] > [β] / [-nas] \_\_), (ii) unconditioned devoicing of voiced stops ([b] > [p]) and (iii) occlusion of voiced fricatives ([β] > [b]) (see §4.1).

This allows us to maintain the long-held position that sound change can only be acoustically or perceptually motivated (Garrett & Johnson 2013, Garrett 2015; for a discussion, see Beguš 2019).<sup>3</sup> This position is challenged by Blust (2005), who lists a number of unnatural sound changes. If, however, these unnatural sound changes result from combinations of natural sound changes (as argued by Beguš 2018, 2019 and Beguš & Nazarov 2018), we can maintain the position that sound change is always phonetically motivated.

Unnatural alternations thus cannot arise from a single sound change. Beguš (2019) additionally argues that unnatural segmental alternations cannot arise from two sound changes either. If a change from '+' to '-' for a given feature [ $F_1$ ] is unnatural and therefore cannot result from a single sound change, some other feature ([ $F_2$ ]) has to change first, so that the change from [ $+F_1$ ] to [ $-F_1$ ] might be natural and motivated. For the full unnatural process to take place, however, that other feature ([ $F_2$ ]) has to change back to its original value. No such requirement exists for unmotivated processes: they can be the result of two sound changes. In sum, the Minimal Sound Change Requirement states that a minimum of three sound changes are required for an unnatural alternation to arise, a minimum of two sound changes for an unmotivated alternation and a minimum of one sound change for a natural alternation (see Beguš 2019).<sup>4</sup>

<sup>3</sup> In some cases, multiple directions of a sound change can potentially be motivated, although such cases are rare.

<sup>4</sup> The proposed model does not account for unnatural processes that result from morphological changes. These are, however, almost always analysed as morphologically conditioned, and can be explained by non-phonological mechanisms. The lack of modelling morphological changes is a shortcoming of the current proposal. Sound change is, however, substantially more unidirectional, and sound-change typology is substantially better understood, than morphological or syntactic change, which is why a model including morphological sources of phonological processes would

### 3 Bootstrapping in order to estimate historical probabilities

The goal of this paper is to develop a model that will provide a quantification of probabilities of natural, unmotivated and unnatural processes which goes beyond the statement that natural processes are the most frequent, unmotivated less frequent and unnatural the least frequent. We can combine the Minimal Sound Change Requirement with the assumption that the probabilities of sound changes influence the probabilities of synchronic alternations. The probability that an alternation arises based on diachronic factors depends both on the number of sound changes that are required for the alternation to arise and on the probability of each individual sound change in the combination. Such probabilities are called HISTORICAL PROBABILITIES ( $P_\chi$ ).

I propose that historical probabilities can be estimated with bootstrapping. Bootstrapping is a statistical technique within the frequentist framework for estimating sampling distribution (and consequently standard errors and confidence intervals for a statistic of interest) by random sampling with replacement (Efron 1979, Efron & Tibshirani 1994, Davison & Hinkley 1997).

#### 3.1 The model

3.1.1 *Individual sound changes.* Probabilities of individual sound changes are estimated from a sample of successes (languages in a sample with a sound change  $S_i$ ) and failures (languages in a sample without the sound change  $S_i$ ), according to (1). If an alternation  $A_k$  requires only one sound change to arise and invariably occurs as a result of that change (i.e.  $A_k$  is natural), then we estimate its  $P_\chi$  as in (1).

$$(1) P_\chi(S_i) = \frac{\text{number of languages with sound change } S_i}{\text{number of languages surveyed}}$$

To estimate the historical probability of a sound change using the bootstrapping technique, we create samples from counts of languages (see §4.2). Languages with a sound change are treated as successes (coded as 1); languages without it as failures (coded as 0). The main advantage of the bootstrapping technique is that it estimates the confidence intervals of a historical probability. To estimate confidence intervals, we sample from this distribution of successes and failures 10,000 times (using ‘sampling with replacement’; Efron 1979, 1987). Each of these 10,000 samples is a probability of a success based on (1) (i.e. the proportion of 1s relative to 0s in the sample). These 10,000 probabilities constitute a

---

first require an elaborate typology of morphologically induced phonological processes.



sampling distribution, based on which standard errors and confidence intervals are calculated.

For example, when estimating the historical probabilities of processes targeting feature [voice], the successes and sample sizes are taken from surveys of sound changes (for exact counts, see §4.2). Sample sizes in our case range from 88 to 294, depending on the sound change (see §4.2). This repeated sampling with replacement yields a sampling distribution of historical probabilities: 10,000 data points for each process. From this sampling distribution, standard error, bias and 95% adjusted bootstrap (BC<sub>a</sub>) confidence intervals (CIs) that adjust for bias and skewness (Efron 1979, 1987) are computed.

The computation is implemented in the statistical software R (R Core Team 2016) with the *boot* package (Davison & Hinkley 1997, Canty & Ripley 2016), using the functions *boot()* and *boot.ci()*. The R code used to implement the proposed technique is available in Appendix C.<sup>5</sup>

3.1.2 *Two or more sound changes.* If an alternation A<sub>k</sub> requires more than a single sound change, then the historical probability of A<sub>k</sub> is estimated as a sum of the historical probabilities of each trajectory T<sub>j</sub> that yields the alternation A<sub>k</sub>, as shown in (2).

$$(2) P_{\chi}(A_k) = P_{\chi}(T_1 \cup T_2 \cup \dots \cup T_n)$$

A trajectory T<sub>j</sub> denotes a combination of sound changes that yields an alternation A<sub>k</sub>. In theory, there are an infinite number of trajectories that yield any given alternation, but for practical purposes, we estimate only the trajectory that involves the least number of sound changes. Historical probabilities of trajectories that require more than three sound changes are assumed to be minor enough to be disregarded for practical purposes.

The historical probability of a trajectory T<sub>j</sub> that requires more than a single sound change is estimated from the joint probability of the individual sound changes required for T<sub>j</sub>, divided by the factorial of the number of sound changes in trajectory T<sub>j</sub> if only one ordering results in the trajectory in question, as shown in (3).

$$(3) P_{\chi}(T_j) = \frac{P_{\chi}(S_1 \cap S_2 \cap \dots \cap S_n)}{n!}$$

Historical probability is a probability that a language L features an alternation A<sub>k</sub>, regardless of the properties of L. In other words, we do not condition historical probabilities on languages that feature a certain property. The historical probability (P<sub>χ</sub>) of the first individual sound change S<sub>1</sub> is thus estimated from the number of successes (languages with S<sub>1</sub>) and the

<sup>5</sup> All appendices are available as online supplementary materials at <https://doi.org/10.1017/S0952675720000263>.

number of failures (languages without  $S_1$ ) according to (1), regardless of the phonemic inventories of languages in the sample.

For example, if the target of the first sound change  $S_1$  in a trajectory that results in an alternation  $A_k$  is a geminate stop, we estimate the historical probability of  $S_1$  from the number of languages with the sound change  $S_1$  divided by the number of all languages surveyed, including those that do not feature geminate stops. The historical probability of an alternation  $A_k$  that requires  $S_1$  is simply the probability that the alternation  $A_k$  arises in a language  $L$ , regardless of whether it features geminate stops.

Once the sound change  $S_1$  has operated, however, we know that language  $L$  necessarily has the target, result and context of  $S_1$ . For this reason, we estimate the historical probability of the subsequent sound changes  $P_\chi(S_2)$  by dividing the number of successes (languages with  $S_2$ ) by the number of languages surveyed that feature the target, result and context of  $S_1$ , if these are also the target of  $S_2$ . The same is true for any subsequent sound change. Once we condition the probability of sound changes and estimate it from samples of sound changes, given that they have the target, result and context of the previous sound change, we can treat the probabilities of individual sound changes as independent events under the channel bias approach, and estimate  $P_\chi$  from the product of the probabilities of individual sound changes, as in (4).

$$(4) \quad P_\chi(T_j) = \frac{\prod_{i=1}^n P_\chi(S_i)}{n!}$$

To estimate standard errors and  $BC_a$  confidence intervals for a historical probability of  $A_k$  that requires more than a single sound change, the proposed technique samples with replacement from  $n$  individual binomial samples (one sample for each individual sound change, constructed as described above) computes the historical probability of each sound change (according to (1)) and then computes the product of the historical probabilities of each individual sound change divided by  $n!$ , according to (4). This process returns 10,000 bootstrap replicates of the historical probability of  $A_k$ , from which the standard errors and  $BC_a$  confidence intervals are computed.

**3.1.3 Comparison.** The proposed technique also allows for the estimation of the difference between the historical probabilities of two alternations, which consequently enables inferential statements on the comparison, as in (5).

$$(5) \quad \Delta P_\chi(A_1, A_2) = P_\chi(A_1) - P_\chi(A_2)$$

The difference between the historical probabilities of two alternations ( $\Delta P_\chi$ ) is estimated with a stratified non-parametric bootstrap, where  $P_\chi$  of each individual alternation  $A_1$  and  $A_2$  is estimated as described in

§3.1.1 and §3.1.2 (depending on whether  $A_1$  and  $A_2$  require trajectories that require one or more sound changes). To compare two historical probabilities, we calculate the difference between  $P_\chi(A_1)$  and  $P_\chi(A_2)$ , which returns 10,000 bootstrap replicates, from which the standard errors and  $BC_a$  confidence intervals are computed.

