Unnatural Phonology:
A Synchrony-Diachrony Interface Approach

A dissertation presented
by
Gašper Beguš
to
The Department of Linguistics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Linguistics

Harvard University
Cambridge, Massachusetts

May 2018

Dissertation Advisor: Professor Kevin Ryan Author: Gašper Beguš

Unnatural Phonology:
A Synchrony-Diachrony Interface Approach

Abstract

This dissertation addresses one of the most contested topics in phonology: which factors influence phonological typology and how to disambiguate these factors. I propose a new framework for modeling the influences of Analytic Bias and Channel Bias on phonological typology. The focus of the dissertation is unnatural alternations — those that operate in precisely the opposite direction from some universal and articulatorily- or perceptually-motivated phonetic tendency. Based on a typological survey of unnatural alternations and gradient phonotactic restrictions, I propose a diachronic device for explaining unnatural processes called the *Blurring Process* and argue that minimally three sound changes are required for an unnatural segmental process to arise (*Minimal Sound Change Requirement*; MSCR). Based on the Blurring Process and MSCR, I propose a new model of deriving typology within Channel Bias. I introduce the concept of *Historical Probabilities of Alternations* (P_{χ}) and propose a method of estimating Historical Probabilities based on the statistical technique *bootstrapping*.

The proposed framework has theoretical implications. The existence of unnatural gradient phonotactic restrictions reveals that both categorical Optimality Theory and weighted-constraint frameworks with restricted Con undergenerate. To address this shortcoming, I propose a formal model of phonological typology that combines estimates of Historical Probabilities with results from the artificial grammar learning experiments. The dissertation adopts the Maximum Entropy model and introduces prior Historical Weights (w_{χ}), which are derived from the Historical Probabilities. Prior variance and Historical Weights allow for a disambiguation between Analytic and Channel Bias influences on typology: both metrics are compared to the observed typology, which yields a quantitative comparison between the two factors. To estimate the contribution of Analytic Bias, I conduct an artificial grammar learning experiment that tests learning of a complex and an unnatural alternation. By combining statistical modeling of diachronic developments with experimental work, the proposed framework allows controlling for Channel Bias factors when testing the Analytic Bias influences and vice-versa and, in turn, provides quantitative means for disambiguating Analytic and Channel Bias influences on typology.

Contents

Page intentionally left blank.

List of Figures

List of Tables

Abbreviations

AB Analytic Bias

BB Batu Belah

BSC Bootstrapping Sound Changes

CB Channel Bias

FD Final devoicing

FN Final nasalization

FV Final voicing

HG Harmonic Grammar

IVD Intervocalic devoicing

IVV Intervocalic voicing

LJ Long Jegan

LRT Likelihood Ratio Test

LTn Long Terawan

LTu Long Teru

MaxEnt Maximum entropy

MSCR Minimal Sound Change Requirement

NGB Natural Gradience Bias

PIE Proto-Indo-European

PND Post-nasal devoicing

PNS Proto-North-Sarawakan

PNV Post-nasal voicing

ROTB Richness of the Base

TMTFS Too Many / Too Few Solutions

TQ Tarma Quechua

UPT Universal phonetic tendency

Acknowledgements

I am grateful for all of the wonderful teachers with whom I have had the privilege of working throughout my education. I would like to especially thank my committee members for all their guidance and assistance and for being incomparable intellectual and personal role models and mentors. I have learned enormously, not only from my teachers, but from my peers and undergraduate students as well. I am grateful to my colleagues and the administrative staff at the Department of Linguistics, the undergraduate students in the courses I taught, and the wonderful community of students, tutors, and residents in Mather House.

The parts of this dissertation that deal with unnatural gradient phonotactics are based on a manuscript that I wrote with Aleksei Nazarov as the second author. The main ideas of these parts are mine and I acknowledge Aleksei's input by citing his 2008 manuscript (Nazarov 2008). Aleksei provided numerous helpful commentaries and improved the manuscript, for which I am very grateful. I would also like to thank Julia Sturm for her experimental stimuli and Willem Adelaar for his permission to analyze his Tarma Quechua recordings. The dissertation also benefited from feedback from audiences at various venues where parts of this work have been presented and from comments by anonymous reviewers.

The research for the dissertation was funded by grants from the Mind Brain Behavior Interfaculty Initiative, the Department of Linguistics, and the Committee on Australian Studies at Harvard University.

This dissertation is dedicated to the most important part of my life — my family. I am deeply grateful to my parents, grandparents, and my sister who supported me unconditionally throughout my life. I owe everything to my wife Nina, and my two boys, Tomaž Kajetan and Lev Vitomil.

Page intentionally left blank.

Chapter 1

Introduction

1.1 Typological discussions

The relationship between phonological GRAMMAR and observed typology has long been a central topic in phonology. Two major lines of thought emerge in this discussion: the Analytic Bias (AB) and Channel Bias (CB) approaches (Moreton 2008). The AB approach assumes that typological patterns emerge because of cognitive biases against certain phonological processes. In other words, some processes are more difficult to learn and these learnability biases result in surface typology (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). The AB approach thus assumes a direct relationship between phonological GRAMMAR and observed typology. The CB approach, on the other hand, assumes that systematic phonetic tendencies or phonetic precursors in the transmission of language from speaker to hearer result in surface typology. In other words, constraints on sound change are responsible for typological patterns: an inherent directionality of sound changes results in a predictable pattern of phonetic processes and phonologizations that ultimately determine surface typology (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 2013 for an overview of the literature). The CB approach therefore does not assume a direct relationship between the GRAMMAR and observed typology: typology is primarily influenced by phonetic forces, and phonological GRAMMAR can be independent from these forces.

Empirical evidence in favor of both approaches exists. The AB approach argues that learning biases influence typology, and substantiates this claim 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 the typological observations result precisely from these differences in learnability. Studies confirming AB involve both artificial grammar learning experiments on adults as well as studies testing phonotactic learning in infants. The generalization that typologically rare processes are more difficult to learn is especially robust when structurally complex alternations are tested against more simple alternations (Moreton and Pater 2012a,b). The results are less uniform for alternations that control for structural complexity: some studies report significant differences in learnability between natural and unnatural processes (Carpenter 2006, 2010, Wilson 2006 via Moreton and Pater 2012a,b), while others do not (Pycha et al. 2003, Kuo 2009, Skoruppa and Peperkamp 2011, via Moreton and Pater 2012a,b, Seidl et al. 2007, Do et al. 2016, Glewwe 2017, Glewwe et al. 2018).

The main evidence in favor of the CB approach is that typologically common processes align with universal articulatory and perceptual tendencies and natural sound changes. In fact, in many cases we can directly observe the origins of synchronic processes through historical developments: phonetically motivated sound changes that lead to phonologization (Hyman 1976, Barnes 2002, Blevins 2004, Garrett 2014).

Both approaches also face challenges in modeling typology. One of the main objections against AB is that speech is an “overlearned skill” (Lofqvist 2006) and that normally developing human infants are able to reproduce language input with a high degree of faithfulness (for an overview of literature, see Dodd et al. 2003). Most studies show that more or less any alternation can be learned given enough exposure (Moreton and Pater 2012a,b, White 2013), even if rates of learnability might vary across processes.[1]

Moreover, to derive typology from learnability differences is not a trivial task. Using computational modeling, Rafferty et al. (2011) suggest that learnability differences might not be sufficient for deriving typological observations. On the other hand, recent models proposed by Morley (2015) and Stanton (2016c) have derived a particular subtype of typological observations from learnability differences. Perhaps the most powerful model so far is the Maximum entropy model of phonological learning (Goldwater and Johnson 2003, Wilson 2006) that computationally encodes learnability differences between different phonological processes. Staubs (2014) proposes a model that derives the observed typological distributions from learnability differences. The ability to derive typological observations computationally does not, of course, constitute sufficient evidence that AB directly influences the observed phonological typology, but it does begin to address the objection raised by Rafferty et al. (2011).

Finally, the AB approach faces difficulties explaining why unnatural processes[2] are substantially less frequent than natural process as well as why some unnatural processes (such as PND)

---

[1] Some studies suggest that some processes might not be learned at all (e.g. Becker et al. 2011). The process tested in Becker et al. (2011), however, is a static phonotactic generalization rather than a productive alternation.

[2] For a new definition of (un)naturalness, see (2).

are attested, whereas others are not (such as FV). 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 the structural complexity of the tested alternation 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[3] 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 learning of unnatural alternations as defined in Section 2.1 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 position of this dissertation is that learnability differences should be estimated for each pair of alternations we are comparing. Learnability differences confirmed for a subset of alternations should not be extended to the entire typology. Experiments that specifically tested the processes discussed in this dissertation 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). While some natural-unnatural pairs of processes might show learnability difference and their typological dif-

---

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

ferences can be explained under the AB approach, deriving substantial typological differences between natural and unnatural alternations based on AB is problematic for processes for which no differences in learning are observed experimentally.

Conversely, a major objection against the CB approach is that it fails to explain why some processes are never attested (Kiparsky 2006, 2008, de Lacy and Kingston 2013). In other words, combinations of sound changes (or a single sound change, if we allow it to be unnatural, as proposed in Blust 2005) could in principle produce a number of unnatural alternations — yet, it seems that some hypothetically available processes are never attested. In fact, Kiparsky (2006, 2008) goes a step further and assumes some processes are impossible in synchronic grammar. This position is also encoded in the classical Optimality Theory (OT; Prince and Smolensky 1993/2004) model, where unnatural constraints are excluded from the universal constraint inventory (Con).[4] On this approach, some output candidates are harmonically bounded and consequently some processes are impossible in synchronic grammars. For example, Kiparsky (2006, 2008) identifies several combinations of sound changes that would lead to final voicing, but the process is never attested (or at least is morphologically limited; see Yu 2004, Blevins 2004, de Lacy 2002). CB faces difficulties explaining this mismatch and AB is invoked to explain it (Kiparsky 2006, 2008, de Lacy and Kingston 2013).

Current models of typology within CB are therefore insufficient. The most common line of thought in deriving the typology within CB has been to assume that rare sound changes produce rare alternations (Blevins 2004); related to this line of reasoning is Ohala's (1989) proposal that the probability of perceptual confusion is proportional to the frequencies of sound changes. Most of the models, however, do not quantify this rarity. Moreton (2008) attempts to

---

[4] For a discussion on the traditional definition of "unnatural" and the definition proposed in this dissertation, see Section 2.1.1.

quantify phonetic precursors with the goal of reaching a more transparent phonetic metric for disambiguating sound changes, but this approach has problems (Yu 2011). As will be shown in this Chapter 3, the results of sound changes operating in combination and phonetic precursors or perceptual confusability do not always align, which makes the quantification of precursors or confusability unsuitable for deriving the typology of all processes (including the unnatural ones).

The low probability of combinations of changes has been occasionally invoked to account for the smaller likelihood of alternations that require more than a single change, especially in morphology (Harris 2005, 2008, Blevins 2004), but also in phonology (Bell 1970, 1971, Greenberg 1978:75-6, Anderson 1981, Morley 2015). Morley (2012) outlines a diachronic explanation for the typology of epenthesis which states that the more diachronic conditions a certain type of epenthesis requires, the less likely it is to arise typologically. The non-existence of some unnatural alternations, Morley (2011, 2015) suggests, might be due to an interplay between diachronic development and the learning of grammar ("hypothesis selection") from surface forms. Morley further argues that one of the scenarios in Kiparsky (2006) that leads to an unnatural process is rare because "the probability of the combined event is the multiplicand of the two probabilities (under assumptions of independence), [and] this number might prove to be small enough to render it unlikely to occur in the given sample" (Morley 2015: e43). None of these models, however, are sufficiently quantified for a typological model, or they fail to yield implementable results. Finally, Cathcart (2015) attempts to quantify the CB influences on typology by automatically identifying the number of combinations of sound changes that produce an unnatural alternation or phonotactic restriction such as final voicing. For each combination of sound changes, that number is then compared to the number of all possible sound changes given the number of permutations and a sample of sound changes. The problem with this approach is that it requires a representative sample of all possible sound changes (which is problematic given current availability of diachronic surveys) and that it is computationally difficult to implement.

As Moreton (2008:84) points out, the AB and CB approaches have in the past often been treated as mutually exclusive, either explicitly or implicitly. A mounting body of research, however, argues that both AB and CB shape typology (Hyman 2001, Myers 2002, Moreton 2008, Moreton and Pater 2012a,b, de Lacy and Kingston 2013). This position is held in this dissertation as well: I maintain that both AB and CB influence typology. However, the goal of phonological theory should be to disambiguate the two influences: what aspects and what proportions of phonological typology are caused by learnability/learning biases (AB) and how much of phonological typology is due to the directionality of sound change (CB)? This is a challenging task, as influences of the two biases are often argued to be conflated (see Section 4.4.1). This dissertation offers a potential solution to this problem by focusing on unnatural processes for which crucial mismatches between the AB and CB approaches can be identified. This dissertation proposes a framework that combines statistical modeling of historical developments with experimental work to disambiguate AB and CB influences on typology. While the dissertation focuses on feature [±voice], the proposed framework is extendable to other processes with the goal of allowing process-specific estimation of the relative contributions of AB and CB on the typology.

1.2 Outline

To disambiguate the two influences, we first need a good understanding of how exactly each of them results in typology. I first propose a new model of typology within the Channel Bias approach. The new model features five crucial components. First, I argue for a new subdivision of naturalness (Section 2.1), whereby processes traditionally labeled as *unnatural* should be subdivided into *unmotivated* and *unnatural* processes. Second, based on a reported unnatural synchronic alternation and sound change, post-nasal devoicing (PND), I contend that sound change is always phonetically motivated and cannot operate in a phonetically unnatural direction. In other words, one of the rare reported cases of an unnatural sound change can be explained through a combination of natural sound changes (Section 2.3), not only in Tswana and Shekgalagari (as proposed in Dickens 1984, Hyman 2001), but in all other cases collected in this paper. I present new evidence from Sogdian, Yaghnobi, and other languages that crucially contributes to this conclusion. This allows us to maintain the long-held position that sound change can only operate in a phonetically natural direction (which has recently been challenged by Blust 2005, 2017; for a discusssion, see Garrett 2014). Third, the dissertation establishes a diachronic model for deriving and explaining *unnatural* alternations called the Blurring Process (Chapter 3) and fourth, argues with a formal proof that a minimum of three sound changes are required for an unnatural process to arise and a minimum of two for an unmotivated process (Minimal Sound Change Requirement) (Section 3.4). I then present a quantitative model that predicts the typology within the Channel Bias by introducing the concept of Historical Probabilities of Alternations, which is based on the two crucial concepts developed in this paper, the Blurring Process and MSCR. Chapter 4 proposes a statistical technique for estimating Historical Probabilities, Bootstrapping Sound Changes (BSC), and illustrates implications of this method for the Channel Bias approach to phonological typology.

Chapter 4 also presents applications of the BSC technique. Bootstrapping Sound Changes allows us to (i) estimate the Historical Probability of any alternation (Section 4.3.1), (ii) compare two alternations, attested or unattested, and perform statistical inferences on the comparison (Section 4.3.2), (iii) predict attestedness in a given sample for any alternation (Section 4.3.3), and (iv) derive a quantitative metric for the Channel Bias influences on typology that can be employed in typological frameworks that model both AB and CB together (Section 5.2). Using the BSC technique, I also identify and quantify several predictions of the Channel Bias approach and compare them with predictions of the Analytic Bias approach (Sections 4.3.4 and 4.4). This allows me to identify crucial mismatches in predictions of the two approaches which in turn allows for disambiguation between AB and CB influences on observed typology (Section 4.4). In other words, unlike previous models, my proposed framework allows us to control for the AB factor when testing the CB influences on typology and vice versa. Chapter 4 shows that a substantial gap in typological frequency between natural and unnatural alternations that show no learnability differences is derivable within the CB using the BSC technique. This means CB influences the observed typology of natural and unnatural processes even when AB is controlled for. In Chapter 6, I do the opposite for the AB influences: I identify those aspects of typology that control for CB influences and present an artificial grammar learning experiment that tests learnability of complex and simple unnatural alternations. The main advantage of the experiment presented in this dissertation is precisely the fact that CB influences are controlled for when testing for the AB factor. The results suggests that rarity of unnatural processes is primarily due to CB, but that the avoidance of complex processes is primarily influenced by AB.

The framework of combining statistical modeling of diachronic developments and experimental work proposed in this dissertation bears several theoretical consequences. In Chapter 5, I show that the existence of unnatural gradient phonotactic restrictions poses a problem for weighted-constraint models that exclude unnatural constraints from the universal constraint inventory. I identify a thus far unobserved prediction of the restricted Con hypothesis in weighted-constraint-based frameworks that states that natural feature values are always more frequent than unnatural feature values (or the so-called Natural Gradience Bias). While most gradient phenomena discussed so far obey the Natural Gradience Bias, Berawan dialects and Tarma Quechua clearly violate it. In the second part of Chapter 5, I propose a MaxEnt model of typology and argue that introducing a historical metric into the model solves the long-standing problem of undergeneration and at the same time keeps typological predictions of OT. Finally, I explore implications of this model for the Too Few / Too Many Solutions problem.

Chapter 2

Unnatural phonology

2.1 Background

2.1.1 Subdivision of naturalness

Directly related to the question of phonological typology is the question of naturalness and how to encode it in the grammar design (Bach and Harms 1972, Stampe 1973, Catford 1974, Hellberg, Hyman 1975, 1978, Donegan and Stampe 1979, Anderson 1981, Westbury and Keating 1986, Archangeli and Pulleyblank 1994, Hayes 1999, Buckley 2000, Hyman 2001, Blevins 2004, 2008a,b, Yu 2004, Blust 2005, Wilson 2006, Hayes et al. 2009, Carpenter 2010, Becker et al. 2011, White 2013, Hayes and White 2013, i.a.). The term *naturalness* itself has received several interpretations in the literature; scholars have argued that unnatural processes are variously possible, dispreferred, or impossible in synchronic grammar. The question of naturalness is relevant to the theory of sound change as well (Catford 1974, Blust 2005). It has long been believed that sound changes are always phonetically motivated, and therefore natural. This position goes back to the Neogrammarian school of thought and posits that "typologies of sound change and possible phonetic precursors correspond perfectly" (Garrett 2014).

Natural and unnatural processes have traditionally been distinguished by specifying that the former are phonetically motivated and typologically common while the latter are unmotivated and typologically rare. However, this division is insufficient, as it fails to capture crucial distinctions within the unnatural group. In the new division I propose, natural processes are phonetically motivated (as in previous proposals): they operate in the direction of universal phonetic tendencies (for the term, see Hyman 1972). I define a *universal phonetic tendency (UPT)* as one which exhibits three crucial properties: (i) passive operation, (ii) cross-linguistic operation, and (iii) the ability to result in common sound patterns.

(1) *Definition of Universal Phonetic Tendency (UPT)*

UPTs are phonetic pressures motivated by articulatory or perceptual mechanisms that passively operate in speech production cross-linguistically and result in typologically common phonological processes.[5]

Passive operation of a phonetic tendency means that the tendency targeting some phonetic feature operates automatically (with no active control by the speakers, Kingston and Diehl 1994), even in languages with full phonological contrast of the equivalent feature in a given position. For example, the observation that voiceless stops have universally more phonetic voicing into closure in post-nasal position compared to the elsewhere position (Hayes and Stivers 2000), even

---

[5] As Moreton (2008) points out, some phonetic precursors do not cause typologically common phonological processes (see also de Lacy and Kingston 2013). This paper focuses on those unnatural processes that operate against phonetic precursors that are robust enough to yield common sound patterns.

for languages with full contrast of voice post-nasally, fulfills the criterion that post-nasal voicing operates passively in world languages.

I argue that within the traditionally labeled "unnatural" group, we find two types of processes. I label *unmotivated* those processes that lack phonetic motivation, but do not operate against any UPT. In other words, while unmotivated processes do not correspond to a particular universal articulatory/perceptual force, they are also not operating specifically against such a force. *Unnatural* processes[6], on the other hand, are those that operate precisely *against* some UPT and are not a UPT themselves.[7]

(2) *A new division of naturalness*

- a. *natural processes*: defined as UPTs
- b. *unmotivated processes*: lack motivation, but do not operate against UPTs
- c. *unnatural processes*: operate against UPTs, are not UPTs

An example of an unmotivated process would be Eastern Ojibwe "palatalization" of /n/ to [ʃ] before front vowels (e.g. [ki-naː**n**-aː] 'you fetch him' vs. [ki-naː**ʃ**-im-i] 'you fetch us'; Buckley

---

[6] Similar, but also crucially different, distinctions have been proposed before. Morley (2014) assumes "anti-natural" processes are those that operate against implicational universals and are unattested: "unattested patterns that do not conform with posited language universals."

[7] As will be shown in Section 3.3 below, some processes might be motivated in both directions by different mechanisms; the voicing and devoicing of stops, or the fricativization of stops and the occlusion of fricatives are two examples of diametrically opposed processes. However, naturalness always needs to be evaluated in a given context; evaluated globally, in all positions, devoicing of stops and occlusion of fricatives are natural, i.e. phonetically motivated, but in a leniting environment, such as post-vocalic position, voicing or fricativization of stops is the natural direction (see Section 3.3 and Ladefoged and Maddieson 1996:137, Kaplan 2010). Different contexts that determine the naturalness of a process can sometimes superficially overlap, e.g. in a language that lacks target segments in those contexts that distinguish the natural from the unnatural direction (as we will see, this is precisely what happens in the case of post-nasal devoicing). It is also possible that two diametrically opposed processes could both be phonetically motivated in a given context: vocalic sound change is usually less unidirectional, but even there, clear principles can be established in which only one direction is the natural one (Labov 1994). To my knowledge, no detailed research is available on sound change that is motivated in both directions in exactly the same context; further research on this topic is a desideratum.

2000). This process lacks phonetic motivation, but its reverse process [ʃ] → [n] / __[+front] is not a UPT. Examples of unnatural processes include final voicing, intervocalic devoicing, and post-nasal devoicing[8]. All of these processes operate directly against clear and well-motivated articulatory phonetic tendencies (see Section 2.3). It has to be noted that phonetic motivation of a process must always be evaluated with respect to the context. For example, devoicing of voiced stops is a natural process word-finally, but unnatural intervocalically (Section 3.3).

Most of the discussions on unnatural phenomena in phonology in fact discuss unmotivated processes, according to the definition in (2) (Buckley 2000, Blevins 2004, 2008a,b, Blust 2005). Some processes that have been labeled as "unnatural", but are in fact unmotivated according to my definition in (2) are collected in (3).

(3) *Some processes labeled as "unnatural"*[9] *in Blevins (2008a) and Hayes and White (2013)*

a. /p/ → [s] / __i (Bantu)

b. /i/ → [u] / d__ (Kashaya)

c. $*\begin{bmatrix}+COR\\ +cont\\ -strid\end{bmatrix}\begin{bmatrix}-stress\\ +round\end{bmatrix}$ ("No [θ,ð] before stressless rounded vowels" in English)

d. $*\begin{bmatrix}+cont\\ +voice\\ -ant\end{bmatrix}\begin{bmatrix}+stress\end{bmatrix}\begin{bmatrix}-son\end{bmatrix}$ ("No [ʒ] before stressed vowel + obstruent" in English)

---

[8] Some proposals motivate post-nasal devoicing articulatorily or perceptually. See Sections 3.1 and 2.2.2 for a discussion of these proposals.

[9] Blevins (2008) gives the following definition of unnaturalness: "Unnatural sound patterns are those with no plausible single phonetic source, origin, or explanation." See Blevins (2008a) for further discussion on her interpretation of (un)naturalness. Hayes and White (2013:47) define unnaturalness within the constraint-based framework. Unnatural are "other constraints" that do not satisfy either phonetic or typological criterion with the following clarification: "The typological criterion can be expressed on the basis of Greenbergian implicational universals (the presence of sequences that violate the constraint implies the presence of closely similar sequences that do not; Greenberg 1966, 1978). The phonetic criterion is that a constraint should be functionally effective, serving to form a phonological system in which words are easier to articulate or in which possible words are perceptually distinct from one another."

While naturalness and how to represent it in the grammar design has been primarily the focus of synchronic studies (Stampe 1973, Hellberg 1978, Anderson 1981, Archangeli and Pulleyblank 1994, Hayes 1999, Hyman 2001, Coetzee and Pretorius 2010), it has been the subject of debate within the theory of sound change as well. The well-accepted Neogrammarian position that sound change is always natural (for an overview of the literature on this position, see Blust 2005:219-223) has recently been challenged. Blust (2005) identifies several unnatural sound changes and argues they had to operate as unitary sound changes. The survey of consonantal sound changes in Kümmel (2007) also lists a number of unnatural sound changes, although they are not labeled as such. A subset of Blust's unnatural sound changes have been explained as a result of a sequence of multiple natural sound changes (Goddard 2007, Blevins 2007, Garrett 2014). However, the most robust cases of unnatural sound change reported in Blust (2005), including post-nasal devoicing, have yet to receive a sufficient explanation. One of the goals of this dissertation is to propose a diachronic explanation for these cases (see Chapter 3).

2.1.2 Sound change

Before we turn to a more detailed discussion of unnatural alternations and sound changes, a clarification of the term *sound change* is necessary (Blevins 2004, Garrett 2014). Despite more than a century of scientific research on sound change, Garrett (2014) in his recent overview of the field acknowledges that sound change has "no generally accepted definition". Many definitions involve the concept of "phonologization" (Hyman 1976, 2013, Barnes 2002) "whereby an automatic phonetic property evolves into a language-specific phonological one" (Garrett 2014) and this process constitutes a sound change. The question remains, however, at what level of abstraction should we define sound change: at a "language-specific phonetic" or "phonological (structured)" level (Hyman 2013, Fruehwald 2016)? Even more relevant for our discussion is the question of what distinguishes a single sound change from a combination of sound changes.

Every model, be it phonological or diachronic, has to operate with some level of abstraction. We could assume that sound change is a change of any articulatory target/gesture (cf. Browman and Goldstein 1992) or, in exemplar theory's terms, any change in a label of a category (Pierrehumbert 2001, Bybee 2001), regardless of how small the difference between the two diachronic stages would be. For example, a minimally higher degree of coarticulation (fronting of [k] in [ki]-sequences) or a minimal change in the F1 target in low vowels would constitute a sound change according to this definition. In such a model, for example, a gradual sound change [a] > [æ] would involve an infinite or at least a very large number of individual sound changes. To be sure, each of these minimal sound changes would have no phonological implications: languages cannot contrast /a/ and /a̝/ with a minor difference in the F1 target or other similarly minimal phonetic differences. While such a radical and phonetically oriented model, in which sound change represents a change in any phonetic specification that is not automatic, might be valid, it would fail to provide results that would be meaningful to phonological theory (for a discussion on the "notoriously fuzzy boundary" between phonetics and phonology, see Kingston and Diehl 1994, Hyman 2013, Cohn 2006, Keyser and Stevens 2006).

The focus of this chapter is on phonological alternations (allophonic or neutralizing) that result, via phonologization, from non-analogical regular sound changes (regardless of the mechanisms of their origin; for an overview, see Garrett and Johnson 2013, Garrett 2014)[10]. For this reason, I adopt the concept of features from phonology, together with the level of abstraction of the phonological feature system (Chomsky and Halle 1968, Hall 2007, Keyser and Stevens

---

[10] I am interested in sound changes that operate categorically throughout the lexicon, regardless of whether they result from lexical diffusion or the Neogrammarian regular sound change.

2006, Hyman 2013) and define sound change as a change in one feature that is phonologized and non-automatic (Hyman 2013, Keyser and Stevens 2006) and that operates throughout the lexicon in a given environment. Synchronic alternations are alternations of at least one feature in a given environment, but can also involve alternations of multiple features simultaneously. I posit that a single instance of sound change can only involve change of a *single* feature in a given environment. That a single sound change involves a change in a *single* "phonetic property" has in fact been previously claimed by Donegan and Stampe (1979) and Picard (1994). Picard (1994) calls this assumption the "minimality" principle: "sound changes are always minimal, and so can involve no more than one basic phonetic property." The minimality principle arises from the assumption that "the substituted sound" should be "as perceptually similar to the original target as possible" (Donegan and Stampe 1979). There is also an articulatory argument for the minimality principle: variation in coarticulation which is the initial mechanism of sound change (cf. Ohala 1983, Lindblom 1990, Lindblom et al. 1995, Beddor 2009) is often continuous and gradual (cf. Garrett 2014, Pierrehumbert 2001), which means that phonologically relevant change proceeds through minimal phonetic changes, and minimal phonetic changes are usually not substantial enough to change two phonological features. I adopt this assumption of minimality (Donegan and Stampe 1979, Picard 1994) and posit that a single sound change is a change in one feature in a given environment. A combination of sound changes is a set of such individual sound changes operating in a given language.

(4) *A single sound change vs. a combination of sound changes*

Sound change is a change of one feature in a given environment; a combination of sound changes is a set of such individual sound changes operating in a given language.

To my knowledge, no detailed studies exist on the question of whether a single sound change can involve a change of more than one feature simultaneously. Direct evidence of such a sound change is hard to obtain because we have only limited access to sound changes in progress and even apparent sound changes in process might not always yield conclusive results: it is, for example, possible that what seems to be a sound change in progress that targets two features simultaneously might in fact be variation between two end-points that result from operation of two individual sound changes. Typological surveys of sound change support the minimality hypothesis, at least as a strong tendency: the substantial majority of sound changes surveyed in Kümmel (2007) do involve a change of only a single feature. Likewise, in a phylogenetic modeling of sound changes in Turkic in Hruschka et al. (2015), 79% of consonantal and 70% of vocalic sound changes involved a change in a single feature. It is reasonable to assume that at least a subset of the two-feature changes involve an additional sound change that is not observed on the surface, although this assumption cannot be proven (see discussion below).

To be sure, universally redundant (referred to as "automatic" here) features (cf. Hyman 2013, Keyser and Stevens 2006) do not contribute to sound change minimality: if stops nasalize (D > N), [±sonorant] is changed along with [±nasal], but [±sonorant] is not contrastive in either nasals or voiced stops and this change does not count as an instance of an additional sound change. More problematic for the minimality assumption are (i) cases of perceptually driven sound changes that involve change of place of articulation, (ii) cases of total assimilation, and (iii) cases of epenthesis or deletion. Some of the recurrent sound changes that appear to target more than a single feature simultaneously are changes that primarily arise due to perceptual similarity, e.g. changes between [ɣ] and [w], [kʷ] and [p], or [pʲ] and [t] (Ohala 1989). It is nevertheless reasonable to assume that a subset of such changes proceed through intermediate stages, e.g. [ɣ] > [β] > [w]. While Ohala (1989) argues strongly against such proposals, there is independent evidence that makes such intermediate stages plausible. Ohala (1989), for example, specifically argues against Whatmough (1937) (via Ohala 1989) who claims that Proto-Indo-European *k^w develops to Greek [pp] through an interstage with [k͡p]. Northwest Caucasian languages, however, confirm that labialized alveolar stops [t^w] or [d^w] can and do develop to doubly articulated bilabio-alveolar stops [t͡p] or [d͡b] (for example in Ubykh), even though these are typologically much rarer than bilabio-velar stops (Catford 1972, Ladefoged and Maddieson 1996, Garrett and Johnson 2013, Beguš 2018). Some dialects of Ubykh then merge these with plain labial stops [p] and [b] (Fenwick 2011), which means we have the full chain of developments [t^w] > [t͡p] > [p] attested. Intermediate stages can thus be motivated at least for a subset of cases that are traditionally used as examples of non-gradual changes involving more than a single feature. In other words, not all changes targeting two features that appear to result from perceptual confusion necessarily lack intermediate stages.

While it has been stipulated that assimilation in more than one feature often (or always) arises through inter-stages or that a segment often (or always) deletes with an intermediate stage with [ʔ], [h], or a glide [j] (Hock 1991 or McCarthy 2008 for a synchronic perspective), one of the most well-studied sound changes in progress, final t/d-deletion in English, points to the contrary as no intermediate stages are observed.[11] Moreover, while epenthesis can be a gradual process (Morley 2012), it would be difficult to represent it with a change of a single feature value. Complete deletion or epenthesis of a segment/feature matrix thus has to count as a single sound change under my approach. Metathesis is another non-canonical sound change: it does

---

[11] One could argue that in synchronic variation between [t/d] and ∅, the potential intermediate stage with *[ʔ] might be lost and that the variation between deleted and undeleted forms is phonologized. To evaluate probability of such an assumption is beyond the scope of this dissertation.

not involve a change in a feature's value, but rather a change in the ordering of features/feature matrices. While usually a sporadic sound change, metathesis *can* be regular (Blevins and Garrett 1998) and it can also be gradual, involving several intermediate stages. Slavic metathesis (/VRC/ > /RVC/), for example, most likely involves an interstage with vowel epenthesis (/V_1RC/ > /V_1RV_2C/) and vowel deletion (/V_1RV_2C/ > /RV_2C/) (Blevins and Garrett 1998).[12]

In sum, I adopt phonological features to represent sound changes that via phonologization result in synchronic alternations. By the word "change" in (4) I mean either a change of one non-automatic feature value in a given environment, a complete deletion/insertion of a feature matrix, or a change of ordering of features/feature matrices in two diachronic stages. In all cases, my assumption is that a single sound change is minimal: it involves a change of a single feature or deletion/epenthesis/reordering of a single feature matrix. Further research on sound changes in progress is needed to confirm the minimality assumption — currently, I can claim that the minimality assumption is at least a strong tendency: it holds for at least the majority of sound changes documented and it is articulatorily and perceptually well-grounded.

2.2 Universal tendencies for voicing

All cases of unnatural phonotactics presented in this dissertation target a single phonological feature: [±voice]. The reason for why I limit the scope of the dissertation is that alternations and phonotactic restrictions targeting this feature are among the most well-studied phenomena in phonology and there exists a substantial body of research on typology as well as phonetics of voicing cross-linguistically. The attention that voicing has received in the literature help us

---

[12] Unless noted otherwise, the following symbols are used in this dissertation: V – vowel, C – consonant, T – voiceless obstruent, D – voiced obstruent, S – voiceless fricative, Z – voiced fricative, N – nasal, R – non-nasal sonorant, V – vowel.

determine in which environments the voice feature is universally dispreferred and vice versa, i.e. which processes targeting [±voice] are "unnatural" according to the definition above.

The table in 2.1 summarizes universal phonetic tendencies that target the feature [±voice], and the paragraphs below summarize the evidence from the literature that these are indeed universal phonetic tendencies as defined in Section 2.1.1.

Table 2.1: Some universal phonetic tendencies for [+voice].

