Bootstrapping Sound Changes*

Gašper Beguš
University of Washington

Abstract

This paper presents a new technique for estimating the influences of channel bias on phonological typology called *Bootstrapping Sound Changes* (BSC). The BSC technique 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. With the BSC technique, we can estimate Historical Probabilities of attested and unattested alternations, compare Historical Probabilities of alternations and perform inferential statistics on the comparison, and compare outputs of the diachronic model against the independently observed synchronic typology to evaluate the performance of the Channel Bias approach. The BSC technique also identifies mismatches in 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.

1 Introduction

Typological literature in phonology has long revolved around the question of which factors influence the observed typology. Two major lines of thought emerge in this discussion: the Analytic Bias (AB) approach and the Channel Bias (CB) approach (Moreton 2008, Yu 2013).[1] The AB approach argues that the observed typology results primarily from differences in the learnability of phonological processes; the CB approach argues that the inherent directionality of sound changes based on phonetic precursors (articulatory and perceptual) results in typology (for further discussion, see Hyman 1975, 2001; Greenberg 1978; Ohala 1981, 1983, 1993; Kiparsky 1995, 2006, 2008; Hayes 1999; Tesar and Smolensky 2000; Blevins 2004, 2006, 2007, 2008a,b; Wilson 2006; Zuraw 2007; Hansson 2008; Moreton 2008; Hayes et al. 2009; Becker et al. 2011; Moreton and Pater 2012a,b; Moreton 2012; de Lacy and Kingston 2013; Garrett and Johnson 2013; Yu 2013; Hayes and White 2013; Cathcart 2015; Kirby and Sonderegger 2015; Greenwood 2016; White 2017; i.a.).

Empirical evidence often supports both approaches equally well. Typologically frequent processes are often shown to directly result from phonologization 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, which supports the CB approach (cf. Hyman 1975; Greenberg 1978; Ohala 1981, 1983, 1993; Lindblom 1986; Barnes 2002; Blevins 2004, 2006, 2007, 2008a,b; Morley 2012; see also Hansson 2008 and Garrett and Johnson

*I would like to thank Kevin Ryan, Jay Jasanoff, Adam Albright, Donca Steriade, Edward Flemming, Patrick Mair, Morgan Sonderegger, 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.

[1]Other names have been used for the two approaches, such as Evolutionary Phonology versus Amphichronic Phonology in Blevins (2004) and Kiparsky (2006, 2008).

2013 for an overview of the literature). On the other hand, typologically rare processes are experimentally shown to be more difficult to learn, which supports the AB approach (Hayes 1999; Tesar and Smolensky 2000; Kiparsky 1995, 2006, 2008; Wilson 2006; Hayes et al. 2009; Becker et al. 2011; de Lacy and Kingston 2013; Hayes and White 2013; White 2017; for an overview of the experimental AB literature, see Moreton and Pater 2012a,b). This ambiguity of evidence poses the primary challenge in typological research. The stance of this paper is that both factors influence the typology (as has been argued by a mounting body of research recently; Hyman 2001, Myers 2002, Moreton 2008, Moreton and Pater 2012a,b, de Lacy and Kingston 2013). There is no doubt that all potential synchronic phonological processes need to be learnable and accommodated by the synchronic grammar and that all synchronic processes arise through some diachronic trajectory. The question that this paper addresses is whether observed typological distributions are influenced primarily by different degrees of the learnability of different processes or primarily by different diachronic trajectories that underlie different processes. The role of phonological research is to quantitatively evaluate which aspects of typology are more likely to result from one factor or the other. For this purpose, we first need detailed quantitative models of both the AB and CB approaches to typology. This paper presents a new technique for deriving typology within the CB approach called *Bootstrapping Sound Changes*.

1.1 Analytic Bias

As already mentioned, the AB approach substantiates the claim that learning biases influence typology with evidence from artificial grammar learning experiments. If typologically infrequent processes are experimentally shown to be more difficult to learn than typologically frequent processes (for an overview, see Moreton and Pater 2012a,b), a reasonable conclusion would be that typological observations result precisely from these differences in learnability.

A challenge that the AB approach faces is that artificial grammar learning experiments testing the learnability of typologically rare or nonexistent unnatural processes frequently fail to show learnability differences compared to typologically frequent natural processes when structural complexity of the tested alternations is controlled for. Influences of Analytic Bias can be subdivided into Substantive Bias and Complexity Bias (Wilson 2006, Moreton 2008, Moreton and Pater 2012a,b). Substantive Bias states that phonetically motivated processes are easier to learn than unmotivated (or unnatural). Complexity Bias[2] states that 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 and Pater (2012a,b) shows that there exist consistent differences in experimental results testing the two biases. While Complexity Bias is consistently confirmed by the majority of studies surveyed, experimental outcomes of the Substantive Bias are mixed. Several studies that test the learning of unnatural alternations as defined in Section 2 found no effect of Substantive Bias (Pycha et al. 2003, Kuo 2009, Skoruppa and Peperkamp 2011, via Moreton and Pater 2012a,b; and more recently Seidl et al. 2007, Do et al. 2016, Glewwe 2017, Glewwe et al. 2018). A comparatively smaller subset of studies, however, do report positive results (Carpenter 2006, 2010; Wilson 2006).

The stance of this paper is that learnability differences confirmed for a subset of alternations should not be extended to the entire typology. Instead, learnability differences should be experimentally tested for each alternation discussed in the AB-CB debate. Experiments that specifically tested the processes discussed in this paper failed to find learning differences between the natural and unnatural processes (Seidl et al. 2007, Do et al. 2016, Glewwe 2017, Glewwe et al. 2018), although there exist substantial typological differences between the two groups. Deriving substantial

[2]Complexity Bias has also been called Structural Bias (Moreton and Pater 2012a,b).

typological differences between natural and unnatural alternations based on the AB approach is problematic for processes for which no differences in learning are observed experimentally.

L1 acquisition and L2 acquisition of word-final stops by speakers of L1s that ban obstruent codas is the only place 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 (overview in Broselow 2018; see also Clark and Bowerman 1986, Kong et al. 2012, and literature therein). It is likely, however, that this type of experiment tests differences in learning of more complex versus less complex articulations (Kong et al. 2012), and not the abstract phonological learning that is, for example, observed 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 and Pater 2012a,b). Articulation of segments that require more articulatory effort in a given position is expected to be learned less successfully: "[c]ross-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 very same mechanism is in fact responsible for final voicing within the CB approach: even adult L1 speakers with full contrast gradiently and passively devoice final voiced stops 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 modeled as a CB influence. It is in fact not trivial to show how differences in L1 articulatory learning would result in phonological typology (cf. Rafferty et al. 2011), given that children reproduce their input with a high degree of faithfulness past some developmental stage (e.g. at about 2–5 years for acquisition of the voicing contrast; see Kong et al. 2012).

1.2 Channel Bias

One of the objections against the CB approach to typology is that it fails to explain why some processes are unattested (Kiparsky 2006, 2008; de Lacy and 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 (2006) invokes grammatical constraints and learnability to explain these typological gaps.

Most of the current models of typology within the CB approach are indeed incapable of 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 CB approach has long been a qualitative observation that common processes are frequent because they are produced by frequent sound changes (Blevins 2013:485, also Greenberg 1978:75–6). Such reasoning does not provide sufficient outputs for a quantitative comparison of different influences on phonological typology.

Despite these objections, mechanisms exist within the CB approach to derive typology beyond the simple statement that rare sound changes produce rare alternations. Based on a typological study of an unnatural process, post-nasal devoicing, Author (2017) argues that unnatural processes require at least three sound changes to arise (as opposed to at least two for unmotivated processes and at least one for natural processes, the so-called Minimal Sound Change Requirement (MSCR)), which explains the relative rarity of processes with different degrees of naturalness. To be sure, the idea that unmotivated processes are rare because they require a complex history is not new

(Bell 1970, 1971; Greenberg 1978:75–6; Cathcart 2015; Morley 2015), but the MSCR is, to the author's knowledge, the first formal proof that explains why unnatural processes are the least frequent (compared to natural or unmotivated processes, see Section 2). The MSCR on its own, however, does not explain why some unnatural processes are attested, while others are not. To quantify the CB influences on typology further, the concept of the MSCR should 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 transitional probabilities from one state to another and the rest probabilities of each state, and therefore is most suitable for modeling the probabilities of various phonotactic restrictions. The probability of each state is determined by the number of previous states from which it can arise and the transitional probabilities between the states. They propose a Markov chain model for determining the probabilities of each state. Modeling the probabilities of transitions (processes) in the instantiation of the model in Bell (1971) involves relative probabilities that only tangentially reflect the frequencies of the processes in the samples. While the main principles of Bell's (1970, 1971) and Greenberg's (1978) model are similar to what will be proposed in this paper, focusing the estimation of probabilities on sound changes, rather than on states, combined with using substantially more elaborate samples of sound changes in our model yields more accurate predictions (Section 4.3). A different model of 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 get an estimate of the probability of certain processes. Due to its design, however, Cathcart's (2015) model needs to rely on representativeness of diachronic surveys for all sound changes, not only for the ones that are estimated (see also Section 3.3) and is computationally demanding, which makes the model difficult to implement. The models in Greenberg (1978) and Cathcart (2015) also do not take into consideration the crucial distinctions made in Author (2017): "the subdivision of unusual rules into unnatural versus unmotivated rules, paired with the proof that the former require at least three sound changes to arise" (the MSCR).[3] The model proposed in this paper has a disadvantage that the trajectories of sound changes that lead to a certain alternation need to be identified manually (similar to Bell's (1971) and Greenberg's (1978) models), but this also means that samples of sound changes need to be representative only for the sound changes we are estimating. This paper shows that the MSCR and a diachronic model of unnatural processes called the Blurring Process (Author 2017) facilitate the identification of trajectories that lead to unnatural processes, which consequently facilitates a quantitative model of typology within the CB approach.

1.3 Goals

The goal of this paper is to propose a quantitative method for estimating the influences of channel bias on phonological typology called Bootstrapping Sound Changes (BSC). The BSC technique es-

[3] 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 and Sonderegger 2013, 2015; Hruschka et al. 2015) do not directly deal with estimating the probabilities of sound changes that operate in combination, but computationally model the initiation and propagation of single sound changes.

timates the so-called Historical Probability, the probability that an alternation arises based on two diachronic factors — the number of sound changes required for an alternation to arise (the MSCR; Section 2), and their respective probabilities, estimated from surveys of sound changes. The probabilities are estimated with the statistical technique *bootstrapping* (Efron 1979); the assumptions of the model are discussed in Section 3.4. This paper argues that with the BSC technique, we can (i) estimate the Historical Probability of any alternation (Section 4.1), (ii) compare two alternations, attested or unattested, and perform statistical inferences on the comparison (Section 4.2), and (iii) compare outputs of the BSC model with independently observed typology to evaluate the performance of the CB approach (Section 4.3). Using the BSC technique, this paper also identifies crucial mismatches in typological predictions between the AB and CB approaches (Sections 4.3 and 5). By testing these mismatched predictions against the observed typology, we can at least partially control for one factor when testing the other and vice versa, which consequently allows for quantitative evaluation of distinct contributions of the AB and CB factors on phonological typology (Section 5).

The BSC technique is applied to three natural-unnatural alternation pairs that target the feature [±voice]: post-nasal (de)voicing (PND, PNV), intervocalic (de)voicing (IVD, IVV), and final (de)voicing (FD, FV). The feature [±voice] is chosen for several reasons. First, phonetic naturalness is probably best understood precisely for this feature. [+voice] is natural intervocalically and post-nasally and unnatural word-initially and word-finally for clear articulatory and perceptual reasons (Aerodynamic Voicing Constraint; Ohala 1983, 2011; Westbury and Keating 1986; for a detailed argumentation, see Author 2017). Second, all three alternations are well-researched typologically: PND and FV are probably two of the most widely discussed alternations in the phonological literature (Hyman 2001; Kiparsky 2006, 2008; Blevins 2004; Yu 2004; Coetzee and Pretorius 2010). Third, the three alternations crucially differ in their synchronic attestedness: PND is reported in thirteen languages as a sound change, which in four languages results in a synchronic alternation (Author 2017; PND has even been confirmed as a productive alternation with wug-tests; Coetzee and Pretorius 2010). IVD is attested once as a morphologically conditioned alternation (Bloyd 2015, 2017) and once as a sound change that results in a gradient phonotactic restriction (Author and Name 2017). FV is arguably never attested as a synchronic alternation (Kiparsky 2006, 2008; cf. Haspelmath 1993, Yu 2004). Most reported cases of FV, such as Lezgian, Latin, and Somali, have been shown not to qualify as FV (Kiparsky 2006, Lipp 2016, Author 2017a, Author 2018).[4] Fourth, the natural counterparts of the three unnatural alternations are recurrent and typologically common phonetic tendencies and alternations. Finally, both attested unnatural processes (PND and IVD) arise from a combination of three natural sound changes (the Blurring Process; see Sections 2 and 4.1) and this development is historically directly or indirectly attested. For FV, Kiparsky (2006) identifies several diachronic trajectories that would yield the alternation, but none appear to be attested. These different degrees of synchronic and diachronic attestedness among PND, IVD, and FV and their respective natural counterparts make good grounds for a comparison of different approaches to phonological typology.