The proposed technique applied to a difference between two alternations enables a comparison of the two alternations with inferential statements. If the 95%  $BC_a$  confidence intervals of the difference both fall either below or above 0, then  $P_\chi(A_1)$  and  $P_\chi(A_2)$  are significantly different, with  $\alpha = 0.05$ .

### 3.2 Sample

Samples used for estimating historical probabilities are created from counts of occurrences of sound changes in typological surveys. The proposed technique is most accurate when typological surveys are large, well-balanced and representative. Sound changes in a survey should always be evaluated with respect to the target of the change, its result and its context. Sound-change occurrences in a typological survey should be properly counted: if two or more daughter languages show the result of a sound change that operated at the proto-stage of the two languages, the sound change should be counted as a single event in the proto-language. For a detailed description of how samples for the six alternations in this paper are created, see §4.2, and for lists of languages, see Appendix A.

The most elaborate currently available survey of sound changes used in the paper is the survey of consonantal sound changes in Kümmel (2007). One major advantage of Kümmel's survey is that it includes language families with a well-reconstructed prehistory and a well-established subgrouping. This allows for a more accurate coding of the occurrence of a sound change than competing surveys.<sup>6</sup>

Sound changes are counted as single events if they operate at a proto-language stage. While it is sometimes difficult to reconstruct whether a sound change in two related languages operated at the proto-stage or independently in individual branches, especially for typologically frequent sound changes, the survey in Kümmel (2007) is the most comprehensive of all available surveys in this respect. Kümmel's subgrouping relies on historical methodology that includes information from phonological as well as morphological and other higher-level evidence.<sup>7</sup>

Kümmel's survey includes approximately 294 languages and dialects of the Indo-European, Semitic and Uralic language families. While it is limited to only three language families, these are families with well-

<sup>6</sup> The only other comparable survey of sound changes known to me is the UniDia database, which surveys 10,349 sound changes from 302 languages (Hamed & Flavier 2009). The UniDia database is, however, less appropriate, because it lacks elaborate diachronic subgroupings of languages.

<sup>7</sup> Additionally, phylogenetic tree analysis does not restrict the direction of sound change and would, for example, incorrectly analyse reported unnatural alternations as resulting from a single sound change.

established subgroupings. This allows for proper coding, and at least partially compensates for the lack of representativeness. The results of the analysis can be affected by the fact that many language families are absent from the survey. However, it is likely that such effects are minor, because types of sound changes do not seem to be radically different across different language families (with recurrent sound changes appearing across all families; see Blevins 2007 and § 5.3). Additionally, I am unaware of reasons to believe that the representativeness (or bias) of a sample is unequal across different sound changes. Because we are primarily interested in the comparison between historical probabilities of various alternations and less so in their absolute values, the model is less prone to influence from biased samples.

### 3.3 Assumptions

As with any diachronic model, the proposed technique has to make some simplifying assumptions. In order to estimate the joint probability of two or more sound changes as a product of the historical probabilities of each individual sound change (see (4)), it is assumed that each sound change is an independent event. The proposed model *does* account for the dependency between sound changes where one sound change alters the target or context of the following sound change. Probabilities of sound changes are estimated on the basis of their targets, results and contexts (§ 3.2) and, crucially, from samples conditioned on the result of the previous sound change (§ 3.1.2). Two crucial assumptions of independence remain: that sound change is (i) independent of previous sound changes when the dependence on targets, results and contexts of the previous sound change is controlled for (§ 3.1.2), and (ii) independent of the global phonemic properties of a language (those properties that do not immediately affect the conditions of sound changes in question).

The first assumption is not controversial when modelling typology within the channel bias approach. The proposed method aims to estimate only the channel bias influences on typology, which is why it has to assume that the probability of sound change is determined only by its frequency of operation, evaluated on a diachronic and unconditioned level.

The second assumption of independence is more problematic: broader phonemic inventories can influence the probabilities of sound changes, especially for vocalic changes (for example due to the effects described in the Theory of Adaptive Dispersion; Liljencrants & Lindblom 1972, Lindblom 1990), but also for consonantal changes. The proposed technique does not model the dependency of sound changes on those phonemic properties that do not immediately affect the targets, results or contexts of the sound changes in question. The sample's representativeness should, however, at least partially cancel out potential dependencies. The sample of sound changes from which the historical probabilities are calculated includes languages with a diverse set of phonemic inventories (see Kümmel 2007). Additionally, we do not condition estimations of historical

probabilities on any specific property of phoneme inventories, which makes the dependency between sound change and more distant phonemic properties less crucial to our proposal. Finally, we are unaware of any properties of phonemic inventories that would affect the rate of the sound changes in question (e.g. intervocalic lenition, occlusion of fricatives, devoicing of stops).<sup>8</sup>

As already mentioned, identification of individual trajectories leading to an alternation  $A_k$  is performed manually in the approach described here. While this task is facilitated by the Blurring Process, which describes mechanisms for unnatural processes to arise, it is nevertheless possible that some trajectories that would potentially influence the final result are missing from the estimation. If we assume that the estimated trajectory  $T_j$  is indeed the most frequent trajectory leading to  $A_k$  and that potential alternative trajectories do not crucially influence the overall historical probability of an alternation, we can generalise the historical probability of that particular trajectory to the historical probability of the alternation. If such an assumption is not met, however, then the proposed technique estimates only the probability that an alternation  $A_k$  arises from a trajectory  $T_j$ .<sup>9</sup> This paper assumes that the estimated trajectories are the most frequent, and that potential alternative trajectories do not crucially influence the results.

The proposed model aims to estimate the channel bias influences on typology. It is possible that learnability influences the probabilities of individual sound changes: learnability can increase or decrease the likelihood of a phonetic precursor being phonologised (as argued by Moreton 2008; for criticism, see Kapatsinski 2011 and Yu 2011). Even if the probabilities of individual sound changes are crucially influenced by learnability (and therefore by analytic bias), and even if learnability causes a higher rate of occurrence of certain sound changes in combination, the requirement that more than one sound change needs to operate in a language for unmotivated alternations and unnatural alternations to arise has to be independent of learnability. This means that at least a portion of the estimated probabilities needs to be influenced by the channel bias.

What is not accounted for in the model is the functional load of individual phonemes (Wedel 2012, Wedel *et al.* 2013, Hay *et al.* 2015), as well as other factors that could potentially influence probabilities of sound changes, such as lexical diffusion or lexical/morpheme frequency during the initial stages of sound change (Bybee 2002), language contact and sociolinguistic factors. The model makes no assumptions about how

<sup>8</sup> The dependence of sound change on broader phonemic inventories is not modelled, primarily because current surveys of sound changes are not sufficiently large and representative. In principle, the proposed technique could model this dependency by estimating the probabilities of sound changes from samples conditioned on some phonemic property of the languages surveyed.

<sup>9</sup> When more representative surveys become available, this assumption could be weakened by using Cathcart's (2015) 'permutation' approach to identify trajectories for each alternation estimated with the proposed technique.

sound change is initiated or spread. These factors can mostly be disregarded, because the goal is to estimate the historical probability of alternation  $A_k$  operating in a language  $L$  with no conditional properties.

Finally, the proposed technique does not directly model the temporal dimension. In the absence of temporal information, we have to make some simplifying assumptions. These are not unique to the present proposal, and are to some extent desirable. The technique estimates historical probabilities within a timeframe that approximates the average timeframe of the languages in the sample. The model also assumes that in order for a resulting alternation to be productive, all sound changes need to operate within one language. While this might be too restrictive, it is desirable to limit the timeframe in which sound changes and corresponding processes have to operate productively for the resulting alternation to be productive. For example, the combination of sound changes (the Blurring Process) that would result in postnasal devoicing in Yaghnobi operates over three languages, and fails to result in a productive synchronic alternation (Beguš 2019). The model also assumes that once a sound change has occurred in a language, it can reoccur. This is a closer approximation to reality than to assume that a sound change cannot operate in daughter languages once it has already operated in the parent language.

The historical probability of an unnatural alternation depends not only on sound changes that are required for the alternation to arise, but also on the probability that the sound change with the opposite effect (in this case, the natural sound change) will operate on the unnatural system and destroy the evidence for it. Influences of potential natural sound changes are not modelled, because the historical probabilities of the natural sound changes are relatively similar for the processes estimated in this paper, and we do not expect this additional factor to alter the results significantly.<sup>10</sup> For other processes not estimated in this paper, the inclusion of the probability of the natural sound change in the model might alter the outcomes significantly.

## 4 Applications

### 4.1 Trajectories

The three natural alternations have obvious origins – three single natural sound changes: postnasal voicing, intervocalic voicing and final devoicing. For the unnatural alternations, we first identify sound changes in the Blurring Process (see §2) that yield the alternation in question. If  $A > B / X$  is a natural sound change, then  $B > A / X$  is unnatural. [Tables Ia–c](#) represent schematically how the unnatural  $B > A / X$  arises via the Blurring Process (see §2; Beguš 2018, 2019, Beguš & Nazarov 2018), together with the actual sound changes that yield the unnatural alternation.