	[+voice] preferred	**[+voice] dispreferred**
Intervocalic voicing	between two vowels	
Postnasal voicing	after [+nasal] consonant	
Voicing agreement	adjacent to [+voice] consonant	adjacent to [−voice] consonant
Final Voicing		word-finally

2.2.1 Intervocalic voicing

Intervocalic devoicing fulfills all criteria to qualify as an unnatural process under the definition in Section 2.1.1: it operates against a universal phonetic tendency that is typologically very common, has clear phonetic motivation (Westbury and Keating 1986), and is active as a passive phonetic tendency (e.g. Davidson 2016). Intervocalic voicing is well attested: the survey in Gurevich (2004) and Kaplan (2010) shows that 26 of 153 (or 17%) languages surveyed have intervocalic voicing as a synchronic alternation. Intervocalic voicing is also well attested as a sound change: the survey in Kümmel (2007) reports over 40 languages with intervocalic voicing as a sound change. In fact, voicing is the most common form of intervocalic stop lenition, followed by spirantization, approximatization and others which are less common (Kaplan 2010).

Moreover, there exists a clear articulatory phonetic motivation for intervocalic voicing. The difference in subglottal and supraglottal pressure is greatest in intervocalic position and is considerably smaller in initial or final position; because a pressure difference is crucial for voicing, voiced stops will be most possible in intervocalic position and dispreferred initially or finally

(Westbury and Keating 1986: 153). Westbury and Keating (1986) argue that intervocalically, voiced stops are articulatorily easier to produce than their voiceless counterparts and continue that any neutralization in the opposite direction (from the expected) would result in "added articulatory cost".

Kaplan (2010) also argues in favor of perceptual motivation for intervocalic voicing. Invoking P-map (Steriade 2001), she claims that intervocalic voicing is the most common lenition (more common than spirantization and approximation) precisely because perceptual differences between voiced and voiceless stops intervocalically are the smallest (i.e. smaller than perceptual differences between intervocalic voiceless stop and voiceless fricatives). Speakers choose the minimal perceptual difference to repair the phonotactic restriction against intervocalic voiceless stops. Finally, intervocalic voicing is a passive phonetic tendency: stops have more voicing going into closure intervocalically compared to other positions (Docherty 1992, Davidson 2016 and literature therein).

2.2.2 Post-nasal devoicing

The claim that PND is an unnatural process according to the definition in (2) is supported by strong articulatory phonetic evidence. Post-nasal voicing (T > D / N__; PNV), the exact inverse process to PND, is a UPT: it is phonetically well motivated, operates passively in world languages, and is typologically common.

The phonetics of PNV are thoroughly investigated in Hayes and Stivers (2000) (cf. Hayes 1999, Pater 1999). Supported by previous work including Rothenberg (1968), Kent and Moll (1969), Ohala and Ohala (1993), Ohala (1983), and others, the authors identify two phonetic factors that render stops in post-nasal position prone to voicing: (i) nasal airflow leak and (ii) expansion of oral cavity volume during velic rising. Both of these factors promote voicing, i.e. counter the anti-voicing effects of closure (Hayes and Stivers 2000). It has long been known that coarticulation occurs in the transition from nasal to oral stops: the velum must rise from a low position to a high position, at which point it closes the nasal cavity. During this process, airflow can leak through the nasal cavity, which means that the airflow necessary to maintain voicing that would otherwise stop during the closure can be maintained longer (Hayes and Stivers 2000). Moreover, when the velum rises from a high position to complete closure, the volume of the oral cavity increases, which again allows a longer period of sufficient airflow that would otherwise stop due to closure (Hayes and Stivers 2000).

Not only is post-nasal voicing phonetically motivated, it is also universally present as a passive phonetic tendency: that is to say, phonetic voicing is found even in languages without phonological PNV, such as English. Hayes and Stivers (2000) show that speakers produce more passive phonetic voicing on voiceless stops in post-nasal position than elsewhere. Speakers produce "significantly more closure voicing" in words like [tampa] than in words like [tarpa] (Hayes and Stivers 2000). PNV thus meets all the criteria for being a UPT.

PNV is commonly attested not only as a phonological and phonetic process, but also as a sound change. Locke (1983) identifies 15 languages, out of a sample of 197, that exhibit PNV as a synchronic process (reported in Hayes and Stivers 2000). Kümmel (2007, 53f.) lists approximately 32 languages in a survey of approximately 200 in which PNV operates as a sound change. By comparison, PND in the same survey is attested only twice: in one instance it targets stops and and in the other affricates.

While post-nasal voicing is a well-motivated and natural process — the opposite process, devoicing of voiced stops in post-nasal position, is unnatural: it operates against a UPT. For reasons discussed above (nasal leakage and increased volume of oral cavity), the post-nasal environment is in all aspects antagonistic to devoicing (compared to other positions, e.g. word-initial and word-final), in the sense that it operates against the voicing-promoting effects of post-nasal position (that counters the anti-voicing effects of the closure). In other words, in the transition from nasal to oral stop, the velum does not close instantaneously; as a result, air leakage occurs into a portion of the following stop closure, prohibiting the "air pressure buildup" necessary to articulate a voiceless stop. Expansion of the oral cavity due to velic rising also has an effect of promoting voicing during the closure: greater volume allows longer period of time before the air pressure buildup (for a discussion on phonetics, see also Hayes and Stivers 2000, Coetzee and Pretorius 2010 and literature therein). Moreover, to my knowledge, PND has not been reported as a passive tendency in any language.

This dissertation also speaks to a long-standing discussion on the role of production vs. perception in sound change and phonology in general. One of the approaches that understands PND as a natural process is the perceptual account outlined in Stanton (2016a). Post-nasal devoicing can be analyzed as a contrast enhancement with the following reasoning: sequences of a nasal and a voiced stop (ND) are perceptually very close to plain nasal stops (N). In order to enhance the perceptual contrast between N and ND, speakers can devoice stops in the latter to NT. There is no doubt that perceptual (or auditory in Johnson and Garrett's 2012 terms) enhancement plays a role in sound change and phonology (for an overview, see Garrett and Johnson 2012, Garrett 2014, Fruehwald 2016). For example, lip protrusion in [ʃ] has no articulatory grounding, but a clear perceptual one: lip protrusion results in lowering of the frequency peak, which in turn results in perceptual enhancement of the contrast [ʃ] ∼ [s]. The majority of cases of perceptual enhancement discussed in Garrett and Johnson (2012) or Garrett (2014), however, involve phonetic properties of phonemes in all positions. Additionally, I am not aware of cases of perceptual enhancement that would operate in the unnatural direction, as would be the case for PND. Below, I present argumentation for the position held in this dissertation that the unnatural processes targeting feature [±voice] do not result from perceptual enhancement, but from a combination of three natural sound changes. To be sure, this does not suggest that perceptual enhancement is not a possible mechanism for sound change, but only that the perceptual enhancement is likely not the underlying cause of the subset of unnatural alternations targeting feature [±voice] discussed in this dissertation.

There are two independent facts that might speak in favor of the perceptual-enhancement approach for the case of PND. Gouskova et al. (2011) observe that post-nasal devoicing in Setswana occasionally result in ejective stops. This distribution would be in line with the perceptual enhancement approach: to maximize the contrast, speaker can use ejection as an additional cue for the original N ∼ ND contrast. Second, a survey in Stanton (2016a:1106) identifies a number of languages that devoice sequences of ND to NT word-finally, but not word-medially: Neverver, Kobon, Naman, Avava, Páez, and Tape. This distribution, too, can be analyzed as contrast enhancement word-finally, where the contrast is additionally reduced.

These proposals, however, face problems. First, there exist no experimental studies that would confirm NT to be perceptually more salient than N or ND in intervocalic position. Kaplan (2008) is the only study known to me that tests this contrast perceptually and does not limit it to final position (Katzir Cozier 2008). While NT is more salient than N and ND word-finally, no such significant effect has been found word-medially: all three stimuli, N, NT, and ND, were perceived with equal rates (Kaplan 2008).[13] Kaplan (2008) indeed argues that post-nasal voiced stops in word-final position are perceptually most confusable, and thus perceptually

---

[13] This result can also be due to a ceiling effect.

motivates the *ND# constraint. First, note that the “motivated” *ND# is limited to word-final position. Second, in English, the repair for *ND# is not devoicing, but deletion (which cannot be interpreted as contrast enhancement). Finally, in the languages surveyed in Stanton (2016a:1106) that devoice sequences of ND# to NT# (Neverver, Kobon, Naman, Avava, Páez, and Tape), voiceless stops contrast with prenasalized voiced stops. In other words, the inventories lack plain voiced stops completely. Final D > T / N__# can thus simply be analyzed as devoicing of voiced stops in word-final position, which is a phonetically well-motivated development and does not lend support for the perceptual enhancement explanation.

While it is true that some stops in Setswana are realized with “weak” ejection, this does not necessarily point to perceptual enhancement either. Setswana has three series of stops: voiced (D), voiceless aspirated (T^h) and voiceless unaspirated that are in variation with ejective articulation (T/T’). While at first glance, the alternation between voiced stops elsewhere and voiceless aspirated/ejective (T/T’) in post-nasal position appears as a fortition, it can also be analyzed simply as a devoicing and is actually expected under the historical approach presented below. From the data described in Gouskova et al. (2011), it appears that the ratio of voiceless unaspirated vs. ejective articulation in the T/T’-series of stops is not much higher in post-nasal position compared to the elsewhere condition. For example, Gouskova et al. (2011) count the number of recorded tokens that have ambiguous acoustics (can be interpreted as either ejective or voiceless aspirated) and the number unambiguous ejective. In post-nasal position, the ratio is 41 (ambiguous) vs. 60 (ejective), in intervocalic position 43 (ambiguous) vs. 54 (ejective). This means that ejection in post-nasal position can be analyzed as part of the common trend in Setswana to realize unaspirated voiceless stops (and in our case, devoiced stops) with ejection. In fact, I will argue below that PND in Tswana results from a combination of sound changes, one of which is devoicing of all voiced stops. Voiced stops are reconstructed to devoice. Because voiced stops have significantly less aspiration than voiceless aspirated stops, one expect a devoiced variant of a stop to be phonetically closest to a voiceless unaspirated T. This series of stops can then get realized as ejective (either when devoicing occurs after that), which would explain the Setswana distribution without resorting to post-nasal fortition of contrast enhancement.

There exists some additional evidence against the perceptual enhancement explanation of PND. PND is not attested as a repair strategy even in cases where we should expect it. For example, many languages disallow NC_1VNC_2 sequences. A recent study in Stanton (2016b) suggests that avoidance of these sequences constitutes a strategy to avoid a difficult contrast with $NVNC_2$. The vowel in NC_1VNC_2 is universally phonetically nasalized, a process which reduces cues for the contrast between NC_1 and N. One way to repair this contrast would be to devoice the first consonant C_1. However, the survey in Stanton (p.c.) shows that NDVNC > NTVNC is not attested. Another challenge for the perception approach is that, as I will argue, strong evidence exists that in all thirteen languages I have examined, PND actually arises through a combination of three natural sound changes rather than through a contrast enhancement. Finally, while perceptual enhancement might explain PND, it cannot, to my knowledge, explain other unnatural processes discussed in this dissertation (especially intervocalic devoicing).

As already mentioned, the dissertation does not argue against the existence of perceptual enhancement as a sound change in general. A number of arguments presented here, however, favor the three-sound-changes approach over the perceptual enhancement approach when dealing with PND. It seems suspect to suggest that a contrast will only be enhanced when a set of three sound changes happen to operate in the pre-history of a system. Future research involving additional processes that target features other than [±voice] should show whether other unnatural processes are better explained by perceptual enhancement or the three-sound-change approach proposed here.

While I claim PND did not arise as contrast enhancement, I, however, acknowledge the possibility that perceptual factors play a role in the phonologization of PND and preservation of the alternation once it arises. It is thus possible that perceptual enhancement helps PND survive as a productive process longer than would be expected otherwise.

2.2.3 Voicing disagreement

The third tendency discussed in this chapter, voicing agreement in clusters of obstruents, is a universal phonetic tendency as well with all three conditions fulfilled. Clusters that agree in voicing are typologically very common and voicing assimilation is one of the most common processes cross-linguistically (Myers 2010). Myers (2010) lists at least 28 such languages and the list is neither exhaustive nor does it result from a survey. The phonetic motivation is straightforward: laryngeal features "overlap and blend" in obstruent clusters which results in passive voicing of preceding voiceless stop due to laryngeal coarticulation (Myers 2010:164ff.). Perceptual factors have also been proposed to promote agreement of laryngeal features (Myers 2010). Finally, agreement in laryngeal features fulfills the third condition of universal phonetic tendencies: voicing before another voiced stop is a passive phonetic tendency[14] (Barry and Teifour 1999) and is attested in languages in which voicing assimilation is not a complete synchronic process. As Myers (2010) claims, "[a]coustic studies have shown that there is a longer voiced interval in an obstruent before a voiced consonant than before a voiceless consonant in English

---

[14] See Section 2.1.1 for a definition on the passive operation of a phonetic tendency.

(Haggard 1978; Docherty 1992: 165; Stevens et al. 1992; Smith 1997; Jansen 2004), French (Snoeren et al. 2006), and Syrian Arabic (Barry and Teifour 1999)" (Myers 2010:164).

2.2.4 Final voicing

It is safe to say that word-final or coda voicing (/T/ → [D] / __#) is one of the most thoroughly discussed phonological process in the literature. It is assumed to be a highly unnatural process that is either impossible or unattested as a synchronic phonological process.

Its opposite process, final devoicing has clear articulatory and perceptual motivations: phonation is difficult to maintain during closure, and this difficulty is even greater word-finally, where stops are produced "with reduced pulmonary pressure" (Iverson and Salmons 2011: 1633, Blevins 2004). Word-final stops also tend to be longer due to final lengthening (Berkovits 1993) which might additionally contribute to devoicing of stops. Moreover, cues for presence or absence of voicing are perceptually impoverished in final position (Steriade 1997, Iverson and Salmons 2011). Passive phonetic devoicing in word-final position is attested even in languages without phonological final devoicing. Word-final devoicing thus fits the bill for a universal phonetic tendency: (a) it has a well-motivated phonetic explanation; (b) there exists a phonetic tendency to devoice final stops even in languages without phonological devoicing; (c) it is common and well-attested cross-linguistically.

Thus, any putative process of word-final voicing would necessarily operate against a universal phonetic tendency. Kiparsky (2006) claims that FV is never attested as a productive synchronic process despite several diachronic scenarios that could lead to it (the scenarios are identified in Kiparsky 2006). In fact, he goes a step further and claims that FV is not only unattested, but also impossible and that cognitive restrictions of synchronic grammar are responsible for this typological gap.

Because of these claims, the (non-)existence of FV has become a battleground for the discussion on factors that influence phonological typology. Non-existence of FV is used as evidence in favor of Analytic Bias approach that claims cognitive restrictions shape the typology: if diachronic explanation (Channel Bias; Moreton 2008) is unable to explain the systematic gap, it has to be Universal Grammar that rules out FV. Blevins (2004) presents several cases of FV, but Kiparsky (2006) argues that none of the apparent cases of final voicing reported there qualifies as a case of synchronic final voicing — or, at least, the described processes have competing alternative explanations.

FV is (among others) reported in Lezgian (Haspelmath 1993, Fallon 1995, Yu 2004, Gajdarov et al. 2009). Haspelmath (1993) and Yu (2004:77) report that Lezgian distinguishes four stop series prevocalically (plain voiced, voiceless ejective, voiceless aspirated, and plain voiceless), which in coda position get reduced to a three-way distinction: the plain voiceless series and voiced series merge into a single voiced series. In other words, Lezgian has a synchronic phonological alternation that targets an unmarked segment, word-final unaspirated voiceless stop, and turn it into a marked segment, voiced stops — final voicing (T → D / __#). This process is limited to monosyllabic words. The examples in Table 2.2 illustrate FV.

Table 2.2: Final voicing in Lezgian (from Haspelmath 1993).

Place	**__#**	**V__V**	
bilabial	rab	rapar	'needle'
dental	pad	patar	'side'
velar	mug	mukar	'nest'
uvular	raq	raɢar	'sun'
dental	warz	wartsar	'moon'
post-alveolar	raʒ	raʧar	'grain'

The phonetic study in Yu (2004), however, shows that underlying voiced and plain voiceless series do not neutralize completely: there exists a statistically significant difference between the outcome of the two series in word-final position. Voiced consonants that derive from underlying plain voiceless stops have a significantly longer closure duration (approximately 25-32ms longer in raw means) as well as a longer duration of voicing into closure than the original voiced series (approximately 24-34ms longer in raw means; Yu 2004). If we wish to maintain that Lezgian voices final stops, we must, at the same time, assume that these consonants receive (at least phonetic) lengthening as well. It is unclear from a synchronic perspective why this should happen.

The fact that the two series do not neutralize completely allows Kiparsky (2006) to propose an alternative analysis. He assumes that the Lezgian synchronic phonological system has four series of stops — but unlike Yu, he proposes that the fourth series consists of voiced geminates. Thus, instead of coda voicing, he assumes that the process in Table 2.2 is in fact onset degemination and devoicing (/Dː/ → [T] / $_{\sigma}$[__). The two analyses are summarized in Table 2.3.

Table 2.3: Different input analyses of Lezgian stops in Yu (2004) and Kiparsky (2006).

Yu (2004)			**Kiparsky (2006)**		
Input	Output		Input	Output	
	__V	__#		$_{\sigma}$[__	__]$_{\sigma}$
D	D	D	D	D	D
T’	T’	T’	T’	T’	T’
T^h	T^h	T^h	T^h	T^h	T^h
T	T	D	Dː	T	Dː

Kiparsky’s (2006) analysis, too, has its shortcomings: like Yu, Kiparsky has to devise a two-step process — devoicing and degemination of voiced geminates in onset position — and onset devoicing is not a particularly common process in its own right. However, this derivation is by no means impossible, and Kiparsky (2008) provides evidence from other languages such as Mordva, Ewondo, and Lac Simon Algonquian (Iverson 1983) demonstrating that initial devoicing is a possible synchronic phonological process. As a sound change, such a development may be attested in Anatolian and in Selkup of the Samoyedic group (Kümmel 2007).

In sum, although Lezgian provides one of the stronger cases of final voicing, this analysis encounters two major problems. First, the voicing process is limited to monosyllabic words. Second, the plain voiceless and voiced series do not neutralize completely in coda position; a phonetic difference remains detectable between the two series. These problems pave the way for alternative proposals that analyze the alternating series as underlyingly voiced and geminate and assume that the synchronic phonological process in Lezgian is in fact onset devoicing rather than final voicing (non-alternating voiced stops are analyzed as voiced and short). Additionally, a lack of speaker data from nonce-word tests makes it difficult to determine how productive this process actually is. Further investigations of this typologically rare case of a "truly" unnatural process are required for more conclusive results: wug-tests would provide information on productivity; dialectal research might reveal dialects that neutralize voiceless and voiced stops completely in word-final position.

In addition to Lezgian, one of the most prominent cases of final voicing has been reported in Proto-Indo-European (PIE).[15] Proto-Indo-European voiceless stop *t (and *p) are reported to surface as /d/ (and /b/) in word-final position, which is most directly reflected in the Italic branch (e.g. PIE *$\mathrm{d^heh_1}$k-e-t > Old Latin ⟨FECED⟩; see Lipp 2016 and literature therein). While the evidence for this sound change is sparse and limited morphologically, it does not seem to have strong counter evidence. A closer treatment of the supposed final voicing, however, reveals that stops, orthographically represented as voiced in the ancient languages, are most likely unreleased

---

[15] While scholars disagree on whether the alleged voicing of word-final voiceless *-t (and *-p) operated only in the Italic branch or at the Proto-Indo-European stage, I will assume that the sound change operated at the proto-stage based on evidence for devoicing in branches, such as Hittite, Italic, Germanic, and indirectly Avestan (see Lipp 2016).

"lenis" and, crucially, voiceless [d̥/t̚] (Martinet 1950, Lipp 2016). In sandhi positions before voiced segments these lenis voiceless unreleased stops receive assimilatory voicing, but are likely voiceless in other positions (Lipp 2016). Several indirect evidence point in this direction. In Avestan, for example, word-final *-t (most likely unreleased [t̚]) is represented by a distinct grapheme ⟨t̰⟩. In closed sandhi positions before voiced segments, this unreleased [t̚] ⟨t̰⟩ gets realized as a voiced stop, e.g. ⟨at̰⟩ 'then' vs. ⟨ad āiš⟩ 'then by them' (Lipp 2016). There is further orthographic evidence that in Italic word-final stops represented by ⟨d⟩ were voiceless. In Oscan, we can find pairs in which final ⟨d⟩ in Latin alphabet is represented by Greek unaspirated voiceless ⟨τ⟩, e.g. ⟨*deded*⟩ vs. ⟨δϵδϵτ⟩ (Martinet 1950). Finally, as Lipp (2016) points out, the voiceless lenis nature of final stops is described by ancient Indic grammarians. He translates Atharva-Prātiśākhya: "words ending in voiceless stops, end according to Śaunaka in voiced stops without having that value, but having lax contact" (Lipp 2016:257).[16] The phonetics of the reported final voicing in Proto-Indo-European and its branches thus most likely does not target feature voice, but is rather a weakening of final stops that surface as unreleased "lenis" voiceless stops.

The current position of this dissertation is that final voicing does not exist as a productive synchronic alternation. The cases discussed so far and in Kiparsky (2006) are either problematic or we lack sufficient phonetic information. Further cases of final voicing have been reported in the literature, but have gone largely unnoticed in the literature on naturalness. Perhaps the most prominent is the case of Lakota. While some studies do not detect final voicing in Lakota

---

[16] Kiparsky (2006) even suggests that word-final orthographic ⟨d⟩ in Latin represents a voiced fricative, but evidence for a fricative articulation is sparse at best. Even if PIE final stops were fully voiced segments, a position that is difficult to verify, there exist a number of diachronic explanations of how they could arise from a combination of sound changes (see Lipp 2016 and discussion therein). Two reasonably possible scenarios involve an interstage in which final stops are either unreleased and glottalized or implosive. These two types of segments would result in surface voiced stops under some reconstructions of the Pre-Proto-Indo-European consonant system (for a detailed discussion, see Lipp 2016).

(Carter 1974, Shaw 1980), Rood (2016) and Riggs (1893) report that voiceless stops alternate with voiced stops in the coda position in Lakota. While a thorough synchronic and diachronic treatment of Lakota is not possible at this point due to the lack of phonetic descriptions, I outline a potential diachronic explanation for the development of FV in Lakota in fn. 58.[17] FV has also been reported as a phonotactic restriction (e.g. in Ho, some varieties of Spanish, and Tundra Nenets; Pucilowski 2013, Burkard and Dziallas 2018, Salminen 1997), but lack of phonetic and phonological evidence in these cases also prevent us to make any firm conclusions (see also Section 3.4). Further studies on these and other languages are needed to reach a more conclusive position on the status of FV.

In sum, based on typological and phonetic evidence, this section argued that IVD, PND, FV, and disagreement in voicing are unnatural processes as defined in (2). In the following, I present a typological study of unnatural alternations and phonotactic restrictions that target feature [±voice]. I discuss historical and synchronic characteristics of the reported cases of unnatural processes that will be relevant for our model of explaining unnatural processes diachronically (Chapter 3).

2.3 Post-nasal devoicing

In this section, I collect all known cases of PND. The collection derives from several sources: a survey of sound changes in Kümmel (2007) that examines approximately 294 languages; a UniDia database of 10,349 sound changes from 302 languages (Hamed and Flavier 2009); a survey of unnatural sound changes in Blust (2005), Blevins (2008a), and Goddard (2007); a

---

[17] Another, but even less certain case of FV is suggested for a Munda language Juang (Donegan and Stampe 2002).

survey of the *NT constraint in Hyman (2001); and a recent description of post-nasal devoicing in Brown (2017).[18] Isolated cases have been reported in Merrill (2014, 2016a,b).

So far, all cases of PND have been treated in isolation, which has led to several opposing explanations of the phenomenon. Explanations of PND rely variously on appeals to hypercorrection (Blust 2005, Xromov 1972), combinations of sound changes (Dickens 1984, Hyman 2001), claims that unnatural sound changes *do exist* (Blust 2005), or claims that PND is articulatorily or perceptually motivated (Solé 2012, Gouskova et al. 2011, Stanton 2016a). I shine light on this murky discussion by showing that, in all thirteen cases[19] of PND I have compiled, there exists either direct or strong indirect evidence that PND emerges as the combined result of three separate instances of single, natural sound changes (as has been argued for Tswana in Dickens 1984 and Hyman 2001). I focus primarily on evidence from Yaghnobi and Sogdian, showing that the two languages present direct historical evidence that PND results from a combination of sound changes. Examining all cases of PND together also allows me to generalize common properties and develop a diachronic model for explaining unnatural phenomena. Furthermore, showing that one of the rare reported unnatural sound changes is in fact a product of a combination of natural sound changes lends support to the position that sound change has to be natural and cannot operate against UPTs (*pace* Blust 2005).

That combinations of sound changes produce unmotivated results is a long-standing and well-known claim. “Telescoping,” for example, describes a phenomenon in which a sound change A

---

[18] It is difficult to estimate how many languages are surveyed in these collections, as authors often fail to report this information. A reasonable guess would be that Blust (2005) surveys a large sample of Austronesian languages and Hyman (2001) surveys a large sample of Bantu languages.

[19] A more conservative count gives nine cases of PND; some scholars count Sicilian and Calabrian and Tswana, Shekgalagari, and Makhuwa as two cases rather than five, due to the close genetic relationships between these languages.

> B in the environment X is followed by B > C, resulting in a sound change A > C that may not be phonetically motivated in environment X (Wang 1968, Kenstowicz and Kisseberth 1977: 64, Stausland Johnsen 2012). This dissertation, however, takes the concept of telescoping one step further, by focusing on alternations that are not only unmotivated, but that operate in exactly the opposite direction of UPTs. I show that for unnatural processes to arise we need a special type of combination of sound changes which I term the Blurring Process.

2.3.1 The data

According to my survey, the existence of PND as a sound change has been reported in thirteen languages and dialects from eight language families.

2.3.1.1 Yaghnobi

PND was first proposed for Yaghnobi by Xromov (1972). Yaghnobi is an Iranian language, spoken by approximately 13,500 speakers in five different areas of Tajikistan (Paul et al. 2010:4). It is the only living descendant of Sogdian, an Eastern Iranian language that was spoken around the fourth century CE. Xromov observes that NT sequences in Yaghnobi correspond to ND sequences in ancestral Sogdian; on the basis of this observation, he posits a sound change D > T / N__ in the development from Sogdian to Yaghnobi. The following table lists cognates from Yaghnobi and Sogdian that confirm this correspondence.[20]

---

[20] Data from older descriptions has been adjusted throughout this dissertation, as accurately as the descriptions allow, to fit IPA conventions.

Table 2.4: PND in Yaghnobi (from Xromov 1972: 128).

Yaghnobi	**Sogdian**	**gloss**
ɣantum	ɣandum	‘wheat’
ʃikampa	əʃkamb	‘stomach’
sank(a)	sang	‘stone’
rankiːna	rang	‘color’
unkuʃt	anguʃt	‘finger’
ʧintir	ʧəndər	postp.
-ant	-and	3rd pl.

Outside of post-nasal position, original voiced stops surface as voiced fricatives [β, ð, ɣ] in Sogdian: this is the result of a sound change that turns voiced stops into voiced fricatives except post-nasally. In Yaghnobi, velar and labial fricatives in the elsewhere condition are preserved as fricatives [w/β] and [ɣ]. The Sogdian dental fricative [ð], on the other hand, surfaces as a stop: [dah] ‘ten’ for Sogdian [ðəsa] (data from Novák, 2010: 31, Novák 2014). Voiceless stops remain unchanged in Sogdian, except post-nasally, where they become voiced (Yoshida 2016) before being devoiced in Yaghnobi.

Regarding the Yaghnobi synchronic system, Novák (2010) reports voiceless stops to be aspirated except pre-consonantally and voiced stops to be phonetically voiced, although no instrumental studies are offered. Synchronic Yaghnobi phonology contrasts voiced and voiceless stops in all positions: initially, finally, intervocalically, in clusters, and post-nasally. The fact that [±voice] contrasts fully in NT and ND sequences in synchronic Yaghnobi means that PND is not an active alternation anymore, but this is likely secondary, introduced late in the language's development through borrowings from Tajik (cf. Xromov 1972). Some examples of borrowed ND sequences from Tajik are given in Xromov (1972: 128): [angiʃt] ‘coal’, [ʧang] ‘dust’, [baland] ‘high’, [lunda] ‘round’. It is even possible to find pairs of native and borrowed words with and without devoicing. For example, the inherited Yaghnobi [vant] ‘tie’ — with an unvoiced stop after a nasal — stands in contrast to the borrowed [band] ‘tie’ with a voiced variant.

2.3.1.2 Tswana, Shekgalagari, and Makhuwa

PND has been reported as a synchronic phonological process in the languages/dialects of the Sotho-Tswana group, especially in Tswana and Shekgalagari (Hyman 2001, Solé et al. 2010), but also in Sotho (Janson 1991/1992), three closely related and mutually intelligible Southern Bantu languages (Makalela 2009). Because PND is least well-described in Sotho, I will focus on Tswana and Shekgalagari. Tswana is spoken by approximately 4–5 million people in Botswana, Namibia, Zimbabwe, and South Africa (Coetzee and Pretorius 2010), and Shekgalagari by approximately 272,000 people in Botswana (Solé et al. 2010).

In Makhuwa, PND is reported as a sound change followed by nasal deletion (Janson 1991/1992). Makhuwa is closely related to Sotho-Tswana too and is spoken by approximately 3 million speakers in Mozambique (Lewis et al. 2015).

Table 2.5 shows that synchronic voiced stops (that surface as a voiced stops in the elsewhere condition) in Shekgalagari become voiceless when preceded by a nasal. Voiceless stops remain unchanged both post-nasally and elsewhere.

Table 2.5: PND in Shekgalagari (table from Solé et al. 2010).

No *N*-prefix	***N*-prefix**	**Gloss**
χʊ-pak-a	χʊ-m-pak-a	‘to praise’
χʊ-tʊt-a	χʊ-n-tʊt-a	‘to respect’
χʊ-cʊ́b-á	χʊ-ɲ-cʊ́b-á	‘to beat’
χʊ-kɛl-a	χʊ-ŋ-kɛl-a	‘to show’
χʊ-bɔ́n-á	χʊ-m-pɔ́n-á	‘to see’
χʊ-dʊʒ-a	χʊ-n-tʊʒ-a	‘to annoint’
χʊ-ɟís-a	χʊ-ɲ-cís-a	‘to feed’
χʊ-at-a	χʊ-ŋ-kat-a	‘like’

Both Tswana and Shekgalagari have unaspirated voiced (D) and voiceless (T) and aspirated voiceless stops (T^h) in their inventories (for phonetic studies, see Coetzee and Pretorius 2010 and Solé et al. 2010): they surface initially, intervocalically, and post-nasally. The only permitted syllable structures are "V, CV, CGV, N" (where G = glide and N = syllabic nasal; Coetzee and Pretorius 2010). Detailed instrumental acoustic studies of the Tswana and Shekgalagari stop system in Coetzee and Pretorius (2010) and in Solé et al. (2010) confirm that post-nasal devoicing is indeed realized as change in the feature [±voice], and not as a change in [±spread glottis].

Several peculiarities need to be noted with respect to Tswana. First, /g/ never surfaces as a voiced stop: while it is devoiced to [k] post-nasally, it gets deleted elsewhere, e.g. [χʊ-at-a] for /χʊ-gat-a/ (cf. [χʊ-ŋ-kat-a], Solé et al. 2010). Second, voiced alveolar stop [d] surfaces as an allophone of /l/ before the high vowels /i/ and /u/ (Coetzee and Pretorius 2010). Third, voiced stops in nasal clusters of secondary origin (after syncope) do not undergo devoicing, but undergo assimilation in Tswana (see Table 2.6). Finally, the nasal in NT sequences Tswana (and Shekgalagari) is retained when stressed and in the 1st person object prefix, but gets deleted elsewhere (see Dickens 1984). Hyman (2001) gives an example of preservation of N-: in the case of class 9 and 10 prefix, when roots are monosyllabic.

In Shekgalagari, secondary ND clusters (from NVD after vowel deletion) remain voiced (Table 2.6); /d/ can surface in the elsewhere condition (not only before /i/ and /u/), but it does not alternate with devoiced [nt] ← /nd/ (because /d/ corresponds to Tswana /tl/). Shekgalagari (unlike Tswana) also features the palatal series of stops that enters PND (Solé et al. 2010). For a detailed discussion of differences between Tswana and Shekgalagari, see Solé et al. (2010) and Dickens (1984).

Table 2.6: PND in Tswana and Shekgalagari (table from Solé et al. 2010).

Ts.& Sh.	/χʊ-m-bɔ́n-á/	→	[χʊmpɔ́ná]
Sh.	/χʊ-mʊ-bɔ́n-á/	→	[χʊmbɔ́ná]
Ts.	/χʊ-mʊ-bɔ́n-á/	→	[χʊmmɔ́ná]

Makhuwa is also reported to undergo PND as a sound change. Sequences of a nasal and a voiced stop (ND) develop to voiceless unaspirated stops (T) (Janson 1991/1992), but the process did not develop into a synchronic alternation in Makhuwa. To my knowledge, no elaborate accounts of the phonetic realization of voiced and voiceless stops in the elsewhere condition exist for Makhuwa.

Because the languages are closely related and there are other instances of common innovation between Sotho-Tswana and Makhuwa, Janson (1991/1992) and Hyman (2001) imply that PND in the three languages is likely a common innovation. For the same reason, I will treat these three languages together in this study. Since the description of PND in Makhuwa (and Sotho) is sparse and lacking in detailed phonetic descriptions, I will focus my discussion on Tswana and Shekgalagari; however in principle, the arguments for these two languages apply to Makhuwa as well.

2.3.1.3 Bube and Mpongwe

Unlike Makhuwa, Bube is not closely related to Tswana and Shekgalagari. Janssens (1993) reports that Bube also features PND. As will be shown below, several aspects of Bube PND are highly reminiscent of the process reported for Tswana, Shekgalagari, and Makhuwa despite the languages not being closely related. Bube is a Northwest A Bantu language, spoken by approximately 51,000 speakers on Bioko island (Lewis et al. 2015). Sequences of a nasal and a voiced stop in Pre-Bube develop to voiceless stops in Bube. The following table illustrates the development.

Table 2.7: PND in Bube (table from Janssens 1993).

Pre-Bube	**Bube**	**gloss**
*è-m-bódì	àpóɽîì	‘goat’
*è-m-bóà	è-pwáà	‘dog’
*è-n-címbá	è-cìppà	‘wild cat’
*-dámb-	-lápáà	‘cook’
*-gɛ̀nd-	-ɛ̀tà-	‘walk’
*-gàŋgà	-àkká	‘root’
-kàŋg- ∼-bàŋg-	-àk	‘attach’

Table 2.7 illustrates that Pre-Bube sequences of a nasal and voiced stop (ND) yield a single voiceless stop (T) with the nasal being lost. In the elsewhere condition, the labial voiced stop surfaces as such, the alveolar develops to [r/l], and the voiced velar stop gets lost (similar to the situation in Tswana and Shekgalagari) (Janssens 1993). Voiceless stops can either delete, develop to [h] (in the labial series), or continue to surface as voiceless stops. NT sequences develop to a plain voiceless stop (T; Janssens 1993). The Bube synchronic system has no alternation between voiceless and voiced stops in the post-nasal vs. elsewhere condition, but the development is intriguing from a diachronic perspective: it appears as if PND operated in Pre-Bube.