The outputs of the BSC model suggest that the typological rarity of unnatural processes discussed here is primarily due to the CB factor: BSC predicts with relatively high accuracy the typological differences between natural and unnatural alternation pairs for which no learnability differences have been observed. On the other hand, the BSC model also shows that a sound change that simplifies a complex alternation to a simple alternation, and consequently simplifies its learn-

[4]There is one possible case of FV that could count as a productive synchronic alternation — Lakota (Rood 2016). Currently, there are no acoustic studies of Lakota word-final stops. Since many reported cases of FV turned out to be cases of word-final unreleased or lax voiceless stops (Kiparsky 2006), I leave Lakota out of this discussion until acoustic data is available.

ing, operates at a significantly higher rate than expected if only the CB factor is modeled. Such higher rates of sound change operation that resolve a complex alternation are most likely influenced by the AB factor. In other words, the model suggests that the typological rarity of unnatural processes that target the feature [±voice] is due to the CB factor, while the rarity of complex alternations is likely influenced by the AB factor.

2 Background

This paper adopts several diachronic concepts from Author (2017). This section discusses their relevance to the BSC technique. In the interest of space, not all details can be presented here; the reader should be directed to Author (2017) for an in-depth discussion of the Blurring Process and the MSCR.

First, this paper adopts the division of phonological processes into *natural*, *unmotivated*, and *unnatural*. Natural processes, such as final devoicing or post-nasal and intervocalic voicing, are cases of phonetically well-motivated universal phonetic tendencies. Universal phonetic tendencies are defined as having articulatory or perceptual motivation, operating passively cross-linguistically, and resulting in typologically common phonological processes (Author 2017). Unmotivated processes lack phonetic motivation, but do not operate against universal phonetic tendencies. Unnatural processes not only lack phonetic motivation, but also operate against universal phonetic tendencies. Examples of unnatural alternations that operate against these universal phonetic tendencies include final voicing (T → D / __#) or post-nasal (D → T / N__) and intervocalic (D → T / V__V) devoicing (for articulatory and perceptual argumentation on the unnaturalness of these processes, see Westbury and Keating 1986; Ohala 1983, 2011; Coetzee and Pretorius 2010; Iverson and Salmons 2011; Author 2017, 2018). An example of an unmotivated process would be Eastern Ojibwe "palatalization" of /n/ to [ʃ] before front vowels (Buckley 2000), which lacks phonetic motivation, but does not operate against universal tendencies.

Sound change is defined as a non-analogical "change of one non-automatic feature value in a given environment" (Author 2017). Features that change automatically along with non-automatic feature values do not count as instances of sound change (e.g. a change in [±nasal] automatically causes a change in [±sonorant]). Additionally, sound change is a completed event that ideally targets all vocabulary items in a given language L. Because the BSC technique models sound changes that, via phonologization, result in phonological alternations, we adopt the level of abstraction from phonology where features encode non-automatic, language-specific, and speaker-controlled phonologized processes that cannot be attributed to universal phonetics (Hyman 2013). Defining sound change is, in itself, a non-trivial task (Garrett 2014) and different definitions are possible or even more appropriate for different purposes. When modeling probabilities of synchronic alternations, it is reasonable to adopt the definition that operates with phonological features (see also discussion in Author 2017). In any case, this problem is not unique to the proposed model: any diachronic model that estimates probabilities of synchronic alternations will need to define sound change at some more or less arbitrary level of abstraction.

This paper also assumes that a single sound change most frequently changes one feature value (or deletes, inserts, or reorders a whole feature matrix in the case of deletion, epenthesis, and metathesis) in a given environment. This "minimality principle", first proposed in Donegan and Stampe (1979) and Picard (1994), is discussed at length in Author (2017). A *combination of sound changes* is defined in this paper as a set of individual sound changes that each target a single feature value.

This paper adopts two key diachronic concepts for the derivation of typology within the Channel

Table 1: Unnatural processes and languages in which they appear.

Unnatural process or sound change	**Language**	**Description**
Post-nasal devoicing	Yaghnobi	Xromov (1972)
	Tswana, Shekgalagari	Hyman 2001, Solé et al. (2010)
	Makhuwa	Janson (1991/1992)
	Bube, Mpongwe	Janssens (1993), Mouguiama-Daouda (1990)
	Konyagi	Merrill (2016a,b)
	Sicilian and Calabrian	Rohlfs (1949)
	Buginese and Murik	Blust (2013)
	Nasioi	Brown (2017)
	Tarma Quechua[a]	Adelaar (1977)
Intervocalic devoicing	Sula	Bloyd (2015)
	Berawan and Kiput	Blust (2005, 2013)

[a]Tarma Quechua features further unnatural distributions (see Author and Name 2017).

Bias approach that have been proposed in Author (2017): the Blurring Process and the Minimal Sound Change Requirement. Typological surveys of unnatural processes targeting the feature [±voice] conducted in Author (2017) and Author and Name (2017) identify thirteen languages in which PND has been reported either as a productive synchronic alternation or as a sound change, and two additional cases of unnatural phonotactic restrictions — the distribution of the feature [±voice] in Tarma Quechua and in the Berawan dialects (summarized in Table 1). Author (2017) argues that PND arises from a combination of three sound changes in all reported cases (as was proposed for Tswana in Dickens 1984 and Hyman 2001) and argues that the three-sound-changes approach is historically directly confirmed by Avestan, Sogdian and Yaghnobi, three languages in ancestral relationship that have all three sound changes attested in written sources. Based on this typological survey, a new model for explaining unnatural processes diachronically is proposed: the Blurring Process. The Blurring Process states that unnatural alternations arise through a combination of a specific set of three natural (phonetically motivated) sound changes: (i) a sound change that causes complementary distribution, (ii) a sound change that targets changed or unchanged segments in the complementary distribution, and (iii) a sound change that blurs the original complementary distribution (Author 2017). All reported cases of unnatural phonotactic restrictions (in the Berawan dialects and Tarma Quechua) also arise through the Blurring Process.[5]

The Blurring Process allows us to maintain the long-held position that the operation of sound changes is limited to phonetically natural directions (Garrett and Johnson 2013, Garrett 2014). This position has recently been challenged by Blust (2005), who lists a number of unnatural sound changes. If these unnatural sound changes can be explained by the Blurring Process as combinations of natural sound changes (as argued by Author 2017 and Author and Name 2017), we can maintain the position that sound change is always phonetically motivated. This restriction is crucial when

[5]The proposed model does not account for unnatural processes that result from morphological changes. These are, however, almost always analyzed as morphologically conditioned and can be explained by non-phonological mechanisms. IVD in Sula (Bloyd 2015) likely results from a combination of morphological changes and is morphologically conditioned. The exact mechanisms of how IVD arises there are not straightforward and are beyond the scope of this paper.

modeling the typology within the CB approach: some processes can only arise from a combination of natural sound changes, which means the probability of an alternation depends not only on the probability of a single (natural or unnatural) sound change, but also on the number of natural sound changes required for a process to arise and their respective probabilities.

The Blurring Process leads to another property of diachronic phonology: the so-called Minimal Sound Change Requirement. As was already argued, a single sound change is always phonetically motivated, which means that unnatural alternations cannot arise from a single sound change. Author (1) additionally shows that unnatural segmental alternations cannot arise from two sound changes either. In sum, the MSCR 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 (for a formal proof of the MSCR, see Author 2017). As will be shown below, the MSCR has implications for the derivation of typology within the Channel Bias approach.

3 Bootstrapping Sound Changes

MSCR predicts that, all else being equal, natural processes will be the most frequent, unmotivated less frequent, and unnatural the least frequent, because the former require one sound change to arise, the second at least two, and the latter at least three ((1) from Author 2017). All else being equal, the more sound changes an alternation requires, the less frequent it will be typologically.

(1) *A scale of decreased probabilities (Author 2017)*
P_χ(natural) > P_χ(unmotivated) > P_χ(unnatural)

The MSCR is a crucial concept in the discussion on different influences on phonological typology because the typological consequences of the MSCR can be ascribed exclusively to CB — learnability is independent of the requirement that some processes need to arise from a number of sound changes (even if the rate of operation of individual sound changes *can* be influenced by learnability, see Section 5). The MSCR, however, only predicts *categorical* relations between alternations with different degrees of naturalness and does not explain why some unnatural processes are attested and others are not. Our goal is to propose a model that would quantify probabilities of natural, unmotivated, and unnatural processes further. We can combine the MSCR with the assumption that the probabilities of sound changes influence the probabilities of synchronic alternations. Crucially, the probability that an alternation arises based on diachronic factors depends on both the number of sound changes that are required for the alternation to arise and the probability of each individual sound change in the combination. Such probabilities are called *Historical Probabilities of Alternations* (P_χ).

(2) *Historical Probabilities of Alternations* (P_χ)
The probability that an alternation arises based on the number of sound changes required (the MSCR) and their respective probabilities, which can be estimated from samples of sound changes.

A challenge in estimating the CB influences on typology with the BSC technique is the possibility that the probabilities of each individual sound change estimated from diachronic surveys are influenced not only by the CB factors, but by learnability (AB) as well (called the *AB-CB conflation* problem henceforth; see Kiparsky 1995, 2008; Moreton 2008). That this is likely not the case for sound changes targeting the feature [±voice] is suggested by sound change typology: the natural sound changes PNV, IVV, and FV are frequent (Kümmel 2007), whereas their unnatural counterparts are never attested as individual *sound changes* (unlike unnatural synchronic alternations, which are rare but attested). Experiments that tested the learnability of these pairs of natural and unnatural processes, however, failed to find significant differences (Seidl et al. 2007, Do et al. 2016, Glewwe 2017; Glewwe et al. 2018). If learnability (AB) factors played the primary role in determining probabilities of individual sound changes, we would expect unnatural sound changes to be as frequent as the natural ones.

It is still possible that learnability influences the probabilities of individual sound changes: learnability can promote or demote the likelihood of a phonetic precursor being phonologized (as argued by Moreton 2008; for criticism, see Yu 2011 and Kapatsinski 2011). Even if learnability indeed influences the probability of the phonologization of a process, its effects are likely minor, promoting or demoting phonologization, rather than affecting the probabilities of sound changes to the degree that we observe in the typology of natural and unnatural processes (Figure 3).

This paper additionally controls for the AB influences by applying the BSC technique on unnatural alternations and by identifying mismatched predictions of the AB and CB approaches. Even if the probabilities of individual sound changes are crucially influenced by learnability (and therefore by AB) and even if learnability causes a higher rate of operation of certain sound changes in combination, the fact that at least two or three sound changes are required to operate in a language (due to the MSCR) for unmotivated alternations and unnatural alternations, respectively, to arise means that CB plays a crucial role in determining the synchronic typological probabilities of these alternations. All else being equal, even if we assumed that learnability is the only factor influencing the probabilities of individual sound changes (even if this stance is highly unlikely), the probability of a single sound change will necessarily be greater than the probability of a combination of three sound changes. This generalization is necessarily influenced by CB because the sound changes need to operate in combination and in the temporal dimension of a given speech community.

Estimating Historical Probabilities is not a trivial task and requires several simplifying assumptions. A detailed discussion of assumptions that the BSC model makes is given in Section 3.4. Many of these assumptions are not limited to the BSC technique, but will pose challenges to any diachronic model. By comparing the outputs of the BSC technique with observed typology, this paper argues that, despite simplifying assumptions, the BSC technique makes the most accurate predictions of typology based on the CB factor (Section 4.3).

3.1 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) from a sample by random sampling with replacement. It was first proposed by Efron (1979) and has seen a wide range of applications ever since (Davison and Hinkley 1997).

The BSC model uses a stratified non-parametric bootstrap technique for estimating Historical Probabilities for several reasons. First, the statistic of interest in BSC is often too complex for an easy analytic solution, especially when we estimate Historical Probabilities of alternations that require more than a single sound change (see (5) below) or when we estimate differences between Historical Probabilities of two alternations (see Section 4.2 below). Second, bootstrapping is a frequentist technique for estimating sampling distribution for a statistic of interest and as such requires no prior beliefs. Finally, bootstrapping allows for inferential statements on the comparison of the Historical Probabilities of two alternations, even when the statistic of interest is complex.