<sup>10</sup> For example, the difference between the historical probabilities of postnasal devoicing and final devoicing is not significant ( $P_{\chi} = 1.4\%$ ,  $BC_a$  CI =  $[-3.4\%, 5.4\%]$ ).

a.	Blurring Process	postnasal devoicing	schematic example
sound change 1	B > C / ¬X	D > Z / [-nas]__	bamba βamba
sound change 2	B > A	D > T	βampa
sound change 3	C > B	Z > D	bampa
result	B > A / X	D > T / [+nas]__	

b.	Blurring Process	intervocalic devoicing	schematic example
sound change 1	B > C / X	D > Z / V__V	baba baβa
sound change 2	C > D	Z > S	baφa
sound change 3	D > A	S > T	bapa
result	B > A / X	D > T / V__V	

c.	Blurring Process	final voicing	schematic example
sound change 1	C > B / X	T: > T / __#	p:ap:ap: p:ap:ap
sound change 2	B > A	T > D / V__	p:ap:ab
sound change 3	C > B	T: > T	papab
result	B > A / X	T > D / __#	

Table I  
Blurring Processes yielding (a) postnasal devoicing, (b) intervocalic devoicing, (c) final voicing.

A combination of the following three natural and well-motivated sound changes yields postnasal devoicing in all known cases: the fricativisation of voiced stops in non-postnasal position, the unconditioned devoicing of voiced stops and the occlusion of voiced fricatives to stops (Table Ia).<sup>11</sup>

<sup>11</sup> Perhaps the most intriguing of these changes is the unconditional devoicing of voiced stops, even in postnasal position. There is historical and phonetic evidence for the operation of unconditioned devoicing of stops in the development of postnasal devoicing. In fact, this sound change is directly attested in Yaghnobi, one of the languages that feature postnasal devoicing. Yaghnobi also contains evidence that this sound change operates unconditionally in voiced stops, as other marginal positions (e.g. after a voiced fricative) also undergo devoicing. The phonetics behind unconditioned devoicing that includes the postnasal position is straightforward: closure is always antagonistic to voicing, and after some period of time speakers need to actively adjust articulators to sustain voicing in all positions. Failure to do

Beguš (2018) argues that intervocalic devoicing results from three sound changes. Voiced stops fricativise intervocalically, voiced fricatives devoice and voiceless fricatives are occluded to stops (see Table Ib). The result is the unnatural intervocalic devoicing ( $D > T / V\_V$ ).

Final voicing is arguably unattested either as a synchronic alternation or as a sound change (Kiparsky 2006, Lipp 2016, Beguš 2018; for an opposing view, see Yu 2004, Rood 2016).<sup>12</sup> A number of diachronic scenarios exist, however, that would yield final voicing; these are identified in Kiparsky (2006). Most of the scenarios either include more than three sound changes or do not result in a phonological alternation, but in a static phonotactic restriction (§4.2). One possible scenario that involves three sound changes and that would result in final voicing is Kiparsky's Scenario 1, which is used here for estimating the historical probability of final voicing.<sup>13</sup> For the sound changes in Scenario 1 to result in synchronic alternations, we need to assume that geminate simplification first operated word-finally, and only later targeted other geminates. Without this assumption, the sound changes in Scenario 1 would result in a phonotactic restriction. The three sound changes operating to yield final voicing in this scenario are geminate simplification in word-final position, voicing of post-vocalic non-geminate stops and unconditioned geminate simplification (see Table Ic).

## 4.2 Counts

Samples of sound changes on the basis of which the estimations of historical probabilities are performed were constructed from counts of occurrences and languages surveyed (from Kümmel's 2007 database). The count of sound-change occurrences was derived from the number of languages listed by Kümmel for each sound change. To avoid counting a single sound change that operates at a proto-stage and is reflected in several daughter languages as independent events, sound changes with exactly same outcome in closely related languages were counted as single

---

so results in devoicing, even postnasally. The strongest empirical evidence in favour of this assumption comes from Davidson (2016). In English, underlyingly *voiced* stops are voiceless (i.e. voicing ceases earlier than during the first 10% of closure duration) or partially voiced (i.e. voicing ceases earlier than during the first 90% of closure duration) in 22% of measured tokens in postnasal prevocalic position in Davidson (2016) (for a detailed discussion, see Beguš 2019).

<sup>12</sup> There is one possible case of final voicing that could count as a productive synchronic alternation, in Lakota (Rood 2016). As the only acoustic study of Lakota word-final stops (Blevins *et al.* 2020) deals primarily with voicing distributions, with very limited data on phonological alternations and no data on productivity of the alternations, I leave Lakota out of this discussion (for discussion, see Beguš 2018).

<sup>13</sup> Kiparsky's (2006) Scenario 2 also includes three sound changes, but the last sound change (apocope after a single consonant) is not attested in Hamed & Flavier (2009). Kümmel (2007) does not include vocalic changes, which is why the UniDia database is used. Because the last sound change is never attested in our surveys, I exclude Scenario 2 from the estimation of  $P_{\chi}(FV)$ .



alternation	sound change	count	surveyed
postnasal voicing	T > D / N __	28	294
intervocalic voicing	T > D / V __ V	38	294
final devoicing	D > T / __ #	24	294

*Table II*

Counts of sound changes in Kümmel (2007) for natural alternations.

events (as grouped together by Kümmel 2007). While it is possible that some dependencies still exist in the data, it is difficult to estimate their number. We assume that potential dependencies do not crucially affect the results.

If a sound change is reported to target a subset of the three major places of articulation (labial, dorsal, velar), rather than the entire set, the counts for alternations that require more than one sound change were multiplied by a coefficient that proportionally penalises the counts. For example, counts of sound changes that target only two places of articulation are multiplied by two-thirds in order to reduce the possibility of final estimated probabilities being inflated: if the first sound change targets two places of articulation and the second sound change targets the third place of articulation, such a combination would not result in an unnatural process.

Postnasal voicing as a sound change is reported in approximately 28 languages by Kümmel, and intervocalic voicing (including postvocalic voicing) in approximately 38. Final devoicing is reported in approximately 24 languages (summarised in Table II). For raw counts, see Appendix A: §1–§3. Postnasal voicing, intervocalic voicing and final devoicing that target a single series of stops are counted without penalisation, together with cases in which these sound changes target more than a single place of articulation.

The first sound change in the Blurring Process that results in postnasal devoicing, the fricativisation of voiced stops, is reported in approximately 66 languages. In 32 languages, the sound change is reported to target all three major places of articulation, in eleven languages the sound change targets two places of articulation and in 23 languages one place of articulation. The final count is thus  $\approx 47$  ( $32 + 11 \times \frac{2}{3} + 23 \times \frac{1}{3}$ ). Instances of intervocalic and postvocalic fricativisation are included in the count (not only cases with fricativisation in all but postnasal position), because the result of such fricativisation after the other two sound changes would also be a system analysed as postnasal devoicing.<sup>14</sup> The probability of the first of

<sup>14</sup> An alternation that resulted from a combination of sound changes in which the first sound change targeted postvocalic stops rather than non-postnasal stops and the other two above-mentioned sound changes would have the same result as in the attested case of postnasal devoicing, and would be analysed as postnasal devoicing with initial devoicing.

alternation	sound change	count	surveyed
postnasal devoicing	D > Z / [-nas]__ or V__(V)	47	294
	D > T	15	263
	Z > D	17	216
intervocalic devoicing	D > Z / V__(V)	42	294
	Z > S	5	216
	S > T	10	248
final voicing	T: > T / __#	3	294
	T > D / V__	23	294
	T: > T	21	≈88

*Table III*

Counts of sound changes in Kümmel (2007) for unnatural alternations.

the sound changes in the Blurring Process that results in postnasal devoicing is estimated on the basis of number of successes (languages in the survey with that sound change) and the total number of languages surveyed (294), without conditioning on the sample. The sample for estimating the probability of the first sound change is unconditioned because the historical probability of  $A_k$  is the probability of  $A_k$  arising in a language, regardless of the properties of its phonemic inventory (see §3.1.2). Once the first sound change has operated, however, we know that the language in question needs to have voiced stops in its inventory. The historical probability of the second sound change, which targets voiced stops, is therefore estimated from the number of successes (languages in the survey with that sound change) and the number of languages with voiced stops. The second change (D > T) is found in  $\approx 15$  ( $13 + 1 \times \frac{2}{3} + 3 \times \frac{1}{3}$ ) languages (including cases of devoicing that are the result of chain shifts). Approximately 31 languages lack voiced stops in Kümmel's survey, meaning that  $P_\chi$  is estimated on the basis of  $294 - 31 = 263$  languages surveyed.<sup>15</sup> After the two sound changes operate, we also know that the language L has voiced fricatives. The historical probability of the last sound change is estimated on the basis of the number of languages with occlusion of voiced fricatives and the number of languages with voiced fricatives (allophonic or phonemic). Approximately 216 languages have voiced (bi)labial, alveolar/dental or velar non-strident fricatives (including the labiodental voiced fricative /v/), according to Kümmel. In  $\approx 17$  ( $1 + 5 \times \frac{2}{3} + 38 \times \frac{1}{3}$ ) languages, occlusion of fricatives is reported as a sound change. The counts for intervocalic devoicing are performed in

<sup>15</sup> One language has only /b/ in its inventory of voiced stops. The low number of inventories that lack voiced stops might be influenced by the area that Kümmel (2007) surveys. In Moran *et al.* (2014), approximately 30% of inventories lack a phonemic labial voiced stop. For consistency, we do not go beyond Kümmel's survey.

the same manner as the counts for postnasal devoicing, and are given in Table III.