In the Mpongwe dialect of Myene (Bantu B language, spoken in Gabon, Lewis et al. 2015), PND is reported marginally for root initial [g] after some prefixes that historically ended in a nasal, e.g. *gàmb- > [i-kamba] (Mouguiama-Daouda 1990). However, because PND is marginal in Mpongwe — it is morphologically limited and does not apply categorically — I will for the most part leave it out of the ensuing discussion.[21]

---

[21] The UniDia survey (Hamed and Flavier 2009) reports that PND targets the labial series of voiced stops in Lembaama. However, I was unable to find a description of this development in the literature.

2.3.1.4 Konyagi

Recently, PND as a sound change has been discovered in Konyagi (also known as Wamey, among others; Merrill 2014, 2016a,b). Konyagi, a member of the Atlantic subfamily of the Niger-Congo group, is spoken by approximately 21,000 speakers in Senegal (Lewis et al. 2015). Note that Konyagi is not part of the Bantu family, which means that it is only very distantly related to the Bantu languages above with PND. Merrill (2014, 2016a,b, p.c.) reconstructs a detailed picture of Konyagi's pre-history. Notable in this reconstruction is a series of voiceless stops in post-nasal position that correspond to voiced stops in the neighboring languages Bedik and Basari of the Tenda group (data in the Table 2.8 is from Merrill 2014, 2016a,b, based on Ferry 1991 and Santos 1996). It thus appears as if Konyagi underwent PND. Voiceless NT sequences develop to a plain voiceless stop (T) in Konyagi. Both voiceless and voiced stops fricativize in the elsewhere condition. Table 2.8 illustrates PND in Konyagi.

Table 2.8: PND in Konyagi (from Merrill 2014, 2016a,b).

Konyagi	**Bedik**	**Basari**	**gloss**
æ̀-jamp	u-jāmb	ɔ-jămb	'millet stalk'
ì-ñə́mp	ɔ-ñʌ̀mb	a-j̃ʌ̌mb	'plunge/immerse'
ì-ntæ̀w̃	gi-ndám	ɑ-ndɑ́w̃	'animal/spirit'
ì-kônt	ɔ-hɔ̀nd	a-xɔ̀nd	'snore'
æ̀-ncæ̀nk	ga-nɟáng	a-nɟɑ̀ng	'Pterocarpus erinaceus (treesp.)'
æ̀-ncə́l̰	gɔ-nɟʌ̀l	ɑ-nɟʌ́n	'caterpillar'
ì-jæ̀nk	u-jáng	a-j̃ɑ̀ng	'be long'
ì-nkòt	gɛ-ngɔ́t	ɛ-ngɔ̀t	'pole'

Synchronically, Konyagi features voiceless and voiced stops as well as voiceless and voiced fricatives. Non-nasal clusters are not permitted. Post-nasally, only voiceless stops are allowed. Elsewhere, voiceless and voiced stops can contrast with voiceless and voiced fricatives, except initially where only fricatives are allowed (Merrill 2014, 2016a,b). PND is part of a synchronic mutation process in Konyagi: devoiced stops after prefixes that historically ended in a nasal alternate with voiced fricatives after prefixes that historically ended in a vowel and voiced stops after prefixes that ended in another consonant (see Table 2.9 and Merrill 2014, 2016a,b). No instrumental phonetic data is available for Konyagi.

Table 2.9: Konyagi PND "mutation" depending on the historical sources of prefixes (table from Merrill 2014, 2016a,b, Santos 1996).

	-V$_{V}$	**-V**$_{C}$	**-N**	**gloss**
bilabial	-wə́nkáák	-bə́nkáák	-mpə́nkáák	'beside'
alveolar	-lə̀mæ̀xə́	-də̀mæ̀xə́	-ntə̀mæ̀xə́	'sweet'
palatal	-jə̀læ̀xə́	-gə̀læ̀xə́	-nkə̀læ̀xə́	'rotten'
velar	-wúnkǽx	-gúnkǽx	-nkúnkǽx	'bitter'

2.3.1.5 South Italian Dialects

Sicilian and Calabrian are dialects of Italian spoken in the corresponding regions of Italy by approximately 4.7 million speakers (Lewis et al. 2015). PND has been reported for these dialects in Rohlfs (1949, 424f.). The peculiarity of South Italian PND is that the sound change targets only the voiced affricate *ʤ, which is devoiced to [ʧ] after the nasal [n] (*ʤ > [ʧ] / N__). Elsewhere, *ʤ develops to [j] in these dialects (Rohlfs 1949, 424f.; Kümmel 2007, 376). Voiced stops are not reported to be devoiced in the post-nasal position: the feature [±voice] is contrastive for stops. The voiceless affricate either remains unchanged or develops to [ʃ] (Rohlfs 1949). Table 2.10 illustrates PND in Sicilian and Calabrian.

Table 2.10: PND in South Italian (from Rohlfs 1949, 424f.).

S.-Ital. dial.	**Standard**	**Gloss**
anʧilu	anʤelo	'angel'
pinʧiri	pinʤere	'push'
kianʧiri	pjanʤere	'to cry'
finʧiri	finʤere	'to feign'
tinʧiri	tinʤere	'to dye'

2.3.1.6 Buginese and Murik

PND has been reported in three Austronesian languages: Buginese, Murik, and the Bengoh dialect of Land Dayak (Blust 2005, 2013). PND in the latter is simply mentioned without accompanying data (Rensch et al. 2006:69; Blust 2013); I therefore leave Land Dayak out of the discussion that follows. Buginese is spoken by approximately 5 million people in Sulawesi (Indonesia); Murik is spoken in Sarawak (Malaysia and Brunei) by approximately 1,000 speakers (Lewis et al. 2015). These three Austronesian languages are not closely related, so we cannot attribute PND to developments in a common ancestor; it is likely that PND developed independently in all three branches (Blust 2005, 2013).

Apparent PND in Buginese is represented in Table 2.11, showing the development of Proto-Malayo-Polynesian (PMP) voiced stops (data from Blust 2013). Velar stops are devoiced post-nasally. Labial stops appear devoiced after nasals, but surface as [w] initially and word-internally (with a sporadic reflex [b] in initial position). The dental stop *d is not implicated in PND; Pre-Buginese *d develops to /r/ in all positions, which does not undergo devoicing, e.g. *dindiŋ > [renriŋ]. Word-initially, however, *d is sporadically preserved as a voiced stop [d] or develops to [l]. The voiced fricative *z is occluded to a voiced palatal stop initially, and develops to a sonorant [r] intervocalically. Post-nasally, *z is devoiced to [c]. Word-finally, all stops develop to [ʔ]. Voiceless NT sequences develop to a geminate voiceless stop (TT) in Buginese (Mills 1975, Blust 2005). Voiceless stops remain unchanged in the elsewhere position (Mills 1975).

Table 2.11: Summary of PND in Buginese.

	#__	**V__V**	**N__**
*b	b/w	w	p
*d	d/r/l	r	r
*g	g	g	k
*z	ɟ	r	c

The data cited as evidence of PND as a sound change in Buginese are as follows:

Table 2.12: PND in Buginese with cognates from reconstructed Proto-South-Sulawesi (Proto-SS; from Blust 2005, 2013).

Proto-SS	**Buginese**	
*bemba	bempa	'water jar'
*lambuk	lampuʔ	'pound rice'
*limboŋ	lempoŋ	'deep water'
*rambu	rampu	'fringe'
*rumbia	rumpia	'sago palm'
*tambiŋ	tampiŋ	'addition to a house'
*barumbun	warumpuŋ	'a color pattern'
*bumbun	wumpuŋ	'heap up'
*geŋgem	geŋkeŋ	'hold in the hand'
*tuŋgal	tuŋkeʔ	'each, single'
*aŋgəp	aŋkəʔ	'price'
*anɟap	ancəʔ	'offerings to spirits'
*ɟanɟi	ɟanci	'to promise'
*punɟuC	ma-poncoʔ	'short'

In Murik, labials, alveolars, and velars undergo devoicing in post-nasal position. In the elsewhere condition, the developments vary according to the place of articulation. Bilabial voiced stops surface as such in the elsewhere condition. Voiced dentals appear as [l] word-initially and [r] word-internally. Velars are devoiced not only post-nasally, but sporadically also in the elsewhere condition. PMP *z develops to the voiced palatal affricate [ɟʝ] initially, to [s] word-internally, and to the voiceless palatal affricate [cç] post-nasally. Table 2.13 illustrates the development of voiced stops from Proto-Kayan-Murik. Voiceless NT sequences develop to a plain voiceless stop (T) in Murik, while plain voiceless stops remain unchanged (Blust 1974, Blust 2005).

Table 2.13: Summary of PND in Murik (as reconstructed in Blust 2005).

	#__	**V__V**	**N__**
*b	b	b	p
*d	l	r	t
*g	g/k	g/k	k
*z	ɟʝ/ɟ	s	cç

PND is confirmed for Murik by the following examples:

Table 2.14: PND in Murik (from Blust 2005: 259f.; Blust 2013: 668).

Proto-KM	**Murik**	
*kelembit	kələmpit	‘shield’
*bumbuŋ	umpuŋ	‘ridge of a roof’
*lindem	lintəm	‘dark’
*-inda	t-inta	‘beneath, below’
*mandaŋ	mantaŋ	‘to fly’
*tundek	tuntuk	‘beak of a bird’
*lindiŋ	lintiŋ	‘wall of a house’
*undik	untik	‘upper course of a river’
*tandab	tantap	‘catch’
*andeŋ	antəŋ	‘deaf’
*pindaŋ	pintaŋ	‘blossom’
*pendan	pəntan	‘small fruit bat’
*nɟʝi	ncçi	‘one’
*menɟʝat	məncçat	‘pull’
*unɟʝuŋ	uncçuŋ	‘tip, extremity’
*anɟʝat	ancçat	‘rattan tote bag’
*tunɟʝuʔ	tuncçuʔ	‘to point, indicate’
*tuŋgan	tuŋkan	‘dibble stick’

Synchronically, PND in Buginese is reported to be an active phonological alternation (Sirk 1983) as well as an active sandhi process (Noorduyn 2012/1955; see Section 2.3.2 and Table 2.15). Buginese contrasts voiceless and voiced stops (Noorduyn 2012/1955), but voiced labial, palatal, and velar stops /b, d, g/ and a glide /w/ are reported to devoice (and change to a stop in the case of /w/) post-nasally in synchronic Buginese (see also Table 2.15 below). In Murik, PND is reported as a synchronic alternation only for the palatal stop series which alternates between [ɟ] in the elsewhere condition and [c] post-nasally. For other places of articulation, Blust (2005) reports PND to be synchronically inactive due to "phonemic restructuring".

2.3.1.7 Nasioi

The most recent report of PND is that in Brown (2017) for the South Bougainville language Nasioi, spoken in Papua New Guinea by approximately 20,000 speakers (Lewis et al. 2015). Brown (2017), based on Hurd and Hurd (1970), claims that PND in Nasioi is a synchronic alternation and supports this claim by showing that while bilabial and alveolar voiced stops /b/ and /d/ contrast in voicing with unaspirated voiceless /p/ and /t/ word-initially and after a glottal stop, only voiceless variants appear in post-nasal position. Elsewhere, stops fully contrast in the feature [±voice], but are not permitted other than word-initially and after a glottal stop: intervocalically, labial and alveolar stops surface as a voiced fricative [β] and a flap [ɾ], respectively. The velar voiced stop /g/ does not occur in the system. This distribution points to a clear case of post-nasal devoicing. PND in Nasioi is not only a phonotactic restriction, but also seems to be an active synchronic alternation, which is illustrated by the example in (5). The voiceless version of the personal pronoun -p/b- surfaces post-nasally; the voiced one surfaces elsewhere.

(5) a. *kara-**b**-ant-∅-in*
talk-him-I-SG-did
'I talked to him.'

b. *tiom-**p**-ant-∅-in*
follow-him-I-SG-did
'I followed him.'

Brown (2017) specifically argues that post-nasal devoicing is an innovation in Nasioi and that the fact that a related language, Nagovisi, shows traces of post-nasal *voicing* (with a speculation that this might be indicative of the proto-language) suggests that Nasioi PND operated as a single sound change. In other words, this suggests that PND is not "derived from the confluence of multiple independent changes" (according to Brown 2017:275).

2.3.2 PND as a synchronic alternation

In the following I argue that Buginese, Konyagi, Nasioi, and especially Tswana and Shekgalagari, confirm that a combination of sound changes *can* and *do* result in productive unnatural synchronic processes. Note that, according to Kiparsky (2008), the third sound change in the series of the three sound changes that result in PND is expected to be blocked by UG, since the combination would result in an unnatural process and the surface pattern does not allow for phonological reanalysis.[22] However, this blocking clearly does not happen: sound change that blurs the original complementary distribution and thus produces a synchronically unnatural process is attested in Buginese, Konyagi, Nasioi, and Tswana and Shekgalagari in particular.

As mentioned above, the combination of three sound changes is reported to yield a synchronic alternation in Buginese derivational morphology. Sirk (1983:35-37) shows that sequences of N + D yield NT (except for dentals), while the sequence N + [w] yields [mp]. He also argues that

---

[22] Kiparsky (2008) posits that either the changes that would cause unnatural alternations are blocked or "the system they appear to give rise to must be reanalyzed". It is unclear to me, however, how the system of PND as is attested, for example, in Tswana could be phonologically reanalyzed. It is impossible to assume that stops in post-nasal position are underlyingly voiceless (and get voiced elsewhere), because voiced and voiceless stops contrast in the elsewhere condition in Tswana. PND could be analyzed as morphologized, but Coetzee and Pretorius (2010) show that the rule extends to nonce words, although they admittedly test only one morphological environment (for a discussion against a morphological analysis, see Hyman 2001). If reanalysis is not available, we would (in line with Kiparsky's 2008 reasoning) expect the last sound change that would result in the unnatural alternation to be blocked.

the only permissible non-geminate clusters in Buginese are NT (to the exclusion of ND) (Sirk 1983:35-37).

Table 2.15: Buginese PND (from Sirk 1983:35-37).

Isolation	**gloss**	**Compound**	**gloss**
wərrə	‘heavy’	sim-pərrə	‘just as heavy’
bone	‘Bone (name)’	arum-pone	‘prince Bone’
gora	‘shouts’	samən-kora	‘loud shouts’
ɟaiʔ	‘root’	maɲ-caiʔ	‘to sew’

Noorduyn (2012/1955) describes PND in Buginese as part of a sandhi phenomenon. A word-final nasal before a word-initial labial approximant [w] in sandhi results in the sequence [mp], e.g. /rilaleŋ wanua/ becomes [rilalempanua]. Noorduyn (2012/1955) also reports that PND in sandhi occasionally targets voiced labial stops as well, e.g. /telluŋ bocco/ → [tellumpoco], but the process is no longer productive. Data on this phenomenon are sparse, however, and detailed phonetic descriptions are lacking.

PND is also reported in Konyagi as part of a synchronic consonant mutation process in adjectives, depending on the prefix and its grade of mutation (Merrill 2014, 2016a,b). Adjectives surface with a sonorant in initial position after vowel-final prefixes that go back to vowel-final prefixes (-V_V), while voiced stops surface in the initial position of these adjectives after vowel-final prefixes that go back to consonant-final prefixes (-V_C). Finally, adjectives surface with voiceless stops in initial position after a nasal-final prefix (-N) (Merrill 2014, 2016a,b).

Table 2.16: Konyagi PND (table from Merrill 2014, 2016a,b, Santos 1996), repeated from Table 2.9.

	-V_V	**-V_C**	**-N**	**gloss**
bilabial	-wə́nkáák	-bə́nkáák	-mpə́nkáák	‘beside’
alveolar	-lə̀mæ̀xə́	-də̀mæ̀xə́	-ntə̀mæ̀xə́	‘sweet’
palatal	-jə̀læ̀xə́	-gə̀læ̀xə́	-nkə̀læ̀xə́	‘rotten’
velar	-wúnkǽx	-gúnkǽx	-nkúnkǽx	‘bitter’

PND is reported as a synchronic alternation in Nasioi as well. Voiceless and voiced labial and alveolar stops contrast initially and after a glottal stop. Intervocalically, voiced stops are lenited to a voiced fricative and a flap, but post-nasally they devoice and merge with voiceless stops. The alternation (PND) that might be morphologically limited is reported for two morphemes: the second person object suffix /-d/ and the third person object suffix /-b/ (e.g. [oo-**d**-a-∅-maan] 'I see you' vs. [manton-**t**-a-∅-maan] 'I feel you'; Brown 2017).[23] Further research is required to establish the synchronic status of PND in Nasioi.

Buginese, Konyagi, and Nasioi have a synchronic process of PND, although the scope of PND in these languages is limited in that either it involves an alternation Z ∼ NT instead of D ∼ NT, or it is morphologically limited, or it targets only a subset of voiced stops. Detailed phonetic descriptions of PND in these languages are lacking. In Tswana and Shekgalagari, on the other hand, PND is not only reported as a productive synchronic alternation, but we also have detailed acoustic and experimental studies of it.

Two major contributions addressing the synchronic status and productivity of PND are Coetzee and Pretorius (2010) and Solé et al. (2010). These papers first show that the process in question in Tswana and Shekgalagari is in fact a phonetic devoicing of stops in post-nasal position. For seven speakers of Tswana (out of twelve total), voiced stops devoice and completely merge with the voiceless series post-nasally (Coetzee and Pretorius 2010); for example, /m-bV/ completely merges with /m-pV/, and the stops in post-nasal position (both the underlyingly voiceless and the devoiced) "agree nearly completely" in all relevant parameters with /re-pV/ and

---

[23] Hurd and Hurd (1970) suggest that /-b/ can be optionally deleted in intervocalic position.

significantly differ from the underlying /re-bV/. The same conclusion is drawn for Shekgalagari based on acoustic and laryngographic data in Solé et al. (2010).[24]

Second, Coetzee and Pretorius (2010: 411) show that PND is fully productive, extending to nonce-words at the same rate as it applies to the native vocabulary in Tswana. The results from nonce-word experiments suggest that PND is not lexicalized, but rather a productive phonological process in the synchronic grammar of Tswana speakers. This means that unnatural alternations (according to the definition in (2)) produced by combinations of sound changes can become part of productive synchronic grammars.

Finally, Coetzee and Pretorius (2010) show that the natural process of PNV operates passively even in cases in which PND is a phonological process. The system containing the unnatural synchronic phonological process PND is attested for seven of the thirteen speakers. For the other five speakers, however, Coetzee and Pretorius (2010: 417) observe that they often realize the whole closure "with voicing" in post-nasal position. This suggests that the five speakers have introduced a new rule into their system, which is the natural, phonetically motivated, and exact inverse process to PND: post-nasal voicing.[25]

2.4 Unnatural trends in the lexicon

In addition to unnatural categorical alternations like PND, my typological study shows that phonotactic restrictions that are not categorical, but gradient can be unnatural as well. This section shows that trends in the lexicon that operate against universal phonetic tendencies are

---

[24] Note that nasals in both Tswana and Shekgalagari are always realized as syllabic.

[25] The new rule, post-nasal voicing, might also be interpreted as dialect mixing.

statistically significant, phonetically real, and point to evidence in favor of productivity of these processes.

2.4.1 Background

In Optimality Theory (OT; Prince and Smolensky 1993/2004) and related theories (Harmonic Grammar: Legendre et al. 2006, Pater 2008, 2009; Maximum Entropy Grammar: Goldwater and Johnson 2003, Hayes and Wilson 2008), two questions have recently received increased attention in the literature: (i) how to represent gradient phonotactic restrictions in the grammar (Frisch et al. 2004, Anttila 2008, Coetzee and Pater 2008, Wilson and Obdeyn 2009) and (ii) whether and how to represent unnatural processes in the grammar (Hayes 1999, Buckley 2000, Hyman 2001, Blevins 2004, 2008, Yu 2004, Blust 2005, Wilson 2006, Hale and Reiss 2008, Samuels 2009, Carpenter 2010, Coetzee and Pretorius 2010, Becker et al. 2011, White 2013, Hayes and White 2013, i.a.). To my knowledge, however, there exists no systematic treatment of the intersection of these two topics: unnatural gradient phonotactics, i.e. phonotactic restrictions that, given a particular environment, target a single (segmental) feature and gradiently favor the value of this feature that is unnatural in that environment.

Phonotactic restrictions and trends in the lexicon that are phonetically unmotivated are especially relevant for current phonological theory. Phonetically unmotivated phonotactics are primarily discussed in light of the "surfeit of the stimulus" problem: there exist unmotivated statistically significant trends in the lexicon, but speakers often fail to generalize these unmotivated trends to nonce words (see, for instance, Becker et al. 2011, 2012, Hayes and White 2013). Hayes and White (2013), for example, identify ten potential "unnatural" phonotactic restrictions based on systematic gaps in the English lexicon and test in a behavioral experiment whether those restrictions are internalized by native speakers. Using Hayes and Wilson's (2008)

Phonotactic Learner to find systematic gaps, they find that, for instance, the English lexicon features an unmotivated restriction against /ʒ/ before stressed vowel followed by an obstruent (*[+continuant, +voice, −anterior] [+stress] [−son] in constraint formalism; Hayes and White 2013). Most studies on this problem agree that speakers do not generalize unmotivated phonotactic restrictions to nonce words, with the suggestion that they might not be represented in grammar (Pycha et al. 2003, Becker et al. 2011, Becker et al. 2012, Hayes and White 2013, Gallagher 2016, Becker et al. 2017). Other studies (for instance, Hayes et al. 2009) suggest, instead, that unmotivated processes can actually be generalized to nonce words, but to a lesser extent than natural processes. Thus, the literature does not agree on the grammatical status of phonetically unmotivated phonotactic restrictions.

The other aspect of the data presented here, beside its unnaturalness, that bears implications for phonological theory is the fact that restrictions are *gradient* rather than categorical. The necessity of encoding gradient phonotactics in the grammar has recently been motivated on the basis of gradient lexicalized cooccurrence restrictions (Berkley 2000, Frisch et al. 2004, Anttila 2008, Coetzee and Pater 2008, Wilson and Obdeyn 2009, Zuraw 2010). Coetzee and Pater (2008), for example, discuss two cases of homorganic consonant co-occurrence restrictions: Arabic and Muna. In both cases, an Obligatory Contour Principle constraint (Leben 1973, McCarthy 1986) creates a pressure that is strong, but fails to reach categorical status because of lexical variation: homorganic consonants are dispreferred within the same root, but not completely ruled out. This dispreference, however, operates in the *natural* direction: restriction against co-occurrence of homorganic consonants is strongly motivated by articulatory (motor planing; see Garrett and Johnson 2013) as well as by perceptual factors (Gallagher 2010).

To my knowledge, no systematic treatment of unnatural gradient phonotactic restrictions exists in the literature. Nazarov's (2008) treatment of Tarma Quechua is in this respect crucial, as the paper discovers that in TQ rates of application of voicing operate in "phonetically unmotivated direction": Section 2.4.2 of this chapter relies heavily on his analysis of TQ lexicon. Likewise, Blust's (2002) discovery of intervocalic devoicing as a sound change helped me identify the phonotactic restriction against intervocalic stops in Berawan: his description and that in Burkhardt (2014) provided grounds for my analysis in Section 2.4.3. However, a unified treatment of unnatural gradience — establishing its existence, statistical significance and phonetic reality, its origins, and discussing theoretical implication that it brings — is lacking and the present section aims to fill this gap.

As already mentioned, most non-natural phonotactic restrictions discussed thus far (including the ten phonotactic constraints in Hayes and White 2013) are unmotivated rather than unnatural (according to my definition in (2)), and we have almost no information on whether gradient phonotactic restrictions can be unnatural. Some gradient restrictions that operate in the unnatural direction have recently been reported in light of the surfeit of the stimulus experiments: a higher rate of alternation in monosyllabic compared to polysyllabic words (Becker et al. 2012) or a stronger preference for antepenultimate stress in LLL words compared to HLL words (i.e. a tendency that operates against the Weight-to-Stress principle; Garcia 2017). Strictly speaking, the first restriction does not contradict any universal phonetic tendency (perhaps just a universal phonological tendency that favors faithfulness in monosyllabic words). The second restriction targets non-segmental features and while the restriction is significant, it is also very subtle in magnitude (see Garcia 2017). Experiments that tested productivity of these two restrictions failed to show positive results (Becker et al. 2012, Garcia 2017). This section presents a novel analysis of the lexicon of two languages, the Tarma dialect of Quechua (data from Adelaar 1977 and Puente Baldoceda 1977, Nazarov 2008) and the Berawan dialects (data from Blust 2002, 2005, 2013, and Burkhardt 2014), with highly unnatural trends in the lexicon that target a single segmental laryngeal feature [±voice] and have a considerably greater magnitude than any cases reported thus far. I will show that these trends run counter to specific universal phonetic tendencies, and argue that the trends are statistically significant, phonetically real, and morphophonologically productive. Based on these findings, I interpret unnatural trends as a gradient phonotactic restriction (in line with Coetzee and Pater 2008 and others).

The two cases presented in this chapter open up a theoretical question for the theory of Markedness (see, e.g., de Lacy 2002, 2004, 2006, de Lacy and Kingston 2013) in weighted-constraint theories like Harmonic Grammar (HG, Legendre et al. 2006, Coetzee and Pater 2008, 2011, Pater 2008, 2009, Albright 2009, Potts et al. 2010). While HG is able to derive gradient phonotactics (see, e.g., Coetzee and Pater 2008), I show in Chapter 5 that, without unnatural Markedness constraints, HG in its current form is unable to derive systems in which, given a context, the unnatural element in that context is more frequent than the natural element. Since the cases presented here are of this exact nature, this opens up the possibility that Markedness constraints might, after all, be able to counter universal phonetic tendencies (cf. Hayes and White 2013 and other work in that direction).

2.4.2 Tarma Quechua

Tarma Quechua is a dialect of Quechua (Quechua I) spoken in Tarma district in the Junín province of Peru. The number of speakers is difficult to establish as the dialect is rapidly being replaced by Spanish (Adelaar 1977). Adelaar (1977) and Puente Baldoceda (1977) report approximately 30,000–40,000 inhabitants of the Tarma district and additional 3,500 inhabitants of La Unión Leticia. However, the number of speakers of the dialect with the particular unnatural phonotactics that are of interest here is much smaller and difficult to estimate.

2.4.2.1 Stop voicing

The Quechua dialect continuum almost uniformly has only voiceless stops in native vocabulary (Adelaar and Muysken 2004). However, Adelaar (1977) and Puente Baldoceda (1977) report that some of these voiceless stops have become voiced in Tarma Quechua (henceforth: TQ), or more precisely, the Quechua dialects spoken in Tarma, Huaricolca, Palcamayo, and La Unión Leticia. Voiceless velar and labial stops (to the exclusion of alveolars) are reported to undergo voicing in intervocalic and post-consonantal positions, but not after a nasal consonant. Adelaar (1977) and Puente Baldoceda (1977) note that voicing does not apply categorically (and that it does not apply post-nasally), but no further analyses on the lexicon are performed. While Adelaar (1977) and Puente Baldoceda (1977) established that voicing differs across phonetic contexts, Nazarov (2008) in his detailed study of TQ native vocabulary established that rates of application differ across phonetic contexts, identified the rates of application and contexts, and showed that the voicing operates in "phonetically unmotivated" direction. Nazarov (2008) discovered that voicing applies almost categorically in consonant clusters, in nearly half of lexical items intervocalically and almost never post-nasally. Note that this variation is confined to the lexicon: lexical items either feature voicing or not. Based on these and my own further analyses, I will show that voicing in Tarma Quechua operates in unnatural direction as defined above: against universal phonetic tendencies and that the unnatural trend is statistically significant.

Nazarov (2008) does not provide any statistical treatments of his observations and does not include loanwords in the analysis. For the purpose of statistical testing the unnatural trends in the lexicon, I re-collected all tokens of a labial or a velar stop from the vocabulary list in Adelaar (1977) following Nazarov's (2008) initial count.

For the purpose of the analysis, I collected all tokens of stops from the vocabulary list in Adelaar (1977). Because alveolars never undergo voicing, they were omitted from the analysis — only labials and velars were kept. In addition, word-final stops and first members of consonant clusters always surface as voiceless, which is why they were also excluded from the analysis. There are only minimal deviations from my counts and counts in Nazarov (2008). A total of 1199 tokens were collected: 910 tokens were from the native TQ vocabulary, 289 are labeled as loans from Spanish in Adelaar (1977). Each data point was annotated for presence or absence of voicing, place of articulation of the stop (labial or velar), and position in the word. Five positions were included in the analysis: word-initial, post-nasal, intervocalic, position after a sonorant, and position after an obstruent. The initial raw data analysis reveals a surprising trend: voicing surfaces almost never post-nasally (9.5%), in almost half of the lexicon intervocalically (42.5%), and almost always post-consonantally, including in positions after a voiceless obstruent (86.1%) (also Nazarov 2008).

Table 2.17: Voiced vs. voiceless labial and velar stops in Tarma Quechua native vocabulary across contexts based on a vocabulary list in Adelaar (1977) (a similar count in Nazarov 2008).

	#__	N__	V__V	R__	T__
voiced	7	7	99	72	68
voiceless	276	67	134	13	11
% voiced	2.5	9.5	42.5	84.7	86.1

To test the statistical significance of this trend, the data were fit to a logistic regression model in R statistical software (R Core Team 2016) using the *glm()* function. The first model includes only native vocabulary. The dependent variable was binary: presence or absence of voicing; the independent variables were Place of articulation (treatment-coded with two levels, labial and velar with labial as the reference level), and Position in the word (treatment-coded with five levels: initial, post-nasal, intervocalic, post-sonorant, post-obstruent and intervocalic as the reference level) with no interactions. The best fitting model was chosen with the step-wise backward model selection technique: higher order interactions were removed step-wise from a full model. If the likelihood ratio tests (LRT) determined an interaction or predictor does not improve fit significantly, they were removed until all predictors in the model significantly improve the fit.

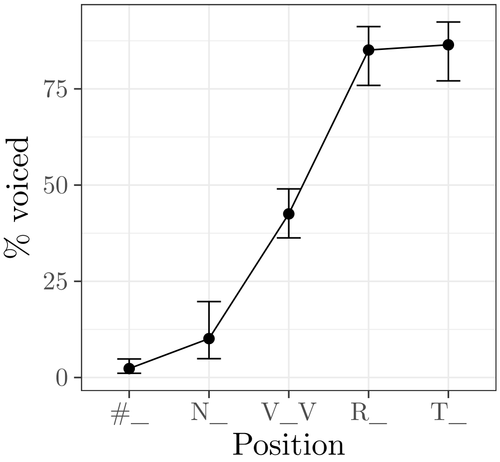


Figure 2.1: Percentage of voiced stops according to position.

Table 2.18: Logistic regression model.

	Est.	SE	z value	Pr(>|z|)
(Intercept)	-0.045	0.172	-0.260	0.7952
V__V vs. R__	2.044	0.332	6.164	0.0000
V__V vs. T__	2.155	0.353	6.101	0.0000
V__V vs. N__	-1.884	0.421	-4.478	0.0000
V__V vs. #__	-3.437	0.407	-8.437	0.0000
velar vs. labial	-0.502	0.214	-2.344	0.0191

As shown in this analysis, [+voice] in labial and velar stops is significantly less frequent word initially and post-nasally compared to intervocalic position. [+voice] is significantly more frequent in post-sonorant and post-obstruent position compared to intervocalic position in TQ native vocabulary.

As argued in Section 2.2, post-nasal and intervocalic position universally prefer voicing, while voiced stops after voiceless obstruents are universally dispreferred. The fact that TQ features less voicing post-nasally and intervocalically than after voiceless obstruents (the first members of a consonant cluster never underwent voicing)[26] is thus highly unnatural. TQ voicing thus operates in a direction opposite to two universal phonetic tendencies: it operates more frequently where it is dispreferred (post-consonantally) and less frequently where it is preferred (post-nasally and intervocalically). These findings are summarized in Table 2.19.

Table 2.19: Unnatural distribution of [+voice].

Universal tendencies for [+voice]	Observed significant trends in TQ
T__ < V__V	V__V < T__
T__ < N__	N__ < V__V < T__

These trends are significant even if we include loanwords in the analysis. Data with loanwords was fit to a model that initially had two independent variables: Position (treatment-coded with same levels as above) and Place of articulation (sum-coded with velar as the reference level). The significance of all main effects remain the same as in the native vocabulary, but now the Position × Place interaction becomes significant.

Furthermore, if we add Loanword status as a predictor to the model, the fit improves significantly, but it is not clear whether loanword status as a predictor is justifiable in a cognitive model: it is likely that speakers are unaware of loanword status for many lexical items; if they are, the distribution of voice for loanwords ceases to be of interest (in this case speakers may

---

[26] The analysis shows post-nasal < intervocalic and intervocalic < post-obstruent, from which post-nasal < post-obstruent can be derived by transitivity.

use two different grammars that govern native and loanword phonologies; see, e.g., Itô and Mester 2002, 2003). The data were fit to a model with three independent variables: Position (coded as above), Place of articulation (coded as above), and Loanword status (sum-coded with native words as the reference level). All two-way interactions are significant and are included in the model. The significance of all main effects remains the same as before (at means of other predictors), except that the V__V vs. T__ ceases to be significant as a main effect: there is, however, a significant interaction Loanword × Position — voicing is more frequent for native vocabulary items in post-obstruent position. Adding loanword status as a predictor also introduces a problem of data scarcity: there are only seven loanwords with labial or velar in post-obstruent position. Other frequency differences remain significant: there is still more voicing intervocalically than post-nasally and, if we refit a model with post-nasal position as the reference level, voicing is more frequent in post-obstruent position than post-nasally ($\beta = 1.5$, z $= 0.7$, $p < 0.05$). Note again that V__V vs. T__ is also significant if loanword status is not in the model as a predictor ($\beta = 1.9$, z $= 5.6$, $p < 0.0001$).

If we isolate loanwords from the native vocabulary, we do not observe any unnatural patterns, which is not surprising as the donor language, Spanish, does not have any of the unnatural patterns in TQ. As will be shown below, however, the unnatural voicing pattern in TQ does apply to a subset of loanwords.

There exists another locus of unnatural gradient phonotactics in Tarma Quechua, which was not analyzed specifically either in Adelaar (1977) or in Nazarov (2008): a restriction against clusters that agree in voicing. If we look into the within-context distribution of voicing in the post-obstruent position, clusters that agree in voicing are gradiently dispreferred in TQ — clusters that disagree in voicing are significantly more frequent.