The computation of BSC is implemented in the statistical software R (R Core Team 2016) with the *boot* package (Canty and Ripley 2016, Davison and Hinkley 1997) using the functions *boot()* and *boot.ci()*. This paper also presents R code that implements the BSC technique and introduces the functions *bsc()*, *summary.bsc()*, *bsc2()*, *summary.bsc2()* (based on the *boot* package) that facilitate the estimation of Historical Probabilities with BSC (available in Appendix A). The functions allow the estimation of Historical Probabilities directly from a vector of counts. The aim of the code is to provide a ready-to-use interface for the estimation of the Historical Probability of any alternation and thus to provide a means for estimating the Channel Bias influence in future discussions on phonological typology.

3.2 The model

As defined in (2), the Historical Probability of an alternation A_k is the probability that a language L features alternation A_k based on the number of sound changes (S_i) alternation A_k requires and their respective probabilities.

3.2.1 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 (3). 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_χ according to (3).

(3)

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

BSC samples with replacement from the sample of successes and failures (based on surveys of sound changes) and calculates the statistic of interest: in our case, the probability according to (3). This is repeated 10,000 times (each sample being of the same length as the sample size), which yields a sampling distribution of Historical Probabilities: 10,000 data points. From this sampling distribution, standard error, bias, and 95% adjusted bootstrap (BC_a) confidence intervals that adjust for bias and skewness (Efron 1979, 1987) are computed.

The analytic equivalent of the BSC technique for an alternation that requires only a single sound change is an empty logistic regression model with the number of successes and failures as the dependent variable and with only the intercept (no predictors included). As the statistic of interest becomes more complex when estimating Historical Probabilities of processes that require multiple sound changes, I shift from the analytic framework to a non-parametric bootstrap. For consistency, I maintain the BSC approach even for alternations that require only a single sound change and could otherwise be estimated using an analytic approach.

3.2.2 Two or more sound changes

If an alternation A_k requires more than a single sound change (i.e. is not natural), then the Historical Probability of A_k is estimated as a sum of the Historical Probabilities of each trajectory T_z that yields the alternation A_k, as shown in (4).

(4)

$$P_\chi(A_k) = P_\chi(T_1 \cup T_2 \cup \ldots \cup T_n)$$

A trajectory T_j denotes a combination of sound changes that yields an alternation A_k. 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_j that requires more than a single sound change is estimated from the joint probability of the individual sound changes required for T_j, divided by the factorial of the number of sound changes in trajectory T_j if only one ordering results in the trajectory in question, as shown in (5).

(5)

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

Estimating the joint probability of individual sound changes ($\mathrm{P}_\chi(S_1 \cap S_2 \cap \ldots \cap S_n)$) is not a trivial task. A number of assumptions are needed in order to compute this joint probability, the most important of which is the assumption that the occurrence of one sound change does not influence the probability of the following sound change. In other words, sound changes are treated as independent events under BSC. This is in fact a desirable assumption when modeling a purely diachronic approach to typology. As will be argued in Section 5, learnability does influence the probabilities of individual sound changes that operate in combination. Because these influences are in the domain of the AB factor, however, they should not be modeled within the CB approach. For a discussion on the assumptions of the BSC model, including a discussion on the dependency of sound changes on phonemic inventories and learnability, see Section 3.4.

As defined in (2), Historical Probability is a probability that a language L features an alternation A_k, 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_χ) of the first individual sound change S_1 is thus estimated from the number of successes (languages with S_1) and the number of failures (languages without S_1) according to (3), 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 stop geminates.

Once S_1 operates, however, we know that language L necessarily has the target/result/context of the sound change S_1. For this reason, we estimate the Historical Probability of the subsequent sound changes $\mathrm{P}_\chi(\mathrm{S}_2)$ by dividing the number of successes (languages with S_2) by the number of languages surveyed that feature the target/result/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/context of the previous sound change, we can treat the probabilities of individual sound changes as independent events under the CB approach and estimate P_χ from the product of the probabilities of individual sound changes (6).

(6)

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

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 BSC 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 (3)), and then computes the product of the Historical Probabilities of each individual sound change divided by $n!$, according to (6). 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.[6]

3.2.3 Comparison

The BSC technique also allows for the estimation of the difference between the Historical Probabilities of two alternations, which consequently enables inferential statements on the comparison.

(7)

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

The difference between the Historical Probabilities of two alternations (ΔP_χ) is estimated with a stratified non-parametric bootstrap, where P_χ of each individual alternation A_1 and A_2 is estimated as described in Sections 3.2.1 and 3.2.2 (depending on whether A_1 and A_2 require trajectories that require one or more sound changes). To compare two Historical Probabilities, BSC additionally calculates 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 BSC technique applied on 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 $\alpha = 0.05$. If, on the other hand, the 95% BC_a confidence intervals of the difference cross 0, then $P_\chi(A_1)$ and $P_\chi(A_2)$ are *not* significantly different with $\alpha = 0.05$.[7]

3.3 Sample

Samples used for estimating Historical Probabilities with BSC are created from typological surveys of sound changes. The BSC 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.

[6] The functions that perform this computation are *bsc()* (performs a stratified non-parametric bootstrap based on the *boot()* function; see A.1) and *summary.bsc()* (computes confidence intervals based on the *boot.ci()* function; see A.3).

[7] The functions that perform this computation are *bsc2()* (performs a stratified non-parametric bootstrap based on the *boot()* function; see A.2) and *summary.bsc2()* (computes confidence intervals based on the *boot.ci()* function; see A.4).

The most elaborate survey of sound changes currently available, on which the BSC analysis is performed, is the survey of consonantal sound changes in Kümmel (2007). One major advantage of Kümmel's (2007) 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, compared to competing surveys (e.g. the UniDia survey; Hamed and Flavier 2009). 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. While subgrouping and probabilities of sound change can be inferred through phylogenetic tree analysis (Hruschka et al. 2015), subgrouping in Kümmel's (2007) survey relies on historical methodology that includes information from both phonological as well as morphological and other higher level evidence. Additionally, phylogenetic tree analysis does not restrict the direction of sound change and would, for example, incorrectly analyze reported unnatural alternations as resulting from a single sound change.

The survey in Kümmel (2007) includes approximately 294 languages and dialects of the Indo-European, Semitic, and Uralic language families. While the survey is not as representative because it is limited to only three language families, the fact that it involves precisely those families that have well-established subgrouping, which allows for proper coding, compensates for the lack of representativeness. Results of the analysis presented in Section 4 are likely not crucially affected by the fact that many language families are excluded from the survey because frequencies and types of sound changes do not seem to be radically different across different language families (with recurrent sound changes appearing across all families; Blevins 2007; see also Section 4.3).

The only other comparable survey of sound changes known to the author is the UniDia database that surveys 10,349 sound changes from 302 languages (Hamed and Flavier 2009). The UniDia database is, however, less appropriate for the BSC technique because it lacks elaborate diachronic subgroupings of languages. The survey appears to list changes from a proto-language to daughter languages irrespective of whether a change occurred at the proto-language stage or independently in the daughter languages. In addition to the lack of subgrouping, the UniDia database is not representative either, focusing primarily on the Bantu language family (83.5% of sound changes are from the Bantu family).

The BSC technique offers some crucial advantages over Cathcart's (2015) proposal of estimating the probabilities of sound changes and their combinations. The requirement of sample representativeness is much weaker under the BSC approach. Cathcart's (2015) model crucially requires surveys of sound changes to be representative for all possible sound changes. Also, the model is based on the UniDia database, which is less appropriate compared to Kümmel's (2007) survey, primarily because of its encoding of sound changes, which lacks subgrouping. Because the identification of historical trajectories that lead to an alternation is performed manually in the BSC model, surveys of sound changes that are used for BSC calculations need not be representative for all possible sound changes, but only for those required for the alternation in question. In fact, elaborate surveys of sound changes can be constructed for each alternation in question even in the absence of a large and representative survey of sound changes.

3.4 Assumptions

As any diachronic model, the proposed BSC 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 (6)), the model assumes that each sound change is an independent event. The BSC 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 based on their targets, results, and contexts (Section 3.3) and, crucially, from samples conditioned on the result of the previous sound change (Section 3.2.2). 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 (Section 3.2.2) and (ii) independent of 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 modeling typology within the CB approach. BSC aims to estimate only the CB influences on typology, which is why it has to assume that the probability of sound change is only determined by its frequency of operation evaluated on a diachronic and unconditioned level. In other words, the BSC model assumes that sound change is blind to AB factors such as the learnability of a process. While the probabilities of individual sound changes are modeled as independent of each other under the CB approach, it is likely that they are not independent: the operation of one change can influence the learnability of the resulting process, which consequently influences the operation of the following sound change. In fact, I will argue in Section 5 that probabilities of sound changes are indeed influenced by learnability factors and that a sound change that simplifies the learning of an alternation operates significantly more frequently than is predicted by only the CB factor. However, the diachronic model proposed here is designed to model only the CB contribution to the typology and should be blind to learnability, which means that the assumption of independence is desired for this purpose.

The second assumption of independence is more problematic: broader phonemic inventories can influence the probabilities of sound changes, especially for vocalic changes (e.g., due to the effects described in the Theory of Adaptive Dispersion, see Liljencrants and Lindblom 1972, Lindblom 1990), but also for consonantal changes. The BSC 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. It is assumed that this dependency is relatively weak and does not crucially affect the probabilities of sound changes targeting the feature [±voice]. The dependence of sound change on broader phonemic inventories is not modeled primarily because current surveys of sound changes are not sufficiently large and representative. In principle, the BSC technique could model this dependency by estimating the probabilities of sound changes from samples conditioned on some phonemic property of the surveyed languages.

As already mentioned, identification of individual trajectories leading to an alternation A_k is performed manually in the current proposal. 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 missed in 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 generalize the Historical Probability of that particular trajectory to the Historical Probability of the alternation. If such an assumption is not met, however, then BSC estimates the probability that an alternation A_k arises from a trajectory T_j.[8] This paper assumes that the estimated trajectories are the most frequent ones and that potential alternative trajectories do not crucially influence the results.

What is not accounted for in the model are the functional load of individual phonemes (Wedel 2012, Wedel et al. 2013, Hay et al. 2015) and other factors that could potentially influence proba-

[8] 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 BSC.

bilities 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 sound change is initiated or spread. These factors can mostly be disregarded because the goal of the BSC technique is to estimate the Historical Probability of alternation A_k operating in a language L with no conditional properties.

Finally, the BSC technique does not directly model the temporal dimension. If more comprehensive typological studies with more detailed temporal information were available, a different model (e.g. a model operating within the Poisson stochastic process) could account for the temporal dimension and estimate probabilities of sound changes given a timeframe. In the absence of temporal information, the BSC technique has to make some simplifying assumptions. These simplifying assumptions are not unique to the present proposal and are to some degree even desirable. The BSC 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 L. While this might be too restrictive, it is, in fact, 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 PND in Yaghnobi operates over three languages and fails to result in a productive synchronic alternation. The model also assumes that once a sound change occurs 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. In other words, sound changes in our model are birth-death events, a view that is substantiated by empirical evidence: sound change operates and then ceases to operate (Chen 1974), at which point it can occur again (e.g. on novel morphological or loanword material).

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 opposite sound change (in our case, the natural sound change) will operate on the unnatural system and destroy the evidence for it. It is relatively unproblematic to include this influence in the model: the product of the estimated Historical Probability and the probability that the natural sound change does not occur would yield a Historical Probability corrected for the potential influence of the natural sound change. Currently, influences of the potential natural sound changes are not modeled because the Historical Probabilities of the natural sound changes (Table 9) are relatively similar for the processes estimated in this paper and we do not expect this additional factor to alter the results significantly. For other processes not estimated in this paper, including the probability of the natural sound change in the model might alter the outcomes significantly.

Most of the influences that are not directly modeled in the current proposal are at least partially accounted for by the fact that the sample size based on Kümmel's (2007) survey is relatively large and relatively representative. If the sample is representative, influences of various linguistic and non-linguistic factors will be reflected already in the sample and the results of the model will not be crucially affected. For practical purposes, these influences can be disregarded, because the effects are likely minor enough not to crucially alter the results. In addition, the BSC technique estimates the Historical Probabilities of alternations in a language L, where L represents a language that has the characteristics of the majority of languages in the sample. The Historical Probability is not conditioned on L's phonemic inventory, functional load of phonemes, or other factors, which is why these factors can be disregarded for practical purposes.