The historical probability of final voicing is estimated on the basis of the one scenario in Kiparsky (2006) that would result in an alternation involving final voicing. Scenarios that would lead to final voicing as a static phonotactic restriction and could involve fewer than three sound changes are excluded. There are three main reasons for distinguishing alternations from static phonotactic restrictions in a diachronic model (Beguš 2019), despite the two phenomena likely being part of the same synchronic grammatical mechanisms (Prince & Smolensky 1993, Hayes 2004, Pater & Tessier 2006). First, unnatural phonotactic restrictions provide considerably less reliable evidence for learners, because the evidence is distributional, rather than appearing within the same morphological unit. This means that it is more likely that a phonotactic process will not be acquired by the learners. Second, alternative analyses of data are often available in the case of phonotactic restrictions, but not for alternations, where evidence for a process comes from within the same morphological unit. Finally, typological surveys of phonotactic restrictions are considerably more difficult to carry out than typological surveys of alternations. In the absence of typological studies, it is difficult to evaluate the predictions that the channel bias model makes for phonotactic restrictions. In fact, final voicing as a phonotactic restriction might not be as rare as has been suggested, with at least two potential phonological systems attested in which voiceless stops do not surface word-finally, but voiced stops do (for example in Ho and some dialects of Spanish; see Beguš 2019).<sup>16</sup>

Counts of the sound changes that lead to final voicing as an alternation are as follows. In three languages, word-final geminates are reported to simplify to singleton stops. (This sound change is necessary if we want the scenario to result in an unnatural alternation, rather than a static phonotactic restriction.) Because this is the first in the series of changes and we do not condition  $P_x$  on any property of the language, as before, the historical probability is estimated from the total number of languages surveyed. The second sound change, postvocalic voicing of voiceless stops, is reported in approximately 23 languages (corrected for place of articulation). The intervocalic condition is excluded from the count, as voicing of intervocalic stops would not target final stops. Because all 294 languages surveyed have voiceless stops, they are all included in the count for estimating the historical probability of the second sound change.<sup>17</sup> Finally,

<sup>16</sup> The scenario that potentially results in final voicing in Lakota is also excluded: fricativisation of voiceless stops word-finally, followed by postvocalic voicing of fricatives and occlusion of fricatives to stops, would potentially result in final voicing. A preliminary estimation of this scenario shows that its historical probability would be very low, because the first sound change is relatively rare (reported only once for one place of articulation in Kümmel 2007).

<sup>17</sup> Since the survey in Kümmel (2007) does not provide the proportion of languages that feature voiceless stops postvocally, we include all languages in the sample. Given the syllabic structure of the languages in the survey, relatively few languages

simplification of geminates is reported in 21 languages. It is difficult to estimate how many languages in Kümmel (2007) allow geminate voiceless stops. While few languages have phonologically contrastive geminates, many more allow sequences of two identical stops at morpheme boundaries (so-called fake geminates; Oh & Redford 2012). To estimate the number of languages that allow such sequences, Greenberg's (1965) survey of consonantal clusters and Ryan's (2019) survey of phonemic geminates were used. At least 30% of languages in Greenberg's survey of approximately 100 languages allow stop–stop final clusters. The number of languages in our sample that allow homorganic stop–stop sequences can be approximated from the proportion of languages that allow phonemic geminates and from the proportion that allow sequences of stops. Languages that allow clusters of stops at morpheme boundaries should in principle allow clusters of homorganic stops: if geminate clusters were simplified, the sound change of simplification would be reported in our sample. The number is thus estimated at 88 (30% of 294 languages). That this estimate is accurate is suggested by a survey of phonemic geminates: Ryan 2019 estimates that approximately 35% of 55 genealogically diverse languages surveyed have phonemic geminates.

## 5 Results

### 5.1 Individual alternations

Table IV shows the Historical Probabilities with estimated 95%  $BC_a$  confidence intervals for the six natural and unnatural alternations discussed above. Figure 1 shows the distributions of bootstrap replicates for the historical probabilities ( $P_{\chi}$ ) of these alternations. The results show a substantial difference between the historical probabilities of the natural and unnatural groups. The model thus predicts that the unnatural alternations will be substantially less frequent than their natural counterparts.

### 5.2 Comparison of alternations

One of the advantages of the proposed model is that inferential statistics can be performed on the comparison between the historical probabilities of any two alternations. Significance testing is performed by estimating a difference between the historical probabilities of two alternations (see §3.1.3).

---

are expected to lack voiceless stops postvocally, but this has the potential to yield anti-conservative results in the estimation of final voicing. On the other hand, the fact that we do not limit the counts for the  $T > D / V \_$  sound change to those that specifically mention word-final position means that the outcomes might be conservative.

		95% BC <sub>a</sub> CI		95% profile CI	
A <sub>k</sub>	P <sub>χ</sub>	lower	upper	lower	upper
PNV	9.5	6.1	12.9	6.5	13.2
PND	0.01	0.006	0.02	—	—
IVV	12.9	9.2	16.7	9.4	17.1
IVD	0.002	0.001	0.007	—	—
FD	8.2	5.1	11.2	5.4	11.7
FV	0.003	0.001	0.01	—	—

Table IV

Estimated P<sub>χ</sub> (in %) for natural and unnatural alternations with 95% BC<sub>a</sub> confidence intervals. We also compute profile CIs from an empty logistic regression, for comparison. The largest difference between the confidence intervals is 0.5%, which suggests that the proposed model estimates CIs with a high degree of accuracy.

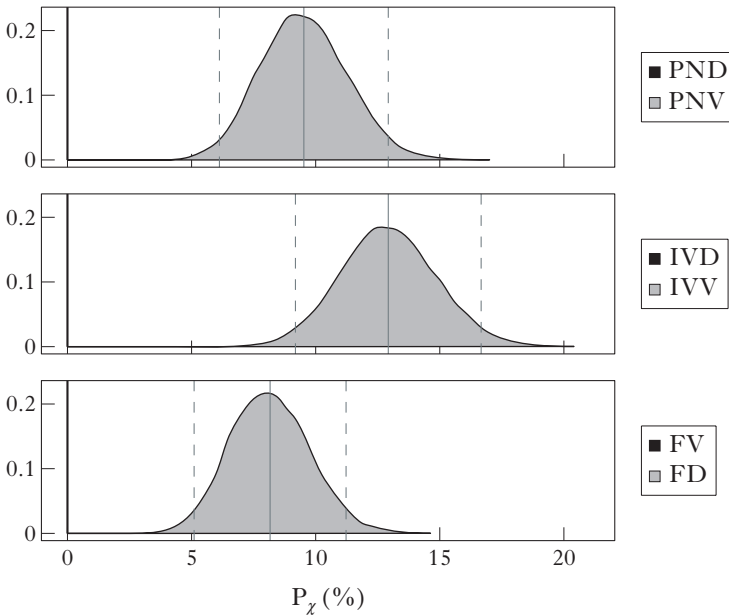


Figure 1

Bootstrap replicates for natural and unnatural alternations. The plots show the observed P<sub>χ</sub> (solid line) and the 95% BC<sub>a</sub> confidence intervals (CIs) (dashed line) for natural alternations (PNV = postnasal voicing, PND = postnasal devoicing, IVV = intervocalic voicing, IVD = intervocalic devoicing, FD = final devoicing, FV = final voicing). The vast majority of bootstrap replicates for unnatural alternations fall outside the limits of the plot.

The historical probabilities of all three natural alternations in Fig. 1 are significantly higher than those of their unnatural counterparts. Table V includes estimates and 95%  $BC_a$  confidence intervals of the difference in historical probabilities ( $\Delta P_\chi$ ) for each natural–unnatural alternation pair.

alternation pair	$\Delta P_\chi$	95% $BC_a$ CI		
		lower	upper	
PNV <i>vs.</i> PND	9.5	6.5	13.3	*
IVV <i>vs.</i> IVD	12.9	9.5	17.0	*
FD <i>vs.</i> FV	8.2	5.4	11.9	*

*Table V*

Estimated  $\Delta P_\chi$  (in %) for natural–unnatural alternation pairs with 95%  $BC_a$  confidence intervals. \* indicates significant differences.

We can also compare alternations within the unnatural group. Figure 2 shows bootstrap replicates of the individual historical probabilities of the three unnatural alternations. The figure shows that the historical probability of postnasal devoicing is higher than those of the other two unnatural alternations. By estimating the difference between two alternations, we can test, for example, whether  $P_\chi(\text{PND})$  and  $P_\chi(\text{IVD})$  or  $P_\chi(\text{PND})$  and  $P_\chi(\text{FV})$  are significantly different, as shown in (6).