We saw that labial and velar stops surface as voiced in non-nasal post-consonantal position (Table 2.17). The following consonants are attested as triggering voicing: [t, ʧ, ʈ͡ʂ, k, s, ʃ, x, l, lʲ, r, j, w] (Nazarov 2008). Note that the list includes voiceless fricatives, affricates and even voiceless stops. The following clusters of two stops are attested: [kb, tb, tg]. The table 2.20 below presents examples of clusters that disagree in voicing after each consonant (data from Nazarov 2008, Adelaar 1977).

Table 2.20: Clusters in TQ (from Nazarov 2008, Adelaar 1977).

1st member	**2nd member**	
	Labial	**Velar**
t	lutbi	mutgi
ʧ	/	aʧga
ʈ͡ʂ	aʈ͡ʂba	maʈ͡ʂga
k	takba	/
s	ʧasbu	ʈ͡ʂasgi
ʃ	kaʃbi	iʃgi
x	saxbi	manexax-gunas
l	ʧilbi	ʧilgi
r	karba	argu
j	ajba	ajga
w	kawbu	awgis

Obstruent clusters that disagree in voicing are much more frequent than clusters that agree in voicing if the second consonant is either a labial or a velar. Table 2.21 shows the number of occurrences of obstruent clusters in which the second element is a labial or a velar. To test the statistical significance of this distribution, the data was fit to a logistic regression model with only voicing as the dependent variable (empty model). The main effect of Place of articulation was not significant. Second-element stops (labial and velar) are significantly more frequently voiced (as opposed to voiceless) in clusters with a voiceless first element in TQ native vocabulary ($\beta = 1.8$, z = 5.6, $p < 0.0001$). This significance remains if we add loanwords into the counts: the best fitting model includes the intercept and a main effect of loanword status (if it is justified cognitively). Voiced stops are more frequent in the second position of obstruent clusters compared to voiceless stops ($\beta = 1.4$ z $= 5.2$, $p < 0.0001$).[27]

Table 2.21: Voice feature in obstruent clusters.

	TT	**TD**	**DT**	**DD**	**Total**
Count	11	68	0	0	79
Percent	13.9%	86.1%	0%	0%	100%

TQ thus features a statistically significant trend that restricts clusters agreeing in voicing in favor of disagreeing clusters. This trend is both gradient and unnatural.

The trend against agreeing obstruent clusters in TQ is unnatural in one additional respect. Table 2.21 shows a preference for TD clusters, compared to DT clusters — which goes against yet another phonetic tendency. Voicing is articulatorily easier to maintain in initial parts of closure than it is to onset voicing after a period of voiceless closure (Ohala and Riordan 1979, Ohala 1997a). The reason for this articulatory dispreference is straightforward and has been identified as Aerodynamic Voicing Constraint: airflow and subglottal-supraglottal pressure difference, necessary for voicing, are sufficient during vowel articulation, but decrease into closure. The reason why voicing is articulatorily difficult to onset after a period of voiceless closure is that it is difficult to reinstantiate airflow and pressure difference — once the closure has caused them to decrease — without releasing the stop closure completely. In addition, there is a typological tendency towards respecting the Syllable Contact Law (Vennemann 1988), which also prefers DT over TD clusters. Finally, decreasing phonation into closure is observed as a passive tendency in several languages (see, for instance, Möbius 2004, Davidson 2016). In other words, voicing has a universal tendency to decrease rather than increase during the closure. The restriction

---

[27] This difference again ceases to be significant if we add loanword status as predictor (sum-coded), but that might be due to the very small number of loanwords with clusters. Also, see above for problems with adding loanword status as predictor.

against DT (decreasing in voicing) clusters in favor of TD clusters (increasing in voicing) in TQ is thus unnatural: it operates against the universal phonetic tendency that decreases voicing into closure.

2.4.2.2 Phonetics

The phonological facts described above clearly indicate unnatural tendencies in the lexicon. However, it is not *a priori* obvious that the phonological transcription used for these facts was faithful to the acoustics. In the following, I present results of my phonetic analysis of Tarma Quechua (of the recordings made by Willem Adelaar). No previous detailed phonetic analyses of the system of voicing in TQ exist: Adelaar (1977), Puente Baldoceda (1977), and Nazarov (2008) are based on qualitative descriptions of recordings and are not supported by phonetic analyses. My analysis confirms the phonetic reality of the TQ voice system as described above, making the case for true unnaturalness in the TQ data.

The analyzed recordings were obtained online[28] in .wav format, sampled at 90 kHz[29] with 16-bit quantization and analyzed with Willem Adelaar's permission in Praat software (Boersma and Weenink 2016). The recordings were made by Willem Adelaar in 1970 in Tarma, in Junín province of Peru. The informant was a 35 year old male speaker of TQ. The recordings are noisy with considerable echo, but the analysis nevertheless reveals important aspects of the unnatural gradient phonotactics and of the phonetic system of TQ in general.

Figure 2.2 shows four waveforms and spectrograms of two TD clusters: [tb] and [kb]. All four spectrograms clearly show that the initial stop of the cluster is voiceless with almost no

---

[28] Accesible online at: https://corpus1.mpi.nl/ds/asv/?0&0%5C&openpath=node:1483874

[29] The original sampling frequency is not known.

phonation into closure and that phonation does not start until the onset of the second stop's closure. First-element stops in the clusters show some echo noise during closure, because the recordings were made in a non-isolated room, but the voicing bar of the second stop is clearly distinguishable from noise vibrations of the first stop in all four cases.

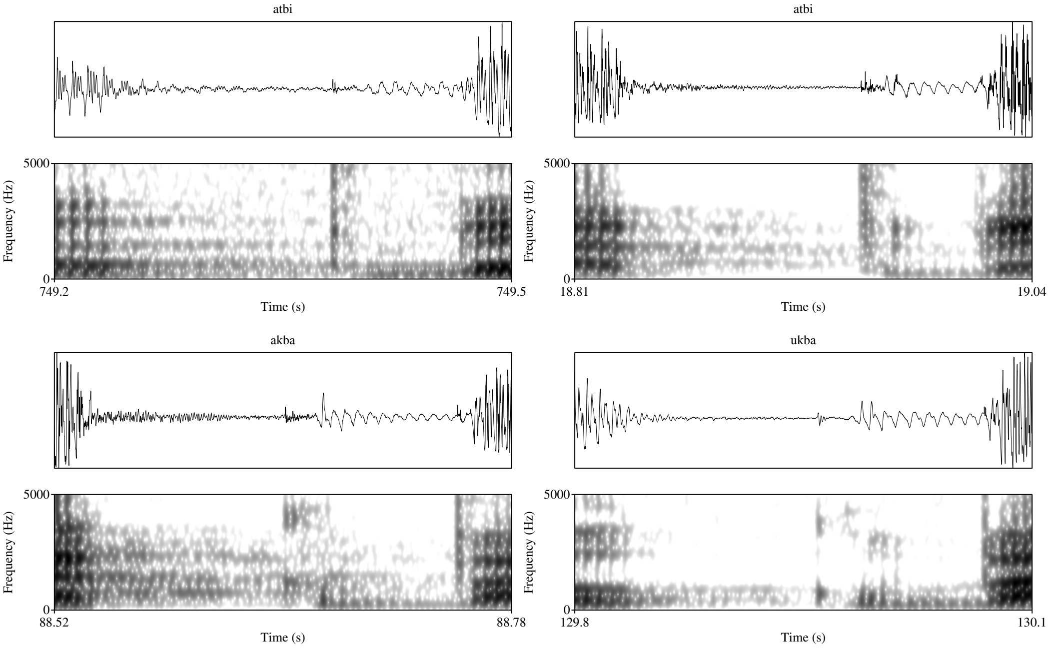


Figure 2.2: Waveforms and spectrograms of four TD clusters: [atbi], [atbi], [akba], and [ukba].

The same situation holds when the first element is an obstruent other than a stop, such as in the [sb] or [sg] sequences in Figure 2.3. The lack of any low frequency energy in the fricative portion of the cluster confirms that the first element, [s], is voiceless. Phonation starts at the onset of the stop.

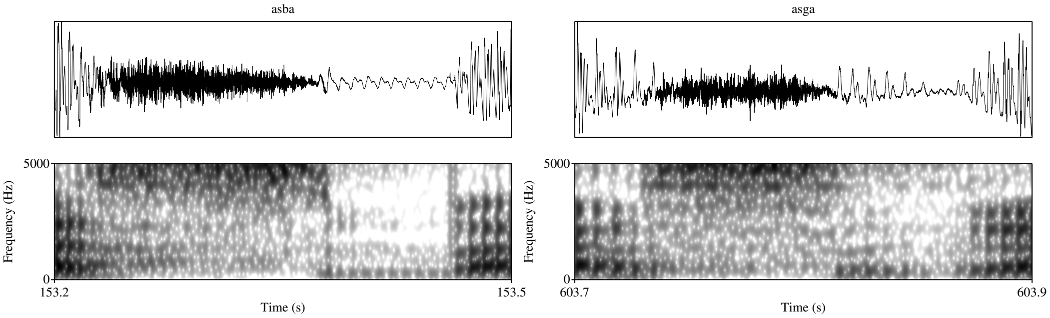


Figure 2.3: Waveforms and spectrograms of two TD clusters: [sb] and [sg].

The exact realization of voiced stops in clusters is not completely uniform and may vary. The exact distribution is difficult to establish with limited data, but a short transitional vocalic element is occasionally found between the voiceless and voiced obstruent, indicating a smaller degree of gestural overlap (Figure 2.4).[30] Occasionally, the voiced element surfaces as a fricative. For details on fricative realization, see discussion in Section 3.5.2.3.

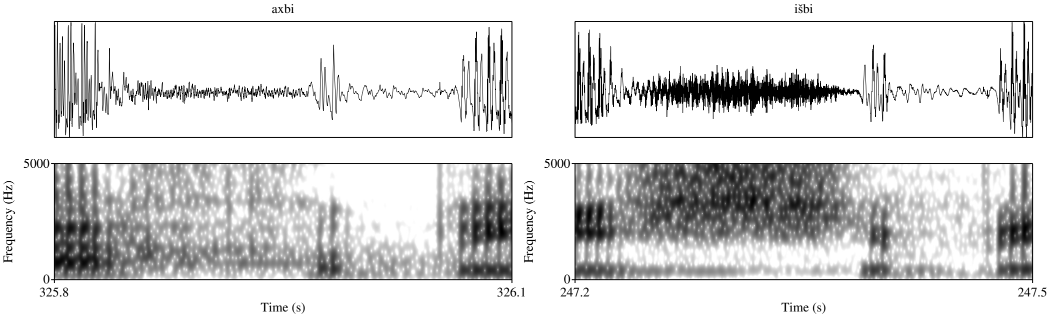


Figure 2.4: Waveforms and spectrograms of voiced labial stops in post-consonantal position with a short vocalic element between the voiceless and voiced element: [xb] (left) and [ʃb] (right).

After nasals, on the other hand, voiceless stops are the preferred variant, as detailed in Table 2.18. Figure 2.5 shows spectrograms with voiceless [k] and [p] after nasals. Also note that voiceless stops in TQ are unaspirated, which means that phonotactic restriction in fact targets the feature [±voice] rather than the feature [±spread glottis].

---

[30] Occasionally, the second element is found to surface as voiceless or deleted.

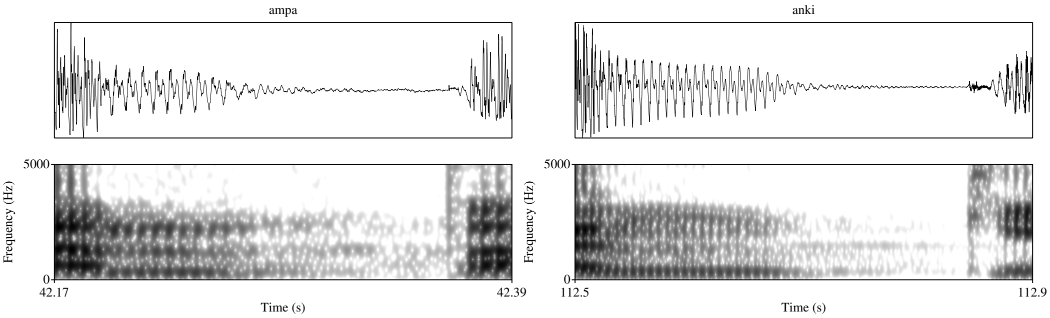


Figure 2.5: Waveforms and spectrograms of voiceless stops in post-nasal position: [mp] and [ŋk] (right).

2.4.2.3 Productivity

The unnaturalness of the gradient phonotactic restriction is phonetically confirmed by the recordings. Corroborating its status as a phonotactic restriction, there exists evidence that it is synchronically active in some morphophonological alternations. Creider (1968) and Adelaar (1977) identify four suffixes with an initial voiced labial stop that feature morphophonemic alternation:

(6) *Alternating suffixes*

a. -ba/-pa ‘genitive’

b. -bax/-pax ‘purposive’

c. -bita/-pita ‘procedentive’

d. -bis/-pis ‘even, too’

The allomorph with voiced initial stops is selected after vowels and non-nasal consonants, including voiceless obstruents; the allomorph with voiceless initial stop is selected after nasals (Creider 1968). The distribution is illustrated in (7).

(7) a. *Intervocalic*

wawxi-gi-**ba** wayi-n

'the house of your brother'

b. *Post-nasal*

wayi-n-**pa** pasa-un

'we're going to walk by way of his house'

c. *Post-obstruent*

tamya-ya-n nuqa-ntik-**baq**

'it is raining now for us' (Creider 1968:12-13)

This process is productive for a subset of suffixes. Other suffixes do not enter the alternation. For example, the highly frequent plural suffix /-guna/ and other suffixes /-bura/, /-gama/, and /-gasga/ have no voiceless allomorphs in post-nasal position (Adelaar 1977:59, also Nazarov 2008). The productivity of this morphophonemic alternation also differs across the dialects. Adelaar (1977:59) reports that voiceless allomorphs are required in Vicora Congas, whereas in Huanuquillo the rate of application varies, i.e. is gradient.

Nevertheless, even if this alternation is morphologically governed, the constraints that motivates the alternation (no voiced stops after a nasal, cf. Coetzee and Pretorius (2010) or no voiceless stop after a voiceless obstruent) are a part of the unnatural phonotactic restriction on the lexicon.

In addition, the behavior of loanwords provides further evidence for the productivity of unnatural gradient phonotactics. Most Spanish loanwords retain their original voicing. Sporadically, however, voicing or devoicing does occur (data from Adelaar 1977). Nazarov (2008) lists four loanwords that undergo voicing. The example in (8-b) is particularly relevant because voicing in this loanword operates in the unnatural direction.

(8) a. Sp. *cuculi* > kuguliː ‘white-winged dove’

b. Sp. *cotpe* > kutbi ‘an animal from the mountains’

c. Sp. *sauco* > sawgu ‘magic tree’

d. Sp. *vaca* > waːga ‘cow’

Nazarov (2008) identifies two loanwords, in which a Spanish voiced intervocalic stop devoices to a TQ voiceless stop (data from Adelaar 1977, Nazarov 2008).[31]

(9) a. Sp. *taruga* > taruka ‘deer’

b. Sp. *dios se lo pague* > jusulpaːki ‘thank you’

The two loanwords with devoicing of intervocalic stop are especially relevant for the discussion on productivity of TQ unnatural gradient phonotactic restriction. Devoicing in TQ [taruka] from Sp. *taruga* cannot be a result of an earlier borrowing, when TQ voicing supposedly did not operate yet: in TQ devoicing never occurs. The historical development of TQ involves only voicing of voiceless stops, not devoicing of voiced stops. Regardless of when Spanish *taruga* was borrowed, sound change could not have produced TQ [taruka]. This means that the gradient phonotactic restriction was likely productive and resulted from the law of frequency effect: because voiced stops surfaced in approximately half of the lexicon, nativization that matches native

---

[31] In other loanwords, Spanish intervocalic voiced stops remain voiced, e.g. TQ [awga] for *ahogar* (Adelaar 1977).

vocabulary frequencies is predicted to occasionally voice voiceless stops of the donor language and devoice voiced ones.[32]

The unnatural phonotactic restriction presented in this paragraph is reported not only for TQ, but also for one other Quechuan dialect. A very similar voicing process whereby Proto-Quechua *p and *k voice in the same positions as in TQ is reported in the dialect of Paccho (Adelaar and Muysken 2004, Adelaar, p.c.). The two dialects, Tarma and Paccho Quechua, are spoken in regions quite distant from each other and are potentially unrelated. Adelaar (p.c.) mentions that the two dialects might have been in contact historically, but the details are unclear. Because there are no descriptions or recordings of Paccho Quechua available, I leave this dialect out of discussion.

2.4.3 Berawan dialects

The Berawan dialects are a group of closely related dialects that belong to the Berawan-Lower Baram group of North Sarawakan languages of the Malayo-Polynesian (Austronesian) language family (Blust 1992). Blust (1992) identifies four dialects of the Berawan dialect group: Long Terawan (LTn), Batu Belah (BB), Long Teru (LTu), and Long Jegan (LJ). They are spoken by approximately 1,000–3,000 speakers in total around the Tutoh and Tinjar tributaries of the Baram river (Lewis et al. 2015, Blust 1992, Burkhardt 2014).

---

[32] Assuming of course that these loanwords were not borrowed to TQ via some other Quechuan dialect without the peculiar voicing process after the voicing was completed in TQ.

2.4.3.1 Restriction against intervocalic voiced stops

According to the description in Burkhardt (2014), the Berawan dialects have two series of stops with respect to laryngeal features: voiced and voiceless, both series being unaspirated.[33] Blust (2013, 2017) and Burkhardt (2014) report that an unnatural sound change, intervocalic devoicing (IVD), operates in Berawan: Pre-Berawan labial and velar stops devoice in Berawan and this devoicing is limited to intervocalic position. Alveolar stops do not devoice, but undergo lenition to [r] in intervocalic position (Burkhardt 2014:249). I will argue that these sound changes resulted in a synchronic pattern that disfavors intervocalic labial and velar voiced stops.

Neither Blust (2013) nor Burkhardt (2014) present any synchronic analysis of Berawan intervocalic devoicing. For the purpose of establishing the existence of an unnatural trend in the lexicon and its statistical significance, I analyzed all native vocabulary items from the vocabulary list in Burkhardt (2014). The list includes 425-466 (mostly morphologically simple) vocabulary items, depending on the dialect. I counted occurrences of voiced and voiceless stops for all three places of articulation in all four dialects according to their position: initially and intervocalically. Clusters are disallowed in Berawan, which means that stops only surface initially, intervocalically, and word-finally. Word-final position is omitted from the count because stops are always voiceless word-finally in accordance with the natural process of final devoicing.[34] I included alveolar stops in the count because, due to intervocalic lenition (*d > r / V__V), they also surface less frequently intervocalically. Geminates are not included in the count. Counts are presented in Table 2.22.

---

[33] The analysis in Burkhardt (2014) is based on recordings made using Sony Minidisc recorder and further analyzed with Toolbox software (Burkhardt 2014:36-8).

[34] For development of voiced stops word-finally, see Section 5.3.1 below.

The raw data analysis reveals that voiced stops are almost categorically prohibited from intervocalic position where their occurrence ranges from 0 to maximally 4. Initially, on the other hand, voiced stops are allowed and surface with slightly lower, but similar frequencies as voiceless stops. Figure 2.6 summarizes the distribution (see also Tables 3.17 and 3.18 for examples of Berawan voiced and voiceless stops).

Stress is on the final syllables in Berawan (Burkhardt 2014). Most words in the vocabulary list in Burkhardt (2014) are disyllabic words with the canonical structure CV.C(C)ˈV(C) and with possible trisyllabic words of the structure CV.CV.C(C)ˈV(C) (Burkhardt 2014:77). This means devoicing could be analyzed as fortition in a prominent position — in the onset of a stressed syllable. Such an analysis, however, fails to account for the data in mono and polysyllabic words. One of the few monosyllabic words in the system with word-initial velar stops surface with a voiced stop: BB [gəm] (< Proto-Malayo-Polynesian *gəm) ‘fist’. Evidence against such a hypothesis can also be found in polysyllabic words. In a disyllabic loanword for ‘tobacco’, devoicing of *g does not apply, even though it appears in the onset of a stressed vowel (BB [siˈɡup] < Brunei Malay sigup). In a trisyllabic word for ‘ashtray’ from the same root, devoicing does occur (BB [səkupˈen] < Brunei Malay pəsigupan; data from Burkhardt 2014). The devoicing in these two cases operates in the exact opposite direction from what would be expected if we assumed onsets of stressed syllables get devoiced. Devoicing of original voiced stops in onsets of unstressed syllables operates in other trisyllabic words, although examples are sparse (the majority of words in Berawan are of the structure CV.C(C)ˈV(C)). Three such examples are BB [səki.oŋ] (< *səɡaquŋ < *səʀaquŋ), BB [kəjjin] (< *dugijjan), and BB [kuβiŋ] (< *bəgβiəŋ < *bəguβaŋ; data from Burkhardt 2014). While it is true that in the only two words with voiced intervocalic labial stops ([təbərroʔ] ‘type of spear’ and BB [təbuttoŋ] ‘festive hat'; Burkhardt 2014), this stop surfaces in an onset of unstressed syllable, the two examples have unclear etymologies and the voiced stops are most likely of secondary origin in these two lexical items.[35]

Table 2.22: Occurrences of stops in Berawan.

		Voiceless		**Voiced**		**% Voiced**	
Dialect	**Place**	#__	V__V	#__	V__V	#__	V__V
Batu Belah	labial	52	10	40	2	43.5	16.7
	alveolar	56	32	22	4	28.2	11.1
	velar	43	54	13	0	23.2	0.0
Long Teru	labial	46	13	38	2	45.2	13.3
	alveolar	54	31	22	2	28.9	6.1
	velar	40	55	11	1	21.6	1.8
Long Jegan	labial	49	10	40	2	44.9	16.7
	alveolar	55	32	22	3	28.6	8.6
	velar	44	58	10	0	18.5	0.0
Long Terawan	labial	41	11	48	1	53.9	8.3
	alveolar	60	25	21	3	25.9	10.7
	velar	50	19	14	0	21.9	0.0

---

[35] If we assumed that devoicing in Berawan was limited to onset positions of stressed syllables, we would still expect *b to yield [g], not [b]. Intervocalic voiced labial stops (*b) in Berawan not only devoice, but also change their place of articulation to [k]. This change of place of articulation needs to precede devoicing, otherwise we would expect original voiceless labial stops (*p) to change to [k] as well, which does not happen (see Section 3.5.2.1.2). Voiced *b in onset of an unstressed syllable that would resist devoicing under the hypothesis that devoicing targets onsets of stressed syllable should thus develop to [g]. Surface [b]s in [təbərroʔ] 'type of spear' and BB [təbuttoŋ] 'festive hat' with unclear etymologies are thus likely secondary. The only scenario under which *b would not be predicted to surface as [g] is if we assumed the change of place of articulation too is limited to onsets of unstressed syllables. Such a condition is, to my knowledge, difficult to motivate: cues for place of articulation should be more salient in the prominent position (see also Section 3.5.2.1.2).

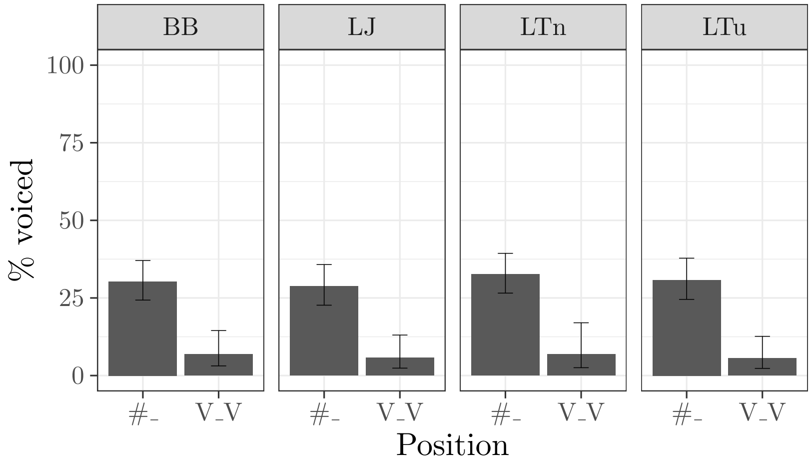


Figure 2.6: Percentage of voiced stops (across places of articulation) according to position — initial vs. intervocalic (from the logistic regression model in Table 2.23).

To test the statistical significance of the restriction against intervocalic voiced stops, the data for each language was fit to a logistic regression model. Presence or absence of the voice feature was the dependent variable and Position (treatment-coded with initial position as the reference) and Place of articulation (sum-coded with velar as the reference) were independent variables. Because of zeros in the count, the full model with all interactions was fit to a logistic regression model using bias-reduction (the model was fit using the *brglm()* function from *brglm* package; Kosmidis 2013). The interaction Place × Position was not significant for any of the four dialects (tested with LRT), which is why the data was refit without this interaction to a logistic regression model without bias reduction (using *glm()*). The best-fitting model was chosen with LRT: for all four dialects it includes main effects Position and Place.

In all four languages the voice feature in stops is significantly less frequent intervocalically compared to initial position. Table 2.23 includes estimates for the main effect Position for all four dialects.[36]

---

[36] In all four dialects the voice feature is also significantly more frequent in alveolars compared to the mean of all stops, but this is not of interest in this dissertation.

Table 2.23: Estimates for the main effect Position (initial vs. intervocalic) from logistic regression models fit to Berawan data.

	Est.	**z score**	**Pr(>|z|)**
BB	-1.773	-3.906	0.0001
LTu	-2.028	-4.139	0.0000
LJ	-1.898	-3.860	0.0001
LTn	-1.893	-3.479	0.0005

Based on these models, I can conclude that the Berawan dialects have a significant trend in the lexicon that restricts voiced stops intervocalically. The statistically significant restriction against intervocalic stops in Berawan is unnatural according to the definition in (2): it operates against the universal phonetic tendency of intervocalic voicing (see Section 2.2).

The distribution of voicing in loanwords suggests that the gradient phonotactic restriction against intervocalic voiced stops was a part of productive alternations at least at some stage of development: devoicing in loanwords operates sporadically, i.e. gradiently. The collection of loanword vocabulary in Burkhardt (2014) includes cases in which devoicing applies regularly, e.g. Brunei Malay [pəsigupan] > Pre-Berawan *səgupan > BB [səkupen] as well as words in which no devoicing applies, e.g. Brunei Malay [sigup] > Pre-Berawan *sigup > BB [sigup].

This non-categorical devoicing in loanwords could also result from different lexical items being borrowed at different stages in the development, i.e. before or after the "intervocalic devoicing" sound change operated.[37] One piece of evidence against this latter scenario is the fact that Batu Belah [səkupen] 'pipe' and [sigup] 'tobacco' both go back to the same Brunei Malay root, yet one undergoes devoicing and the other does not. It is difficult to argue two lexical items of the same root were borrowed at different times, although of course not impossible. In a list of 15 loanwords in Burkhardt (2014), there are six cases in which a voiced velar or labial stop surfaces

---

[37] See section 3.5 for arguments against intervocalic devoicing being a single sound change.

in intervocalic position in the donor language. In one case, devoicing occurs; in the remaining five cases stops remain voiced.

The dispreference for intervocalic voicing remains significant even if we add loanwords to the count. In all four dialects, voicing is significantly less frequent intervocalically compared to initial position when loanwords are added to the count. Thus, I have demonstrated that the Berawan dialects exhibit a statistically significant trend in the lexicon — a dispreference for voiceless stops between vowels, which I interpret as a gradient phonotactic restriction as per the logic in Coetzee and Pater (2008). This gradient phonotactic restriction opposes a universal phonetic tendency: intervocalic voicing (see Section 2.2), and is thus unnatural by my definition in Section (2). This means that Berawan exhibits yet another case of gradient unnatural phonotactics (see Section 5.1 for a discussion of unnatural gradient phonotactic restrictions and their theoretical implications).

The data presented in this chapter seem to suggest, at first glance, that unnatural processes arise from a single sound change in the development of languages. However I will demonstrate below that a thorough investigation reveals a common pattern of complementary distribution in all cases along with other pieces of evidence, strongly suggesting that a combination of natural sound changes operated in place of a single unnatural sound change.

Page intentionally left blank.

Chapter 3

The Blurring Process

In this chapter, I propose a new model for explaining unnatural processes diachronically, called the Blurring Process. I first discuss all thirteen cases of PND and argue that PND did not operate as a single unnatural sound change, but as a combination of natural sound changes in all thirteen languages in which PND is reported either as a synchronic alternation or as a sound change. This typological study provides a basis for establishing a new diachronic model for explaining unnatural phenomena, the Blurring Process. Using the Blurring Process, I also provide a proof that at least three sound changes are required for an unnatural process to arise: the Minimal Sound Change Requirement.

Just as for post-nasal devoicing, it appears on the surface (and it has been claimed in the literature for Berawan and Kiput, Blust 2005, and TQ, Nazarov 2008) that single instances of unnatural sound changes operate in the development of Kiput, Berawan and Tarma Quechua. The fact that the unnatural phenomena in these languages are gradient is another argument in favor of a single sound change hypothesis: gradience is the prominent property of sound changes in progress. Using the Blurring Process model, however, I argue in the second part of the chapter that the seemingly unnatural sound changes and the resulting phonotactic restrictions in Kiput, Berawan, and Tarma Quechua arise from a combination of three natural sound changes, and I point to advantages that this explanation bears over the alternative single-sound-change approaches.

3.1 Explanations of PND

Several accounts in the literature understand PND as a single sound change; explanations for this process run the gamut from appeals to hypercorrection (Xromov 1972, Blust 2005), to arguments that PND is actually a phonetically plausible or even natural process (Solé et al. 2010, Solé 2012). Three problems arise with such accounts. First, they all struggle to explain why devoicing would operate in post-nasal position, where nasal leakage and volume expansion (Hayes and Stivers 2000, Coetzee and Pretorius 2010) militate against the anti-voicing pressure of closure (Ohala and Riordan 1979), whereas in other contexts, where speakers have to accommodate for voicing considerably more than post-nasally, voicing is preserved (see also the discussion in Section 3.3). Second, they each examine and account for only a single instance of PND, examined in isolation without the relevant cross-linguistic data. Finally, most of the existing proposals fail to explain recurrent patterns that are observed in all thirteen languages (e.g. fricativization of voiced stops elsewhere, see Section 3.2) and why only a subset of places of articulation enter into post-nasal devoicing. In Section 3.2, I argue that all cases of reported PND (that have until now been studied in isolation) in fact result from a combination of sound changes. This explanation was proposed for Tswana already in Dickens (1984) and Hyman (2001).

Xromov (1972) appears to invoke hypercorrection to explain PND in Yaghnobi. He postulates that in Sogdian and Yaghnobi, all stops voice post-nasally, but that this sound change is in progress, which results in variation between [NT]∼[ND] for the underlying /NT/. He further claims that the underlying /ND/ joined this pattern (probably through hypercorrection), which means that at some stage variation between [NT]∼[ND] existed for /ND/ as well. Once the voicing process was finished, the voiceless variant became more frequent for both inputs. No motivation is given for why the voiceless variant becomes prevalent, except for an assumption that devoicing might have been influenced by morphological analogy in dentals and then spread to other places of articulation. Xromov's explanation might be more convincing if we assumed hypercorrection arose from an interaction between two hypothetical dialects, where one voices all stops, and the other preserves the contrast; speakers of the neutralizing dialect could then hypercorrect their voiced stops to voiceless ones in post-nasal position. Not only do we lack evidence for the stage with the two dialects, but this assumption also fails to explain the connection between post-nasal devoicing and the development of voiced stops to voiced fricatives in the elsewhere condition.

Blust (2005, 2013) offers three possible explanations for the emergence of PND in Buginese, Murik, and Land Dayak. First, he notes that, much as PNV can be understood as an assimilation of stops to a voiced environment, PND can be explained as *dissimilation*. This assumption, however, lacks explanatory power: it simply restates that PND is the opposite process from PNV and does not specify whether such dissimilation would be driven by perceptual or other factors, why it fails to target some places of articulation, and why it fails to target other contexts, such as the intervocalic position. Blust (2013: 668) himself notes that "this does little to explain why a change of this type would occur."

Blust's second explanation for Austronesian PND postulates that the three languages in question first underwent PNV: voiceless stops became voiced in post-nasal position, thus eliminating

NT sequences. According to Blust (2013), after PNV, “voice was free to vary” post-nasally, and the “voiceless variant of postnasal obstruents prevailed over time.” In other words, Blust (2005) seems to suggest that the lack of contrast in [±voice] enables devoicing in this position. There are three major issues with this approach. First, it is not parsimonious to assume the independent occurrence of PNV three times without any comparative evidence. Second, it is difficult to explain why a voiceless variant would prevail in an environment in which voicing is preferred compared to other environments (e.g. word-initially), but would fully contrast in those other environments, where voicing is dispreferred. While it is true that the functional load/frequency of phonemes can influence the probability of a merger (Wedel 2012, Wedel et al. 2013, Hay et al. 2015), it is unclear why a merger would first happen to voiced stops and then to voiceless stops. On the other hand, Blust is right in that diminished functional load of post-nasal stops in the languages that do not contrast [±voice] in post-nasal position can influence the operation of devoicing in the sense that the unconditioned devoicing (proposed in Section 3.2) does not get blocked for functional reasons (Wedel et al. 2013). Finally, this line of reasoning, like many others proposed thus far, fails to explain some common patterns, such as fricativization in the elsewhere condition or the absence of devoicing for some places of articulation (Section 2.3.1).

Somewhat related to Blust’s (2005) second proposal is the idea that the loss of contrasts could influence the operation of sound change (cf. Keyser and Stevens 2001, Hyman 2013). PND is, however, attested both in languages in which [±voice] likely does not contrast post-nasally (e.g. in Yaghnobi, Konyagi) as well as in languages in which [±voice] fully contrasts post-nasally at the time of devoicing (e.g. in Tswana, where the nasal is not lost before stops in all environments; see Table 2.5 and Section 2.3.1.2). Similarly, an analysis of devoicing as a loss of contrast in [±voice] in the elsewhere condition is not appealing. Devoicing is attested in languages in which voiced stops fricativize elsewhere (which means that the contrast is that of T ∼ Z, e.g. in the Austronesian languages) and there are no voiceless fricatives (S) in the systems, meaning [voice] could be redundant (and [±continuant] contrastive), as well as in languages in which the contrast in the elsewhere position is between voiceless and voiced fricatives (S ∼ Z, e.g. in Konyagi or Tswana) and stops surface only post-nasally (and contrast in voice), meaning [±voice] cannot be redundant. Finally, the fact that complete devoicing occurs in some subdialects of Tswana without any contrast being lost beforehand and that in Yaghnobi devoicing targets all stops, not only post-nasal ones, speaks against this line of reasoning (see also Sections 3.2.1 and 3.2.2 for further discussion).