4 Applications

4.1 Estimation of Historical Probabilities

The BSC technique enables the estimation of Historical Probabilities for natural, unmotivated, and unnatural alternations, both attested and unattested (according to Section 3.2). For the purpose of illustrating the method, this paper presents estimates of the Historical Probabilities (P_χ) of the natural alternations — post-nasal voicing (PNV), intervocalic voicing (IVV), and final devoicing (FD) — and their unnatural counterparts — post-nasal devoicing (PND) intervocalic devoicing (IVD), and final voicing (FV). To construct samples of sound changes for these six alternations, the survey of consonantal sound changes in Kümmel (2007) is employed.

4.1.1 Trajectories

The three natural alternations have obvious origins — the single natural sound changes PNV, IVV, and FD, respectively. For the unnatural alternations, we first identify sound changes in the Blurring Process (Section 2) that yield the alternation in question. If A > B / X is a natural sound change, then B > A / X is unnatural. Tables 2, 3, and 4 represent schematically (left column) how the unnatural B > A / X arises via the Blurring Cycle or the Blurring Chain (two subtypes of the Blurring Process; see Section 2, Author 2017, and Author and Name 2017). The actual sound changes that yield the unnatural alternation are identified in the right columns.

The origins of the unnatural alternations PND, IVD, and FV are well-established. Author (2017) demonstrates that PND always results from the Blurring Cycle. A combination of the following three natural and well-motivated sound changes yields PND in all known cases: the fricativization of voiced stops in non-post-nasal position, the unconditioned devoicing of voiced stops, and the occlusion of voiced fricatives to stops. Table 2 illustrates the development.[9]

Table 2: Blurring Cycle (schematic; left) yielding PND (right).

	Blurring Cycle	**PND**
1.	B > C / ¬X	D > Z / [−nas]__
2.	B > A	D > T
3.	C > B	Z > D
Result	B > A / X	D > T / [+nas]__

Author and Name (2017) argue that IVD results from the Blurring Chain. Voiced stops fricativize intervocalically, voiced fricatives devoice, and voiceless fricatives get occluded to stops (see Table 3). The result is the unnatural intervocalic devoicing (D > T / V__V).

Table 3: Blurring Chain (schematic; left) yielding IVD (right).

	Blurring Cycle	**IVD**
1.	B > C / X	D > Z / V__V
2.	C > D	Z > S
3.	D > A	S > T
Result	B > A / X	D > T / V__V

[9] T represents voiceless stops, D voiced stops, S voiceless fricatives, and Z voiced fricatives.

FV is arguably unattested both as a synchronic alternation and as a sound change (Kiparsky 2006, Lipp 2016, Author 2017a, fn. 4, cf. Yu 2004, Rood 2016). A number of diachronic scenarios exist, however, that would yield FV and 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 instead (Section 6). One possible scenario that would result in FV is Scenario 1[10] in Kiparsky (2006), which is used here for estimating the Historical Probability of FV. 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 FV in this scenario are geminate simplification in word-final position, voicing of post-vocalic non-geminate stops, and unconditioned geminate simplification (see Table 4).

Table 4: Modified Blurring Chain (schematic; left) that would yield FV (right).

	Modified Blurring Cycle	IVD
1.	C > B / X	T: > T / __#
2.	B > A	T > D / V__
3.	C > B	T: > T
Result	B > A / X	T > D / __#

4.1.2 Counts

Based on the trajectories identified here that result in natural and unnatural alternations, counts of sound changes and languages surveyed are performed based on Kümmel (2007). Sound change occurrences are counted from the number of languages that Kümmel (2007) lists for each sound change. If Kümmel (2007) lists more than one language per exact realization of a sound change, the occurrences are treated as independent events, even though the languages might be closely related. While it is likely that some of the sound changes counted as independent events in related languages operated as a single event at the pre-stage, we do not expect this to be the case in many occurrences and therefore we do not expect the results to be crucially affected by such counts.

PNV that targets labials, dental/alveolars, or velars is reported in approximately 42 languages in Kümmel (2007). IVV is reported in approximately 28 languages if only occurrences that strictly require intervocalic (as opposed to post-vocalic) context are counted. FD is reported in approximately 33 languages. PNV, IVV, and FD that target a single series of stops are counted together with cases in which these sound changes target more than a single place of articulation. In fact, sound changes for all six natural and unnatural alternations are counted as successes even if they target only a single place of articulation because the resulting alternation would count as natural/unnatural, even if it targeted only a single place of articulation. Unclear cases marked with "?" in Kümmel (2007) are excluded from the count. Table 5 summarizes the counts of languages with sound changes that result in natural alternations.

[10]Kiparsky's (2006) Scenario 2 also includes three sound changes, but the last sound change (apocope after a single consonant) is never attested in the UniDia database of sound changes (Hamed and Flavier 2009). Kümmel's 2007 survey does not include vocalic changes, which is why the UniDia database that surveys 10,349 sound changes from 302 languages is used. Because the last sound change is never attested in our surveys, I exclude Scenario 2 from the estimation of P_χ(FV).

Table 5: Counts of sound changes in Kümmel (2007) for natural alternations.

Alternation	**Sound change**	**Count**	**Surveyed**
PNV	T > D / N__	42	294
IVV	T > D / V__V	28	294
FD	D > T / __#	33	294

For the unnatural alternations that require more than a single sound change, counts are performed for each individual sound change in the corresponding Blurring Processes. The first sound change in the Blurring Chain that results in PND, the fricativization of voiced stops, is reported in approximately 97 languages. Instances of intervocalic and post-vocalic fricativization are included in the count (not only cases in which fricativization occurs in all but post-nasal position) because the result of such fricativization after the other two sound changes would be a system analyzed as PND as well.[11] The probability of the first sound change in the Blurring Cycle that results in PND is estimated based on the 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 that A_k arises in a language L, regardless of the properties of its phonemic inventory (see Section 3.2.2). Once the first sound change operates, however, we know that the language in question needs to have voiced stops in its inventory. The Historical Probability of the second sound change that 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 sound change (D > T) is reported in approximately 18 languages (also counting cases of devoicing that are the result of chain shifts). Approximately 31 languages lack voiced stops in the survey in Kümmel (2007),[12] which means that P_χ is estimated based on $294 - 31 = 263$ languages surveyed. 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 based on the number of languages with occlusion of voiced fricatives and the number of languages surveyed with voiced fricatives (allophonic or phonemic). Approximately 217 languages in the survey have voiced (bi)labial, alveolar/dental, or velar non-strident fricatives,[13] according to Kümmel (2007). In approximately 27 languages, occlusion of fricatives is reported as a sound change. The counts for IVD are performed in the same manner as the counts for PND and are given in Table 6.

The Historical Probability of FV is estimated based on the one scenario in Kiparsky (2006) that would result in FV as an alternation. The scenarios that would lead to FV as a static phonotactic restriction and could involve fewer than three sound changes are excluded. There are three main reasons for why it is justified to distinguish alternations from static phonotactic restrictions in a diachronic model (Author 2017) despite the two phenomena likely being part of the same synchronic grammatical mechanisms (Prince and Smolensky 1993/2004, Hayes 2004, Pater and Tessier 2006). First, unnatural phonotactic restrictions that do not result from a Blurring Process provide

[11] An alternation that resulted from a combination of sound changes in which the first sound change targeted post-vocalic stops rather than non-post-nasal stops and the other two aforementioned sound changes have the same result as in the attested case of PND, and would be analyzed as PND with initial devoicing.

[12] One language has only /b/ in its inventory. The low number of inventories that lack voiced stops might be influenced by the areal that Kümmel (2007) surveys. Based on the PHOIBLE database (Moran et al. 2014), approximately 30% of inventories lack a phonemic labial voiced stop. For consistency purposes, we stay within Kümmel's (2007) survey with this acknowledgement.

[13] The labiodental voiced fricative /v/ is included in the count.

considerably less reliable evidence for learners because the evidence is distributional rather than appearing within the same morphological unit. This means that the likelihood of a process not being acquired by the learners is considerably greater when it does not arise from a Blurring Process. Second, alternative analyses of data are often available in the case of phonotactic restrictions that do not result from a Blurring Process. Alternative explanations are not available in the case of alternations, where evidence for a process comes from within the same morphological unit. Finally, typological surveys of phonotactic restrictions are considerably more difficult to establish (compared to typological surveys of alternations). In the absence of typological studies, it is difficult to evaluate predictions of the CB model for phonotactic restrictions. In fact, FV as a phonotactic restriction might not be as rare, with at least two potential phonological systems attested in which voiceless stops do not surface word-finally, but voiced stops do (Ho and some dialects of Spanish; see Author 2017). For a further discussion on the differences between phonotactic restrictions and alternations, see Author (2018).[14]

Counts of the sound changes that lead to FV as an alternation are as follows. In approximately six 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 as opposed to a static phonotactic restriction.) Because this is the first in the series of changes and we do not condition P_χ on any property of language L, as before, the Historical Probability is estimated from the total number of languages surveyed. The second sound change, post-vocalic voicing of voiceless stops, is reported in approximately 32 languages. Because all languages have voiceless stops, all 294 languages surveyed are included in the count for estimating the Historical Probability of the second sound change. Finally, simplification of geminates is reported in 27 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 must allow allophonic geminates at morpheme boundaries. To estimate the number of languages that allow allophonic geminates, Greenberg's (1965) survey of consonantal clusters and Ryan's (to appear) survey of phonemic geminates are used. At least 30% of languages in Greenberg's (1965) survey of approximately 100 languages allow stop + stop final clusters. The number of languages in our sample that allow allophonic homorganic stop-stop sequences (geminates) can be approximated from the proportion of languages that allow phonemic geminates and from the proportion of languages 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 of course 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 (to appear) estimates that approximately 35% of 55 genealogically diverse languages surveyed have phonemic geminates.

The estimates of Historical Probabilities are computed with a stratified non-parametric bootstrap as described in Sections 3.2.1 and 3.2.2 (see also A.1 and A.3). Table 7 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_χ) of these natural and unnatural alternations. Table 7 and Figure 1 illustrate a

[14]The scenario that potentially results in FV in Lakota is currently also excluded: the fricativization of voiceless stops before clusters and word-finally, followed by post-vocalic voicing of fricatives and occlusion of fricatives to stops, would potentially result in FV. A preliminary estimation of this scenario shows that its Historical Probability is even lower than the probability of the scenario estimated in Table 7: $P_\chi = 0.003\%$ [0.001%, 0.01%]. The low P_χ is likely a consequence of the first sound change being relatively rare. Because this additional P_χ is approximately 1/10 of the probability estimated in Table 7, we do not expect the absence of this scenario in our model to alter the result substantially.

Table 6: Counts of sound changes in Kümmel (2007) for natural alternations.

Alternation	Sound change	Count	Surveyed
PND	D > Z / [−nas]/V__(V)	97	294
	D > T	18	263
	Z > D	27	216
IVD	D > Z / V__ (V)	83	294
	Z > S	7	216
	S > T	34	248
FV	T: > T / __#	6	294
	T > D / V__	32	294
	T: > T	27	≈88

Table 7: Estimated P_χ (in %) for natural and unnatural alternations with 95% BC_a and analytic profile confidence intervals.

		95% BC_a CI		95% Profile CI	
A_k	P_χ	Lower	Upper	Lower	Upper
PNV	14.3	10.2	18.4	10.6	18.6
PND	0.05	0.02	0.09	—	
IVV	9.5	6.1	12.9	6.5	13.2
IVD	0.02	0.008	0.05	—	
FD	11.2	7.8	15.0	8.0	15.2
FV	0.01	0.004	0.03	—	

substantial difference in Historical Probabilities between the natural and unnatural group. The Channel Bias approach estimated with the BSC technique thus predicts that the unnatural alternations (PND, IVD, and FV) will be substantially less frequent than their respective natural alternations (PNV, IVV, and FD).

The Historical Probabilities and confidence intervals of the natural alternations PNV, IVV, and FD can also be estimated analytically (see Section 3.2). To illustrate the accuracy of the BSC technique, the 95% BC_a bootstrap confidence intervals are compared with confidence intervals computed using an analytic solution (see Section 3.2.1). Analytic profile confidence intervals are computed from an empty logistic regression model with a binomial distribution based on the number of successes and failures (languages with and without sound change S_i) and with only an intercept. Table 7 compares the two sets of confidence intervals. The highest difference between the analytic profile CIs based on a logistic regression model and the BC_a bootstrap CIs is 0.4%, which suggests that the BSC model estimates CIs with high accuracy.

4.2 Comparison of alternations

One of the advantages of the BSC method is that inferential statistics can be performed on the comparison between the Historical Probabilities of any two alternations. In other words, the BSC technique enables significance testing of the Historical Probabilities of pairs of alternations. The difference between the Historical Probabilities of two alternations (ΔP_χ) is estimated by a stratified non-parametric bootstrap as described in Section 3.2.3 (see also A.2 and A.4).