- (6) a.  $\Delta P_\chi(\text{PND}, \text{IVD}) = P_\chi(\text{PND}) - P_\chi(\text{IVD}) = 0.010\% [0.003\%, 0.02\%]$   
 b.  $\Delta P_\chi(\text{PND}, \text{FV}) = P_\chi(\text{PND}) - P_\chi(\text{FV}) = 0.009\% [0.001\%, 0.02\%]$

Because the 95%  $BC_a$  CIs of the difference in historical probability between postnasal devoicing and final voicing and between postnasal devoicing and intervocalic devoicing lie above zero, it can be concluded that the historical probability of postnasal devoicing ( $P_\chi(\text{PND})$ ) is significantly higher than the historical probabilities of final voicing ( $P_\chi(\text{FV})$ ) and intervocalic devoicing ( $P_\chi(\text{IVD})$ ) (with  $\alpha = 0.05$ ).

The proposed technique makes some simplifying assumptions that introduce confounds to the estimation of historical probabilities (see §3.2, §3.3 and §4). Because differences in the historical probabilities among unnatural alternations are considerably smaller than differences between natural and unnatural pairs (Fig. 1), estimation of these differences is substantially more likely to be influenced by these confounds, and therefore to be less reliable. Until more comprehensive surveys are available, however, the proposed model makes, as far as I know, the most accurate approximations of historical probabilities of alternations, both for natural–unnatural alternation pairs and for alternations within the unnatural group.

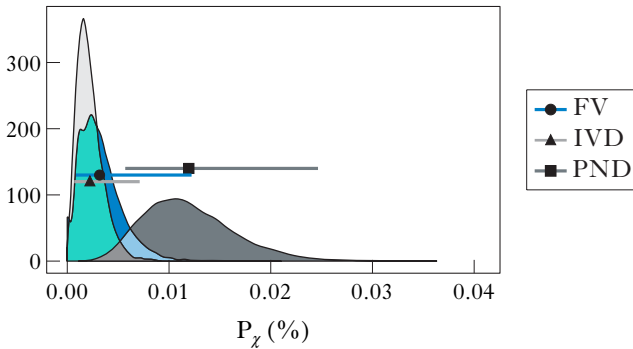


Figure 2

Bootstrap replicates unnatural alternations. The plots show the observed  $P_\chi$  (circle, triangle and square) and the 95%  $BC_a$  CI (solid lines).

### 5.3 Comparing historical probabilities with observed synchronic typology

We can evaluate the model's predictions by comparing historical probabilities with independently observed typology of synchronic alternations.<sup>18</sup> Table VI compares historical and observed synchronic probabilities. Historical probabilities are estimated as described above (see §4 and Table IV). The synchronic typology is estimated with a non-parametric bootstrap technique, in the same way as described in §3.1.1, except that the estimation is based on the numbers of languages in a sample with and without a particular synchronic alternation. The database used for estimating synchronic typology is PBase (Mielke 2019), which surveys 629 languages in total. Postnasal voicing is attested in 28 languages, intervocalic voicing in 51 and final devoicing in 31 (the languages on the basis of which the count was performed are listed in Appendix B). All three alternations were counted, even if they target only one place of articulation. Neither the historical (Kümmel 2007) nor the synchronic sample was constructed specifically for the purpose of establishing typology of processes that target feature [voice], making them less prone to biases.

The estimations of the synchronic typology of unnatural alternations can be computed (as in Table VI) from the surveys of unnatural processes in Beguš (2018, 2019) and Beguš & Nazarov (2018). Postnasal devoicing has been confirmed as a fully productive synchronic alternation in two related languages (Tswana and Shekgalagari) and as a morphophonological alternation in a number of others (e.g. Buginese and Nasioi; see Beguš 2019). For the purpose of comparison, only fully productive alternations

<sup>18</sup> Estimation of synchronic typological probabilities faces even more difficulties than estimation of historical probabilities, such as dependency of typological estimates on areal or historical factors. A comparison of historical probabilities and observed synchronic typology can only be qualitative at this point, especially until more comprehensive and well-balanced surveys are available.

$A_k$	$P_\chi$	95% $BC_a$ CI		typology	95% $BC_a$ CI	
		lower	upper		lower	upper
PNV	9.5	6.1	12.9	4.5	2.9	6.2
PND	0.01	0.006	0.02	0.5	0.0	1.2
IVV	12.9	9.2	16.7	8.1	6.0	10.2
IVD	0.002	0.001	0.007	0.2	0.0	0.5
FD	8.2	5.1	11.2	4.9	3.3	6.7
FV	0.003	0.001	0.01	0.0	0.0	0.0

*Table VI*

A comparison of historical probabilities ( $P_\chi$ ) and observed synchronic typology with 95%  $BC_a$  CIs for natural and unnatural processes.

were counted in the synchronic typology. Because Tswana and Shekgalari are closely related, postnasal devoicing here was counted as a single occurrence. Intervocalic devoicing is attested only once as a morphologically conditioned synchronic process (Bloyd 2015), although detailed descriptions are lacking. To the best of my knowledge, final voicing is not attested as a productive phonological alternation in any language, which is why its synchronic typological probability is estimated as less than  $P = 1/600$  (an approximate estimate of languages surveyed in these surveys of unnatural alternations is 600).<sup>19</sup>

Table VI and the corresponding plot of estimated historical and synchronic probabilities with 95%  $BC_a$  CIs in Fig. 3 suggest that the model correctly predicts natural alternations to be considerably more frequent than their unnatural alternations. Historical probabilities and observed synchronic typology also match to the extent that the 95%  $BC_a$  confidence intervals of both historical and synchronic typological probabilities overlap for the compared processes. It needs to be stressed here that, for unnatural processes, historical probabilities and observed synchronic typology are completely independent. In other words, the model estimates the probability of a combination of three sound changes, none of which is individually related to the unnatural synchronic alternation from which synchronic typological probabilities are estimated.

## 6 Implications

The technique employed here also helps in the identification of mismatches in predictions between the analytic bias and channel bias

<sup>19</sup> If we counted only the best candidate for final voicing, Lakota, as featuring fully productive unnatural alternations (Rood 2016, Blevins *et al.* 2020), the typological probabilities of final voicing would be estimated at  $P(1/600) = 0.17\%$ .



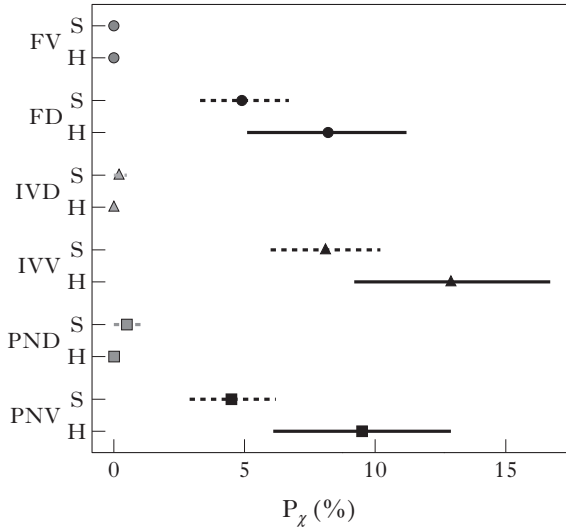


Figure 3

Observed historical (H; solid line) and synchronic (S; dashed line) probabilities (in %) with 95% BC<sub>a</sub> CIs from Table VI.

approaches to typology. If two typologically different alternations show no learnability differences, but have significantly different historical probabilities, it is reasonable to assume that the differences in the observed typology between the two alternations is influenced by the channel bias factor. On the other hand, if two typologically different alternations have the same historical probabilities and show differences in learnability, it is reasonable to assume that these differences result from the analytic bias factor.

The proposed model suggests that the observed rarity of unnatural alternations targeting feature [voice] is primarily influenced by the channel bias factor. The typology is predicted with relatively high accuracy (§5.3 and Fig. 3), whereas learning experiments found no differences between the natural and unnatural alternations for any of the three pairs (Seidl *et al.* 2007, Do *et al.* 2016, Glewwe 2017, Glewwe *et al.* 2018).

The model predicts not only that unnatural alternations will be rare (§4), but also that complex alternations will be less frequent than simple alternations. The minimality principle (cf. §2), which is at least a strong tendency, states that sound change is a change in one feature (or the deletion or reordering of feature matrices) in a given environment. This means that featurally complex alternations that change more than a single feature can only arise from the phonologisation of more than one sound change. Because the probability of a combination of more than one sound changes will be lower than the probability of one sound change, all else being equal, featurally complex alternations are predicted to be typologically less frequent within the channel bias approach. Exactly the same

sound change	alternation	$P_{\chi}$	lower	upper	F	$P_{\chi}$	$P_{\text{complex}}$
	<i>no alternation</i>	83.5	—	—	0		
D > Z / [-nas]__	D → Z / [-nas]__	16.0	11.9	20.1	1	↓	↓
D > T	Z → T / [+nas]__	0.5	0.3	0.8	2	↓	↓
Z > D	PND	0.01	0.006	0.02	1	↓	↑

*Table VII*

Mismatches in predictions (framed) between the channel bias approach ( $P_{\chi}$ ) and the complexity bias approach ( $P_{\text{complex}}$ ) for postnasal devoicing. The sound-change column represents the three sound changes from which the unnatural process postnasal devoicing results, and the alternation column represents the synchronic alternation after each of the three sound changes. The  $P_{\chi}$  column gives the estimated probability of each alternation with 95% BC<sub>a</sub> lower and upper 95% CIs. The features (F) column gives the number of features a learner has to learn for each synchronic alternation.

generalisation is, however, also predicted by the analytic bias approach to typology: numerous studies have confirmed that featurally complex alternations are consistently underlearned compared to featurally simple alternations (COMPLEXITY BIAS; Moreton & Pater 2012a, b).