A third explanation offered in Blust (2005) invokes dissimilation by hypercorrection. Blust notes that NT sequences in Buginese and Murik develop either to T or TT. This means that, at a certain point, NT sequences were absent from the language and only voiced stops surfaced after nasals (ND). At this point, according to Blust, speakers "may have assumed that prenasalized obstruents had acquired voicing by assimilation" and then "undid" that assumed voicing. This account faces three major difficulties. As already pointed out by Blust (2005), it is unclear what would "prompt speakers to assume that voicing assimilation had taken place in earlier clusters" of ND. Second, even if they had made this assumption, the speakers would still have to apply dissimilation, i.e. devoicing in a context that strongly promotes voicing compared to other positions. Finally, this particular approach with hypercorrection lacks broader explanatory power, since it cannot be extended to cases of apparent PND in other languages, where the sound change NT > TT, T is not attested (e.g. in Tswana).

Some analyses have attempted to account for PND by motivating the process phonetically. Solé et al. (2010) and Solé (2012) specifically identify PND as a "historical process," suggesting that they assume PND operated as a single instance of sound change. Moreover, these authors claim that PND is not necessarily an unnatural process and may in fact have a phonetic explanation. The main evidence for this claim comes from Shekgalagari, which is assumed to feature “early velic rising” in NT sequences. This process is supposed to follow from the fact that (i) speakers do not show any passive voicing in the NT sequences in Shekgalagari, and (ii) underlying nasal-fricative sequences /nz/ yield a nasal affricate [nts]. This process of early velic rising, which is argued to account for both of these observations, would also have caused a “long stop closure” in ND sequences. Because voicing is difficult to maintain, especially during longer closure, the result would be devoicing of the stop (Solé et al. 2010: 612).

This explanation has three major drawbacks. First, secondary ND sequences surface as NN in Tswana and ND in Shekgalagari. The following two examples illustrate this distribution (repeated Table 2.6):

Table 3.1: Secondary ND sequences in Tswana and Shekgalagari (table from Solé et al. 2010).

Ts.&Sh.	/χʊ-m-bɔ́n-á/	→	[χʊmpɔ́ná]
Sh.	/χʊ-mʊ-bɔ́n-á/	→	[χʊmbɔ́ná]
Ts.	/χʊ-mʊ-bɔ́n-á/	→	[χʊmmɔ́ná]

If early velic rising in Shekgalagari were indeed a phonetic process, we should expect to see devoicing in secondary ND sequences as well. The fact that the stops in secondary ND sequences surface as voiced speaks against the proposal in Solé et al. (2010) and Solé (2012). Of course, one could assume that early velic rising operated prior to the period during which secondary ND sequences arose. However, there is a flaw in this assumption: Solé et al. (2010) provide evidence for early velic rising from synchronic phonetic data. If we postulate that early velic rising is responsible for PND as a synchronic phonetic process, we should expect secondary sequences to undergo devoicing as well. Conversely, if we posit that early velic rising should have been completed by the time secondary ND sequences were introduced, we should not expect to find continuing evidence for this process in the current phonetic data. The only reasoning that could explain why secondary ND sequences do not undergo devoicing under the early velic rising approach would be to assume that devoicing is blocked in order to prevent the merger of two grammatical morphemes, /m-/ and /mʊ-/. Even if this is indeed the case, it still means that early velic rising is not completely automatic in synchronic Shekgalagari, but can be overridden for functional reasons. Also, Solé et al. (2010) note that Shekgalagari has ND sequences in approximately ten items that do not arise from the /mʊ-/ morpheme. While "most" of the ten items are assumed to be borrowings, it is unclear if all items indeed are borrowings (at least one of them is a grammatical marker [ǹde-]). This reinforces the view that early velic rising is not a completely automatic phonetic tendency in Shekgalagari, but can be blocked by at least functional and loanword factors. Second, realization of /nz/ as [nts] is typologically common and proceeds from general phonetic tendencies cross-linguistically (cf. Ohala 1997, Kümmel 2007, Ali et al. 1979, Steriade 1993, Busà 2007; see also Section 3.3). The fact that Shekgalagari /nz/ surfaces as [nts] hardly tells us anything about early velic rising. Third, an explanation along these lines fails to account for other cross-linguistic cases of apparent PND where no traces of early velic rising can be found. If PND were indeed motivated by early velic rising, we should expect to find evidence that early velic rising is similarly responsible for reported cases of PND outside Tswana and Shekgalagari (to my knowledge, there exist no studies of this phenomenon). Finally, the explanation proffered in Solé et al. (2010) fails to explain the connection between PND and the recurrent pattern of fricativization in the elsewhere condition or the pattern of unconditioned devoicing of voiced stop found in subdialects of Tswana (see Section 3.2.2 and Table 3.6).

Finally, Dickens (1984) and Hyman (2001) propose an explanation for PND in Tswana that assumes a set of three non-PND sound changes that conspire to produce apparent PND: fricativization, devoicing of stops, and occlusion of fricatives. I will argue in the remainder of this chapter that this is in fact the correct explanation not only for Tswana and Shekgalagari, but that an essentially similar historical scenario played out in all thirteen cases of reported PND described above. Unfortunately, at the time of Dickens' and Hyman's work, no historical parallels existed in the literature that would support their explanation, which led other authors to propose alternative accounts of the data. Admittedly, in the absence of typological parallels, one might judge an explanation that operates with a single (albeit unnatural) sound change more parsimonious and justified than an explanation that requires three separate sound changes. By bringing numerous cases of PND from disparate language families together and arguing that diachronic developments very similar to that proposed for Tswana in Dickens (1984) and Hyman (2001) operated in all thirteen languages, the next section dispels this concern and validates the three-sound-change analysis on typological grounds. I provide new evidence in favor of the three-sound-change approach and argue that the three sound changes are directly historically attested in Avestan, Sogdian, and Yaghnobi, three languages in ancestral relationship.

3.2 A combination of sound changes

A closer look into the collected data from Section 2.3.1 reveals an important generalization: for all cases of PND, either direct evidence or clear indirect evidence can be found that, at some stage of development, voiced stops surfaced as voiced fricatives except in post-nasal position. In other words, in the first stage of the development of PND, a natural sound change[38] operates

---

[38] see Section 3.3 on a discussion of naturalness of the sound changes involved in my proposal.

that fricativizes voiced stops except post-nasally (D > Z / [−nas]__). This sound change results in a complementary distribution: voiced stops surface post-nasally, voiced fricatives elsewhere. I argue that in all thirteen cases PND is a result of this complementary distribution plus the unconditioned devoicing of voiced stops: because voiced stops surface only post-nasally, the unconditioned devoicing results in apparent PND. While this explanation has been proposed for Tswana in Dickens (1984) and Hyman (2001), this section argues that the same diachronic scenario plays out in all thirteen cases by pointing out to thus far unobserved pieces of evidence in favor of the three-sound-change approach. It is also argued that Yaghnobi offers the most direct evidence in favor of the three-sound-change approach, since all three diachronic stages are historically attested in written sources. Yaghnobi also provides crucial evidence in favor of the assumption that the second sound change, devoicing of voiced stops, is unconditioned.

3.2.1 Yaghnobi

Yaghnobi is a descendant of Sogdian, a Middle Iranian language spoken in the first millennium CE and preserved in documents from that period. In Sogdian, all voiced stops surface as voiced fricatives except post-nasally. This complementary distribution is directly attested in Sogdian and confirmed by the writing system (cf. Sims-Williams 1987:178). It is equally clear that the Sogdian pattern developed through a sound change D > Z / [−nas]__ in an earlier stage of the language. In Avestan, which is closely related to Sogdian (but more archaic) and can be used to represent the parent language of Sogdian, voiced stops correspond to Sogdian voiced fricatives except post-nasally, e.g. Avestan [dasa] vs. Sogdian [ðasa] ‘ten’, Avestan [gari-] vs. Sogdian [ɣarí] ‘mountain’ (Kümmel 2006), Avestan [asənga-] vs. Sogdian [sang] ‘stone’ (Bartholomae 1961:210).[39]

Given the complementary distribution attested in Sogdian, I propose that, on the way to Yaghnobi, an additional sound change operated: unconditioned devoicing of voiced stops (D > T; for phonetic motivation and attestations of this change, see Section 3.3). Evidence for this sound change in Yaghnobi are attested as directly as a sound change can be attested diachronically: Stage 2 (Sogdian) features voiced stops which surface as voiceless at Stage 3 (Yaghnobi). Because voiced stops surface after nasals, this combination of sound changes results in apparent PND. The development is summarized in Table 3.2.

Table 3.2: Development of PND from Avestan to Yaghnobi.

Stage	**Sound change**	**Language**	**Example**
1		Avestan	band
2	D > Z / [−nas]__	Sogdian	βand
3	D > T	Yaghnobi	vant

Sogdian thus provides direct historical evidence showing that the apparent case of PND in Yaghnobi is a side effect of two natural and well-attested sound changes: (i) fricativization of voiced stops except in post-nasal position (D > Z / [−nas]__), and (ii) unconditioned devoicing of voiced stops (D > T).

To get PND, we need a third sound change: occlusion of voiced fricatives to stops (Z > D). Yaghnobi provides additional evidence for this process too: the original voiced labial and velar stops (Avestan [b] and [g] that develop to Sogdian [β] and [ɣ] in the elsewhere position) still surface as voiced fricatives [β] and [ɣ] in the elsewhere position in the modern Yaghnobi (e.g. Yagh. [vant] ‘tie’ from Sogd. [βand]; Yagh. [ɣar] ‘mountain’ from Sogd. [ɣarí]). Nevertheless,

---

[39] Yaghnobi examples are primarily taken from Xromov (1972) and Novák (2010, 2013, 2014).

the voiced alveolar fricative (Sogdian [ð]) in the elsewhere position gets occluded in Yaghnobi and surfaces as a voiced stop [d], thus blurring the original complementary distribution (Xromov 1972: 123). Apparent PND in Yaghnobi thus fully holds only for the dental series of stops, because only this series of stops underwent a sound change that turned the original voiced fricatives "back" to stops (Z > D). Table 3.3 illustrates the three sound changes that operated on the alveolar series of stops to produce PND.

Table 3.3: Development of coronals from Avestan to Yaghnobi.

Stage	**Sound change**	**Language**	**N__**	**Elsewhere**
1		Avestan	band	dasa
2	d > ð/[−nas]__	Sogdian	βand	ðasa
3	d > t	Yaghnobi	vant	*ðasa
4	ð > d	Yaghnobi	vant	das

Yaghnobi and Sogdian contain yet more crucial evidence in favor of the three-sound-change approach outlined above. One of the most serious objections to the three-sound-change approach in Yaghnobi and elsewhere is that devoicing should not be analyzed as unconditioned, precisely because it operates only in post-nasal position. This being the case, devoicing should indeed be considered unnatural according to the definition in (2). However, Sogdian and Yaghnobi provide strong and unique evidence to suggest that devoicing of voiced stops is in fact unconditioned. I show below that devoicing targets all surface voiced stops in Yaghnobi, not just those in post-nasal position.

It is true that voiced stops surface primarily in post-nasal position in Sogdian, but they also surface in one additional, more marginal context: after voiced fricatives. The origin of such a distribution is straightforward: original voiced stops did not fricativize (i) after nasals (as already mentioned above) and (ii) after another fricative in order to avoid clusters of two fricatives. This property of Sogdian and Yaghnobi, which has up to now been a less discussed fact (Novák 2013), essentially means that voiced stops surface as fricatives in Sogdian in all positions except post-nasally (ND) and occasionally after a voiced fricative (ZD). We already saw that ND sequences devoice to NT in Yaghnobi. Crucially, Sogdian sequences of a voiced fricative and a voiced stop (ZD), too, devoice in the development from Sogdian to Yaghnobi — to ST (in central and western dialects),[40] e.g. Sogd. [pəzda], Yagh. [past] or [pajst] 'smoke'; potentially also Sogd. [ðəɣʷda], Yagh. [duxtar], Sogd. [suɣd-], Yagh. [suxta]; Sogd. [əχʃɨβdi], Yagh. [xiʃift][41] (from Novák 2013, 2014 and Xromov 1987). This devoicing is automatically explained by my proposal that assumes voiced stops devoice unconditionally in Yaghnobi. ZD is thus predicted to first develop to ZT and then to ST when the voiced fricative assimilates in voicing to the following voiceless stop. That this scenario indeed took place is suggested by evidence from Western Yaghnobi, in which forms like [avd] and [avt] for Sogd. [əβdá] are attested (data from Novák 2014). While the first form [avd] dialectally lacks devoicing, the second form [avt] shows that only voiced stops devoice: the fricative is not yet assimilated to the following voiceless stop (compare the assimilated form from Eastern Yaghnobi [aft]).

Conversely, my model assumes voiced fricatives and consequently clusters of voiced fricatives do not devoice. This assumption is borne out by the data: Sogdian clusters of two voiced fricatives (that arise from a cluster of a voiced fricative and an approximant [dw]), [ðβ], do not devoice in Yaghnobi:[42] Sogd. [ðβəri] 'door' or [əzβaːk] 'tongue, language' vs. Yagh. [zⁱvoːk] or

---

[40] It appears that devoicing post-nasally is dialectally more regular than devoicing after voiced fricatives (Novák 2013, 2014), but this distribution can also result from borrowings.

[41] While it is possible, although unlikely, that Yagh. [suxta] and [xiʃift] preserve an earlier (pre-Sogdian) stage where the cluster of a fricative and a stop was not yet voiced, this cannot be the case in Yagh. [past] < Sogd. [pəzda] (and potentially also Yagh. [duxtar]), where the cluster is voiced already in Proto-Iranian (see reconstructions in Novák 2013).

[42] Sequences of two fricatives are also realized with an automatic epenthetic vowel

[d^{i}var] (Novák 2013). Table 3.4 illustrates the development of voiced fricative-stop and fricative-fricative sequences.

Table 3.4: Correspondences of ZD and ZZ sequences.

Sogdian	**Yaghnobi**
zd	st
zβ	z^{i}v
ðβ	d^{i}v
ɣd	xt
βd	ft

Sequences of a voiced fricative and a voiced stop ZD that devoice in Yaghnobi to ST strongly support the assumption that voiced stops undergo unconditioned (not only post-nasal) devoicing in Yaghnobi: all voiced stops that surface in the Sogdian and pre-Yaghnobi system get devoiced in the development of Yaghnobi. The development of ZD clusters in Yaghnobi is summarized in Table 3.5.

Table 3.5: Development ZD sequences from Sogdian to Yaghnobi.

Stage	**Sound change**	**Language**	**Sequence**	**Example**
1		Sogdian	ZD	zd
2	D > T		ZT	*zt
3	assimilation	Yaghnobi	ST	st

In the other twelve languages with PND, non-nasal clusters are generally not allowed (or they became simplified before the emergence of PND), so we do not see devoicing anywhere other than in post-nasal position. Devoicing in clusters in Yaghnobi thus offers a crucial piece of evidence in favor of the proposal that the devoicing of voiced stops that occurs in the development of PND is *unconditioned*; if it were not, we would not be able to unify our account of the devoicing of ND and ZD.

3.2.2 Tswana, Shekgalagari, and Makhuwa

If Yaghnobi provides crucial evidence for the three-sound-change approach on historical grounds, Tswana offers crucial *dialectal* evidence for this hypothesis. There are at least three different systems of stops in the micro-dialects of Tswana. Among one set of speakers, voiced stops get devoiced in all environments: no voiced stops are allowed in the system. Speakers of this system have been labeled "devoicers" (Coetzee et al. 2007). Another set of speakers changes voiced stops into fricatives in all positions but post-nasally (these speakers are called "leniters"). A third set of speakers use the so-called PND system: for these speakers, voiced stops surface as voiceless only post-nasally (Zsiga et al. 2006, Coetzee et al. 2007, Coetzee and Pretorius 2010). The three systems of Tswana are represented below.

Table 3.6: Microdialects of Tswana (from Zsiga et al. 2006, Coetzee et al. 2007).

	***#ba**	***aba**	***mba**
devoicers	#pa	apa	mpa
leniters	#βa	aβa	mba
PND	#ba	aba	mpa

As Hyman (2001) argues, the PND system arises precisely through the combination of two other (devoicing and leniting) systems: leniters take on fricativization except after nasals (D > Z / [−nas]__), while devoicers undergo unconditioned devoicing (D > T). Following Dickens (1984), Hyman (2001) argues that post-nasal devoicers undergo both sound changes. In other words, Dickens (1984) and Hyman (2001) argue that PND in Tswana results from a combination of three sound changes, the first of which is fricativization of voiced stops except post-nasally (the sound change that operates in the leniters' system). Unconditioned devoicing happens next (the sound change that operates in the devoicers' system), and because voiced stops surface only after nasals, the result is apparent PND. This interstage is confirmed by Kutswe, a variety of Sotho in which the alternation is between [p] post-nasally and [β] elsewhere (as mentioned in Dickens 1984). This pattern is obscured in Tswana and Shekgalagari, however, by an additional change in the dialect with PND: unconditioned occlusion of fricatives (Z > D). After this change, voiced stops surface as voiceless after nasals and, crucially, the feature [±voice] is fully contrastive in the elsewhere condition (Hyman 2001). Recall that this final sound change also occurred in Yaghnobi, but only for the alveolar series of stops.

The attestation of the subdialect of Tswana that has unconditioned devoicing of voiced stops also speaks against analyses that assume post-nasal devoicing results from the loss of contrast in the feature [±voice], when voiced stops develop to voiced fricatives elsewhere (see Section 3.1). In the "devoicer" subdialect, unconditioned devoicing occurs without the development of voiced stops to fricatives (i.e. when the feature [±voice] is fully contrastive).

The pattern of development of Tswana voiceless stops provides a structural parallel that speaks in favor of the proposed explanation. As Hyman (2001) points out, voiceless stops underwent fricativization (except post-nasally) along with voiced stops. Table 3.7 shows this development.

Table 3.7: Development of voiceless stops in Tswana with examples (table from Hyman 2001).

	#__	**V__V**	**N__**	**#__**	**N__**	**Gloss**
*p	ɸ	ɸ	pʰ	ɸeɲa	m-pʰeɲa	'conquer (me)'
*t	r̥	r̥	tʰ	rátá	n-tʰátá	'love (me)'
*k	h, x	h, x	kxʰ, kʰ	xátá	ŋ-kxʰátá	'trample (me)'

These data provide yet another piece of evidence that complementary distribution occurred first in Tswana and neighboring dialects, in both the voiced and voiceless series of stops (Hyman 2001). The voiceless and voiced series underwent lenition except in post-nasal position (as in the leniters dialect), and then voiced stops underwent further changes (unconditioned devoicing, as in the devoicers dialect) to produce PND, whereas voiceless stops retained the complementary distribution.[43]

3.2.3 Bube and Mpongwe

Despite the fact that Bube and Mpongwe are not closely related to Tswana, Shekgalagari, and Makhuwa, we find striking similarities between these languages that uniformly point to the conclusion that complementary distribution (with lenition of voiced stops except post-nasally) operated in prehistory of Bube and Mpongwe, too. The data for the two languages, however, is sparse and detailed phonetic descriptions are lacking.

Several indicators in Bube data clearly point to a pre-stage with complementary distribution. Janssens (1993:37) reports that in the labial series, the voiced stop and voiced fricative [b] and [β] are in free variation in medial position. Also, exactly parallel to Tswana, the voiced velar stop is lost in Bube, pointing indirectly to an interstage with [ɣ] (also reconstructed independently in Janssens 1993:27). Finally, the alveolar series of stops undergoes lenition as expected: in the elsewhere condition, [d] develops to [l] or [r] (likely through an interstage *[ð]), depending on the vowel quality, e.g. *-dɔ́b- > [-lɔ́bâ] (Janssens 1993:23). In fact, Janssens independently reconstructs a proto-stage of Bube with exactly the complementary distribution we observe in other languages with PND: voiced stops surface as fricatives except post-nasally.

In Mpongwe, PND that targets only the velar series also arises through a combination of three sound changes. Mouguiama-Daouda (1990) independently reconstructs a stage in which [g] surfaced as a voiced velar fricative [ɣ] except post-nasally, citing as evidence the fact that the fricative is still realized as [ɣ] by some older speakers (Mouguiama-Daouda 1990). PND in Bube

---

[43] Because other Sotho-Tswana languages and Makhuwa are closely related to Tswana, the same analysis can be applied to these languages as well.

and Mpongwe thus follows the usual trajectory: complementary distribution and unconditioned devoicing of stops that surface only post-nasally.

3.2.4 Konyagi

Konyagi and related languages, too, provide strong evidence in favor of a stage with complementary distribution, just like in Yaghnobi or Tswana. Proto-Tenda, an ancestor of Konyagi, is reconstructed on the basis of Konyagi and its neighboring languages Basari and Bedik (Merrill 2014, 2016a,b). The first stage of Proto-Tenda features a phonemic inventory with voiced and voiceless stops and fricatives. In Stage II, all stops fricativized everywhere but in post-nasal position. This fricativization in Proto-Tenda was quite radical, targeting both voiced and voiceless stops, as well as nasal stops and, as I reconstruct, geminates/clusters[44]. The fricativization is directly confirmed by the Basari dialect, which preserves the Proto-Tenda II stage with minor deviations. Table 3.8 shows the development of consonants in non-post-nasal position from Proto-Tenda I and II to Basari and Konyagi. Just like in Basari, original voiced stops surface as fricatives in Konyagi (further developing to voiced sonorants in some places of articulation), except after nasals.

Table 3.8: Development of non-post-nasal consonants from Proto-Tenda to Konyagi (table from Merrill 2014, 2016a,b, Santos 1996).

Proto-Tenda I	p	t	c	k	f	ʃ	x	b	d	ɟ	g	w	ɣ	l	m	n	ɲ	ŋ
Proto-Tenda II	f	r̥	ʃ	x	f	ʃ	x	w	r	j	ɣ	w	ɣ	l	w̃	l̃	j̃	ɣ̃
Basari	f	s	ʃ	x	f	ʃ	x	w	r	j	ɣ	w	ɣ	l	w̃	n	j̃	ɣ̃
Konyagi	f	r	s	x	f	s	x	w	l	j	/	w	/	l	w̃	l̃	j̃	/

---

[44] I reconstruct that voiced geminate fricatives later undergo occlusion to voiced stops.

In post-nasal position, Konyagi voiced stops remain stops. The following table summarizes the development of consonants in post-nasal position in the descendants of Proto-Tenda (from Merrill 2014, 2016a,b, p.c.).[45] Post-nasally, no fricatives occur in either Konyagi or Basari.

Table 3.9: Development of post-nasal consonants from Proto-Tenda to Konyagi (table from Merrill 2014, 2016a,b, Santos 1996).

Proto-Tenda I	p	t	c	k	f	ʃ	x	b	d	ɟ	g	w	ɣ	l	m	n	ɲ	ŋ
Proto-Tenda II	p	t	c	k	p	c	k	b	d	ɟ	g	b	g	l	m	n	ɲ	ŋ
Basari	p	t	c	k	p	c	k	b	d	ɟ	g	b	g	l	m	n	ɲ	ŋ
Konyagi	p	t	c	k	p	c	k	p	t	c	k	p	k	l	m	n	ɲ	ŋ

While in Bedik and Basari, original post-nasal voiced stops remain voiced, they devoice in Konyagi. Here too, however — like in all other cases of PND — the voiced stops that devoice post-nasally are in fact the only voiced stops in the language (the stage confirmed in today's Basari). Thus, again, the sound change operating in Konyagi is in fact unconditioned devoicing of voiced stops. The apparent PND is once more the result of a combination of sound changes.

The third sound change, occlusion back to stops, is lacking in Konyagi. As a result, it is not voiceless and voiced stops that alternate at some stage of development in Konyagi, but rather voiced fricatives and voiceless stops. Konyagi, however, does not lack voiced stops completely: original fricative geminates were later occluded to stops and simplified, resulting in voiced stops being reintroduced into the synchronic inventory.[46] All these changes result in a synchronic mutation system that involves voiced and voiceless stops and voiced fricatives or approximants (see Section 2.3.2).

---

[45] The nasal stops are lost before voiceless stops.

[46] Likewise, a Proto-Tenda sequence of a nasal and an implosive yields an ND sequence in Konyagi.

So far, I have argued that all cases of apparent PND can be accounted for through a combination of two or three well-motivated sound changes. I now turn to a case of PND from the South Italian dialects to illustrate that such sound change combinations involving complementary distribution are not limited to stops, but can apply to other segments as well.

3.2.5 South Italian

On the surface, the data in South Italian suggest that *ʤ devoices to [ʧ] only in post-nasal position. However, if we look at the development of *ʤ elsewhere, we observe that it gets de-occluded and further develops to [j] except after nasals (e.g. *faʤina > Calabrian [fajina], *leʤere > Calabrian [lejere], Sicilian [lejiri]; Rohlfs 1949:358). Given the cases of post-nasal devoicing discussed so far, de-occlusion of the affricate and the development to the glide provides evidence for a stage with complementary distribution. At the point when *ʤ appears post-nasally and [j] (or probably *ʒ) elsewhere, an unconditioned devoicing of voiced affricates occurs. This, too, is a well-attested and motivated sound change: voiced affricates are highly dispreferred, which is why devoicing is a natural tendency in these segments. Voice is difficult to maintain, especially in affricates, which combine two articulations that are highly antagonistic to voicing: closure and frication (Ohala 1983, 2006). Table 3.10 illustrates the reconstructed development: Stage 2 shows a period of complementary distribution, and Stage 3 represents the development after the unconditioned devoicing of voiced affricates.

Table 3.10: Devoicing of post-nasal affricates in South Italian.

Stage	**Sound change**	**Elsewhere**	**N__**
1		faʤina	pinʤere
2	ʤ > j / [−nas]__	fajina	pinʤere
3	DZ > TS	fajina	pinʧiri

Note that this set of sound changes is in principle the same as in the previous cases, but here complementary distribution targets affricates instead of stops. Also, Stage 4 is absent from South Italian: the change that would occlude [ʝ] to [ʤ] is absent.

3.2.6 Buginese and Murik

The emergent pattern that we have seen in all the apparent cases of PND so far can be generalized as follows: (1) a set of segments enters complementary distribution; (2) a sound change occurs that operates on the unchanged subset of those segments; (3) optionally, another sound change occurs that blurs the original complementary distribution environment.

Let us now turn to the three Austronesian languages. On the surface, the data from Buginese and Murik seem to point to PND operating as a single sound change. Moreover, there is no direct historical or dialectal evidence to suggest otherwise, as is the case for Yaghnobi, Tswana, and Konyagi. If the only attested instances of PND were those found in Austronesian languages, we would likely be forced to assume the operation of a single sound change — PND. However, these languages do, at least, show clear traces of a stage with complementary distribution between voiced stops post-nasally and voiced fricatives elsewhere. Below, I argue that the three-sound-change explanation again better captures the data, despite the lack of direct historical or dialectal evidence.

The main evidence against PND as a single sound change in Austronesian comes from the voiced labial stop in Buginese. Already in Proto-South-Sulawesi (PSS; from which Buginese developed), *b had developed to [w] except word-initially and post-nasally (Mills 1975: 547). Later, the change *b > [w] also targeted the word-initial position, as is clear from initial stop in cases like *bumbun > [wumpun]. Thus, at one stage in the language's development, voiced stops surfaced only post-nasally: again, we have clear evidence for complementary distribution.

From there, the development followed the trajectory described above: unconditioned devoicing of voiced stops occurred, but produced apparent PND because voiced stops surfaced only post-nasally. The development is illustrated in Table 3.11.

Table 3.11: Development of PND in bilabials in Buginese.

Stage	**Sound change**	**Language**	**Example**
1		PSS	*bumbun
2	D > Z / [−nas]__	Pre-Buginese	*wumbun
3	D > T	Buginese	wumpun

In Buginese, /w/ continues to surface as a non-obstruent (since only two changes operated in the labial series). Based on the development of the labial series, we can reconstruct that the velar series likewise undergoes complementary distribution and unconditioned devoicing. Unlike the labial series, however, the voiced velar fricative *ɣ undergoes occlusion to [ɡ] (all three sound changes operated in the velar series), thus obscuring evidence for an inter-stage with complementary distribution.[47] Note that this is precisely the same scenario attested in Yaghnobi, with the only difference being that, in Yaghnobi, it was the alveolar series of fricatives that underwent occlusion, whereas in Buginese, it was the velar series. The development leading to apparent PND in velars (with all three sound changes operating) is illustrated in Table 3.12.

Table 3.12: Development of PND in velars in Buginese.

Stage	**Sound change**	**Language**	**N__**	**Elsewhere**
1		PSS	*aŋɡəp	*ɡiliŋ
2	D > Z / [−nas]__	Pre-Buginese	*aŋɡəp	*ɣiliŋ
3	D > T	Pre-Buginese	*aŋkəp	*ɣiliŋ
4	Z > D	Buginese	aŋkəp	ɡiliŋ

The dental series of stops in Buginese escapes PND because *d developed to [r] in all positions (*dindiŋ > [renriŋ]) and, as such, became ineligible for the devoicing of voiced stops. The

---

[47] The reconstruction of the interstage with [ɣ] thus relies on the labial series.

development of *z also conforms to the proposal above: intervocalically, it undergoes rhotacism to [r] (possible through an interstage with *ʝ); post-nasally, I reconstruct it occludes to *ɟ and devoices to [c] (according to the unconditioned devoicing of voiced stops which predictably targets the palatals); initially, it remains a fricative *ʝ and later occludes together with *ɣ to [ɟ] (in other words, there was likely a stage of complementary distribution: *ɟ post-nasally, *ʝ/r elsewhere).

The initial stage with complementary distribution in Buginese stops is additionally confirmed by a recent description of Buginese by Valls (2014) which reports that voiced stops [b, d, g] have voiced fricatives [β, ð/ɾ, ɣ] as allophones in apparent free variation in non-post-nasal position. This suggests that the last sound change that turns voiced fricatives to voiced stops is likely in progress in Buginese.

In Murik, complementary distribution is likewise attested and can be found in the development of voiced dental stops. On the surface, Proto-Kayan-Murik *d surfaces as [l] initially, [r] intervocalically, and [t] post-nasally (see Table 3.13). These data point to a stage with complementary distribution (PKM in the Table 3.13). I propose a historical development with the following trajectory: Proto-Kayan-Murik *d lenited to [l] or [r] (probably through an inter-stage with *ð) initially and intervocalically. After a nasal, however, *d remained a stop.

Table 3.13: Development of coronals in Murik.

Stage	**Language**	**#__**	**V__V**	**N__**
1	Pre-PKM	d	d	d
2	PKM	ð	ð	d
3	Murik	l	r	t

Next, I argue that unconditioned devoicing of voiced stops occurred. Because devoicing operated during a period when voiced stops surfaced only after nasals, apparent PND is the result. Based on the development of the alveolar series, I reconstruct that fricativization targeted labials and velars too: original voiced stops fricativized to *β and *ɣ (except after a nasal), at which point voiced stops (surfacing only post-nasally) got devoiced (e.g. *b > [p], *g > [k]). Unlike the alveolars, *β and *ɣ then underwent occlusion "back" to stops, resulting in apparent PND. Table 3.14 traces the proposed trajectory from (Pre-)Proto-Kayan Murik to Murik for the labial series of stops.

Table 3.14: Reconstructed development of labials in Murik.

Stage	**Language**	**#__**	**V__V**	**N__**
1	Pre-PKM	b	b	b
2	PKM	β	β	b
3	PKM	β	β	p
4	Murik	b	b	p

A peculiarity of the development in Murik is that it combines the "PND of stops" that we saw, for example, in Tswana, with the "PND of affricates" that we saw in Sicilian and Calabrian — i.e., whereas the previously discussed languages devoice either stops or affricates, Murik devoices both. The development of stops in this language is straightforward — it follows the usual trajectory of PND: complementary distribution, unconditioned devoicing, and then optional occlusion to stops (as we saw above). The development of affricates is more complicated, but nevertheless revealing. PMP *z develops to *s intervocalically already in Proto-Kayan-Murik (Blust 2005); this development cannot be considered part of PND, because it happens at an earlier stage. Elsewhere, *z is preserved as voiced and palatalizes to *ʝ. Post-nasally, occlusion to an affricate *ɟʝ takes place.[48] The affricate gets devoiced, exactly as in Sicilian and Calabrian,

---

[48] Alternatively, *z gets occluded to an affricate *ɟʝ in all positions and later undergoes deocclusion to *ʝ in initial position, parallel to the South Italian development.

together with devoicing of other voiced stops. The initial fricative *ʝ then gets occluded to a voiced stop, together with other voiced fricatives in Murik.[49]

In sum, even though Buginese and Murik offer neither dialectal nor historical evidence for complementary distribution, there is enough language-internal evidence to posit that, at one stage, voiced stops (and affricates) surfaced only after nasals: the original voiced labial stops surface as fricatives even today in Buginese, whereas in Murik voiced dental stops surface as lenited [r] or [l], in accordance with the reconstructed complementary distribution.

3.2.7 Nasioi

That PND in Nasioi results from a combination of three natural sound changes is strongly suggested by Nasioi internal evidence and by diachronic development of its related languages. Evidence for a stage with complementary distribution comes from Nasioi itself: voiced stops /b/ and /d/ surface as a fricative [β] and a flap [ɾ] intervocalically. Lenition of voiced stops in related languages strongly suggests that a complementary distribution consisting of voiced stops developing to voiced fricatives except post-nasally, operated already in pre-Nasioi. In Buin and Nagovisi, for example, [ɾ] and [d] are in complementary distribution whereby [d] surfaces post-nasally and [ɾ] elsewhere (Brown 2017). In Motuna, the complementary distribution is not limited to alveolars, but targets labials as well. The stops [b] and [d] are in complementary distribution with [w] and [r] such that the stops surface only post-nasally and the approximants surface elsewhere, including in word-initial position (Onishi 2012). Table 3.15 summarizes the development of Motuna voiced labial and avleolar stops.

---

[49] The affricate articulation in initial position (Blust 1974) is likely secondary: there exists variation between affricate and stop articulation. It is well known that palatal stops often develop into affricates.

Table 3.15: Summary of the development of Motuna labial and velar stops (Onishi 2012).

	#__	**V__V**	**N__**
*b	w	w	b
*d	r	r	d

It is reasonable to assume that Motuna represents the proto-stage of Nasioi. In the development of Nasioi, voiced stops devoice, but because they surface only post-nasally, the change looks like post-nasal devoicing. Finally, Nasioi undergoes occlusion of fricatives to stops (like Tswana and Shekgalagari, Yaghnobi, and Buginese and Murik), but this occlusion is contextually limited to the initial and post-obstruent positions (see Section 3.2.7). Intervocalically, voiced fricatives still surface as fricatives. Occlusion of fricatives that is limited to initial position or to clusters is a well-motivated and common sound change (see Kümmel 2007 and the discussion in Section 3.3 below).