The Historical Probabilities of all three natural alternations in Figure 1 are significantly higher

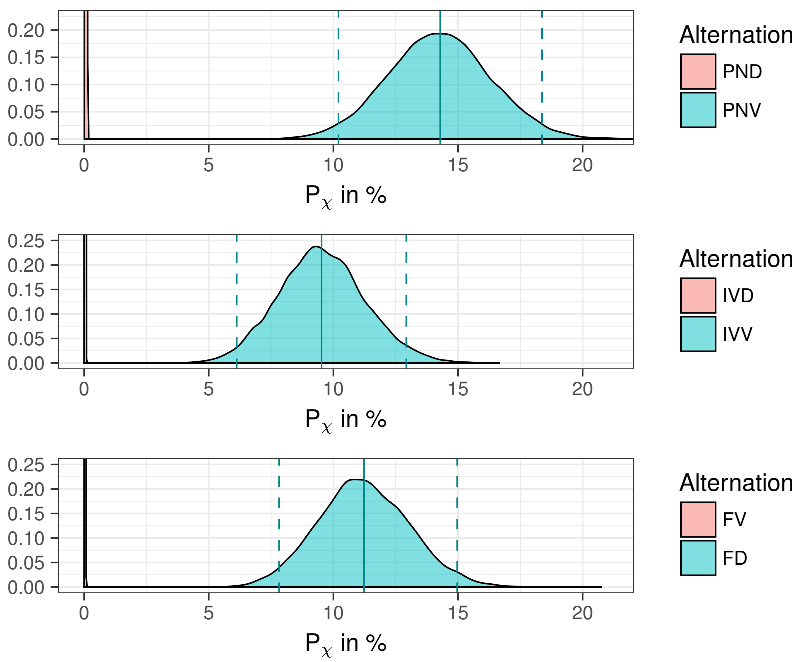

Figure 1: Bootstrap replicates for natural and unnatural alternations. The plots show the observed P_χ (solid line) and the 95% BC_a CI (dashed line). The distribution of bootstrapped P_χ for unnatural alternations does not feature confidence intervals because the probabilities are too small to be visible. For the purpose of representation, the vast majority of bootstrap replicates for unnatural alternations fall outside the limits of the plot.

Table 8: Estimated ΔP_χ (in %) for natural-unnatural alternation pairs with 95% BC_a confidence intervals.

		95% BC_a CI		
Alternation pair	Δ P_χ	**Lower**	**Upper**	
PNV vs. PND	14.2	10.5	18.7	*
IVV vs. IVD	9.5	6.4	13.2	*
FD vs. FV	11.2	8.1	15.3	*

than the Historical Probabilities of their unnatural counterparts. Table 8 includes estimates and 95% BC_a confidence intervals of the difference in Historical Probabilities (ΔP_χ) for each natural-unnatural alternation pair.

Historical Probabilities of all three natural alternations in Figure 1 are significantly higher than Historical Probabilities of their unnatural counterparts. Table 8 includes estimates and 95% BC_a confidence intervals of the difference in Historical Probabilities (ΔP_χ) for each natural-unnatural alternation pair.

The BSC technique also enables the comparison of alternations within the unnatural group. Figure 2 shows bootstrap replicates of the individual Historical Probabilities of the three unnatural alternations, PND, IVD, and FV. The figure shows that the Historical Probability of PND is higher compared to the Historical Probabilities of the other two unnatural alternations. By estimating the difference between two alternations with BSC, we can test, for example, whether P_χ(PND) and P_χ(IVD) or P_χ(PND) and P_χ(FV) are significantly different. The estimated $\Delta P_\chi(\text{PND}, \text{IVD})$ and $\Delta P_\chi(\text{PND}, \text{FV})$ are computed as described above and in Section 3.2, and are given in (8).

(8) a. $\Delta P_\chi(\text{PND}, \text{IVD}) = P_\chi(\text{PND}) - P_\chi(\text{IVD}) = 0.026\%\ [-0.004\%, 0.064\%]$
b. $\Delta P_\chi(\text{PND}, \text{FV}) = P_\chi(\text{PND}) - P_\chi(\text{FV}) = 0.036\%\ [0.011\%, 0.074\%]$

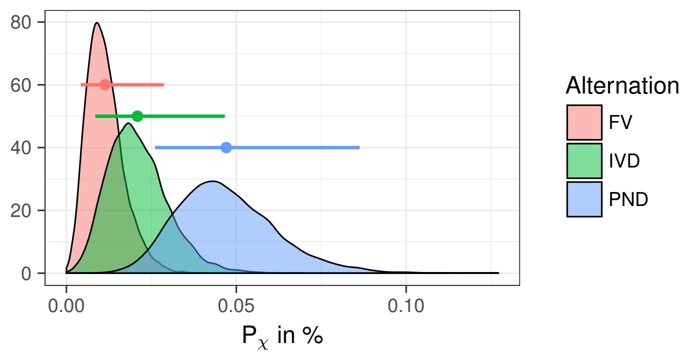

Figure 2: Bootstrap replicates for unnatural alternations with observed P_χ (colored dot) and 95% BC_a confidence intervals (colored lines).

Because the 95% BC_a CIs of the difference in Historical Probability between PND and FV lie above zero, it can be concluded that P_χ(PND) is significantly higher than P_χ(FV) (with $\alpha = 0.05$). The Historical Probabilities of IVD and PND are, however, not significantly different, because the BC_a CIs cross zero (see (8-b)).

Certainly, the BSC technique makes some simplifying assumptions that introduce confounds to the estimation of Historical Probabilities (see Sections 3.3, 3.4, and 4.1). Because differences in the Historical Probabilities between unnatural alternations are considerably smaller than differences between natural-unnatural pairs (Figure 1), estimation of these differences is substantially more prone to be influenced by these confounds. Until more comprehensive surveys are available, however, the BSC technique makes, to the author's knowledge, the most accurate approximations of Historical Probabilities of alternations, both for natural-unnatural alternation pairs as well as for alternations within the unnatural group.

4.3 Comparing P_χ to observed synchronic typology

4.3.1 Natural vs. unnatural processes

Predictions of the BSC model can be evaluated by comparing Historical Probabilities with independently observed typology of *synchronic* alternations. Estimation of synchronic typological probabilities faces even more difficulties and problematic assumptions than estimation of Historical Probabilities. The presence of an alternation that results from a sound change in two related languages cannot be counted as independent, although it is often treated as such in synchronic typological surveys. Moreover, language contact and linguistic areas likely influence observed synchronic typology to a greater degree compared to the typology of sound changes, although this observation would need a more elaborate evaluation.

For all these reasons, 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. Nevertheless, Historical Probabilities estimated with the BSC technique match the observed synchronic typology relatively well and, to the author's knowledge, better than alternative approaches (see Section 1.2). Table 9 compares Historical and observed synchronic probabilities. Historical Probabilities (P_χ) are estimated with the BSC technique as described above (see Section 4.1 and Table 7). The synchronic typology is estimated with a non-parametric bootstrap technique in the same way as described in Section 3.2.1, except that the estimation is based on the number of languages in a sample with a synchronic alternation and the number of languages in a sample without the synchronic alternation. To be sure, synchronic typology is estimated from surveys of synchronic alternations, rather than from diachronic surveys of sound changes. For PNV and IVV, the surveys in Locke (1983) (reported in Hayes and Stivers 2000) and Gurevich (2004) (reported in Kaplan 2010) are used. Because no systematic typologies of FD exist, this process is left out of the comparison. PNV is attested in 15 of 197 languages (Locke 1983, reported in Hayes and Stivers 2000), and IVV in 26 of 153 languages (Gurevich 2004, reported in Kaplan 2010).

The synchronic typology of unnatural processes is challenging to estimate because considerably higher standards of evidence are required to confirm a productive synchronic status for an unnatural alternation compared to a natural alternation, and because typological surveys of unnatural alternations are usually not performed in a systematic and controlled manner. Despite these obstacles, the following estimations of the synchronic typology of unnatural alternations can be computed (summarized in Table 9) based on surveys of unnatural processes in Author (2017) and Author and Name (2017). PND has been confirmed as a fully productive synchronic alternation in two related languages (Tswana and Shekgalagari) and as a morphophonological alternation in a few others (Buginese, Nasioi; see Author 2017). For the purpose of comparison, only fully productive alternations are counted in the synchronic typology. Because Tswana and Shekgalagari are closely related, PND here is counted as a single occurrence. IVD is attested only once as a morphologically conditioned synchronic process (Bloyd 2015), although detailed descriptions are lacking. FV is, to the author's knowledge, not attested as a productive phonological alternation in any language, which is why its synchronic typological probability is estimated below P($\frac{1}{600}$).[15] An approximate estimate of languages surveyed in these surveys of unnatural alternations is 600.

Table 9: A comparison of Historical Probabilities (P_χ) and observed synchronic typology (Typol.) with 95% BC_a CIs for natural and unnatural processes.

		95% BC_a CI			95% BC_a CI	
A_k	P_χ	Lower	Upper	Typol.	Lower	Upper
PNV	14.3	10.2	18.4	7.6	4.1	11.2
PND	0.05	0.02	0.9	0.17	0.0	0.5
IVV	9.5	6.1	12.9	17.0	11.1	22.9
IVD	0.02	0.008	0.05	0.17	0.0	0.5
FD	11.2	7.8	15.0	—	—	
FV	0.01	0.004	0.03	<0.17	0.0	0.5

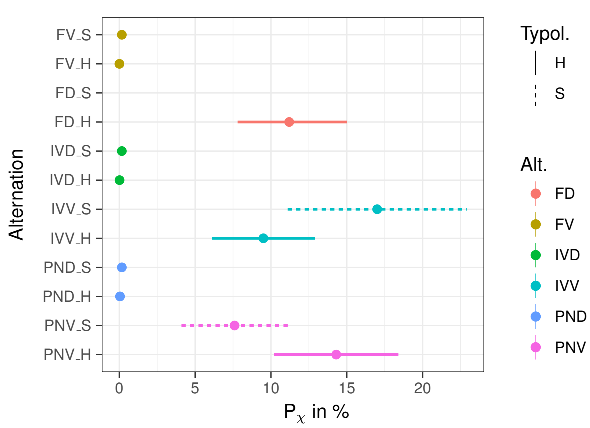

Figure 3: Observed Historical (H, solid line) and synchronic (S, dashed line) probabilities (in %) with 95% BC_a CIs from Table 9 estimated with BSC.

[15]If we counted the best candidate for FV, Lakota, as featuring fully productive unnatural alternations (Rood 2016), the typological probabilities of FV would be estimated at P($\frac{1}{600}$) = 0.17%.

Table 9 and the corresponding plot of estimated Historical and synchronic probabilities with 95% BC_a CIs in Figure 3 show that BSC correctly predicts natural alternations to be considerably more frequent than their unnatural alternations. Historical Probabilities and observed synchronic typology also match to the degree that the 95% BC_a confidence intervals of both Historical and synchronic typological probabilities always overlap for all five processes compared. It needs to be stressed here that for unnatural processes, the comparison of Historical Probabilities and observed synchronic typology is completely independent. In other words, the BSC technique estimates the probability of a combination of three sound changes, none of which are, by themselves, related to the unnatural synchronic alternation, from which synchronic typological probabilities are estimated.

4.3.2 Within the unnatural group

Inferential statements and predictions of the BSC method can be compared against the observed typology not only across natural-unnatural alternation pairs, but also within the unnatural group.

The BSC model predicts PND to be significantly more frequent than FV (see (8) and Figure 2). Based on a survey in Author (2017) that aims to collect all reported cases of PND, the Blurring Process that leads to PND is attested in thirteen languages. In at least two languages, the Blurring Process results in a productive synchronic alternation, but it is likely that other languages, such as Buginese and Nasioi, feature the process as a productive alternation as well. IVD results from a Blurring Process in two languages. In one language, the Blurring Process results in a gradient phonotactic restriction (according to the survey in Author and Name 2017). In one additional case, IVD is attested as a morphologically conditioned synchronic alternation (Bloyd 2015). Finally, FV is, to the author's knowledge, never attested as a combination of sound changes or as a synchronic phonological alternation (for reasons why Lezgian and other cases are not analyzed as featuring FV, see Kiparsky 2006, Lipp 2016, and Author 2018). Although the comparison is currently qualitative, the typology suggests that PND is indeed more frequent than FV, just as predicted by BSC.[16]

As already mentioned, alternations are distinguished from static phonotactics for the purposes of diachronic modeling (Section 4.1). Kiparsky (2006) lists a number of scenarios that would result in a static phonotactic restriction against voiceless stops with voiced stops surfacing word-finally. It is possible that some of his scenarios that require fewer than three sound changes would result in a productive unnatural phonotactic restriction, although this is less likely than in the case of unnatural alternations. It is also possible that CB alone cannot explain the relative rarity of FV and other unnatural phonotactic restrictions (as opposed to alternations). A further study is required to answer this question. A preliminary survey, however, suggests that FV as a static phonotactic restriction (as opposed to an alternation) might not be so rare, which would be expected under the BSC approach. Two languages might qualify as featuring this process: Ho and some varieties of Spanish (see Section 4.1 and Author 2017).