There is a crucial mismatch between the predictions of the analytic bias and channel bias approaches with respect to unnatural alternations. The channel bias approach predicts that the more sound changes an alternation requires, the lower the historical probability of that alternation, regardless of its complexity (see Table VII). In other words, the prediction that complex alternations will be rare is violable: if the three sound changes of a Blurring Process result in a simple unnatural alternation, it will still be predicted that the simpler alternation will be less frequent than an unmotivated complex alternation, because the first requires three sound changes to arise and the latter only two (§2).

We can estimate the historical probabilities for each step in the Blurring Process that leads to unnatural alternations. Take as an example postnasal devoicing. The historical probabilities of each resulting alternation (after the first, second and third sound changes) were estimated as described in §3.<sup>20</sup> The  $P_{\chi}$  columns in Table VII illustrate that each additional sound change decreases the historical probability of the resulting alternation.

On the other hand, the analytic bias approach predicts that structurally more complex alternations will be typologically less frequent because they are more difficult to learn than structurally simple alternations (complexity bias has been confirmed in many studies, almost without exception; e.g. Moreton & Pater 2012a, b). While many criteria for complexity in phonological alternations can be invoked, I focus on the measure of complexity in

<sup>20</sup> The probability of the initial stage before the first sound change operates is calculated simply as  $1 - P_{\chi_{1,2,3}}$ , where  $P_{\chi_{1,2,3}}$  is the sum of the historical probabilities of the first, first and second, and all three sound changes.

Moreton *et al.* (2017), which accounts for learnability asymmetries in several earlier studies (for an overview, see Moreton & Pater 2012a, b), and is based on ‘concept learning’. The complexity of an alternation is derived primarily from the number of features manipulated by an alternation.

If we analyse each step in the Blurring Process in terms of such synchronic complexity, the first two sound changes indeed increase the complexity of the resulting alternation, but the third change decreases its complexity.<sup>21</sup> The alternation  $Z \rightarrow T / [+nas] \_$  manipulates two feature values,  $[\pm\text{continuant}]$  and  $[\pm\text{voice}]$ , while  $D \rightarrow T / [+nas] \_$  (postnasal devoicing) manipulates only  $[\pm\text{voice}]$ . From a phonological perspective, the first is more complex than the latter (Moreton *et al.* 2017).

Complexity bias thus predicts that the alternations that arise from the first and second sound changes in the Blurring Process will be increasingly rare, but predicts that the structurally simpler alternations resulting from the combination of all three sound changes will be more frequent than the complex alternation requiring only two sound changes.

The mismatched predictions illustrated in Table VII provide new information for disambiguating analytic bias and channel biases. The analytic bias–channel bias complexity mismatch can be directly evaluated against the observed typology: if unmotivated structurally complex alternations that require two sound changes are typologically more common than structurally simpler unnatural alternations, channel bias has to be the leading cause of this particular typological observation. If, on the other hand, structurally more complex unmotivated alternations that require two sound changes are typologically less frequent than what would be predicted by the channel bias approach, we have a strong case in favour of the analytic bias influence, and more precisely in favour of complexity bias within the analytic bias approach to typology.

In fact, typological observations suggest that the complex synchronic alternation  $Z \rightarrow T / [+nas] \_$  that results from the first two sound changes in a Blurring Process might be attested less frequently than would be predicted by channel bias, suggesting that complexity bias influences this distribution. The historical probability of  $Z \rightarrow T / [+nas] \_$  is significantly higher than the historical probability of postnasal devoicing. The difference is estimated at  $\Delta P_x(Z \rightarrow T / [+nas] \_) = 0.4\%$ ,  $[0.2\%, 0.8\%]$ . In other words, the historical probability of the alternation

<sup>21</sup> The fact that the first two sound changes in the Blurring Process occur relatively frequently, despite increasing the complexity of the alternations, argues against the radical approach to the analytic bias–channel bias conflation problem that states that sound-change probabilities are primarily influenced by learnability, and hence that estimated channel bias influences are crucially conflated with analytic bias influences. If anything, analytic bias influences would militate against the first two sound changes operating in combination, because the resulting alternations would be more difficult to learn. Because the Blurring Process does occur, it means that the driving force behind the sound changes in question is not crucially influenced by analytic bias (although analytic bias can of course still influence the relative frequencies of sound change).

$Z \rightarrow T / [+nas] \_$  arising through two sound changes is predicted to be approximately fifty times as frequent as the historical probability of postnasal devoicing (see Table VII). Surface synchronic typology, however, does not conform to this generalisation.

A system in which postnasal devoiced stops contrast with voiced fricatives elsewhere (a complex alternation that arises via the combination of two sound changes) is synchronically confirmed in Konyagi, Punu, Pedi, Sie and potentially Nasioi (Dickens 1984, Santos 1996, Hyman 2001, Merrill 2014, 2016a, b, Brown 2017).<sup>22</sup> Other languages are more difficult to classify, because some of them appear to feature full postnasal devoicing only for a subset of places of articulation. While  $Z \rightarrow T / [+nas] \_$  indeed appears to be more frequent than postnasal devoicing, the magnitude of the difference is smaller than predicted by the channel bias.

Even more intriguing is the high frequency at which the third sound change in the Blurring Process, the occlusion of voiced fricatives to stops ( $Z > D$ ), operates on synchronic systems that feature the alternation  $Z \rightarrow T / [+nas] \_$  (after the first two changes). The historical probability of the third sound change that leads to postnasal devoicing, occlusion of voiced fricatives for languages that have voiced fricatives in the system, estimated independently of the Blurring Process (i.e. estimated from an unconditioned diachronic sample), is  $P_x(Z > D) = 20.4\%$  [14.8%, 25.5%]. Of the languages in the survey in Beguš (2019) that undergo the first two sound changes in the Blurring Process, which leads to postnasal devoicing, six (out of ten) feature occlusion of stops for at least one place of articulation or in at least one position in the word.<sup>23</sup> If we take only cases in which the occlusion of fricatives targets more than two places of articulation, only Tswana, Shekgalagari, Makuwa and Murik would count. It does appear, however, that the occlusion of voiced fricatives in a synchronic system that undergoes the first two sound changes is more frequent than the model predicts for the occlusion of voiced fricatives in general.

To test the hypothesis that the last sound change operates with higher frequency than would be predicted by only the channel bias approach, we can compare the unconditioned historical probability of the occlusion of fricatives with the historical probability of the occlusion of fricatives

<sup>22</sup> We count Kutswe and Pulana together with Pedi, because they are closely related. Even if we counted them separately (Kutswe and Pulana as one language), the distribution is still significant ( $p = 0.016$ ). Punu is a language that undergoes a different development from the one described in §2. In Punu, the resulting alternation is not postnasal devoicing, but a complex alternation between voiceless stops postnasally and voiced fricatives elsewhere. For a discussion, see Hyman (2001).

<sup>23</sup> Postnasal devoicing occurrences in Tswana, Shekgalagari and Makuwa are counted as only one occurrence. South Italian dialects that devoice affricates are not counted. I also exclude Mpongwe from the count because of the limited description and marginal status of postnasal devoicing there. I include Pedi, which features  $Z \rightarrow T / [+nas] \_$  (Dickens 1984), and Sie based on counts of the synchronic database of phonological rules in Mielke (2019).

in those languages that have already undergone the first two sound changes in the Blurring Process that lead to postnasal devoicing. In other words, we compare the probability of the occlusion of fricatives regardless of whether it simplifies the alternation (assuming only channel bias influences) with the probability of the occlusion of fricatives operating in the Blurring Process, where it simplifies the alternation and consequently its learnability. Counts for the unconditioned historical probability of the occlusion of fricatives are based on the survey of sound changes in Kümmel (2007). 44 languages with voiced fricatives (out of 216 surveyed) undergo the occlusion of voiced fricatives. As already mentioned, under the less conservative count, six out of ten languages with occlusion and fricative devoicing show occlusion for at least one place of articulation or for at least one context (word-initially in Nasioi).<sup>24</sup> The difference between the two counts is statistically significant ( $p = 0.009$ ; Fisher's Exact Test). This means that the last sound change in the Blurring Process that decreases the complexity of the resulting alternation operates at significantly higher rates than would be predicted if we assumed only channel bias influences.<sup>25</sup>

This suggests that the high rate of occurrence of the third sound change in the Blurring Process (in the case of postnasal devoicing, the occlusion of fricatives) is likely an influence of complexity bias within the analytic bias approach. While analytic bias probably does not crucially influence the probabilities of the first two sound changes in the Blurring Process in the direction that interests us, because they increase complexity and therefore would be predicted to reduce learnability, it is likely that the frequency of the third sound change, and therefore the lower probability of the more complex unmotivated alternation, is influenced precisely by complexity bias.<sup>26</sup> This paper has identified and described one such instance; investigation of further such cases should yield a better understanding on how learnability and sound-change frequency interact.