3.3 Naturalness of the three sound changes

All sound changes assumed by the proposals above are natural, that is, both phonetically well-motivated and well-attested, with clear phonetic precursors. A survey of sound changes in Kümmel (2007) lists approximately 56 cases in which voiced stops undergo fricativization in either post-vocalic or intervocalic position (in four cases the non-nasal environment is specifically mentioned), plus an additional two cases in which voiced fricatives occlude to stops post-nasally. Moreover, only two of the 17 surveyed languages with NC sequences permit sequences of nasal + continuant phonotactically (Maddieson 1984, reported in Steriade 1993). The articulatory reasons for the occlusion of post-nasal fricatives are clear: if in the transition from nasal stop to oral fricative the velum rises early, it causes "denasalization of the final portion of the nasal consonant" (Busà 2007:157) and results ultimately in a period of oral (denasalized) occlusion (see also Ohala 1993, 1997b, Ali et al. 1979). Moll and Daniloff (1971) show that in 83% of English NC(CC) and NCN sequences, the gesture toward velar closure onsets during the nasal consonant: "movement toward velar closure is initiated long before the articulatory contact for the non-nasal consonant, usually during the preceding nasal" (Moll and Daniloff 1971:681). Because the majority of velar closure movements onsets simultaneously with the "moment of nasal contact", Moll and Daniloff (1971) speculate that precisely the nasal contact articulation might trigger the early velar movement (also supported by Kent et al. 1974).[50] While the mechanisms of epenthetic occlusion in NS clusters are well described, there are relatively few phonetic studies of the process available. The few existing studies observe that epenthetic closure (in NS) is shorter than underlying closure (in NTS), although in some studies this difference fails to reach statistical significance (Fourakis and Port 1986, Yoo and Blankenship 2003, Recasens 2012 and a discussion in Warner and Weber 2001). A number of properties of epenthesis in NS clusters suggest that the processes is a passive phonetic tendency: epenthesis is often not a categorical process, operates gradiently, and phonetic properties of epenthetic oral closure differ from the closure of underlying stops.[51] Further phonetic studies on languages other than English are desired (to my knowledge, the only proper phonetic account of this phenomena outside of English is Recasens' 2012 study of Valencian Catalan). Despite the lack of phonetic studies, it

---

[50] As Moll and Daniloff (1971) point out, this speculation is not without difficulties as gesture onsets for velar movement appear elsewhere on the temporal continuum. For a discussion, see Moll and Daniloff (1971).

[51] Fourakis and Port (1986) report that in some dialects of English, oral closure never surfaces in NS sequences. The fact that epenthesis can be language specific does not mean that occlusion in NS sequences is not a passive phonetic tendency: speakers of these dialects apparently feature a high level of articulatory precision for the timing of gestures in NS sequences or even delay the velic movements after the release of the nasal stop. This means velar closure coincides or happens later than stop release. In other dialects and languages that feature less articulatory precision in the timing of NS sequences, the articulatory mechanism outlined above will cause velar closure movement to onset earlier, which in turn results in epenthetic closure. It is also possible that different mechanisms are responsible for dialectal differences. Phonetic data from languages other than English is desired for this discussion.

is safe to assume dispreference of frication post-nasally is a UPT that is typologically common and articulatorily well motivated.

Unconditioned devoicing of voiced stops is a well-motivated process too. Closure is in all respects antagonistic to voicing for clear aerodynamic reasons: the closure causes air pressure buildup in the oral cavity, which results in an equalization of subglottal and oral pressure. When this happens, the vocal folds are unable to vibrate due to the lack of airflow and voicing ceases. This mechanism has long been known as the Aerodynamic Voicing Constraint (Ohala 1983, 2011). The antagonism of closure to voicing is also confirmed by typology: of 706 languages surveyed, 166 have only voiceless stops (Ruhlen 1975, reported in Ohala 1983), among others "Cantonese, Hawaiian, Zuñi, Ainu, and Quechua" (Ohala 2011:64). As a comparison, only four languages are reported to have only voiced stops (Ohala 1983), and even in those languages, it appears that stops are voiceless initially.[52] Unconditioned devoicing is also attested as a sound change. In Tocharian and in some subdialects of Tswana (Gouskova et al. 2010), in addition to other examples listed in Kümmel (2007), all voiced stops devoice in all positions, including the post-nasal position.

One might question the naturalness of unconditioned devoicing when voiced stops surface exclusively in post-nasal position: what mechanisms motivate devoicing if it operates only in the post-nasal position which itself favors voicing? As defined in (2), naturalness must always be evaluated with respect to a given context. Post-nasally, voicing of voiceless stops is indeed a UPT: voicing passively continues into closure in post-nasal position because nasal leakage and volume expansion counter the anti-voicing effect of closure (compared to stops in other

---

[52] Initial stops in Yidiny have been reported to surface as voiceless, or at least partially voiced (Dixon 1977). My preliminary analysis of recordings made by Dixon and obtained from AIATSIS confirm these claims: stops tend to be voiceless utterance-initially.

positions). The increased amount of passive voicing, which is a phonetic precursor based on coarticulation, can result in a sound change of post-nasal voicing. However, voiced stops absent a specific context are nevertheless articulatorily dispreferred when compared to voiceless stops: closure is always antagonistic to voicing for clear aerodynamic reasons. Closure causes airflow to cease because of pressure build up (Ohala and Riordan 1979). To counter this effect, speakers must passively or actively accommodate for voicing by expanding the volume of the oral cavity; otherwise vocal folds cease to vibrate after 5–15 ms (without any accommodation) or after approximately 70 ms (with only passive accommodation) from the onset of closure (Ohala and Riordan 1979). Because of nasal leakage and volume expansion due to velum rising, speakers need to accommodate for voicing the least in post-nasal condition; much less than, for example, word-initially or word-finally (Westbury and Keating 1986, Steriade 1997, Iverson and Salmons 2011). This means that devoicing that only targets the post-nasal position in a language that has voiced stops in other positions would indeed be unnatural. Despite nasal leakage and volume expansion, however, speakers still need to accommodate for voicing (all else being equal), in order to counter the anti-voicing effect of closure — even in post-nasal position. Failure to accommodate for voicing will result in devoicing, which means that devoicing is motivated post-nasally when it targets other positions as well. If voiced stops happen not to surface other than post-nasally due to an earlier sound change, devoicing is still phonetically motivated precisely because closure is antagonistic to voicing and speakers have to accommodate for voicing in all positions (despite the fact that some environments counter the antagonistic effect of closure on voicing).

Three pieces of evidence from the data independently support the position that unconditioned devoicing operates in the development of PND. None of the languages with reported PND, with the exception of Yaghnobi, have non-nasal clusters or permit voiced stops to surface in any other position than post-nasally (at some stage of development). This means that the effect of unconditioned devoicing can only be observed post-nasally. Yaghnobi, however, shows that devoicing of voiced stops operated in its prehistory, not only post-nasally (N__), but also in the other position where voiced stops surfaced: after voiced fricatives (Z__). In other words, devoicing in Yaghnobi had to be unconditioned rather than limited to post-nasal position, which is clear from the fact that devoicing targeted all positions where voiced stops surfaced: both ND *and* the more marginal ZD clusters.

The second piece of evidence in support of the articulatory motivation of unconditional devoicing comes from Tswana. As already mentioned, there are at least three microdialects within Tswana: a devoicing dialect, a leniting dialect, and a PND dialect (Table 3.6).[53] Phonetic studies of the dialect that devoices all stops in all positions show that devoicing is complete and unconditioned: devoiced stops that go back to voiced stops are voiceless in all positions, including the post-nasal position (for instance, the devoiced post-nasal labial and alveolar stop have only 11% of closure voiced, compared to 7% and 12% intervocalically[54]; Gouskova et al. 2011). These instrumental results support my assumption that unconditioned devoicing is an attested sound change, motivated by the anti-voicing effect of closure, even if it appears to operate only post-nasally. The devoicing microdialect of Tswana confirms that where unconditioned devoicing happens, it targets stops in all positions, including the post-nasal position, precisely because

---

[53] Some other systems of stops within Tswana microdialects are also reported: e.g. positional devoicers that have voiceless stops in all positions but initial position, where voiced stops remain voiced (Gouskova et al. 2011). Such systems are likely the result of fricativization except post-nasally, devoicing of stops, occlusion of initial fricatives and devoicing of fricatives. Such systems are thus a combination of developments that we see in PND and intervocalic devoicing that we see in Berawan dialects. The treatment of this development is, however, beyond the scope of this work.

[54] For one speaker alveolars have 7% of closure voiced post-nasally and 5% intervocalically (Gouskova et al. 2011).

closure is always antagonistic to voicing. Furthermore, the fact that unconditioned devoicing, confirmed by instrumental phonetic studies, operates precisely in the subdialect of the language that features PND, additionally strengthens the assumption that the second sound change in the three-sound-change combination leading to PND was *unconditioned* devoicing.

Third, that voicing of stops is a phonetic tendency post-nasally, while devoicing is a tendency evaluated on a global, unconditioned level (even for post-nasal position) is strongly suggested by data in Davidson (2016). Davidson (2016) acoustically analyzes percentage of phonation into closure in English voiced stops. To my knowledge, the study is one of the largest of the kind including 37 speakers. Davidson (2016) measures voicing into closure of English voiced stops [b], [d], and [g] and divides the stops into three categories: stops are *devoiced* if phonation ceases within the first 10% of the closure; stops are *fully voiced* if phonation continues at the 90% of the closure. Stops that fall between these two categories are considered *partially voiced*. The data show that stops have the highest proportion of fully voiced closure precisely in post-nasal position. In other words, evaluating contextually, post-nasal position favors voicing more strongly than any other position. Despite this distribution, devoicing of stops is still a phonetic tendency when evaluated globally. English voiced stops [b], [d], and [g], despite being phonemically voiced, are partially voiced or devoiced in 22% of cases in word-internal post-nasal pre-vocalic position (-N__V-). In other words, in 22% of recorded underlying voiced stops in word-medial post-nasal prevocalic position (N = 272),[55] voicing into closure ceases somewhere before the 90% of the closure. This tendency reflects the anti-voicing effect of the closure that operates in all environments, although it is more suppressed post-nasally compared to other positions for clear articulatory reasons.

---

[55] Davidson (2016) presents data for all following contexts. I am grateful to Lisa Davidson (p.c.) for providing me with the counts in pre-vocalic and word-internal position.

Finally, unconditioned occlusion of non-sibilant fricatives to stops is also well attested and motivated.[56] Non-sibilant fricatives are typologically and articulatorily dispreferred (Maddieson 1984). There are many languages without fricatives in their inventories, but none without stops (21 or 6.6% of languages surveyed in Maddieson 1984 lack any fricative, even a strident fricative). Moreover, fricatives require a greater level of articulatory precision than other manners of articulation: compared to stops, the articulatory targets and shape of the vocal tract require greater precision (Ladefoged and Maddieson 1996: 137). Deviation from precise articulatory targets can thus lead to the occlusion of fricatives to stops. Kümmel (2007) identifies at least six languages in which a sound change turned voiced fricatives to voiced stops for all places of articulation, as well as several more in which occlusion is limited to a single place of articulation. Kümmel (2007) also identifies cases of occlusion of fricatives that is limited to the initial position or to clusters (as is reconstructed for Pre-Nasioi).

3.4 The Blurring Process

In Section 3.2 above, I argued that one of the rare reported cases of an unnatural sound change, PND, did not operate as a single unnatural sound change, but as a combination of natural sound changes in all thirteen languages in which PND is reported either as a synchronic alternation or as a sound change. This typological study provides a basis for establishing a new diachronic model for explaining unnatural phenomena, the Blurring Process, which I present in this section. Using the Blurring Process, I also provide a proof that at least three sound changes are required for an unnatural process to arise: the Minimal Sound Change Requirement.

---

[56] The occlusion of fricatives is another example of naturalness needing to be evaluated with respect to a given context. While unconditioned occlusion of fricatives is well-motivated (see the discussion in this paragraph), intervocalically its opposite process, fricativization of stops, is the natural direction (see also Ladefoged and Maddieson 1996: 137, Kaplan 2010).

We saw above that all cases of PND proceed along a common trajectory of development. For all cases, I reconstruct a stage with complementary distribution, which is followed by an unconditioned sound change. To get a synchronic unnatural alternation, another sound change has to operate that blurs the original complementary distribution. I label this historical development a *Blurring Process.*

(10) *Blurring Process*

a. A set of segments enters complementary distribution

b. A sound change occurs that operates on the changed/unchanged subset of those segments

c. Another sound change occurs that blurs the original complementary distribution

Based on (10), I can identify several trajectories that result in unnatural processes: subtypes of the Blurring Process. Let us assume that A → B / X is a natural alternation and a UPT, whereby A and B represent feature matrices that select one or more segments in a given language and X specifies the environment of the alternation. Because A → B / X is a UPT, A and B differ in exactly one feature, ϕ_1, such that the value of ϕ_1 in B is universally preferred in the environment X. This means that its inverse process, B → A / X, is an unnatural process: the value of ϕ_1 in A is universally dispreferred in X. How does B → A / X arise? There are a number of possible trajectories, but I will focus on two main trajectories, i.e. combinations of sound changes, that are attested as historical developments. I will refer to the first development in (11) as the *Blurring Cycle* and the second development in (11) as the *Blurring Chain.* The crucial difference between the two is that in the Blurring Cycle, the sound change B > A does operate, but because it is unconditioned, i.e. not limited to the unnatural environment

X, it can be phonetically motivated. In the Blurring Chain, on the other hand, B > A never operates. Instead a "chain" of developments occurs: B > C (/X) > D > A. The motivation for the term "cycle" is clear: the last sound change in the Blurring Cycle reverses the first sound change (although the two differ in their contexts). In other words, the last sound change targets the outcome and results in the target of the first sound change. The term "chain" is likewise motivated: the outcomes of a sound change in Blurring Chain become targets for following sound changes. Both developments "blur" the original complementary distribution, resulting in an alternation that operates against universal phonetic tendency (B > A / X).[57] The Blurring Cycle and Blurring Chain are schematized in (11) with the unnatural alternation (B > A / X) that results from a combination of the three sound changes appearing under the line.

	Blurring Cycle	*Blurring Chain*
	B > C / ¬X	B > C / X
(11)	B > A	C > D
	C > B	D > A
	B > A / X	B > A / X

PND in all thirteen cases is a result of the Blurring Cycle.

	Blurring Cycle	*PND*
	B > C / ¬X	D > Z / [−nas]
(12)	B > A	D > T
	C > B	Z > D
	B > A / X	D > T / [+nas]

---

[57] Although some aspects of the Blurring Process resemble rule ordering in opacity (Kiparsky 1971, 1973), I avoid the opacity terminology, because the end result of the Blurring Process is a non-opaque simple alternation. Moreover, opacity concerns synchronic alternations, whereas the Blurring Process is a diachronic development. I likewise avoid the Duke-of-York terminology (Pullum 1976): the sound changes in the Blurring Process never operate in the opposite (unnatural) direction, as would be the case in the Duke-of-York derivation.

The Blurring Process approach explains further unnatural data beyond PND. In Section 3.5 of this chapter, I argue that the Blurring *Chain* explains the unnatural voicing of voiceless stops in Tarma Quechua (Adelaar 1977, Puente Baldoceda 1977, Nazarov 2008) and unnatural intervocalic devoicing in Berawan dialects (Blust 2005, Blust 2013, Burkhardt 2014).

The Blurring Process thus serves as a diachronic model for explaining seemingly unnatural sound changes and synchronic processes and should be considered as an alternative to other strategies. The most common strategy for explaining unnatural sound changes thus far is Ohala's (1981) hypercorrection approach. Indeed, most studies that have tried to explain PND in isolation have invoked hypercorrection: speakers analyze sequences of ND as voiced from NT and mentally "undo" this voicing. However, in the absence of any restriction, hypercorrection as an explanation for unnatural phenomena leads to overgeneration. Unrestricted hypercorrection leads to the conclusion that every unnatural sound change should be possible (but perhaps very rare) — since, by definition, unnatural sound changes operate against UPTs. Speakers can analyze the surface data as having undergone a UPT and "undo" the UPT to get the unnatural process with a single sound change. This is the reason why Ohala (1981) himself restricts the operation of hypercorrection to "those consonantal features [...] which have important perceptual cues spreading onto adjacent segments". He specifically notes that the voice feature is unlikely to undergo dissimilation based on hypercorrection (see also Blust 2005). I have argued here that the Blurring Process approach is superior to hypercorrection in the case of PND; in Section 3.5, I argue that the Blurring Process better explains other unnatural processes, such as intervocalic devoicing. While this chapter does not argue against the existence of hypercorrection, further research should reveal what processes are better explained by one or the other strategy.

The Blurring Process has another advantage: it provides the groundwork for establishing the minimal number of sound changes required for natural, unmotivated, and unnatural processes to arise. Specifically, I argue that we can prove formally what we observe typologically: that the emergence of an unnatural process requires at least three sound changes to operate in combination. As per the definition in (4), we assume that a single instance of sound change means a change of one feature in a given environment (Section 2.1.2). For a natural sound change (A > B / X), A and B differ in exactly one feature ϕ_1 (for example [±voice] in the case of PNV or final voicing), so that a given value of ϕ_1 in B is universally preferred in environment X and its opposite value $\neg\phi_1$ in A is dispreferred in the same environment X. How do we get the unnatural B > A / X? With one single natural sound change, it is impossible, because B > A / X is by definition unnatural. Moreover, a combination of two natural sound changes also cannot yield B > A / X. Why? A and B differ in one feature only (ϕ_1). For a B > A / X sound change to arise, therefore, we first need B to change into something other than A (it cannot change to A directly because such a sound change is unnatural). So, let B change to C, where B and C differ in one feature, ϕ_2, but, to be sure, a different feature from the one that separates A and B (ϕ_1). From this point, it is impossible for an unnatural sound change to arise without a third sound change. Indeed, C cannot develop directly to A, since the two segments differ in two features: feature ϕ_1, which distinguishes A and B, and feature ϕ_2, which distinguishes B and C, with $\phi_1 \neq \phi_2$. Since, by definition, two sound changes are required in order to change two features (see the discussion in Section 2.1.2), it follows that at least three sound changes must take place in order for an unnatural process to arise. This proof can be formalized as the *Minimal Sound Change Requirement*:

(13) *Minimal Sound Change Requirement (MSCR)*

Natural processes arise through a minimum of one sound change. A minimum of two sound changes have to operate in combination for an unmotivated process to arise. A minimum of three sound changes have to operate in combination for an unnatural process to arise.

The MSCR is derived even more clearly if we use feature notation to represent the Blurring Process. Let ϕ_1 and ϕ_2 be two features in a feature matrix. Let us assume that a change in the direction $-\phi_1 > +\phi_1$ is a universal phonetic tendency, given a value of some other feature in the feature matrix ϕ_{1+n} (and other participating features) and given an environment X. How does the unnatural change in the direction $+\phi_1 > -\phi_1$ arise? According to the definition, sound change cannot produce $+\phi_1 > -\phi_1$ in a single step, given the constant value α of ϕ_{1+n}. The change from $+\phi_1 > -\phi_1$, however, *can* be phonetically motivated with different values of ϕ_{1+n} (e.g. in a Blurring Chain) or when $[+\phi_1, \alpha\phi_{n+1}]$ only appears in a given environment, which means the context becomes irrelevant ($\varnothing$) for evaluating naturalness (as is the case in the Blurring Cycle). In other words, we cannot change $[+\phi_1,\alpha\phi_{n+1}]$ to $[-\phi_1,\alpha\phi_{n+1}]$, but it is possible that $+\phi_1 > -\phi_1$ is motivated under the $\neg\alpha$ value of ϕ_{n+1} or $\varnothing$ (motivated under a different context). This means that first, a sound change that targets ϕ_{n+1} has to operate and change its value, either in a given environment X (Blurring Chain) or in the elsewhere condition ¬X (Blurring Cycle). Under the changed ϕ_{n+1}, the $+\phi_1 > -\phi_1$ can be motivated (which is the second sound change in the Blurring Process). Finally, in order for the change $+\phi_1 > -\phi_1$ to appear unnatural, the value of ϕ_{n+1} has to change to the initial stage (the third sound change). Feature values of each stage in the Blurring Process that produce the unnatural $[+\phi_1, \alpha\phi_{n+1}] >$

$[-\phi_1,\ \alpha\phi_{n+1}]$ are illustrated in Table 3.16. At least three changes are needed to get from Stage 1 to Stage 4 (MSCR).

Table 3.16: Changes in feature values in a blurring process.

Stage	1.	>	2.	>	3.	>	4.
$\begin{bmatrix}\boldsymbol{\phi_1}\\ \vdots\\ \boldsymbol{\phi_{1+n}}\end{bmatrix}$	$\begin{bmatrix}+\\ \vdots\\ \alpha\end{bmatrix}$		$\begin{bmatrix}+\\ \vdots\\ \neg\alpha\end{bmatrix}$		$\begin{bmatrix}-\\ \vdots\\ \neg\alpha\end{bmatrix}$		$\begin{bmatrix}-\\ \vdots\\ \alpha\end{bmatrix}$

As already mentioned, rule telescoping (Wang 1968, Kenstowicz and Kisseberth 1977: 64, Stausland Johnsen 2012) has long been known to produce unmotivated results. However, to derive unnatural alternations, we need a special combination of sound changes: the Blurring Process. To my knowledge, this chapter also presents the first proof establishing the minimal number of sound changes required for different degrees of naturalness. The MSCR is a crucial concept when deriving typology within the CB approach.

It is important to note that fewer than three sound changes are required to produce unnatural static phonotactic restrictions. For instance, Kiparsky (2006) provides several trajectories of two or more sound changes leading to the unnatural process of final voicing. Most of his scenarios, however, would result in a phonotactic restriction against voiceless stops word-finally and not in FV as an alternation. Even a single sound change might produce an unnatural phonotactic restriction: if a sound change targets voiceless stops to the exclusion of voiced stops, the resulting phonological system would allow only voiced oral stops in word-final position. Similarly, post-vocalic voicing of stops would result in a system that would restrict word-final voiceless stops and allow voiced stops in word-final position. Such a system is better analyzed as a natural restriction against post-vocalic voiceless stops, but another sound change that would reintroduce post-vocalic voiceless stops into the system would create an unnatural phonotactic restriction where voicing of stops would be contrastive initially and word-medially, but not word-finally.

For the purposes of modeling diachronic development of unnatural processes, this dissertation distinguishes between static phonotactic restrictions and alternations. The difference is that the latter require a segment to alternate in a given phonetic environment within the same morphological unit. In other words, MSCR holds for alternations in which a segment B surfaces as a segment A in an environment X *within* the same morpheme.

One of the advantages of constraint-based approaches to phonology is that static phonotactics and alternations are modeled simultaneously with the same theoretical devices. It is likely that productive phonotactic restrictions and productive alternations have the same underlying grammatical mechanisms and should therefore not be distinguished in synchronic grammars. There are three main reasons for why such a distinction is justified when modeling unnatural processes diachronically. Alternations provide considerably more and more reliable evidence for learners. In case of alternations, the learners are faced with surface forms within the same morpheme that contain both A and B. In case a combination of sound changes does not result in an alternation, speakers are faced with either categorical or gradient distributions. There is no doubt that phonotactic restrictions can be active phonological processes, but the likelihood of learners not acquiring a distribution of segments as an active process is much higher if evidence for the distribution does not surface in the same morpheme. Because of this increased likelihood that a combination of sound changes that provides only distributional evidence will not result in a productive process, combined with the fact that typology of unnatural phonotactics is more challenging to establish (see below), I make a distinction between alternations and static phonotactics when modeling diachronic development of unnatural processes. In fact, the two unnatural phonotactic restrictions that I argue show signs of productivity (Berawan and TQ; Sections 2.4.2 and 2.4.3) did arise from a combination of sound changes (Sections 3.5.2.1, and 3.5.2.3) that would also give evidence for an alternation if all changes operated categorically in the two languages. The fact that evidence for the unnatural restriction was morphological (which resulted from the Blurring Process) likely contributed to the productivity of the unnatural phonotactic restrictions.

Second, phonotactic restrictions that do not arise from a Blurring Process might allow alternative analyses. In the scenario outlined above in which post-vocalic voicing of stops is followed by a sound changes that reintroduces voiceless stops in post-vocalic (but not final) position, a potential analysis could treat the newly introduced stops as a distinct series. While such an analysis is undesirable, it is available in the case of phonotactic restrictions and impossible in the case of alternations: one cannot assume a different phoneme within the same morphological unit. Finally, I distinguish phonotactic restrictions from alternations for practical purposes. It appears that many unnatural phonotactic restrictions go unnoticed in the literature or are not analyzed as unnatural, probably precisely due to the reasons described above. Identifying unnatural distributions/phonotactic restriction is a harder task not only for the learner, but also for linguists describing languages. For this reason, it is considerably more difficult to create typological studies of unnatural phonotactic restrictions. For example, I have found three languages that potentially allow only voiced but not voiceless stops in word-final position. Most of these languages have not been discussed in the literature on unnatural processes. Ho, some dialects of Spanish, and Tundra Nenets might have a restriction against final voiceless stops while allowing word-final voiced stops for at least a subset of places of articulation or at least in phonetic variation. In Ho, for example, voiceless stops do not appear word-finally (except in some loanwords), but voiced labial and retroflex stop do surface word-finally (Pucilowski 2013). Word-final voiceless stops and labial and velar voiced stops are in Spanish restricted to loanwords. Word-final /d/, however, surfaces in native vocabulary. In a variety of Madrid Spanish, word-final /d/ is most frequently deleted or surfaces as a voiceless fricative, but in citation form, /d/ is in 30% recorded to surface as a voiced stop [d], although the experimental design was limited to one lexical item and includes 13 speakers (Burkard and Dziallas 2018). In Tundra Nenets, voiced stops are reported to surface word-finally to the exclusion of voiceless stops (Salminen 1997). All these cases are preliminary descriptions and detailed phonetic data is needed to confirm these restrictions. If the stops in word-final position are indeed phonetically real, these cases suggest that final voicing as a phonotactic restriction might be more common than previously thought. This would align well with the distinction between unnatural alternations that follow the MSCR and unnatural phonotactics that do not.

MSCR is also limited in scope to segmental phenomena. Prosodic processes cannot be represented by feature matrices as in Table 3.16 and require an independent tier under most approaches to prosodic phonology. Because the changes on a prosodic tier can be independent from changes in the segmental tier, unnatural processes can arise from less than three sound changes. Unnatural prosodic phenomena are beyond the scope of this dissertation. A more detailed investigation of what constitutes an unnatural change and whether the lack of the MSCR restriction in the prosodic domain translates into a higher rate of unnatural processes would be desired.

In sum, the MSCR states that unnatural segmental alternations always arise from at least three sound changes. Unnatural phonotactic restrictions or prosodic alternation can arise from fewer sound changes. The dissertation distinguishes between alternations and phonotactic restrictions because the first are more likely to result in productive processes, lack alternative analyses, and are easier to detect. While MSCR is not a requirement on unnatural phonotactic restrictions, the two cases presented here (Berawan and TQ) that did result in productive unnatural phonotactic restrictions actually did arise from three sound changes. A further study of unnatural phonotactics and their diachronic developments would reveal how readily restrictions arising from fewer than three sound changes result in productive processes.

A combination of more than one sound change can in some rare cases also result in what would usually be analyzed as a natural alternation. This is primarily true for processes that allow multiple intermediate stages in their development. For example, the natural process /k/ → [ʧ] / __[+front] can arise through a combination of individual sound changes as defined in Section 2.1.2 (with intermediate stages such as [c] > [cç]). The resulting process, /k/ → [ʧ] / __[+front] is nevertheless natural because it operates in the direction of a universal phonetic tendency of fronting of velars before front vowels (Guion 1996).

There are two potential exceptions to the MSCR. First, unnatural alternations might arise from less than three sound change when the change from $+\phi_1$ to $-\phi_1$ (Stage 2 to 3 in Table 3.16 automatically entails the change from $\neg\alpha$ to α (in ϕ_{n+1}; Stage 3 to 4 in Table 3.16). In other words, if a change in one feature (such as a change from [−nas] to [+nas]) entails change of another, automatic (redundant) feature (e.g. [−son] to [+son]) or if one of the changes in the Blurring Process is a perceptual change that can target two features simultaneously and precisely those two features play a role in the Blurring Process, then the unnatural alternation might arise from two changes, as we know that the minimality principle (see Section 4.4.2) holds only for non-automatic non-redundant features and might be violated in some cases (see Section

4.4.2). We expect such cases, if at all existent, to be extremely rare because of the restriction that the change of two features in the Blurring Process be automatic/redundant or simultaneous.

Second, MSCR might not hold in rare cases in which a phonetic change allows a reinterpretation of phonological status of certain segments. Such cases are, however, expected to be rare and, more importantly, probably always allow alternative analyses. Let us imagine a language in which voiceless stops shorten in word-final position or in coda position more generally. Postvocalically, voiceless stops universally have some degree of voicing into closure as it is difficult to cease voicing immediately after the vowel (Westbury and Keating 1986, Davidson 2016). If closure duration of such voiceless stops shortens significantly in the coda position, the percentage of voicing into closure will increase along with the shortening of closure duration. Such shortened voiceless stops can now be analyzed as either shortened voiceless stops or as voiced stops, because of the higher percentage of voicing into closure due to their shorter duration.[58]

---

[58] This scenario is assumed to have occurred in Lakota (Adam Albright, p.c., for a description, see Rood 2016). Lakota voiceless aspirated and unaspirated stops surface as "at least partially" voiced stops in coda position (Rood 2016; see also Riggs 1893). On the surface, this development appears as an unnatural alternation — coda voicing. Due to the lack of phonetic studies, however, it is unclear how much voicing into closure these stops actually have and whether they feature more voicing into closure in absolute terms compared to voiceless stops. In any case, comparative data suggest that FV in Lakota is in fact a result of three sound changes — a Blurring Process. That Lakota final voicing has a complex history is argued by Rankin's (2001) unpublished work, summarized in Rood (2016), who assumed Lakota development did not proceed through a single stage. While Rankin (in Rood 2016) assumes that voiced stops go back to nasals, I will argue here for a slightly different development. Lakota voiceless /p/, /t/, and /k/ are reported to surface as [b], [l], and [g] word-finally after apocope of final vowels or in coda position after reduplication or compounding and in other morphological environments. The coronal series reveals that voiceless stops not only voice in coda position, but in the case of the coronal /t/ also develop to a sonorant [l]. On the surface, this reflex seems easy to explain: /t/ first voiced to /d/ after which the sound change that turned [d] to [l] occurred. This seems reasonable at first glance because Lakota [l] corresponds to Dakota [d] and to [n] in other related languages. This assumption is, however, misleading: there never was a sound change [d] > [l] in Lakota. The sound that [l] goes back to is reconstructed as Proto-Sioux *R that variously yields [r], [l], [d], [nd], or [t] (with similar reflexes to the original *r) in daughter languages (Rankin et al. 1998). *R was possibly a voiced fricative with a high degree of sonority, although its phonetic status is unclear. This means that the voiceless /t/ most likely first fricativized and voiced to *R in coda position in morpheme boundaries, where articulation is generally weakened. Such a development has a parallel in the development of Spanish voiceless stops that have various realization syllable-finally, e.g. /k/ can be realized as [x], [ɣ], or [g] before a stop like /t/ (Quilis 1993). In Lakota, such lenition appears to be limited to morpheme boundaries. From here, the coronal (most likely voiced fricative) *R regularly develops to Lakota [l] and to [n] in other languages (Rankin et al 1998). We can assume the same development for the other two series, /p/ and /k/ as well. There exist comparative evidence for fricativization and voicing in these two places of articulation. Cognates of Lakota voiced word-final stops in related languages are either voiceless stops or nasals (Riggs 1893, Rankin et al. 2015). We saw that voiced fricatives word-finally yield nasals in related languages. It is thus likely that /p/ in word-final/coda

In such cases that do not undergo the Blurring Process, however, the conditioning factor for a change is still present in the synchronic system. In the presence of this conditioning factor, alternative analyses are possible. Finally, even if unnatural patterns can arise outside of the Blurring Process, they are comparatively rare, as this dissertation argues. All unnatural processes known to me for which no alternative explanations exist result from the Blurring Process. Under this less strong position, then, MSCR is at least a strong tendency and the rest of the argumentation that follow from MSCR still holds.

We saw that, while a single sound change is constrained to follow phonetic naturalness, a combination of sound changes appears unconstrained: any number of single instances of sound change can operate on each other (limited of course by the timeframe of active operation), and even if such a combination results in unnatural alternation, the synchronic grammar can still incorporate it (Section 2.3.2); the final sound change will not be blocked (*pace* Kiparsky 2006, 2008). However, if we assume that the combination of sound changes is unconstrained, we still need to explain why unnatural processes (B > A / X) are rare — as in the case of PND — or even unattested — as in the case of final voicing. Chapter 4 presents a new model for deriving typology within the CB approach that crucially relies on the Blurring Process and MSCR.

---

position, too, fricativizes and devoices to *W (another proto-phoneme that was likely a labial voiced fricative). Instead of developing to a sonorant, *W (< *p) undergoes occlusion to [b]. In related languages, it yields a nasal [m], e.g. *topa yields Lakota [tob] and [tom] elsewhere (Riggs 1893). There is independent evidence that suggest bilabial sonorants occludes to [b] in pre-consonantal positions: Proto-Siouan *waRó·ke > blokétu (Rankin et al. 2015, thus also Rankin 2001). Finally, there is comparative evidence that voiceless stops in the velar series (/k/) also fricativize in clusters: Proto-Dakota *kri yield Dakota [hdi] and Lakota [glí] 'come or arrive at home' (Rankin et al. 2015). A more elaborate account of the development of FV is beyond the scope of this dissertation, but the preliminary analysis suggest that a combination of sound changes might underlie the unnatural alternation in this case too.

3.5 Origins of unnatural phonotactics

In this section, I show that the Blurring Process explains further unnatural processes beyond PND and propose a new and unified treatment of historical developments leading to the cases of unnatural gradient phonotactics presented in Section 2.4. I argue that what appears to be a clear case of a single sound change that operates in the unnatural direction is better explained as a combination of three natural sound changes (the so-called Blurring Cycle). This approach automatically derives several unusual aspects of the data including the unnatural rates of application of voicing in TQ and the intriguing change of place of articulation in Berawan.