In sum, the CB factor estimated with the proposed BSC technique (Figure 3) correctly predicts natural alternations targeting the feature [±voice] to be significantly more frequent than unnatural alternations. Moreover, the BSC predictions match the observed synchronic typology to the degree that 95% BC_a confidence intervals overlap for all alternations. These are, to the author's knowledge, the most accurate predictions so far, especially considering the fact that the Historical Probabilities of unnatural alternations are estimated not directly from observed surface typology, but from the typology of natural sound changes that are independent of the unnatural result.

Kiparsky (2006, 2008) and others (de Lacy and Kingston 2013) claim that the CB approach to typology fails to explain why some processes, such as FV, are non-existent. BSC offers a potential

[16] For evaluation of further typological predictions of the BSC technique, see Sections 4.3 and 5.

solution to this problem. As was argued above, the BSC method predicts that some unnatural processes will be significantly less frequent than others (or even unattested). FV is, for example, predicted to be very rare or possibly unattested, whereas PND is predicted to be significantly more frequent. This prediction seems to match the observed synchronic typology.

5 Implications

The BSC technique not only estimates Historical Probabilities of different alternations, but also helps identify mismatches in predictions between the AB and CB approaches to typology. This section compares the outputs of the BSC model to the predictions of the AB approach to identify mismatches that help quantitatively estimate the influences of the two factors on typology.

If two typologically unequal 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 CB factor. On the other hand, if two typologically unequal alternations have equal Historical Probabilities and show differences in learnability, it is reasonable to assume that the differences result from the AB factor. In the case of the unnatural alternations PND, IVD, and FV, the BSC technique suggests that the observed typology is primarily influenced by the CB factor. The CB factor estimated with BSC predicts the typology with relatively high accuracy (Section 4.3.1 and Figure 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 BSC technique further enables identification of mismatches in the predictions of the AB and CB approaches, especially with respect to the complexity of alternations and their typological attestedness. BSC predicts not only that unnatural alternations will be rare (Section 4), but also that, all else being equal, complex alternations will be less frequent than simple alternations. The minimality principle (Donegan and Stampe 1979, Picard 1994, Author 2017, and Section 2), which is at least a strong tendency, states that sound change is a change in one feature (or the deletion/reordering of feature matrices) in a given environment. This means that featurally complex alternations that change more than a single feature need to arise from the phonologization of more than one sound change. Because the probability of a combination of two 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 CB approach. Exactly the same generalization is, however, also predicted by the AB approach to typology: numerous studies have confirmed that featurally complex alternations are consistently underlearned compared to featurally simple alternations (Complexity Bias; Moreton and Pater 2012a,b).

There is a crucial mismatch in predictions between the AB and CB approaches with respect to unnatural alternations. The BSC technique makes the following predictions: the more sound changes an alternation requires, the lower the Historical Probability of that alternation, regardless of its complexity (see Table 10). In other words, the BSC prediction that complex alternations will be rare is violable: if the three sound changes of a Blurring Process result in a simple unnatural alternation, BSC still predicts 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 (the MSCR).

We can estimate the Historical Probabilities for each step in the Blurring Process that leads to unnatural alternations. Let us take as an example PND. The Historical Probabilities of each resulting alternation (after the first, second, and third sound changes of the Blurring Process

Table 10: Mismatches in predictions (framed) between the Channel Bias approach (P_χ) and the Complexity Bias approach ($P_{\mathbf{cplx}}$) for PND. The *Sound Change* column represents the three sound changes from which the unnatural process PND results, and the *Alternation* column represents the synchronic alternation after each of the three sound changes. The P_χ column gives the estimated probability of each alternation based on the BSC technique with 95% BC_a Lower and Upper CIs (Lo. and Up.). The *Features* column counts the number of features a learner has to learn for each synchronic alternation.

Sound change	**Alternation**	$\mathbf{P}_\chi$	**Lo.**	**Up.**	**Features**	$\mathbf{P}_\chi$	$\mathbf{P_{cplx}}$
	No alternation	65.9	61.6	72.4	0		
D > Z / [−nas]__	D → Z / [−nas]__	33.0	27.6	38.4	1	↓	↓
D > T	Z → T / [+nas]__	1.1	0.7	1.8	2	↓	↓
Z > D	PND	0.05	0.03	0.09	1	↓	↑

operate) were estimated with BSC as described in Section 3.[17] Table 10 (column P_χ) clearly shows that each additional sound change decreases the Historical Probability of the resulting alternation.

On the other hand, the AB 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 almost without exception in many studies; Moreton and Pater 2012a,b.) If we analyze each step in the Blurring Process in terms of synchronic complexity, the first two sound changes in the Blurring Process indeed increase the complexity of the resulting alternation,[18] but the third sound change decreases its complexity. 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 comparatively more frequent than the complex alternation requiring only two sound changes. Let us call this prediction the *AB-CB complexity mismatch.* Since no learnability differences are observed for natural-unnatural alternation pairs, the AB approach makes no predictions about the relative rarity of unnatural alternations as opposed to natural alternations.

The mismatched predictions of BSC and Complexity Bias illustrated in Table 10 provide crucial new information for disambiguating AB and CB biases. The AB-CB 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, CB 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 CB approach compared to structurally simpler unnatural alternations, we have a strong case in favor of the AB influence, and more precisely in favor of Complexity Bias within the AB approach to typology.

In fact, typological observations suggest that the complex synchronic alternation Z → T / [+nas]__ that results from the first two sound changes in a Blurring Process might be attested less frequently than would be predicted by CB, suggesting that Complexity Bias influences this distri-

[17]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.

[18]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 AB-CB conflation problem that states that sound change probabilities are primarily influenced by learnability and hence that estimated CB influences are crucially conflated with AB influences. If anything, AB 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 operating are not crucially influenced by AB (although AB can of course still influence the relative frequencies of sound change).

bution. The Historical Probability of Z → T / [+nas]__ is significantly higher than the Historical Probability of PND. The difference is estimated with BSC at $\Delta\mathrm{P}_\chi(\mathrm{Z} \rightarrow \mathrm{T} \ / \ [+\mathrm{nas}]\_\_, \mathrm{PND}) = 1.1\%, \ [0.6\%, 1.7\%]$. In other words, the Historical Probability of the alternation Z → T / [+nas]__ that arises through two sound changes is predicted to be approximately twenty times more frequent than the Historical Probability of PND (see Table 10). Surface synchronic typology, however, does not conform to this generalization.

A system in which post-nasal 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, and potentially Nasioi (Dickens 1984; Hyman 2001; Merrill 2014, 2016a,b; Santos 1996; Brown 2017).[19] Other languages are more difficult to classify because some of them appear to feature full PND only for a subset of places of articulation. While Z → T / [+nas]__ indeed appears to be more frequent than PND, the magnitude of the difference appears to be smaller than predicted by BSC.

Even more intriguing is the high frequency at which the third sound change in the Blurring Process, occlusion of voiced fricatives to stops (Z > D), operates on synchronic systems that feature the alternation Z → T / [+nas]__ (after the first two changes in the Blurring Process). The Historical Probability of the third sound change in the Blurring Cycle that leads to PND, 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 $\mathrm{P}_\chi(Z > D) = 12.5\%, \ [7.9\%, 17.1\%]$. Of the languages in the survey in Author (2017) that undergo the first two sound changes in the Blurring Process, which leads to PND, six languages (out of eight, or approximately 75%)[20] feature occlusion of stops for at least one place of articulation or in at least one position in the word. If we count 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 of the Blurring Cycle is more frequent than BSC 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 CB approach, we can compare the unconditioned Historical Probability of the occlusion of fricatives with the Historical Probability of the occlusion of fricatives in those languages that have already undergone the first two sound changes in the Blurring Cycle that lead to PND. In other words, we compare the probability of the occlusion of fricatives regardless of whether it simplifies the alternation (assuming only the CB influences) with the probability of the occlusion of fricatives operating in the Blurring Process, where it simplifies the alternation and consequently its learnability. Because both of these estimations involve a single sound change and because the second sample is small (eight observations), the significance of the difference is tested using Fisher's Exact Test. Counts for the unconditioned Historical Probability of the occlusion of fricatives is based on the survey of sound changes in Kümmel (2007). Twenty-seven 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 eight languages in the Blurring Cycle show occlusion for at least one place of articulation or for at least one context (word-initially in Nasioi). The difference between the two counts is statistically significant ($p < 0.001$, Fisher's Exact Test).

[19] Punu is a language that undergoes a different development from the one described in Section 2. In Punu, the resulting alternation is not PND but the complex alternation between voiced fricatives elsewhere and voiceless stops post-nasally. For a discussion, see Hyman (2001).

[20] PND occurrences in Tswana, Shekgalagari, and Makhuwa 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 PND there. I include Pedi that features Z → T / [+nas] (Dickens 1984).

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 only assumed CB influences.[21]

This suggests that the high occurrence of the third sound change in the Blurring Process (in the case of PND, the occlusion of fricatives) is likely an influence of Complexity Bias within the AB approach. While AB likely does not crucially influence the probabilities of the first two sound changes in the Blurring Process because they increase complexity and therefore lower learnability, it is likely that the occurrence of the third sound change, and therefore the lower probability of the more complex unmotivated alternation, is influenced precisely by Complexity Bias.

Mismatches in predictions between the AB and CB approaches identified by BSC can thus shed new light on the discussion of AB vs. CB influences on typology. Based on the comparison of estimated Historical Probabilities with the observed synchronic typology, this paper suggests that the typological rarity of unnatural alternations targeting the feature [±voice] likely results from CB (Figure 3). On the other hand, the typological rarity of complex processes and the higher rate of sound changes that simplify an alternation is likely influenced by AB (Table 10).

These results have immediate theoretical consequences. One of the advantages of the Optimality Theory family of approaches to phonology (Prince and Smolensky 1993/2004) is the derivation of the so-called factorial typology, with which the theory predicts possible processes and rules out impossible ones. This paper suggests that some alternations are not impossible, but are simply unattested due to the CB factor. Further, the BSC model suggests that some alternations are rare not due to any grammatical constraints, but because they have no phonetic precursors and because a number of sound changes, each with some probability, need to operate to produce those processes (the MSCR). In fact, when the learnability of the unnatural alternations PND, IVD, and FV was tested against their natural counterparts, no significant differences were observed, suggesting that there are no universal synchronic grammatical constraints against these processes. On the other hand, this paper also shows that some typological observations, such as the avoidance of complex alternations, cannot be explained only within the CB approach and that these preferences should indeed be encoded in synchronic grammar. Applying the BSC technique on further natural and unnatural alternations combined with experimental work should yield further results informative for phonological theory. Finally, the BSC technique provides a quantitative input that can be used in theoretical models to combine the AB influences with the CB factor (see Author 2017b).

6 Conclusion

This paper proposes a technique for estimating channel bias influences on phonological typology called Bootstrapping Sound Changes. The BSC technique estimates Historical Probabilities of alternations that are based on two diachronic factors: the number of sound changes required for an alternation to arise (the MSCR) and their respective probabilities. The paper provides a detailed description of the statistical model and discusses its assumptions, properties of the sample, and implementation. This paper also includes functions in the statistical software R (R Core Team 2016) for performing the BSC analysis.

Several applications of the BSC technique are presented. The BSC technique (i) estimates the Historical Probability of any synchronic alternation, both attested and unattested, (ii) compares the Historical Probabilities of two alternations and performs inferential tests on the comparison, and (iii) compares the Historical Probabilities to independently observed synchronic typology to

[21]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 get blocked by the grammar.

evaluate the CB influences on typology. Finally, the BSC technique identifies mismatches in predictions between the AB and CB approaches, which yields new insights into the discussion of different influences on phonological typology.

The Historical Probabilities of three natural-unnatural alternation pairs that target the feature [±voice] are estimated using the BSC technique. First, it is shown that BSC predicts the observed typology with relatively high accuracy — to the degree that at least 95% BC_a confidence intervals overlap for all alternations. This is especially relevant for explaining typological differences between natural and unnatural alternations, which pose a problem for the AB approach to typology — three previous studies that tested the learnability of these natural-unnatural alternation pairs found no significant differences. No other proposal known to the author predicts with significance that unnatural alternations will be substantially less frequent than natural alternations and at the same time predicts that some unnatural alternations (such as FV) will be significantly less frequent than others (such as PND), a situation that is substantiated by the independently observed typology. In other words, the BSC technique derives the observed typology relatively accurately not only for the natural-unnatural alternation pairs, but also within the unnatural group.