## 7 Conclusion

This paper has proposed a technique for estimating channel bias influences on phonological typology using the statistical technique of bootstrapping. Historical probabilities of alternations were estimated on the basis of two diachronic factors: the number of sound changes required for an alternation to arise and their respective probabilities.

The model was applied to six natural and unnatural alternations targeting the feature [voice]. It can (i) estimate the historical probability of any

<sup>24</sup> Cases with variation are counted as involving the sound change.

<sup>25</sup> This is exactly the opposite of what is proposed by Kiparsky (2008), who claims that the sound change that would result in an unnatural alternation would be blocked by the grammar.

<sup>26</sup> If both approaches make the same predictions, it is difficult to distinguish between the two. We are primarily interested in mismatches in predictions, because we can then test the hypothesis against the observed typology.

synchronic alternation, both attested and unattested, (ii) compare the historical probabilities of two alternations and perform inferential tests on the comparison and (iii) compare the historical probabilities to independently observed synchronic typology, in order to evaluate the channel bias influences on typology. Finally, we identified mismatches in predictions between the analytic bias and channel bias approaches, which yields new insights into the discussion of different influences on phonological typology. The results suggest that the typological difference between natural and unnatural alternations targeting the feature [voice] is primarily due to channel bias, but that the relatively low frequency of complex alternations and the higher rate of operation of sound changes that simplify an alternation are due to analytic bias.

These conclusions have direct theoretical implications. Synchronic grammar should ideally derive all observed patterns, and at the same time exclude impossible processes. Typological observations often prompt adjustments in grammar design. The proposed framework suggests that some typological gaps are historical accidents that need not be encoded in synchronic grammars, and quantifies these gaps. On the other hand, this paper has also suggested that some typological observations, such as the avoidance of complex alternations, cannot be explained only within the channel bias approach, and that these preferences should indeed be encoded in synchronic grammar. The paper limits the application of the technique to six alternations targeting feature [voice]. Estimation of channel bias and analytic bias influences should be performed on further alternations, in order to gain a better understanding of which observations result from constraints in synchronic grammar and which from diachronic development.

## REFERENCES

- Barnes, Jonathan (2002). *Positional neutralization: a phonologization approach to typological patterns*. PhD dissertation, University of California, Berkeley.
- Becker, Michael, Nihan Ketrez & Andrew Nevins (2011). The surfeit of the stimulus: analytic biases filter lexical statistics in Turkish laryngeal alternations. *Lg* **87**. 84–125.
- Beguš, Gašper (2018). *Unnatural phonology: a synchrony–diachrony interface approach*. PhD dissertation, Harvard University.
- Beguš, Gašper (2019). Post-nasal devoicing and the Blurring Process. *ŹL* **55**. 689–753.
- Beguš, Gašper & Aleksei Nazarov (2018). Gradient trends against phonetic naturalness: the case of Tarma Quechua. *NELS* **48:1**. 73–86.
- Bell, Alan (1970). *A state-process approach to syllabicity and syllabic structure*. PhD dissertation, Stanford University.
- Bell, Alan (1971). Some patterns of occurrence and formation of syllable structures. *Stanford Working Papers on Language Universals* **6**. 23–137.
- Blevins, Juliette (2004). *Evolutionary Phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.
- Blevins, Juliette (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* **32**. 117–166.
- Blevins, Juliette (2007). The importance of typology in explaining recurrent sound patterns. *Linguistic Typology* **11**. 107–113.

- Blevins, Juliette (2008a). Natural and unnatural sound patterns: a pocket field guide. In Klaas Willems & Ludovic De Cuyper (eds.) *Naturalness and iconicity in language*. Amsterdam & Philadelphia: Benjamins. 121–148.
- Blevins, Juliette (2008b). Consonant epenthesis: natural and unnatural histories. In Jeff Good (ed.) *Linguistic universals and language change*. Oxford: Oxford University Press. 79–107.
- Blevins, Juliette (2015). Evolutionary Phonology: a holistic approach to sound change typology. In Patrick Honeybone & Joseph C. Salmons (eds.) *The Oxford handbook of historical phonology*. Oxford: Oxford University Press. 485–500.
- Blevins, Juliette, Ander Egurtzegi & Jan Ullrich (2020). Final obstruent voicing in Lakota: phonetic evidence and phonological implications. *Lg* **96**. 294–337.
- Bloyd, Tobias (2015). Toward a phonological reconstruction of Proto-Sula. *University of Hawai'i at Mānoa Working Papers in Linguistics* **46:8**. 1–23.
- Blust, Robert (2005). Must sound change be linguistically motivated? *Diachronica* **22**. 219–269.
- Bouchard-Côté, Alexandre, David Hall, Thomas L. Griffiths & Dan Klein (2013). Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences* **110**. 4224–4229.
- Broselow, Ellen (2018). Laryngeal contrasts in second language phonology. In Larry M. Hyman & Frans Plank (eds.) *Phonological typology*. Berlin: de Gruyter Mouton. 312–340.
- Brown, Jason (2017). Post-nasal devoicing in Nasioi. *Oceanic Linguistics* **56**. 267–277.
- Buckley, Eugene (2000). On the naturalness of unnatural rules. *UCSB Working Papers in Linguistics* **9**. 16–29.
- Bybee, Joan (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* **14**. 261–290.
- Canty, Angelo & Brian Ripley (2016). *Boot: bootstrap functions*. R package version 1.3-18. [cran.r-project.org/web/packages/boot/boot.pdf](https://cran.r-project.org/web/packages/boot/boot.pdf).
- Carpenter, Angela C. (2006). *Acquisition of a natural versus an unnatural stress system*. PhD dissertation, University of Massachusetts, Amherst.
- Carpenter, Angela C. (2010). A naturalness bias in learning stress. *Phonology* **27**. 345–392.
- Cathcart, Chundra A. (2015). A probabilistic model of Evolutionary Phonology. *NELS* **45:1**. 145–150.
- Clark, Eve V. & Melissa Bowerman (1986). On the acquisition of final voiced stops. In Joshua A. Fishman, Andrée Tabouret-Keller, Michael Clyne, Bh. Krishnamurti & Mohamed Abdulaziz (eds.) *The Fergusonian impact: in honor of Charles A. Ferguson on the occasion of his 65th birthday*. Vol. 1: *From phonology to society*. Berlin: De Gruyter. 51–68.
- Davidson, Lisa (2016). Variability in the implementation of voicing in American English obstruents. *JPh* **54**. 35–50.
- Davison, A. C. & D. V. Hinkley (1997). *Bootstrap methods and their applications*. Cambridge: Cambridge University Press.
- de Lacy, Paul & John Kingston (2013). Synchronic explanation. *NLLT* **31**. 287–355.
- Dickens, Patrick (1984). The history of so-called strengthening in Tswana. *Journal of African Languages and Linguistics* **6**. 97–125.
- Do, Youngah, Elizabeth Zsiga & Jonathan Havenhill (2016). Naturalness and frequency in implicit phonological learning. Paper presented at the 90th Annual Meeting of the Linguistic Society of America, Washington, D.C. Slides available (December 2020) at <https://jhavenhill.com/presentations/DoEtAILSA90.pdf>.
- Donegan, Patricia J. & David Stampe (1979). The study of natural phonology. In Daniel A. Dinnsen (ed.) *Current approaches to phonological theory*. Bloomington: Indiana University Press. 126–173.
- Efron, Bradley (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics* **7**. 1–26.

- Efron, Bradley (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association* **82**. 171–185.
- Efron, Bradley & Robert J. Tibshirani (1994). *An introduction to the bootstrap*. Boca Raton: Chapman Hall/CRC.
- Garrett, Andrew (2015). Sound change. In Claire Bowerman & Bethwyn Evans (eds.) *The Routledge handbook of historical linguistics*. London & New York: Routledge. 227–248.
- Garrett, Andrew & Keith Johnson (2013). Phonetic bias in sound change. In Yu (2013a). 51–97.
- Glewwe, Eleanor (2017). Substantive bias in phonotactic learning: positional extension of an obstruent voicing contrast. Paper presented at the 53rd Annual Meeting of the Chicago Linguistic Society. Handout available (December 2020) at <https://erglewwe.files.wordpress.com/2018/02/cls-handout.pdf>.
- Glewwe, Eleanor, Jesse Zymet, Jacob Adams, Rachel Jacobson, Anthony Yates, Ann Zeng & Robert Daland (2018). Substantive bias and word-final voiced obstruents: an artificial grammar learning study. Paper presented at the 92nd Annual Meeting of the Linguistic Society of America, Salt Lake City.
- Greenberg, Joseph H. (1965). Some generalizations concerning initial and final consonant sequences. *Linguistics* **18**. 5–34.
- Greenberg, Joseph H. (1978). Diachrony, synchrony, and language universals. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.) *Universals of human language*. Vol. 1: *Method and theory*. Stanford: Stanford University Press. 61–92.
- Hamed, Mahé Ben & Sébastien Flavier (2009). UNIDIA: a database for deriving diachronic universals. In Monique Dufresne, Fernande Dupuis & Etleva Vocaj (eds.) *Historical linguistics 2007: selected papers from the 18th International Conference on Historical Linguistics, Montreal, 6–11 August 2007*. Amsterdam & Philadelphia: Benjamins. 259–268.
- Hay, Jennifer B., Janet B. Pierrehumbert, Abby J. Walker & Patrick LaShell (2015). Tracking word frequency effects through 130 years of sound change. *Cognition* **139**. 83–91.
- Hayes, Bruce (2004). Phonological acquisition in Optimality Theory: the early stages. In René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in phonological acquisition*. Cambridge: Cambridge University Press. 158–203.
- Hruschka, Daniel J., Simon Branford, Eric D. Smith, Jon Wilkins, Andrew Meade, Mark Pagel & Tanmoy Bhattacharya (2015). Detecting regular sound changes in linguistics as events of concerted evolution. *Current Biology* **25**. 1–9.
- Hyman, Larry M. (1975). *Phonology: theory and analysis*. New York: Holt, Rinehart & Winston.
- Hyman, Larry M. (1976). Phonologization. In Alphonse Juillard (ed.) *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*. Vol. 2: *Phonology*. Saratoga: Anma Libri. 407–418.
- Hyman, Larry M. (2001). On the limits of phonetic determinism in phonology: \*NC revisited. In Elizabeth Hume & Keith Johnson (eds.) *The role of speech perception in phonology*. San Diego: Academic Press. 141–185.
- Kapatsinski, Vsevolod (2011). Modularity in the channel: the link between separability of features and learnability of dependencies between them. In Lee & Zee (2011). 1022–1025.
- Kiparsky, Paul (1995). The phonological basis of sound change. In John A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 640–670.
- Kiparsky, Paul (2006). The amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics* **32**. 217–236.