3.5.1 Previous accounts

A potential explanation for TQ data is given in Nazarov (2008). The most elaborate historical treatment of the alleged unnatural sound changes in Section 2.4 is given in Blust's (2005) analysis of Kiput and Berawan. Both explanations are in many ways similar, which is why they will be discussed together in this section. The discussion of historical development of Kiput and Berawan in Blust (2005) is closely related to Blust's hypothesis that unnatural sound changes *do* exist. He first specifically rejects the possibility that intervocalic devoicing could be anything but a single sound change: "intervocalic devoicing affected a single feature value. There is thus no possibility of considering a concatenation of natural changes which cumulatively produced an unnatural result" (Blust 2005:243). According to Blust, the Berawan data directly attest to existence of unnatural sound changes precisely because the unnatural intervocalic devoicing had to operate as a single sound change.

As already mentioned for PND, the most common strategy for explaining unnatural sound changes thus far is invoking Ohala's (1993) hypercorrection. Blust (2005) proposes that IVD in

Berawan can be considered a dissimilation based on hypercorrection. Nazarov (2008) makes a similar claim for Tarma Quechua. Because the opposite process, intervocalic voicing, is common, "the listener assumes wrongly that an assimilation has taken place and mentally 'undoes' it" (Blust 2005:243). Blust (2005) also acknowledge the problems that such an explanation brings. First, [±voice] is, according to Ohala (1993), a feature less commonly prone to dissimilation (Blust 2005:244). Moreover, as Blust (2005) acknowledges, the dissimilation by hypercorrection hypothesis fails to explain why devoicing operates only on a subset of places of articulation (e.g. alveolars undergo lenition instead of voicing).[59] Finally, hypercorrection is not well-suited for explaining Tarma Quechua: explaining of different rates of voicing in Tarma Quechua becomes problematic under the hypercorrection hypothesis. It is unclear why hypercorrection would operate more frequently post-nasally than intervocalically or why it would operate more frequently post-consonantally than intervocalically.[60]

Blust's (2005) argument against the possibility that multiple sound changes operated in the pre-history is also problematic. The fact that sound change targets only one feature value provides no evidence that would exclude the possibility of multiple sound changes operating in combination. In fact, I present a body of evidence arguing to the opposite: that one single sound change could not have operated in the history of Berawan dialects (see Table 3.22).

---

[59] Other proposals that invoke dissimilation as perceptual enhancement or that claim intervocalic devoicing is phonetically motivated are also discussed in Blust (2005). All proposals face similar problems: they fail to account for asymmetries in voicing across different places of articulation and fail to derive the peculiar voicing distribution in Tarma Quechua. Due to the problems that all current proposals of intervocalic devoicing face, Blust (2005) leaves the question of how exactly the unnatural sound change arose, open.

[60] One potential explanation for lack of voicing post-nasally within the hypercorrection approach could come from contacts with varieties of Quechua with post-nasal voicing (Adelaar and Muysken 2004). However, it is not clear that this contact occurred and how this hypercorrection could have occurred. Additionally, this leaves the difference in voicing rates between post-consonantal and intervocalic unexplained.

3.5.2 A new explanation

Just like for post-nasal devoicing, it appears on the surface (and it has been claimed in the literature for Berawan and Kiput, Blust 2005) that single instances of unnatural sound changes operate in the development of Kiput, Berawan and Tarma Quechua. The fact that the unnatural phenomena in these languages are gradient (lexically diffusing) is another supposed argument in favor of a single sound change hypothesis: gradience is the prominent property of sound changes in progress. Using the Blurring Process model, however, I will argue that the seemingly unnatural sound changes and the resulting phonotactic restrictions in Kiput, Berawan, and Tarma Quechua arise from a combination of three natural sound changes, and I will point to advantages that this explanation bears over the alternative single-sound-change approaches.

3.5.2.1 Berawan

3.5.2.1.1 Diachronic development

As already mentioned, Berawan dialects have been reported to undergo intervocalic devoicing (Blust 2013, 2017, Burkhardt 2014). Berawan bilabial and velar stops *b and *g devoice intervocalically, but remain voiced word-initially.[61] Bilabial stops undergo an additional change intervocalically: in addition to devoicing, they also change their place of articulation from bilabial to velar (*b > k / V__V). Table 3.17 lists some cases of intervocalic devoicing reported in Burkhardt (2014) and Blust (2013: 667-8).[62]

---

[61] Voiceless stops remain largely unchanged in all positions with some secondary developments (Burkhardt 2014). For a treatment of voiced stops word-finally (where they develop to nasal stops), see Section 5.3.1. For argumentation that devoicing is likely not conditioned by stress, see Section 2.4.3.1.

[62] Proto-Malayo-Polynesian *ʀ and *g developed to *g in Pre-Berawan (Burkardt 2014) and this change is applied to the reconstructed forms for the purpose of clarity.

Table 3.17: Examples of intervocalic devoicing in Berawan (data from Blust 2013 and Burkhardt 2014).

Sound change	PMP/Pre-Berawan	Batu Belah
*b > k / V__V	*abiəŋ	akiŋ
	*bibi	biki
	*bəlibiəw	bəlikiəw
	*bibuj	bikuj
	*dibiən	dikin
*g > k / V__V	*bigiu	bikiw
	*gigiəq	gikiʔ
	*magi	maki
	*igiəŋ	ikiŋ
	*ugat	ikit

The list in Table 3.17 is merely an illustration of intervocalic devoicing; it is far from exhaustive. In fact, IVD in Berawan is well-documented and almost exceptionless. A comprehensive study of Berawan dialects in Burkhardt (2014) includes between 425 and 466 vocabulary items for each of the four languages and Pre-Berawan reconstructions for each cognate (489 in total). Based on our counts, *b or *g appears intervocalically in 36 of these reconstructed words, and in all 36 cases the Berawan dialects show a voiceless stop, the regular reflex of *b and *g in intervocalic position.[63]

In contrast to intervocalic position, *b and *g remain unchanged in initial position. There are 46 reconstructed words with initial *b in Pre-Berawan. In all but one word the initial *b remains unchanged.[64] A similar distribution holds for the velar voiced stop in initial position

---

[63] Long Terawan undergoes further changes that do not interact with our analysis (see Burkhardt 2014).

[64] In the one exception, devoicing occurs initially in all four dialects: *bəlippiəŋ > pəlipiŋ. According to Burkhardt (2014:144), this development is sporadic in a word that already exhibits another sporadic development: degemination of -pp-. There is only one other example in which devoicing initially occurs only in Long Terawan: *buraq > [purăh] (Burkhardt 2014).

as well: *ɡ is reconstructed in twelve lexical items of Pre-Berawan and in all of them voicing is retained.[65] Table 3.18 lists some examples of initial voiced stops in Pre-Berawan and Berawan.

Table 3.18: Initial voiced stops (data from Blust 2013 and Burkhardt 2014).

PMP/Pre-Berawan	Batu Belah
*ɡəm	ɡəm
*ɡiɡun	ɡikuŋ
*ɡimot	ɡimok
*bitok	bitok
*buliən	bulin
*busak	busek

A peculiar fact about the diachronic development of Berawan is that, while velar and bilabial stops undergo devoicing, alveolars undergo lenition in the same word-internal position. Pre-Berawan voiced alveolar *d remains a voiced stop initially, but develops to [r] word-internally.

In addition to the unexpected medial devoicing change, there is another quite natural type of devoicing operating in Berawan: devoicing of voiced geminates. Because geminates only appear intervocalically, this devoicing change is seemingly restricted to intervocalic position as well. Devoicing of geminates, however, is well-motivated as a context-free sound change. Since voicing is articulatorily difficult to maintain during the closure due to decreased airflow (Ohala 1983, 1997a), and geminates have longer closures, voiceless geminates are universally preferred over voiced ones. Berawan geminates arose after schwa, from consonant clusters, and after "h-accretion": addition of [h] at the end of words which caused shortening of vowels and consequently lengthening of consonants (Burkhardt 2014:260, 282-286). Unlike simple alveolar stops, geminate alveolar stops did undergo devoicing (7-c). Some examples of the development of geminates are given in (14).

---

[65] There is only one case of sporadic devoicing in Long Terawan.

(14) *Origins of geminates in Berawan*

a. *bunbun > *bubbun > buppuŋ

b. *taɡʀaŋ > *taɡɡaŋ > takkiŋ

c. *m-iddəm > mittăm

Geminate devoicing, too, contributes to the restriction against intervocalic voiced obstruents, precisely because geminates surface only intervocalically. However, because geminate devoicing is not unnatural and because voiceless geminates are preferred to voiced ones in all positions, I do not consider geminate devoicing to be a case for or against unnatural sound change, and I will not discuss these cases any further.[66]

3.5.2.1.2 A Blurring Chain

Stage 1 and 2 in a Blurring Chain is the development of a complementary distribution (Section 3.4). The material presented in Sections 2.4 and 3.5.2.1.1 provides several pieces of indirect evidence for the assumption that stops in the three languages entered complementary distribution at some stage of the development. Table 3.19 summarizes the development of Pre-Berawan voiced stops.

---

[66] Labial geminate stops arising after schwa and from consonant clusters do not undergo a change in place of articulation (unlike simple stops), e.g. *təbu > *təbbu > [təppu], *mə-bənnən > *mə-ppənnən > *ppənnən > [pənnən] (after the loss of *mə- and initial degemination) or *əbbıs > *əppıq > [pıʔ] (after the loss of initial schwa and initial degemination). Geminates arising via "h-accretion", however, do undergo a change in place of articulation: they develop to voiceless velar geminate stops. The relative chronology of gemination and devoicing is difficult to establish. We have two possible scenarios: either (i) gemination precedes devoicing (*tuba > *tuɡa > *tuɡɡah > [tukkih], argued for in Burkhardt 2014), or (ii) devoicing precedes gemination (*tuba > *tuɡa > *tukah > [tukkih]). Because the exact development cannot be reconstructed or is at best based on relative chronology, we will not discuss the geminate cases any further.

Table 3.19: Summary of developments in Berawan (based on Burkhardt 2014 and Blust 2013).

Pre-Berawan	**Berawan**	
	#__	**V__V**
*b	b	k
*d	d	r
*g	g	k

The intriguing aspect of the development of Berawan is that, while the labial and velar series undergo intervocalic devoicing (fortition or a decrease in sonority), the alveolar series of stops undergoes intervocalic lenition, i.e. increase in sonority. This asymmetry is hard to explain under other accounts. Under the Blurring Chain approach, the asymmetry is actually expected. Lenition of alveolars in intervocalic position suggests an earlier stage with complementary distribution. Pre-Berawan *d develops to [r] intervocalically and remains a voiced stop [d] initially. It is likely that the increase in sonority intervocalically followed a gradual path via fricativization of [d]: *d > *ð > [r] (which is a common sound change, cf. Kümmel 2007:60, 79). In other words, I reconstruct that voiced alveolar stops underwent intervocalic lenition to [r], likely through an interstage with [ð], which means that at some point in the development [d] was in complementary distribution: the voiced stop surfaced as a fricative intervocalically.

Based on the development of alveolars, I can reconstruct that such complementary distribution underlies the other two series of stops as well. Let us posit that Pre-Berawan first undergoes intervocalic lenition in all series of stops, not just in alveolars. Intervocalic fricativization of voiced stops is a common and phonetically motivated (Kirchner 2001, Kaplan 2010) — natural sound change. As already mentioned, the alveolar series preserves this initial stage of complementary distribution in today's system: intervocalically, *d surfaces as [r] < *ð and does not undergo devoicing, while initially it is preserved as a voiced stop. Stage 2 of the development is illustrated in Table 3.20.

Table 3.20: Stage 2 in the development of Berawan.

Pre-Berawan	**Berawan**	
	#__	**V__V**
*b	b	β
*d	d	ð > r
*g	g	ɣ

I propose that at the stage of complementary distribution in Pre-Berawan, another sound change (the second in the blurring cycle) occurred that targeted the changed subset of segments (10-b): unconditioned devoicing of voiced fricatives. Voicing in fricatives is highly dispreferred and articulatorily difficult to maintain — requirements for voicing and for frication are diametrically opposite which is the source of articulatory dispreferrence: "one condition requires oral pressure to be as low as possible, the other to be as high as possible" (Ohala 2006:688; see also Ohala 1983, 1997a, Smith 1997). Unconditioned devoicing of fricatives is thus a natural, motivated, and common sound change. Because voiced fricatives at this stage surface only intervocalically, the result is an apparent intervocalic devoicing. Note also that because *ð further develops to [r], it escapes fricative devoicing and the original complementary distribution is in the alveolar series still preserved. Stage 3 is illustrated in Table 3.21.

Table 3.21: Stage 3 in the development of Berawan.

Pre-Berawan	**Berawan**	
	#__	**V__V**
*b	b	ɸ
*d	d	r
*g	g	x

The Blurring Chain hypothesis has several advantages. The labial series of stops in Berawan not only underwent devoicing, but also change of place of articulation. The first advantage of the Blurring Chain approach is that this change of place of articulation is easier to motivate than under other approaches. The sound change [ɸ] > [x] or [β] > [ɣ] (if it happened prior to devoicing) is much more common than [p] > [k] or [b] > [g]. In fact, the only two cases of change of place of articulation from labial to velar in the survey of consonantal sound changes in Kümmel (2007) involve precisely fricatives: none are reported to involve stops. This distribution might be the result of greater perceptual similarity between [ɸ] vs. [x] or [β] vs. [ɣ] than between [b] vs. [g], although extensive studies on the perceptual aspects of this change are lacking — many studies that test perceptual confusability involve differences between non-strident and strident fricatives (e.g., Miller and Nicely 1955, Alwan et al. 2011). There exists some evidence in favor of perceptual motivation of the [ɸ] > [x] change: Redford and Diehl's (1999) data suggests that [f] vs. [θ] (another non-strident fricative) is perceptually more confusable compared to [p] vs. [t]. Regardless of how we motivate it, the change from labial to velar place of articulation is a much more common sound change when the target is a fricative than when the target is a stop and an explanation that invokes the first is more desirable than an explanation that invokes the latter.

The change in place of articulation that operated in Pre-Berawan reveals another crucial piece of evidence in favor of the Blurring Chain approach: if we assume that intervocalic devoicing operated as a single sound change, we cannot chronologically order the change in place of articulation with respect to intervocalic devoicing (cf. Blust 2017). Let us consider the option that intervocalic devoicing operated as a single sound change. There are two logical chronological orders of intervocalic devoicing and the change of place of articulation: either one precedes the other or vice versa.

Table 3.22: Two possible relative chronologies under the assumption that IVD operates as one sound change.

Chronology 1		**Chronology 2**	
1. intervocalic devoicing	b > p	1. change of place	b > g
2. change of place	p > k	2. intervocalic devoicing	g > k

If devoicing happened first, we would expect original [p] from Pre-Berawan voiceless *p to change its place of articulation as well. This does not happen: Pre-Berawan *apuj yields [apoj] and not **akuj in all four dialects. If change in place of articulation happened first, we would expect the change of place of articulation to operate in word-initial position as well. This does not happen: Pre-Berawan *bibi yields [biki], not **giki. The only possibility to chronologically order the two sound changes and derive the Berawan data with a single-sound-change approach is to limit the already unusual sound change — change of place of articulation in stops (b > g) — to an even more unusual environment, intervocalic position. This would be highly unexpected: stops are perceptually better cued internally than initially where formant transitions into closure are lacking. In the survey of consonantal sound changes in Kümmel (2007) there are no cases reported of a change of [b] to [g] in intervocalic position.

In fact, precisely the change of place of articulation that targets only intervocalic [b] while initial [b] remains unchanged strongly suggests that the two were at some point distinct sounds and that the sound change of intervocalic devoicing and change of place of articulation operated on one of the two sounds in complementary distribution.

Finally, the last sound change under the blurring chain approach (10-c) that operated in Pre-Berawan was occlusion of the velar voiceless fricative *x to [k]. Occlusion of fricatives is a natural and motivated sound change as well, although not as unidirectional as the other two in the blurring chain. Kümmel (2007) reports at least two cases of unconditioned sound change [x] > [k]. The sound change is also phonetically motivated: fricatives require more articulatory precision than stops (Ladefoged and Maddieson 1996: 137). The occlusion of fricatives can be motivated as reducing this articulatory precision, i.e. laxing of articulatory targets.

The sound change, *x > [k], blurs the original complementary distribution and the result is intervocalic devoicing, as it is attested in Berawan today. The Blurring Chain in Berawan that results in D > T / V__V is summarized in (15).

(15) *Blurring Chain in Berawan*

D > Z / V__V

Z > S

S > T

The reconstructed trajectory can be illustrated on a lexical item that includes both an initial and an intervocalic stop: Berawan [bikuj] 'pig' from Proto-Austronesian *babuj > *bibuj.

(16) *babuj > *biβuj > *biɸuj > *bixuj > [bikuj]

In sum, there exist several advantages of the blurring chain explanation in Berawan. First, lenition of the alveolar series of stops automatically follows from the new analysis: it reveals an earlier stage of complementary distribution. Likewise, change in place of articulation becomes well-motivated and consequently, the chronology problem outlined above in Table 3.22 is solved. Finally, all sound changes posited are natural and well-motivated.

3.5.2.2 Kiput

3.5.2.2.1 Diachronic development

In addition to Berawan, intervocalic devoicing as a sound change has been reported for another Austronesian language: Kiput (Blust 2005).[67] There, however, the sound change does not result in a significant unnatural trend in the lexicon. The language nevertheless provides insights into the historical development of unnatural phonotactics and helps us better understand intervocalic devoicing in Berawan (see Section 3.5.2.2 for how this evidence factors into our proposal).

Kiput is a Malayo-Polynesian and, more specifically, North Sarawakan, Berawan-Lower Baram language of the Austronesian family, spoken by approximately 450 speakers in northern Sarawak in Borneo, Malaysia (Blust 2002). It features several peculiar developments which have been extensively discussed in Blust (2002). The most unusual of these is intervocalic devoicing, detailed in Blust (2002, 2005, 2013).

Blust (2002) establishes that the Pre-Kiput voiced velar stop *g, palatal affricate *ɟ͡ʝ and labiodental fricative *v devoiced to Kiput [k], [c͡ç], and [f], respectively in intervocalic position.[68]

---

[67] Recently, intervocalic devoicing has been reported as a synchronic alternation for Sula in Bloyd (2017). It is clear from the data that the intervocalic devoicing there cannot be result of a sound change: devoicing operates exclusively at morpheme boundaries, whereas elsewhere voiced stops remain voiced intervocalically (Bloyd 2015). The existence of intervocalic devoicing as a synchronic process there does not speak against our proposal. The alternations are nevertheless interesting from a synchronic perspective: it seems that there is indeed synchronic intervocalic devoicing in Sula. Because the data are sparse and the language is poorly described, I leave Sula out of our discussion. Further investigations into the prehistory of Sula and its synchronic alternations are a desideratum.

[68] All three consonants that devoice have transparent origins in Proto-North-Sarawakan (PNS; the direct predecessor of Pre-Kiput). Pre-Kiput *g goes back to a PNS voiced velar stop *g, whereas Pre-Kiput *ɟ͡ʝ and *v have various different sources in PNS. Pre-Kiput*ɟ͡ʝ continues PNS *ɟ͡ʝ or goes back to a PNS glide *j that is both phonemic and also automatic in hiatus sequences with high front vowel and a following vowel. By the same token, *v goes back to *w which can be either phonemic, or automatic in hiatus between a high back vowel and any following vowel (Blust 2002).

Word-initial obstruents remain voiced. Word-final stops devoice by final devoicing; clusters are not allowed. Obstruents do not appear in other positions (Blust 2002).

Table 3.23 provides examples of intervocalic devoicing in Kiput. For the voiced velar stop series, the list is exhaustive: of 307 items on the vocabulary list with reconstructions in Blust (2002), four lexical items have intervocalic [g] in Proto-North-Sarawakan (Proto-North-Sarawakan (PNS)). In three cases, devoicing occurs. The fourth case is an exception to this rule: PNS *tegeʀaŋ yields Kiput [təgəriə]. For the developments [ɟ͡ʝ] > [c͡ç] / V__V and [v] > [f] / V__V the table lists only a subset of all cases from the list. There are altogether 19 and 9 cases of devoicing of [ɟ͡ʝ] and [v], respectively, in the same 307-word vocabulary list.

Table 3.23: Examples of intervocalic devoicing from Kiput (data from Blust 2002 and 2005).

Sound change	**Pre-Kiput**	**Kiput**
*g > k / V__V	*agem	akəm
	*pager	pakəl
	*tugal	tukin
*ɟ͡ʝ > c͡ç / V__V	*puɟ͡ʝut	puc͡çut
	*taɟ͡ʝem	tac͡çəm
	*kaju > kaɟ͡ʝu	kac͡çəw
	*lia > lija > liɟ͡ʝa	ləc͡çih
*v > f / V__V	*ɟ͡ʝawaj > *ɟ͡ʝavaj	dafiəy
	*sawa > *sava	safəh
	*dua > *duwa > *duva	dufih

As mentioned above, the obstruents *g and *ɟ͡ʝ remain voiced word-initially. There are seven lexical items with Proto-North-Sarawakan initial [g] in the 307-word Kiput vocabulary list. [g] remains voiced in all but one lexical item: Kiput [ketaan] for PNS *guta-an ‘able to endure pain’ (Blust 2002: 411).

The palatal affricate likewise remains voiced word-initially, but also loses its frication and develops to a voiced stop [d]. This occurs in three of four cases, e.g. *ɟ͡ʝawaj > [dafiəj]. In one word the affricate retains its frication: PNS *ɟ͡ʝauq yields Kiput [ɟ͡ʝəuʔ].[69]

The voiced bilabial fricative [v] does not appear word-initially.

The data presented here (from Blust 2002) thus confirms Blust's (2002) claim that devoicing occurs exclusively intervocalically. Devoicing targets only the velar stop, palatal affricate, and labiodental fricative: voiced labial and alveolar stops remain voiced in all positions. The developments are summarized in Table 3.24.

Table 3.24: Summary of developments in Kiput (data from Blust 2002 and 2005).

Pre-Kiput	**Kiput**	
	#__	**V__V**
*b	b	b
*d	d	d
*g	g	k
*ɟ͡ʝ, *j	d	c͡ç
*v, *w	/	f

Devoicing sometimes also operates in loanwords. Blust (2002) provides a list of 130 loanwords, mostly from Malay. In three cases, a borrowed voiced velar stop devoices, while it remains voiced in the remaining four, e.g. [sigup] → [sikup] vs. [bagi] → [bagiʔ]. The voiced palatal affricate devoices in three loanwords and remains voiced in six loanwords, e.g. [piɟ͡ʝit] → [pic͡çit] vs. [raɟ͡ʝin] → [raɟ͡ʝɪn].

To sum up, the data presented above suggests that unnatural intervocalic devoicing occurred as a sound change from Pre-Kiput to Kiput. Blust (2005: 243) goes a step further and claims that intervocalic devoicing had to occur as a single sound change because it targets only one

---

[69] Blust (2002) claims that in two cases initial [ɟ͡ʝ] remains an affricate. However, Kiput [ɟ͡ʝəj] goes back to PNS [aɟ͡ʝaj] in which [ɟ͡ʝ] appears intervocalically. [ɟ͡ʝəj] is therefore not a case of preservation of an initial affricate.

feature and because there exists “no possibility of considering a concatenation of natural changes which cumulatively produced an unnatural result.”

Diachronic intervocalic devoicing is likely to have created an unnatural phonotactic restriction against intervocalic voiced velar, labiodental, and palatal obstruents in favor of the unnatural element in this position: voiceless obstruents. The fact that devoicing has happened in some loanwords provides evidence for this. However, the restriction was probably only active for a limited period of time, after which novel vocabulary was introduced in the language via borrowings and the alleged intervocalic devoicing ceased to operate. The fact that devoicing has only applied to loanwords sporadically, i.e. gradiently, may be evidence the the unnatural phonotactic restriction was gradient, as it is in the case of Tarma Quechua and Berawan. However, an alternative explanation is that loanwords were introduced to the language at different stages in its development.

In synchronic Kiput there is no significant restriction against intervocalic voiced [g] any longer. I analyzed Blust’s (2003) 932-word vocabulary list, which altogether contains 10 words with intervocalic [g] and 63 words with intervocalic [k]. It is true that the voiced velar stop occurs much less often than its voiceless pair in intervocalic position. However, this is likely to be a consequence of the fact that voiced velar stops in Kiput are in general less frequent that their voiceless pairs. The vocabulary list in Blust (2003) includes 121 instances of word-initial [k], but only 21 instances of word-initial [g]. The number of occurrences of [k] and [g] are represented in Table 3.25. Statistical significance is calculated using Fisher’s Exact Test and Pearson’s Chi-squared Test. The ratio of voiced vs. voiceless stops is almost identical across the environments. The restriction against voiced intervocalic stops is not statistically significant in Kiput, with p-value equaling 1.0.

Table 3.25: Voiceless vs. voiced stops in Kiput word-initially and internally.

	Voiceless		**Voiced**		**% Voiced**	
Place	#__	V__V	#__	V__V	#__	V__V
velar	121	63	21	10	14.8	13.7

In sum, an unnatural sound change IVD is reported to operate in Kiput and the data seemingly suggest that IVD indeed operated in its pre-history. While the Kiput data do not provide direct evidence for the existence of unnatural gradient phonotactics, they offer important insights for the diachronic treatment of unnatural phenomena, as will be discussed below.

3.5.2.2.2 A Blurring Chain

In Kiput we also find clear traces of a stage with complementary distribution (stage 2 of the Blurring Chain). I claim that Kiput intervocalic devoicing, too, results from a Blurring Chain.

Sounds, targeted by devoicing in Kiput are summarized in Table 3.26.

Table 3.26: Devoiced sounds in Kiput (data from Blust 2002 and 2005).

Pre-Kiput	**Kiput**	
	#__	**V__V**
*g	g	k
*ɟ͡ʝ, *j	d	c͡ç
*v, *w	/	f

Note that, while *ɟ͡ʝ undergoes devoicing intervocalically, it also undergoes change in initial position: the affricate *ɟ͡ʝ loses its frication part and develops to [d]. In other words, *ɟ͡ʝ in Kiput enters complementary distribution: *ɟ͡ʝ surfaces as [d] initially and remains *ɟ͡ʝ intervocalically (stage 2). Let us reconstruct that, like in Berawan, the velar stop enters a similar complementary distribution (10-a): it surfaces as voiced fricative intervocalically and remains a stop initially. The voiced fricative [v] surfaces only intervocalically. Stage 2 is summarized in Table 3.27.

Table 3.27: Stage 2 in the development of Kiput.

Pre-Kiput	**Kiput**	
	#__	**V__V**
*g	g	ɣ
*ɟ͡ʝ, *j	d	ɟ͡ʝ
*v, *w	/	v

At this point, we can posit that the second sound change of the Blurring Chain operated (10-b): an unconditioned devoicing of voiced fricatives and affricates (stage 3). Fricative and affricate devoicing is a well-motivated natural sound change (see Section 3.5.2.1 above). Voiced palatal affricate devoices to [c͡ç], while voiced labiodental fricative *v devoices to [f] and voiced velar stop *ɣ devoices to *x. Stage 3 is summarized in Table 3.28.

Table 3.28: Stage 3 in the development of Kiput.

Pre-Kiput	**Kiput**	
	#__	**V__V**
*g	g	x
*ɟ͡ʝ, *j	d	c͡ç
*v, *w	/	f

That fricatives indeed devoice in Kiput is confirmed precisely by the attested development *v > [f]. While *x further develops to [k] via occlusion (just like in Berawan, (10-c)), [f] is still preserved as a fricative and directly shows that devoicing of fricatives operated in Pre-Kiput. Because affricates and fricatives only surface intervocalically, the Blurring Chain results in an apparent intervocalic devoicing.

3.5.2.3 Tarma Quechua

3.5.2.3.1 Diachronic development

No explicit treatments of the pre-history of TQ stop voicing exist in the literature (for Nazarov's 2008 proposal, see Section 3.5.1 above). We know, however, that Proto-Quechua and Pre-Tarma Quechua only had voiceless stops, or at least had no voicing contrast. Adelaar (1977) describes neighboring dialects as featuring only voiceless stops; my qualitative acoustic analysis of spectrograms of these dialects with no voiced stops suggests that voiceless stops are indeed realized as phonetically voiceless in all positions.[70] Therefore, it seems that voiced stops in TQ had to result from a sound change: voicing of voiceless stops.

If we posit that a single sound change produced this unnatural phonotactic restriction, we would have to assume that the sound change operates in a highly unnatural direction. The sound change would voice Pre-TQ voiceless labial and velar stops *p and *k to [b] and [g] intervocalically and post-consonantally, but not post-nasally. Such a sound change would thus operate in precisely the exact opposite direction of the universal phonetic tendencies described in Section 2.2. Voiceless alveolar stop *t would, however, have to remain unchanged in all positions (Adelaar 1977). Initial stops of any place of articulation would likewise have to resist voicing. These hypothetical developments along with reconstructions are given in Table 3.29.

---

[70] Acoustic analysis was made on recordings of dialects of Tarma Quechua with no voiced stops. Recordings by Willem Adelaar are noisy and there exists some dialectal mixing, which makes the analysis difficult. Nevertheless, spectrogram analysis shows that a substantial low energy "voicing bar" is lacking in voiceless stops across phonetic positions.

Table 3.29: Stop voicing in Tarma Quechua (data from Adelaar 1977).

Context	**Voicing**	**Labial**		**Velar**	
		Pre-TQ	TQ	Pre-TQ	TQ
#__	✗	*pirwa	pirwa	*kawa	kawa
N__	✗	*wampu	wampu	*ʧiŋka	ʧiŋka
V__V	✓	*kupa	kuba	*ʧaki	ʧagi
R,T__	✓	*takpa	takba	*kuʧka	kuʧga

The most intriguing aspect about this hypothetical sound change is that this unnatural voicing operates gradiently rather than categorically with different rates of application across different environments. If this is indeed a single sound change, it thus appears as if at some stage of development it operated as sound change in progress and did not operate categorically. At the same time, the fact that voicing would have to operate with greater frequency in post-obstruent position than post-nasally and intervocalically makes the hypothetical sound change even more unnatural.

For these two reasons — its highly unnatural direction and its apparent in-progress operation — voicing in Tarma Quechua appears to constitute one of the more compelling cases of seemingly unnatural sound changes in the literature.

3.5.2.3.2 A Blurring Chain

In the following, I argue that the unnatural voicing sound change and the resulting unnatural phonotactic restriction against post-nasal voiced stops and agreeing obstruent clusters in Tarma Quechua result from a Blurring Chain development as well. The distribution of voicing across lexical items in TQ is repeated in Table 3.30 (repeated from Table 2.17).

Table 3.30: Voiced vs. voiceless labial and velar stops in Tarma Quechua native vocabulary across contexts based on a vocabulary list in Adelaar (1977) (a similar count in Nazarov 2008).

	#__	N__	V__V	R__	T__
voiced	7	7	99	72	68
voiceless	276	67	134	13	11
% voiced	2.5	9.5	42.5	84.7	86.1

I argue that TQ voicing did not operate as a single sound change, but that the development, like in the cases described in Chapter 2 above, proceeded through a combination of three sound changes. I argue that the Blurring Chain approach automatically explains several peculiarities in the TQ development that other approaches are unable to derive.

Stage 1 in the development of Blurring Chain is complementary distribution (10-a). Based on phonetic facts (for details, see below), I argue that the first sound change in the development of the Blurring Chain in Tarma Quechua was fricativization of labial and velar voiceless stops [p] and [k] to *ɸ and *x. The alveolar series of stops fail to undergo fricativization, likely because fricativization would result in a dispreferred dental voiceless fricative [θ]. Similar systems in which fricativization targets only the labial and velar series of stops to the exclusion of alveolars are reported in Nepali and Taiwanese (Kaplan 2010).

I argue that the reconstructed Pre-Tarma-Quechua fricativization operates with different rates of application across different phonetic environments. Fricatives are universally dispreferred in post-nasal position. We therefore expect the rate of application of fricativization to be the lowest in post-nasal position. Intervocalically, fricativization is expected to operate more frequently than post-nasally: intervocalic fricativization is a common natural and phonetically well motivated sound change (Kirchner 2001, Kaplan 2010). Finally, in post-consonantal position fricativization is motivated by stop cluster avoidance: clusters of two stops are universally dispreferred and fricativization is employed to avoid these clusters. Kümmel (2007), for example, lists approximately 27 cases of sound change that fricativizes a voiceless stop in a stop-stop cluster.[71]

Fricativization thus creates the first condition for the Blurring Chain: complementary distribution (10-a) in which voiceless stops surface as voiceless fricatives intervocalically and in clusters. The second sound change in the development of TQ Blurring Chain targets the changed subset in the complementary distribution (10-b): I argue that non-sibilant voiceless fricatives undergo voicing in pre-vocalic position. Voiced fricatives are universally dispreferred (Ohala 1997a, 2006), but in a vocalic environment fricative voicing is a common sound change. The voicing of fricatives in TQ is almost exactly parallel to the voicing process in two languages: Catalan and Avestan. In Catalan, voiceless fricatives in final position undergo voicing before a following vowel or a sonorant, even when the fricative is preceded by a voiceless stop: in *cops amagats* ‘hidden blows’, for example, the realization of /-ps/ is [-bz] (Wheeler 2005, cited in Strycharczuk 2012). The difference between Catalan and TQ is that in the former, voicing is restricted to final position and the preceding stop undergoes voicing assimilation, while in TQ this does not happen. In Avestan, the parallel is even more striking. The Avestan dental fricative *θ develops to a voiced dental fricative ⟨δ⟩ [ð] after voiceless fricatives ⟨f⟩ [f] and ⟨x⟩ [x] in prevocalic and pre-sonorant position, e.g. ⟨uxδa-⟩ [uxða-] < *uxθa- < *ukt^{h}a- (Hoffman and Forssman 2004). There are numerous further cases of voicing of fricatives in a vocalic environment. Kümmel (2007) reports at least 13 cases where fricative voicing in a vocalic or [+voice] environment is a sound change that operated on more than one fricative in a language: there are 31 more cases in which one of the fricatives in the system is targeted. The sound change of fricative voicing in a vocalic environment is also a phonetically motivated sound change with

---

[71] In all reported cases, the stop that gets fricativized is the first member of the consonant cluster. See discussion below for an explanation of why fricativization in TQ occurs in the second member of the cluster.

similar phonetic motivations as intervocalic voicing of stops — the difference in subglottal and supraglottal pressure necessary for voicing is greatest in intervocalic position (see discussion in Section 2.2.1 and in Westbury and Keating 1986, Wheeler 2005, Strycharczuk 2012, Strycharczuk and Simon 2013, Davidson 2016). Data in Möbius (2004) also suggest that fricative voicing is a phonetic tendency in the context between a voiceless and a voiced sound: voiced fricatives in German TZ sequences receive gradual increase in voicing towards in the second half of the fricative. Likewise, data in Davidson (2016) show that fricatives have the highest proportion of voicing when preceded by voiceless stops.

The unnatural distribution of voicing in TQ becomes phonetically motivated under the Blurring Chain explanation: voicing distribution in fact goes back to the distribution of fricativization. As already mentioned, fricativization is expected exactly as attested: least frequently post-nasally, more frequently intervocalically, and most frequently in clusters.