As already mentioned, BSC also identifies crucial mismatches in predictions between the AB and CB approaches to typology. Both AB and CB approaches predict that complex alternations will be less frequent than simple alternations, but within the CB approach this prediction can be violated in the case of unnatural vs. unmotivated alternations. This paper shows that occlusion of voiced fricatives operates significantly more frequently as the last sound change in the Blurring Process, where it simplifies an alternation, compared to its operation in an unconditioned sample (where no simplification occurs). In other words, the sound change that simplifies a complex alternation operates significantly more frequently than it would as predicted by only the CB approach, suggesting that the AB factor is responsible for the typological distribution.

The results suggest that the typological difference between natural and unnatural alternations targeting the feature [±voice] is primarily due to CB, but that the relatively low frequency of complex alternations and the higher rate of the operation of sound changes that simplify an alternation are due to AB.

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. Estimation of the CB and AB influences should thus 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. The BSC model hopes to provide a step in this direction.

References

Adelaar, Willem F. H. 1977. *Tarma Quechua: Grammar, Texts, Dictionary.* Lisse: The Peter de Ridder Press.

Barnes, Jonathan. 2002. Positional neutralization: a phonologization approach to typological patterns. Ph.D. dissertation, University of California, Berkeley.

Becker, Michael, Nihan Ketrez, and Andrew Nevins. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87(1): 84-125.

Bell, Alan. 1970. A state-process approach to syllabicity and syllabic structure. Ph.D. Dissertation, Stanford University.

Bell, Alan. 1971. Some patterns of the occurrence and formation of syllabic structure. *Working Papers on Language Universals* 6: 23-138.

Blevins, Juliette. 2004. *Evolutionary Phonology.* Cambridge: Cambridge University Press.

Blevins, Juliette. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32: 117-165.

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 *Naturalness and Iconicity in Language*, edited by Klaas Willems and Ludovic De Cuypere. 121-148. Amsterdam: Benjamins.

Blevins, Juliette. 2008b. Consonant epenthesis: natural and unnatural histories. In *Language Universals and Language Change*, edited by Jeff Good. 79-107. Oxford: Oxford University Press.

Blevins, Juliette. 2013. Evolutionary Phonology: A holistic approach to sound change typology. In *Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons. Oxford: Oxford University Press.

Bloyd, Tobias. 2015. Toward a Phonological Reconstruction of Proto-Sula. *University of Hawai'i at Mānoa: Working Papers in Linguistics* 46(8): 1-23.

Bloyd, Tobias. 2017. Synchronic intervocalic fortition in Sula: a counter-universal. Talk at the *91st Annual Meeting of the Linguistic Society of America*, January 5-8, 2017, Austin, TX.

Blust, Robert. 2005. Must sound change be linguistically motivated? *Diachronica* 22(2): 219-269.

Blust, Robert. 2013. *The Austronesian Languages.* Canberra: Asia-Pacific Linguistics.

Bouchard-Côté, Alexandre, David Hall, Thomas L. Griffiths, and Dan Klein. 2013. Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences*, 110(11): 4224-4229.

Broselow, 2018. Laryngeal contrasts in second language phonology. In *Phonological typology*, edited by Larry M. Hyman and Frans Plank. 312-340. Berlin: De Gruyter.

Brown, Jason. 2017. Postnasal devoicing in Nasioi. *Oceanic Linguistics* 56(1): 267-277.

Buckley, Eugene. 2000. On the naturalness of unnatural rules. In *Proceedings from the Second Workshop on American Indigenous Languages. UCSB Working Papers in Linguistics* 9: 1-14.

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 and Brian Ripley. 2016. *boot: Bootstrap R (S-Plus) Functions.* R package version 1.3-18.

Carpenter, Angela C. 2006. Acquisition of a natural versus an unnatural stress system. Ph. D. Dissertation, University of Massachusetts, Amherst.

Carpenter, Angela C. 2010. A naturalness bias in learning stress. *Phonology* 27(3): 345-392.

Cathcart, Chundra A. 2015. A probabilistic model of Evolutionary Phonology. *Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society.* Vol. 1, edited by Thuy Bui and Deniz Özyıldız. 145-150. Amherst, MA: GLSA.

Chen, Matthew Y. 1974. Natural phonology from the diachronic vantage point. In *Papers from the Parasession on Natural Phonology*, edited by A. Bruck, R. A. Fox, and M. W. LaGaly. 21-29. Chicago, IL: Chicago Linguistic Society.

Clark, Eve V., and Melissa Bowerman. 1986. On the acquisition of final voiced stops. In *The Fergusonian Impact: In Honor of Charles A. Ferguson on the Occasion of his 65th Birthday.*, edited by J. A. Fishman. 51-68. Volume 1: From Phonology to Society. Berlin: De Gruyter.

Coetzee, Andries W. and Rigardt Pretorius. 2010. Phonetically grounded phonology and sound change: The case of Tswana labial plosives. *Journal of Phonetics* 38: 404-421.

Davison, A. C. and D. V. Hinkley. 1997. *Bootstrap Methods and Their Applications.* Cambridge: Cambridge University Press.

de Lacy, Paul and John Kingston. 2013. Synchronic explanation. *Natural Language and Linguistic Theory.* 31(2): 287-355.

Dickens, Patrick J. 1984. The history of so-called strengthening in Tswana. *Journal of African Languages and Linguistics* 6:97-125.

Do, Youngah, Elizabeth Zsiga, and Jonathan Havenhill. 2016. Naturalness and frequency in implicit phonological learning. Talk presented at the *90th Annual Meeting of the Linguistic Society of America*, Washington, DC, January 7-10, 2016.

Donegan, Patricia J. and David Stampe. 1979. The study of Natural Phonology. In *Current Approaches to Phonological Theory*, edited by Daniel Dinnsen. 126-173. Bloomington: Indiana University Press.

Efron, Bradley. 1979. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* 7(1): 1-26.

Efron, Bradley. 1987. Better bootstrap confidence intervals. *Journal of the American Statistical Association* 82(397): 171–185. doi:10.2307/2289144.

Garrett, Andrew. 2014. Sound change. In *The Routledge Handbook of Historical Linguistics*, edited by Claire Bowern and Bethwyn Evans. 227-248. New York: Routledge.

Garrett, Andrew and Keith Johnson. 2013. Phonetic bias in sound change. In *Origins of Sound Change: Approaches to Phonologization*, edited by Alan C. L. Yu. 51-97. Oxford: Oxford University Press.

Glewwe, Eleanor. 2017. Substantive bias in phonotactic learning: Positional extension of an obstruent voicing contrast. Talk presented at the *53rd meeting of Chicago Linguistic Society*, Chicago, IL, May 25-27, 2017.

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, UT, January 4-7, 2018.

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 *Universals of Human Language*, edited by Joseph H. Greenberg. Volume I: Method & Theory. 61-92. Stanford, CA: Stanford University Press.

Greenwood, Anna. 2016. An experimental investigation of phonetic naturalness. Ph.D. Dissertation, University of California, Santa Cruz.

Gurevich, Naomi. 2004. *Lenition and contrast: The functional consequences of certain phonetically conditioned sound changes.* New York: Routledge.

Hamed, Mahé B. and Sébastien Flavier. 2009. UNIDIA: A database for deriving diachronic universals. In *Historical Linguistics 2007: Selected Papers from the 18th International Conference on Historical Linguistics, Montreal, 6–11 August 2007*, edited by Monique Dufresne, Fernande Dupuis, and Etleva Vocaj. 259-268. Amsterdam: John Benjamins.

Hansson, Gunnar. 2008. Diachronic explanations of sound patterns. *Language and Linguistics Compass* 2:859-893.

Haspelmath, Martin. 1993. *Lezgian grammar*. Berlin: Mouton de Gruyter.

Hay, Jennifer B., Janet B. Pierrehumbert, Abby J. Walker, and Patrick LaShell. 2015. Tracking word frequency effects through 130 years of sound change. *Cognition* 139: 83-91.

Hayes, Bruce. 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In *Functionalism and Formalism in Linguistics, Volume I: General Papers*, edited by Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly. 243-285. Amsterdam: John Benjamins.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: The early stages. In *Constraints in Phonological Acquisition*, edited by R. Kager, J. Pater and W. Zonneveld. 153-208. Cambridge: Cambridge University Press.

Hayes, Bruce and James White. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44(1): 45-75.

Hayes, Bruce, Péter Siptár, Kie Zuraw, and Zsuzsa Londe 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85(4): 822-863.

Hayes, Bruce, and Tanya Stivers. 2000. *Postnasal Voicing.* Ms. Available May 9, 2018 at: `http://linguistics.ucla.edu/people/hayes/Phonet/NCPhonet.pdf`

Hruschka, Daniel J., Simon Branford, Eric D. Smith, Jon Wilkins, Andrew Meade, Mark Pagel, and Tanmoy Bhattacharya. 2015. Detecting regular sound changes in linguistics as events of concerted evolution. *Current Biology* 25(1): 1-9.

Hyman, Larry M. 1975. *Phonology: Theory and Analysis.* New York: Holt, Rinehart & Winston.
Hyman, Larry M. 2001. The limits of phonetic determinism in phonology: *NC revisited. In *The Role of Speech Perception in Phonology*, edited by Elizabeth Hume and Keith Johnson. 141–186. San Diego, CA: Academic Press.
Hyman, Larry M. 2013. Enlarging the scope of phonologization. In *Origins of Sound Change: Approaches to Phonologization*, edited by Alan C. L. Yu. 3-28. Oxford: Oxford University Press.
Iverson, Gregory K. and Joseph C. Salmons. 2011. Final devoicing and final laryngeal neutralization. In *The Blackwell Companion to Phonology: Suprasegmental and Prosodic Phonology.* Volume 2, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice. 1622-1643. Malden: Wiley-Blackwell.
Janson, Tore. 1991/1992. Southern Bantu and Makua. *Sprache und Geschichte in Afrika* 12/13: 1-44.
Janssens, Baudouin. 1993. Doubles réflexes consonantiques: quatre études sur le bantou de zone A (bubi, nen, bafia, ewondo). Ph.D. Dissertation, Université libre de Bruxelles, Faculté de Philosophie et Lettres, Bruxelles.
Kapatsinski, Vsevolod. 2011. Modularity in the channel: The link between separability of features and learnability of dependencies between them. In *Proceedings of the XVIIth International Congress of Phonetic Sciences*, edited by W. S. Lee and E. Zee. 1022-1025. Hong Kong.
Kaplan, Abby. 2010. Phonology shaped by phonetics: The case of intervocalic lenition. Ph.D. Dissertation, University of California, Santa Cruz.
Kiparsky, Paul. 1995. The phonological basis of sound change. In *Handbook of Phonological Theory*, edited by John Goldsmith. 640-670. Oxford: Blackwell, 1995.
Kiparsky, Paul. 2006. Amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics* 32(2): 217-236.
Kiparsky, Paul. 2008. Universals constrain change, change results in typological generalizations. In *Linguistic universals and language change*, edited by Jeff Good. 23-53. Oxford: Oxford University Press.
Kirby, James and Morgan Sonderegger. 2013. A model of population dynamics applied to phonetic change. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, edited by Markus Knauff, Michael Pauen, Natalie Sebanz, and Ipke Wachsmuth. 776-781. Austin, TX: Cognitive Science Society.
Kirby, James and Morgan Sonderegger. 2015. Bias and population structure in the actuation of sound change. *arXiv*: 1507.04420.
Kong, Eun Jong, Mary E. Beckman, and Jan Edwards. 2012. Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese. *Journal of Phonetics* 40(6): 725-744.
Kümmel, Martin. 2007. *Konsonantenwandel.* Wiesbaden: Reichert.
Kuo, Li-Jen. 2009. The role of natural class features in the acquisition of phonotactic regularities. *Journal of psycholinguistic research* 38(2): 129-150.
Labov, William. 1994. *Principles of Linguistic Change.* Volumes 1 and 2. Oxford: Blackwell.
Liljencrants, Johan and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48(4): 839-862.
Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In *Speech Production and Speech Modelling*, edited by William J. Hardcastle and Alain Marchai. 403-439. Dordrecht: Kluwer.
Lindblom, Björn. 1986. Phonetic universals in vowel systems. In *Experimental Phonology*, edited by John J. Ohala and Jeri J. Jaeger. 13-44. Orlando: Academic Press.
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.
Locke, John. 1983. *Phonological Acquisition and Change.* New York: Academic Press.
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. Talk presented at the *90th Annual Meeting of the Linguistic Society of America*, Washington, DC, January 7-10, 2016.
Merrill, John. 2016b. Konyagi post-nasal devoicing? Ms., University of California, Berkeley.