- Kiparsky, Paul (2008). Universals constrain change; change results in typological generalizations. In Jeff Good (ed.) *Linguistic universals and language change*. Oxford: Oxford University Press. 23–53.
- Kirby, James & Morgan Sonderegger (2013). A model of population dynamics applied to phonetic change. In Markus Knauff, Michael Pauen, Natalie Sebanz & Ipke Wachsmuth (eds.) *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Austin: Cognitive Science Society. 776–781.
- Kirby, James & Morgan Sonderegger (2015). Bias and population structure in the actuation of sound change. <http://arxiv.org/pdf/1507.04420.pdf>.
- Kong, Eun Jong, Mary E. Beckman & Jan Edwards (2012). Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: the cases of Greek and Japanese. *JPh* **40**. 725–744.
- Kümmel, Martin Joachim (2007). *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert.
- Kuo, Li-Jen (2009). The role of natural class features in the acquisition of phonotactic regularities. *Journal of Psycholinguistic Research* **38**. 129–150.
- Labov, William (1994). *Principles of linguistic change*. 2 vols. Oxford & Cambridge, Mass.: Blackwell.
- Lee, Wai-Sum & Eric Zee (eds.) (2011). *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong 2011*. Hong Kong: City University of Hong Kong.
- Liljencrants, Johan & Björn Lindblom (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Lg* **48**. 839–862.
- Lindblom, Björn (1986). Phonetic universals in vowel systems. In John J. Ohala & Jeri J. Jaeger (eds.) *Experimental phonology*. Orlando: Academic Press. 13–44.
- Lindblom, Björn (1990). Explaining phonetic variation: a sketch of the H&H theory. In William J. Hardcastle & Alain Marchal (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer. 403–439.
- Lipp, Rainer (2016). Final stops in Indo-European: their phonological classification as a key to the Proto-Indo-European root structure constraint. *Slovo a Slovesnost* **77**. 251–299.
- Merrill, John (2014). A historical account of the Fula and Sereer consonant mutation and noun class systems. Ms, University of California, Berkeley.
- Merrill, John (2016a). Consonant mutation and initial prominence: the historical loss of lexical contrastiveness. Paper presented at the 90th Annual Meeting of the Linguistic Society of America, Washington, D.C. Slides available (November 2020) at [https://drive.google.com/file/d/1Bbr\\_JfjlcnaW5oUiOKOtYal7bHzMzvm/view](https://drive.google.com/file/d/1Bbr_JfjlcnaW5oUiOKOtYal7bHzMzvm/view).
- Merrill, John (2016b). Konyagi post-nasal devoicing? Ms, University of California, Berkeley.
- Mielke, Jeff (2019). PBase: a database of phonological patterns. Available (December 2020) at <https://pbase.phon.chass.ncsu.edu>.
- Moran, Steven, Daniel McCloy & Richard Wright (2014). *PHOIBLE online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available (December 2020) at <https://phoible.org>.
- Moreton, Elliott (2008). Analytic bias and phonological typology. *Phonology* **25**. 83–127.
- Moreton, Elliott (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* **67**. 165–183.
- Moreton, Elliott & Joe Pater (2012a). Structure and substance in artificial-phonology learning. Part 1: Structure. *Language and Linguistics Compass* **6**. 686–701.
- Moreton, Elliott & Joe Pater (2012b). Structure and substance in artificial-phonology learning. Part 2: Substance. *Language and Linguistics Compass* **6**. 702–718.
- Moreton, Elliott, Joe Pater & Katya Pertsova (2017). Phonological concept learning. *Cognitive Science* **41**. 4–69.

- Morley, Rebecca L. (2015). Can phonological universals be emergent? Modeling the space of sound change, lexical distribution, and hypothesis selection. *Lg* **91**. e40–e70.
- Myers, Scott (2002). Gaps in factorial typology: the case of voicing in consonant clusters. Ms, University of Texas, Austin. Available as ROA-509 from the Rutgers Optimality Archive.
- Oh, Grace E. & Melissa A. Redford (2012). The production and phonetic representation of fake geminates in English. *JPh* **40**. 82–91.
- Ohala, John J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick & M. F. Miller (eds.) *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistic Society. 178–203.
- Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.) *The production of speech*. New York: Springer. 189–216.
- Ohala, John J. (1993). The phonetics of sound change. In Charles Jones (ed.) *Historical linguistics: problems and perspectives*. London & New York: Longman. 237–278.
- Ohala, John J. (2011). Accommodation to the aerodynamic voicing constraint and its phonological relevance. In Lee & Zee (2011). 64–67.
- Pater, Joe & Anne-Michelle Tessier (2006). L1 phonotactic knowledge and the L2 acquisition of alternations. In Roumyana Slabakova, Silvina A. Montrul & Philippe Prévost (eds.) *Inquiries in linguistic development: studies in honor of Lydia White*. Amsterdam & Philadelphia: Benjamins. 115–131.
- Picard, Marc (1994). *Principles and methods in historical phonology: from Proto-Algonkian to Arapaho*. Montreal: McGill-Queen's University Press.
- Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.
- Pycha, Anne, Pawel Nowak, Eurie Shin & Ryan Shosted (2003). Phonological rule-learning and its implications for a theory of vowel harmony. *WCCFL* **22**. 423–435.
- R Core Team (2016). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Rafferty, Anna N., Thomas L. Griffiths & Marc Ettliger (2013). Greater learnability is not sufficient to produce cultural universals. *Cognition* **129**. 70–87.
- Rood, David (2016). The phonology of Lakota voiced stops. In Catherine Rudin & Bryan J. Gordon (eds.) *Advances in the study of Siouan languages and linguistics*. Berlin: Language Science Press. 233–255.
- Ryan, Kevin M. (2019). *Prosodic weight: categories and continua*. Oxford: Oxford University Press.
- Santos, Rosine (1996). *Le Mey: langue ouest-atlantique de Guinée*. PhD dissertation, Université Paris III.
- Seidl, Amanda, Eugene Buckley & Alejandrina Cristia (2007). Complexity trumps naturalness. Paper presented at the 81st Annual Meeting of the Linguistic Society of America, Anaheim. Slides available (December 2020) at <https://web.ics.purdue.edu/~aseidl/SeidlBuckleyCristia.pdf>.
- Skoruppa, Katrin & Sharon Peperkamp (2011). Adaptation to novel accents: feature-based learning of context-sensitive phonological regularities. *Cognitive Science* **35**. 348–366.
- Wedel, Andrew (2012). Lexical contrast maintenance and the organization of sub-lexical contrast systems. *Language and Cognition* **4**. 319–355.
- Wedel, Andrew, Abby Kaplan & Scott Jackson (2013). High functional load inhibits phonological contrast loss: a corpus study. *Cognition* **128**. 179–186.
- Westbury, John R. & Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *JL* **22**. 145–166.
- White, James (2017). Accounting for the learnability of saltation in phonological theory: a maximum entropy model with a P-map bias. *Lg* **93**. 1–36.

- Wilson, Colin (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* **30**. 945–982.
- Yu, Alan C. L. (2004). Explaining final obstruent voicing in Lezgian: phonetics and history. *Lg* **80**. 73–97.
- Yu, Alan C. L. (2011). On measuring phonetic precursor robustness: a response to Moreton. *Phonology* **28**. 491–518.
- Yu, Alan C. L. (ed.) (2013a). *Origins of sound change: approaches to phonologization*. Oxford: Oxford University Press.
- Yu, Alan C. L. (2013b). Individual differences in socio-cognitive processing and the actuation of sound change. In Yu (2013a). 201–227.