Moran, Steven, Daniel McCloy, and Richard Wright. 2014. *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at phoible.org on May 6, 2018.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25(1): 83–127.

IMoreton, Elliott. 2012. Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* 67(1): 165-183.

Moreton, Elliott, and Joe Pater 2012a. Structure and substance in artificial-phonology learning. Part I, Structure. *Language and Linguistics Compass* 6 (11): 686-701.

Moreton, Elliott, and Joe Pater 2012b. Structure and substance in artificial-phonology learning. Part II, Substance. *Language and Linguistics Compass* 6 (11): 702-718.

Morley, Rebecca L. 2012. The emergence of epenthesis: An incremental model of grammar change. *Language Dynamics and Change* 2: 59-97.

Morley, Rebecca L. 2015. Can phonological universals be emergent? Modeling the space of sound change, lexical distribution, and hypothesis selection. *Language* 91(2): e40-e70.

Mouguiama-Daouda, P. 1990. Esquisse d'une phonologie diachronique du Mpongwe. *Pholia* 5: 121-146.

Myers, Scott. 2002. *Gaps in factorial typology: the case of voicing in consonant clusters.* Ms., University of Texas, Austin. ROA.

Ohala, John J. 1981. The listener as a source of sound change. In *Papers from the Parasession on Language and Behavior*, edited by Carrie S. Masek, Robert A. Hendrick, and Mary Frances Miller. 178–203. Chicago: Chicago Linguistic Society.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, edited by Peter F. MacNeilage. 189-216. New York: Springer.

Ohala, John J. 1993. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, edited by Charles Jones. 237-278. London: Longman.

Ohala, John J. 2011. Accommodation to the aerodynamic voicing constraint and its phonological relevance. In *Proceedings of the 17th International Congress of Phonetic Sciences*. 64-67.

Pater, Joe and Anne-Michelle Tessier. 2006. L1 phonotactic knowledge and the L2 acquisition of alternations. In *Inquiries in Linguistic Development: In honor of Lydia White*, edited by Roumyana Slabakova, Silvina A. Montrul and Philippe Prévost. 115-131. Amsterdam: John Benjamins.

Picard, Marc. 1994. *Principles and Methods in Historical Phonology: From Proto-Algonkian to Arapaho.* Montreal: McGill-Queen's University Press.

Prince, Alan. and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell. First published in 1993, Tech. Rep. 2, Rutgers University Center for Cognitive Science.

Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted. 2003. Phonological rule-learing and its implications for a theory of vowel harmony. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, edited by Gina Garding and Mimu Tsujimura. 101-114. Somerville, MA: Cascadilla Press.

R Core Team 2016. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rafferty Anna N., Thomas L. Griffiths, and Marc Ettlinger. 2013. Greater learnability is not sufficient to produce cultural universals. *Cognition* 129(1): 70-87.

Rohlfs, Gerhard. 1949. Lautlehre. Vol. 1 of *Historische Grammatik der Italienischen Sprache und ihrer Mundarten*. Bern: Francke.

Rood, David. 2016. The phonology of Lakota voiced stops. In *Advances in the Study of Siouan Languages and Linguistics*, edited by Catherine Rudin and Bryan J. Gordon. 233-255. Berlin: Language Science Press.

Ryan, Kevin. To appear. *Prosodic weight: categories and continua.* Oxford University Press.

Santos, Rosine. 1996. *Le Mey: langue ouest-atlantique de Guinée.* Paris: Université Paris III.

Seidl, Amanda, Eugene Buckley, and Alejandrina Cristià. 2007. Complexity trumps naturalness. Talk presented at the *81st Annual Meeting of the Linguistic Society of America*, Anaheim, CA, January 4-7, 2007.

Skoruppa, Katrin and Sharon Peperkamp. 2011. Adaptation to novel accents: feature-based learning of context-sensitive phonological regularities. *Cognitive Science* 35: 348-366.

Solé, Maria-Josep, Larry M. Hyman, and Kemmonye C. Monaka. 2010. More on post-nasal devoicing: The case of Shekgalagari. *Journal of Phonetics* 38(4): 299-319.

Tesar, Bruce and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Wedel, Andrew. 2012. Lexical contrast maintenance and the organization of sublexical contrast systems. *Language and Cognition* 4(4): 319-355.

Wedel, Andrew, Abby Kaplan, and Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2): 179-186,

Westbury, John R. and Patricia A. Keating 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22(1): 145-166.

White, James. 2017. Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language* 93(1): 1-36.

Wilson, Colin. 2006. Learning Phonology with Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science* 30: 945-982.

Xromov, Al'bert. 1972. *Jagnobskij jazyk*. Moscow: Nauka.

Yu, Alan C. L. 2004. Explaining final obstruent voicing in Lezgian: Phonetics and history. *Language* 80: 73–97.

Yu, Alan C. L. 2011. On measuring phonetic precursor robustness: a response to Moreton. *Phonology* 28(3): 491-518.

Yu, Alan C. L. 2013. Individual differences in socio- cognitive processing and sound change. In *Origins of Sound Change: Approaches to Phonologization*, edited by Alan C. L. Yu. 201-227. Oxford: Oxford University Press.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83(2): 277-316.

A Supplementary materials

A.1 bsc()

The function *bsc()* takes two vectors of equal length as arguments: a vector with counts of languages with a sound changes required for an alternation A_k, and a vector of languages surveyed for each sound change. The function internally transforms the vectors with counts into a binomial distribution of successes and failures for each sound change in the count. It returns R bootstrap replicates of the Historical Probability of A_1, computed according to (3), (4), (5), and (6). Stratified non-parametric bootstrapping is performed based on the *boot* package: the output of *bsc()* is an object of class "boot". The output of *bsc()* should be used as an argument of *summary.bsc()* (see A.3), which returns the observed P_χ and 95% BC_a CIs. Two optional arguments of *bsc()* are *order* (if True, Historical Probabilities are divided by $n!$) and R, which determines the number of bootstrap replicates.

```r
bsc <- function (counts, surveyed, order = T, R = 10000) {
  library(boot)
  if (length(counts) != length(surveyed)) {stop
    ("Vectors must be of equal length.")
  }
  binom <- unlist(mapply(c,
                         lapply(counts, function(x) rep(1, x)),
                         lapply(surveyed - counts, function(x) rep(0, x)),SIMPLIFY=F)
  )
  snumb <- paste("s", 1:length(surveyed), sep="")
  ident <- rep(snumb, surveyed)

  scsample <- data.frame(binom,ident)

  if (order == TRUE) {n <- factorial(length(counts))}
  if (order == FALSE) {n <- 1}

  bsc <- function(x, id) {
    sc1 <- tapply(x[id,1], x[id,2], mean)
    sc <- prod(sc1) / n
    return(sc)
  }

  boot.scsample <- boot(scsample, statistic = bsc, R, strata = scsample[, 2]
                        )
  return(boot.scsample)
}

# Example:
pnd.counts <- c(97, 18, 27)
pnd.surveyed <- c(294, 263, 216)

pnd <- bsc(pnd.counts, pnd.surveyed)
summary.bsc(pnd)

# Output:
##BOOTSTRAPPING SOUND CHANGES
##
##Observed P = 0.04704 %
##Estimated 95 % BCa CI = [ 0.0261 %, 0.0862 %]
```

A.2 bsc2()

The function *bsc2()* compares the Historical Probabilities of two processes with BSC. It takes as an input the output of *bsc()* for the process in question. The function transforms the counts into a binomial distribution of successes and failures. It returns R bootstrap replicates of the difference in Historical Probability between the two alternations, computed according to (3), (4), (5), (6), and (7). Stratified non-parametric bootstrapping is performed based on the *boot* package: the output of *bsc2()* is an object of class "boot". The output of *bsc2()* should be used as an argument of *summary.bsc2()* (see A.4), which returns the observed ΔP_{χ} and 95% BC_a CIs for the difference. If 95% BC_a CIs fall above or below zero, it spells out that the difference is significant, and that it is not otherwise. Two optional arguments of *bsc()* are *order* (if True, Historical Probabilities are divided by $n!$) and R, which determines the number of bootstrap replicates.

```
bsc2 <- function(bsc.alt1a, bsc.alt2a, order = T, R = 10000){
  library(boot)
  bsc.alt1 <- bsc.alt1a$data
  bsc.alt2 <- bsc.alt2a$data
  bsc.alt1$scid <- "first"
  bsc.alt2$scid <- "second"
  bsc.diff.df <- rbind(bsc.alt1,bsc.alt2)
  bsc.diff.df$comb <- as.factor(paste(bsc.diff.df$scid,bsc.diff.df$ident, sep = ""
    ))

  bsc.diff.df$scid <- NULL
  bsc.diff.df$ident <- NULL

  if (order == TRUE) {  n1 <- factorial(length(unique(bsc.alt1$ident)))
  n2 <- factorial(length(unique(bsc.alt2$ident)))}
  if (order == FALSE) { n1 <- 1
  n2 <- 1}

  l <- length(unique(bsc.alt1$ident))
  m <- length(unique(bsc.alt2$ident))

  bsc.diff <- function(x, id) {
    sc1 <- tapply(x[id,1], x[id,2], mean)
    sca <- (prod(sc1[1:l]) / n1)
    scb <- (prod(sc1[(l+1):(l+m)]) / n2)
    sc <- sca - scb
    return(sc)
  }

  boot.diff <- boot(bsc.diff.df, statistic = bsc.diff, R, strata = bsc.diff.df[,
    2]
                  )
  return(boot.diff)
}

# Example:
pnd.counts <- c(97,18,27)
pnd.surveyed <- c(294,263,216)

fv.counts <- c(6,32,27)
fv.surveyed <- c(294,294,88)

pnd <- bsc(pnd.counts, pnd.surveyed)
fv <- bsc(fv.counts, fv.surveyed)

pndfv <- bsc2(pnd, fv)
summary.bsc2(pndfv)
```

```r

#Output:
##BOOTSTRAPPING SOUND CHANGES - COMPARE
##
##Observed Delta P = 0.03568 %
##Estimated 95 % BCa CI = [ 0.0114 %, 0.0744 %]
##
##P(A1) is significantly higher than P(A2).
```

A.3 summary.bsc()

The function *summary.bsc()* computes the 95% BC_a CI for the bootstrap replicates based on the *bsc()* function (see A.1) using the *boot.ci()* function from the *boot* package and returns the observed and estimated Historical Probabilities. For details, see A.1.

```r
summary.bsc <- function (bsc.alt) {
  bsc.ci.alt <- boot.ci(bsc.alt, type="bca")
  title <- "BOOTSTRAPPING SOUND CHANGES"
  prob <- paste("Estimated P =", round(bsc.alt$t0*100, digits = 5), "%")
  bca <- paste("Estimated 95 % BCa CI = [", round(bsc.ci.alt$bca[4]*100, digits =
      4),"%,",
               round(bsc.ci.alt$bca[5]*100, digits = 4),"%]")
  #rnsc <- paste(pasteR, n.sc.paste, countsp, surveyed, sep = "\n")
  probbca <- paste(prob, bca, sep = "\n")
  cat(title, probbca, sep = "\n\n")
}
```

A.4 summary.bsc2()

The function *summary.bsc2()* computes the 95% BC_a CI for the bootstrap replicates based on the *bsc2()* function (see A.2) using the *boot.ci()* function from the *boot* package and returns the observed and estimated differences in Historical Probabilities of two alternations. For details, see A.1.

```r
summary.bsc2 <- function (bsc2.alt) {
  bsc2.ci.alt <- boot.ci(bsc2.alt, type="bca")
  title <- "BOOTSTRAPPING SOUND CHANGES - COMPARE"
  prob <- paste("Estimated",expression(Delta), "P =", round(bsc2.alt$t0*100,
      digits = 5), "%")
  bca <- paste("Estimated 95 % BCa CI = [", round(bsc2.ci.alt$bca[4]*100, digits =
       4),"%,",
               round(bsc2.ci.alt$bca[5]*100, digits = 4),"%]")
  if (bsc2.ci.alt$bca[4] > 0 & bsc2.ci.alt$bca[5] > 0) {
    sig <- "P(A1) is significantly higher than P(A2)."
  }
  else if (bsc2.ci.alt$bca[4] < 0 & bsc2.ci.alt$bca[5] < 0) {
    sig <- "P(A1) is significantly lower than P(A2)."
  } else {
    sig <- "P(A1) and P(A2) are not significantly different."
  }
  probbca <- paste(prob, bca, sep = "\n")
  cat(title, probbca,sig, sep = "\n\n")
}
```