The last sound change that operated in the Blurring Chain of TQ was occlusion of voiced fricatives "back" to voiced stops (10-c), just like in the development of Berawan and Kiput. This natural sound change (see discussion in 3.5.2.1) blurs the original complementary distribution and the result is the peculiar voicing of Tarma Quechua. The development is summarized in (17).

(17) *Blurring Chain in Tarma Quechua*

T > S / [-nas,-#]__

S > Z / __V

Z > D

There exists strong indirect dialectal evidence and direct phonetic evidence in favor of the proposed Blurring Chain explanation.

Dialectal data show that fricativization of voiceless stops is a common process across Quechua. Fricativization is reported “prominent” in dialects near Cusco and in Bolivia (Adelaar and Muysken 2004:199). In Cusco Quechua, for example, voiceless stops fricativize in clusters (from Adelaar and Muysken 2004:199):

(18) *Fricativization of voiceless stops in Cusco Quechua*

a. *aptaj > [haxʷtaj]
b. *upjaj > [uxjaj]

Fricativization of voiceless aspirated stops also occurs in Imbabura Quechua (Adelaar and Muysken 2004:199).

(19) *Aspiration and fricativization in Imbabura Quechua*

Proto-Quechua *paki > *pʰaki > Imbabura Quechua [faki]
Proto Quechua *qipa > *kʰipa > Imbabura Quechua [xipa]

Note the exact parallel between TQ and Imbabura Quechua with respect to place of articulation of fricativization: only the labial and velar series undergo aspiration and fricativization, while alveolars retain the stop manner of articulation.

The fact that in TQ, it is the second element in clusters that fricativizes is also not unmotivated. Most frequently, first elements (i.e. coda stops) undergo fricativization, but cases in which the second element fricativizes are attested too, e.g. Nivx (Shiraishi 2006, Kingston 2008).

In fact, fricativization of the second element likely finds internal motivation: fricativization of the second element results from the fact that only labials and velars fricativize in TQ. The following clusters of two stops are attested: [kb, tb, tg] (going back to *kβ, *tβ, *tɣ < *kɸ, *tɸ, and *tx). Note that in two of the three clusters, the first element is an alveolar stop that cannot undergo fricativization. To avoid clusters of two stops, either the labial or the velar stop had to undergo fricativization and they happen to surface in the second position. For the [kb]-cluster, it is possible that both elements underwent fricativization, but that only the second one underwent voicing because it surfaced in pre-vocalic position. Later, fricatives occluded to stops in Tarma Quechua and the result is the regular [kb] < *xβ.

Finally, there exists strong direct phonetic evidence within Tarma Quechua itself that the proposed Blurring Chain is the most likely trajectory of TQ development, i.e. that the development proceeded through an interstage with fricatives. Phonetic analyses of TQ show that voiced stops in fact occasionally still surface as voiced fricatives in apparent free variation. Figure 3.1 presents waveforms and spectrograms of phonemic voiced stops that surface as voiced fricatives. Spectrograms clearly show that the manner of articulation of sounds in question is frication. Formants are present throughout the consonantal part, even in cases where formants do not result from echo noise in the room: after a voiceless stop. Moreover, fricatives feature gradual increase in amplitude and lack burst, characteristic of stop consonants: both these features are confirmed in the spectrograms.

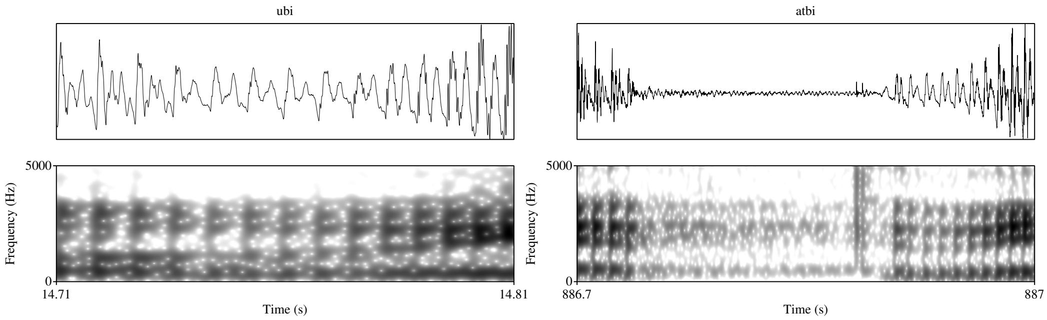


Figure 3.1: Waveforms and spectrograms of voiced stops that surface as voiced fricatives: intervocalically (left) and after a voiceless stop (right).

TQ phonetics thus preserves direct evidence that the unnatural voicing pattern developed through an interstage with fricatives and that the last sound change, occlusion of voiced fricatives to stops, does not operate categorically. In other words, the last sound change in the Blurring Chain seems to be in progress in the language recorded in Tarma in 1970.

There exists further phonetic evidence within TQ pointing to the validity of the Blurring Chain proposal. We saw that the last sound change reconstructed for TQ is occlusion of voiced fricatives to voiced stops. That fricatives in TQ indeed undergo occlusion is suggested by the *voiceless* series of stops/fricatives. Puente Baldoceda (1977:9) reports that voiceless labial fricatives and labial stops are in free variation in Tarma Quechua. In other words, original voiceless fricatives are in the process of undergoing occlusion. This development is confirmed by cases such as Tarma Quechua [ɸlawta] ∼ [plauta] (in free variation), borrowed from Spanish *flauta* with an original voiceless fricative. Exactly the same sound change in progress, occlusion of fricatives, is reconstructed by Blurring Chain for the voiced series of fricatives.

In sum, the Blurring Chain explanation bears several advantages over the alternative strategies for explaining unnatural gradient phonotactics in TQ. Rates of voicing across different environments that are highly unnatural under other explanations now receive straight-forward phonetic motivation: rates of voicing in fact go back to rates of fricativization. The asymmetries in voicing across places of articulation are also explained under the new explanation: fricativization targets only alveolars and velars. Evidence in favor of an interstage with fricatives is observed in synchronic acoustic analysis. Finally, all sound changes in the Blurring Chain explanation are phonetically motivated and typologically frequent.

Page intentionally left blank.

Chapter 4

Bootstrapping Sound Changes

4.1 Historical Probabilities of Alternations

The aim of this chapter is to propose a probabilistic model of typology within the CB approach. The standard explanation offered for typology within CB is that "common sound patterns often reflect common instances of sound change" (Blevins 2013:485, also Greenberg 1978:75–6). In other words, the more common a sound change is, the more common the synchronic alternation it will produce. However, by assuming only this factor, we face the crucial problem: if sound change has to be natural, why do we nevertheless see unnatural alternations? Of course, we can assume that unnatural alternations arise through combinations of sound changes (as is argued for in Chapter 3). Under this hypothesis, however, we are faced with the problem that was raised by Kiparsky (2006, 2008) and de Lacy and Kingston (2013): if any combination of sound changes is possible, why are some patterns very common, others less common, and even more importantly, some non-existent.

Below, I propose a model that crucially relies on the distinction between a single sound change and a combination of sound changes, the Blurring Process, and MSCR, and adds a probabilistic dimension to the derivation of typology. I argue that this approach derives surface typology more accurately than the proposals entertained so far: the new model not only explains why natural alternations are the most frequent, unmotivated alternations less frequent, and unnatural alternations the least frequent, but also estimates historical probabilities of individual synchronic alternations, thus quantifying the Channel Bias contribution to phonological typology.

The Blurring Process model allows us to maintain the long-held position that a single sound change is always phonetically motivated (natural) and that a single sound change cannot operate against a UPT. The Blurring Process, however, also allows derivation of unnatural alternations through a combination of natural sound changes. In other words, natural alternations are phonologized instances of at least one sound change. Unmotivated alternations are phonologized combinations of at least two sound changes. Unnatural alternations are phonologized combinations of minimally three sound changes. Crucially, the number of sound changes required for a process to arise determines that process's relative frequency: all else being equal, the probability of a single sound change occurring will be greater than the probability of two or three particular sound changes occurring in sequence, which translates into a scale of probability in which natural alternations are the most likely to occur, followed by unmotivated changes, and then unnatural alternations. Even if we do not assume that a sound change is strictly minimal, but can in some instances involve changes of more than one feature simultaneously (Section 2.1.2), typology in (20) still holds: even if we admit simultaneous sound changes of multiple features in the set of possible sound changes, they are still considerably less frequent than sound changes that target only one feature (see Section 2.1.2). The overall probability of an unnatural alternation will thus nevertheless be smaller than the probability of an unmotivated alternation, all else being equal.

(20) *A scale of decreased probabilities*

$$P_\chi(\text{natural}) < P_\chi(\text{unmotivated}) < P_\chi(\text{unnatural})$$

Their low probability is not the only reason why unnatural processes are rare. Crucially, as soon as an unnatural process operating against a universal phonetic tendency does arise and becomes fully and productively incorporated into the synchronic phonological grammar (B → A / X), the inverse universal phonetic tendency (A > B / X) will begin operating against it. As a result, the probability that an unnatural alternation will survive is even further reduced by the fact that a common sound change (and universal passive phonetic tendency) operates progressively against its existence. This erosion is precisely what we see happening in Tswana: in a system with an unnatural alternation (PND), a single instance of natural and opposite sound change (PNV) is in the process of operating against the unnatural alternation (see section 2.3.2 above).

One could argue that, even though combinations of sound changes are less frequent than any single instance of sound change, over the course of an almost unlimited timespan, sound changes ought to "stack up," yielding multiple unnatural alternations in any given language. In other words, given that every language has a several-thousand-year history during which sound changes have occurred continuously, we should perhaps expect many more unnatural alternations than are actually attested. Consider, however, that any given sound change has a limited time of operation. In other words, a sound change becomes active at one point in time and ceases to operate at another point in time. This is primarily evident from the fact that, at some point in any language, certain sound changes cease to apply to novel vocabulary, loanwords, and morphological alternations. For an unnatural phonological alternation to arise, all the sound changes that play a part in the Blurring Process must be active as ordered: a sound change in the Blurring Process must be active and exceptionless, but can cease to operate by the time of the following sound change, provided that the substantial novel/loan vocabulary or morphological contexts are not introduced that would fail to undergo the change.[72] Thus, the timespan available to produce such unnatural alternations is not unlimited, but rather limited by the time in which single sound changes that combine to yield the alternation in question are affect a sufficient proportion of lexical items and morphological alternations in a given language such that the resulting alternation is productive. In probabilistic terms, we would say that language history is not a pure-birth process, but rather a birth-death process.

The idea that unusual rules are rare because they require complex history is not new. The low probability of combinations of changes has been previously relied on to account for the rarity of certain morphological processes, and some attempts have been made to use this reasoning in phonology as well (Bell 1970, 1971, Greenberg 1978:75–6, Cathcart 2015, Morley 2015). Blevins (2004:310) briefly mentions that the rarity of certain morphological processes (such as tense marking on pronouns in Gurnu) might be explained by the low probability of co-occurrence of the factors that led to this system. This approach is employed more thoroughly by Harris (2005, 2008), who notes: "the more changes are involved, the less likely all will happen to co-occur" and "it is an idea that is necessary to discuss, because the role of probability has not been included in previous discussions of rare phenomena." However, the proposals often do not go much further than stating the generalization that combinations of changes produce rare patterns. To my knowledge, none of the proposals so far actually explain the mechanism for why the least frequent processes are also phonetically unnatural (and less frequent phonetically

---

[72] The first sound Change in the Blurring cycle in fact has to be inactive at the time of the second sound change for the unnatural process to arise.

unmotivated): a generalization that follows automatically from the formal proof that unnatural alternations require at least three sound changes (MSCR).

Two models attempt to quantify probabilities of occurrence of various, primarily static phonotactic processes and explain the relative rarity of some processes. 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 the other and rest probabilities of each state, and is as such most suitable for modeling probabilities of various phonotactic restrictions. A probability of each state is determined by the number of previous states it can arise from and transitional probabilities between states. They propose a Markov Chain Model for determining probabilities of each state. Their modeling of probabilities of transitions (processes) in the one instantiation of the model in Bell (1971), however, involves relative probabilities that only tangentially reflect frequencies of processes in samples. The most elaborate model of calculating probabilities of combination of sound changes is offered by Cathcart (2015), who calculates combinations of sound changes that lead to a certain process (in this case, final voicing) and compares that to combinations of all sound changes in a given survey to get an estimate of the probability of certain processes. The models in Greenberg (1978) and Cathcart (2015), however, do not take into consideration the crucial distinctions made in this dissertation: 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. This means the models need to rely on representativeness of diachronic surveys for all sound changes, not only for the ones that are estimated (see also Section 4.2.1). Cathcart's (2015) model is also computationally difficult to implement. The model in Bell (1971), on the other hand, is not elaborate enough to provide applicable results.

Finally, the model of automated reconstruction in Bouchard-Côté et al. (2013) estimates 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, Hruschka et al. 2015) do not directly deal with estimating probabilities of sound changes that operate in combination, but computationally model initiation and propagation of single sound changes. I show that MSCR and the new division of naturalness facilitate the development of a quantifiable model of typology within CB and point to the novel predictions that the proposed model brings.

We saw that the MSCR predicts natural processes will be the most frequent, unmotivated less frequent, and unnatural the least frequent (20). MSCR, however, only predicts *categorical* relations between alternations with different degrees of naturalness. Our goal is to propose a model that would quantify probabilities of natural, unmotivated, and unnatural processes further. We can combine MSCR with the assumption that frequencies of sound changes influence the frequencies of synchronic alternations. Crucially, the probability that an alternation arises depends on the number of sound changes it requires to arise and the probability of each individual sound change in the combination. I call such probabilities *Historical Probabilities of Alternations* (P_χ).

(21) *Historical Probabilities of Alternations* (P_χ)

The probability that an alternation arises based on the number of sound changes required (MSCR) and their respective probabilities that can be estimated from samples of sound changes.

In other words, Historical Probabilities are estimated based on two diachronic (CB) factors: the number of sound changes required for an alternation and the respective probabilities of those sound changes, estimated from samples of sound changes. The concept of Historical Probabilities of Alternations thus provides a tool for estimating the Channel Bias contribution to phonological typology for any given alternation. Its applications are discussed in the rest of this chapter.

4.2 The model

As defined in (21), the Historical Probability of an alternation A_k is the probability that a language L features A_k based on the number of sound changes (S_i) the alternation A_k requires and their respective probabilities.

4.2.1 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 occurrence 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.

The most elaborate survey of sound changes currently available based on which I perform the BSC analysis 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 well-reconstructed prehistory and well-established subgrouping. This allows the survey to properly code the occurrence of a sound change, where 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 or 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 sound change as well as from higher level evidence (e.g. morphology). Additionally, phylogenetic tree analysis does not restrict the direction of sound change and would crucially analyze 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 as it might be 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 my analysis are likely not crucially affected by the fact that most 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 (see also Section 4.3.4).

The only other comparable survey of sound changes known to me 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 other quantitative approaches to estimating probabilities of sound changes or their combinations, especially over the one proposed in Cathcart (2015). The requirement for samples to be representative 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. Because identification of historical trajectories that lead to an alternation is performed manually in my model, surveys of sound changes that I use 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. Additionally, Cathcart (2015) uses the UniDia database for his model, which is less appropriate compared to Kümmel's (2007) survey, primarily because of its encoding of sound changes, which lacks subgrouping.

4.2.2 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 in Efron (1979) and has seen a wide range of applications ever since (Davison and Hinkley 1997).

The 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 (24) or when we estimate differences between two Historical Probabilities (see Section 4.3.2 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 Historical Probabilities of two alternations, even when the statistic of interest is complex (as will be shown below).

The computation of BSC is implemented in the R Statistical Software (R Core Team 2016) with the *boot* package (Canty and Ripley 2016, Davison and Hinkley 1997) using functions *boot()* and *boot.ci()*. This dissertation also presents R code that implements the BSC technique and introduces functions *bsc()*, *summary.bsc()*, *bsc2()*, *summary.bsc2()*, *plot.bsc()* and *plot.bsc2()* (based on the *boot* package) that facilitate the estimation of Historical Probabilities with BSC (available in Appendix A). The functions allow estimation of Historical Probabilities directly from a vector of counts and should be easy to use even for researchers without substantial statistical knowledge. The aim of the code is to provide an 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.

4.2.2.1 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 (22). If an alternation A_k requires only one sound change to arise and invariably occurs as a result on that change (i.e. A_k is natural), then we estimate its P_χ according to (22).

(22)

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

The BSC samples with replacement from the sample of successes and failures and calculates the statistic of interest: in our case, the probability according to (22). 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 I compute standard error, bias, and 95% adjusted bootstrap (BC_a) confidence intervals that adjust for bias and skewness (Efron 1979, 1987).

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 with no predictors. 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.

4.2.2.2 Two or more sound changes

If an alternation A_k requires more than a single sound change, 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; see (23).

(23)

$$P_\chi(A_k) = P_\chi(T_1 \cup T_2 \cup T_3 \cup \ldots \cup T_n)$$

A trajectory T_j denotes a combination of sound changes that yields an alternation A_k. In theory, there is 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 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 a 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; see (24).

(24)

$$P_\chi(T_j) = \frac{P_\chi(S_1 \cap S_2 \cap S_3 \cap \ldots \cap S_n)}{n!}$$

Estimating the joint probability of individual sound changes ($P_\chi(S_1 \cap S_2 \cap S_3 \cap \ldots \cap S_n)$) is not a trivial task. We need to make a number of assumption 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, I treat sound changes as independent events under the BSC. For a discussion of assumptions of the BSC model, see Section 4.2.3 below.

As defined in (21), 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 number of failures (languages without S_1) according to (22), regardless of the phonemic inventories of languages in the sample.

For example, if the target of the first sound change S_1 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 the probability that the alternation A_k arises in a language L (regardless of whether it features stop geminates) and *not* the probability that the alternation A_k arises in a language L that features geminate stops.

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 $P_\chi(S_2)$ from the number of successes (languages with S_2) divided 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 and estimate P_χ from a product of probabilities of individual sound changes:

(25)

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

To estimate standard errors or 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 number of individual binomial samples (one sample for each individual sound change, constructed as described above), computes the Historical Probability of each sound change (according to (22)), and then computes the product of Historical Probabilities of each individual sound change, divided by $n!$ according to (25). This process returns 10,000 bootstrap replicates of the Historical Probability of A$_k$, based on which standard errors and BC$_a$ confidence intervals are computed.

4.2.2.3 Comparison

BSC not only estimates Historical Probabilities of individual alternations, but also allows for the estimation of the difference between the Historical Probabilities of two alternations.

(26)

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

The difference between the Historical Probabilities of two alternations (ΔP_χ) is estimated with a stratified non-parametric bootstrap where P$_\chi$ of each individual alternation A$_1$ and A$_2$ is estimated as described in Sections 4.2.2.1 and 4.2.2.2 (depending on whether A$_1$ and A$_2$ require trajectories that require one or more sound changes). Then using BSC on the difference additionally calculates the difference between $P_\chi(A_1)$ and $P_\chi(A_2)$, which returns 10,000 bootstrap replicates, based on which standard errors and BC_a confidence intervals are computed.

The BSC technique applied on a difference between two alternations enables 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 $\alpha = 0.05$.

4.2.3 Assumptions

The model presented in Section 4.2 makes some crucial assumptions that are discussed in this section. In order to estimate the joint probability of two or more sound changes as a product of Historical Probabilities (see (25)), the model assumes that each sound change is an independent event. This is not a controversial assumption for a diachronic model: there is no reason to believe occurrence of one sound change affects the probability of the following sound changes, unless the first sound change crucially alters phonemic inventory of the language in question. The BSC model does, however, account for at least some of the dependency between sound changes and phonemic inventories by estimating probabilities from samples conditioned on the result of previous sound change and by evaluating sound changes according to the target, result, and context.

While probabilities of individual sound changes in the Blurring Process are expected to be independent of each other under the CB approach, it is possible that they are not independent due to learning (AB) factors. In fact, I will argue in Section 6.5 that probabilities of sound changes can indeed be “catalyzed” by learning if the resulting alternation is less complex. However, the historical model proposed here is designed to model only CB contribution to the typology, which means that the assumption of independence is desired for this purpose.

What is not accounted for in the model are the functional load of individual phonemes and the dependency of sound changes on broader phonemic inventories that do not immediately affect the target, result, or context of the sound changes in question. Broader phonemic inventories can influence probabilities of sound changes, especially for vocalic changes (due to the effects described in the Theory of Adaptive Dispersion, see Liljencrants and Lindblom 1972, Lindblom 1990). BSC also does not model other factors that could potentially influence probabilities of sound changes, such as language contact or sociolinguistic factors, and makes no assumptions about how sound change is initiated or spread.

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 assumptions. First, the probabilities in BSC are estimated 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 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 in its daughter languages. This is a closer approximation to reality than to assume that 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 which is substantiated by empirical evidence: sound change operates and then ceases to operate, at which point it can occur again (e.g. on novel morphological or loanword material). These assumptions about the temporal dimension are not limited to our model: any model of sound change probability will be faced with this problem because of the lack of more comprehensive surveys that include temporal information. The assumptions in the BSC model approximate the reality reasonably well and, to my knowledge, better than other proposals.

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: a 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, I do not model the influence of the potential natural sound change because the Historical Probabilities of the natural sound changes (Table 4.5) are relatively similar for the processes I estimate and we do not expect this additional factor to alter the results significantly. For other processes, 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 our proposal are at least partially accounted for by the fact that the sample size in our case 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, we can disregard these influences, first because the effects are likely minor enough not to crucially alter our results, and second because current typological surveys do not allow for models that would account for these minor influences. In addition, the BSC technique estimates Historical Probabilities of alternations in a language L, where L represents a language that has the characteristics of the majority of languages in our sample. We do not condition Historical Probability on its phonemic inventory, functional load of phonemes, or other factors, which is why we can disregard these other factors for practical purposes.

4.3 Applications

4.3.1 Estimation of Historical Probabilities

The BSC technique enables estimation of Historical Probabilities for any synchronic alternation. We can estimate Historical Probabilities of natural, unmotivated, and unnatural alternations, both attested and unattested (according to Section 4.2). For the purpose of illustrating the method, I estimate the Historical Probabilities (P_χ) of the natural alternations post-nasal voicing (PNV), final devoicing (FD), and intervocalic voicing (IVV), and their unnatural counterparts post-nasal devoicing (PND), final voicing (FV), and intervocalic devoicing (IVD). These processes have received a substantial amount of attention in phonological literature, show different degrees of historical and synchronic attestedness, and are typologically, phonetically, and experimentally well researched (see Section 2.2).

To construct samples of sound changes for these six alternations, I use the survey of consonantal sound changes in Kümmel (2007). The three natural alternations have the obvious origins: the single natural sound changes PNV, IVV, and FD, respectively. For the unnatural alternations, I first identify sound changes in the Blurring Process (Chapter 3) that yield the alternation in question. If A > B / X is a natural sound change, B > A / X is unnatural. (27), (28), and (29) 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 Chapter 3) and identify the actual sound changes that yield the unnatural alternation (right column).

Chapter 3 demonstrates that PND results from the Blurring Cycle. A combination of the following three natural and well-motivated sound changes yield PND: fricativization of voiced stops that operates in non-post-nasal position, unconditioned devoicing of voiced stops, and occlusion of voiced fricatives to stops. (27) illustrates the development.[73]

	Blurring Cycle — schematic	*PND*
(27)	B > C / ¬X	D > Z / [−nas]__
	B > A	D > T
	C > B	Z > D
	B > A / X	D > T / [+nas]__

I also argued in Chapter 3 that IVD results from the Blurring Chain. Voiced stops fricativize intervocalically, voiced fricatives devoice, and voiceless fricatives get occluded to stops (see (28)). The result is the unnatural intervocalic devoicing (D > T / V__V).

	Blurring Chain — schematic	*IVD*
(28)	B > C / X	D > Z / V__V
	C > D	Z > S
	D > A	S > T
	B > A / X	D > T / V__V

---

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

FV is arguably unattested both as a synchronic alternation as well as a sound change (Kiparsky 2006, cf. Yu 2004). A number of diachronic scenarios, however, exist 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 rather in a static phonotactic restriction. One possible scenario that would result in FV is Scenario 1[74] in Kiparsky (2006) which I use here for estimating the Historical Probability of FV (Kiparsky 2006).[75] 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 (29)).

	Modified Blurring Cycle — schematic	*FV*
(29)	C > B / X	Tː > T / __#
	B > A	T > D / V__
	C > B	Tː > T
	B > A / X	T > D / __#

Based on the trajectories identified here that result in natural and unnatural alternations, I perform counts of sound changes and languages surveyed from Kümmel (2007). Sound changes 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,

---

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

[75] For a discussion on Scenario 2, see Section 3.4.

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 we count only contexts that strictly require intervocalic (as opposed to post-vocalic) context. FD is reported for 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 4.1 summarizes counts of languages with sound changes that result in natural alternations.

Table 4.1: 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, I perform counts for each individual sound change in the Blurring Process. Fricativization of voiced stops is reported in approximately 97 languages. I include instances of intervocalic and post-vocalic fricativization in the count as well (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.[76] I estimate the probability of the first sound change in the Blurring Cycle that results in PND based on the number of successes (languages in the survey with that sound change) and the total number of language surveyed (294) without conditioning on the sample. The sample for estimating the probability of the first sound changes is unconditioned, because the Historical Probability of A_k is the probability that A_k arises in a language L, regardless of properties of its phonemic inventory (see Section 4.2.2.2). Once the first sound change operates, however, we know that the language in question needs to have voiced stops in its inventory. I therefore estimate the Historical Probability of the second sound change that targets voiced stops 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 which are the result of chain shifts). Approximately 31 languages lack voiced stops in the survey in Kümmel (2007),[77] which means that I estimate P_χ based on $294 - 31 = 263$ languages surveyed. After the two sound changes operate, we also know that the language L has voiced fricatives. I estimate P_χ of the last sound change 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-

---

[76] 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.

[77] 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 PHOIBLE database (Moran et al. 2014), approximately 30% of inventories lack a phonemic labial voiced stop. For consistency purposes, I stay within the Kümmel's (2007) survey with this acknowledgement.

strident fricatives,[78] according to Kümmel (2007). In approximately 27 languages occlusion of fricatives is reported as a sound change.

The first sound change in the Blurring Chain that yields IVD is fricativization of voiced stops post- or intervocalically, which is attested in approximately 83 languages. Because fricativization of voiced stops is the first sound change in the combination, I estimate its probability based on the total number of languages in the survey. The second sound change, unconditioned devoicing of voiced fricatives, is attested in approximately 7 languages out of approximately 216 languages that have voiced fricatives in their inventories. Finally, occlusion of voiceless fricatives to stops is reported in approximately 34 languages out of 248 languages with non-strident voiceless fricatives in their phonemic inventories.

To estimate the Historical Probability of FV, I take the one scenario from Kiparsky (2006) that would result in FV as an alternation. For a discussion on why I exclude other scenarios that would result in phonotactic restrictions, see below and Section 3.4.[79] Counts of the sound changes that lead to FV as an alternation are the following: in approximately 6 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, I estimate its Historical Probability from the total number of languages surveyed. The second sound change, post-vocalic voicing of voiceless stops,

---

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

[79] I currently also exclude the scenario that potentially results in FV in Lakota: 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 4.3: $P_\chi = 0.003\%$ [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 4.3, I do not expect the absence of this scenario in our model to alter the result substantially.

Table 4.2: 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	Tː > T / __#	6	294
	T > D / V__	32	294
	Tː > T	27	≈88

is reported in approximately 32 languages. Because all languages have voiceless stops, I include all 294 languages surveyed in 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, I use Greenberg's (1965) survey of consonantal clusters and Ryan's (to appear) survey of phonemic geminates. At least 30% of languages in Greenberg's (1965) survey of approximately 100 languages allow stop + stop final clusters. I assume that 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 sequences of stops or from the proportion of languages that allow phonemic geminates. 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. I thus estimate the number at 88 (30% of 294 languages). That my 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.

To compute the estimates of Historical Probabilities I use the *bsc()* function (see A.1) that transforms two vectors of length n (number of sound changes), where the first vector includes counts of languages in a sample with sound changes in a given trajectory and the second vector includes counts of languages surveyed for each sound change, into a series of successes and failures from which bootstrap replicates are sampled. Based on the *boot* package (Canty and Ripley 2016, Davison and Hinkley 1997), the *bsc()* function performs bootstrapping for the statistic in (22) (if trajectory T_j requires a single sound change) or for the statistic in (25) (if trajectory T_j requires more than one sound change) and returns 10,000 bootstrap replicates. The *summary.bsc()* function computes 95% BC_a confidence intervals based on the bootstrap replicates (using the *boot.ci()* function from the *boot* package; Canty and Ripley 2016, Davison and Hinkley 1997). Table 4.3 shows the Historical Probabilities with estimated 95% BC_a confidence intervals for the six natural and unnatural alternations discussed above. Figure 4.1 shows distributions of bootstrap replicates for the Historical Probabilities (P_χ) of these natural and unnatural alternations. Table 4.3 and Figure 4.1 illustrate a 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 the respective natural alternations (PNV, IVV, and FD).

The Historical Probabilities and confidence intervals of the natural alternations PNV, IVV, and FD could also be estimated analytically (see Section 4.2). To illustrate the accuracy of the BSC technique, I compare the 95% BC_a bootstrap confidence intervals with confidence intervals computed with an analytic solution (see Section 4.2.2.1). 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 4.3 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 the BSC model estimates CIs with high accuracy.

Table 4.3: Estimated P_χ (in %) for natural and unnatural alternations with 95% BC_a and 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		—

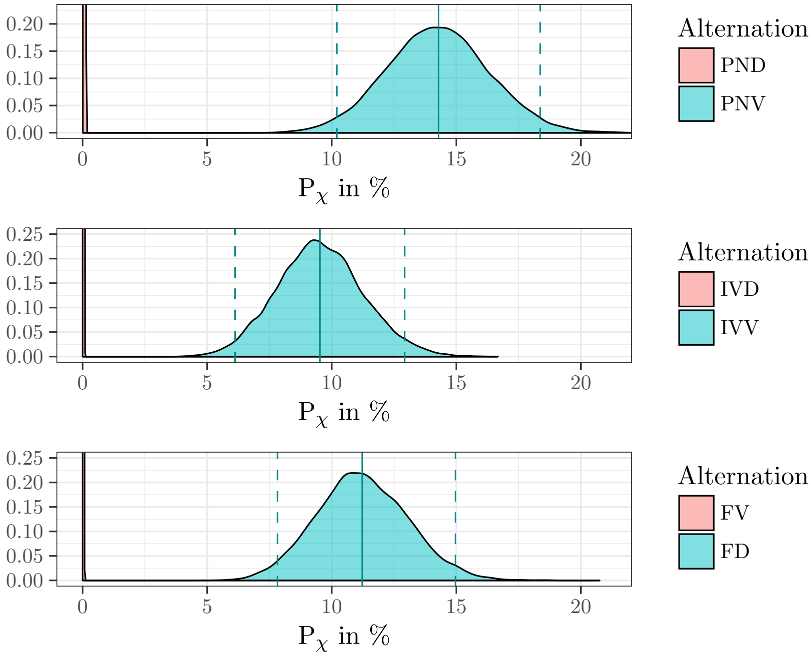


Figure 4.1: Bootstrap replicates for natural and unnatural alternations. The plots also show the observed P_χ (solid line) and 95% BC_a CI (dashed line). The distribution of bootstrapped P_χ for unnatural alternations does not feature confidence intervals because the probabilities are to small to be visible. For the purpose of representation, the vast majority of bootstrap replicates for unnatural alternations fall outside the limits of the plot.

4.3.2 Comparison of alternations

One of the advantages of the BSC method is that we can perform inferential statistics on comparisons between the Historical Probabilities of any two alternations. In other words, BSC allows us to test whether the Historical Probabilities of two alternations are significantly different. I estimate the difference between the Historical Probabilities of two alternations (ΔP_χ) by stratified non-parametric bootstrap. Samples of sound changes for estimating differences in Historical Probabilities are the same as the samples for estimating Historical Probabilities of individual alternations. BSC samples with replacement and computes the statistic in (26): the difference between two Historical Probabilities. Each individual P_χ is computed as outlined in Section 4.2. If the 95% BC_a confidence intervals for $\Delta\mathrm{P}_\chi$ of two alternations fall either below or above 0, the Historical Probabilities of the two alternations are significantly different (with $\alpha = 0.05$). Functions that perform this computation are *bsc2()* (performs stratified non-parametric bootstrap based on the *boot()* function; see A.2) and *summary.bsc2()* (performs computation of confidence intervals based on the *boot.ci()* function; see A.4).

Figure 4.2 shows bootstrap replicates of individual Historical Probabilities of 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, I can test, for example, whether $\mathrm{P}_\chi(\mathrm{PND})$ and $\mathrm{P}_\chi(\mathrm{IVD})$ or $\mathrm{P}_\chi(\mathrm{PND})$ and $\mathrm{P}_\chi(\mathrm{FV})$ are significantly different. I estimate the following $\Delta P_\chi(\mathrm{PND}, \mathrm{IVD})$ and $\Delta P_\chi(\mathrm{PND}, \mathrm{FV})$ as described above and in Section 4.2.

(30) a. $\Delta P_\chi(\mathrm{PND}, \mathrm{IVD}) = P_\chi(\mathrm{PND}) - P_\chi(\mathrm{IVD}) = 0.026\%\ [-0.004\%, 0.064\%]$

b. $\Delta P_\chi(\mathrm{PND}, \mathrm{FV}) = P_\chi(\mathrm{PND}) - P_\chi(\mathrm{FV}) = 0.036\%\ [0.011\%, 0.074\%]$

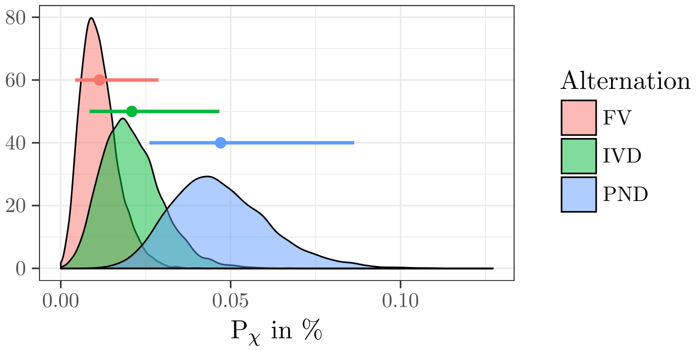


Figure 4.2: Bootstrap replicates for unnatural alternations with observed P_χ (colored dot) and 95% BC_a confidence intervals (colored lines).

Because the 95% BC_a CIs of the difference in Historical Probability between PND and FV lie above zero, I can claim 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: the 95% BC_a CIs cross zero. Like any method, the BSC technique has confounds (see Sections 4.2.1, 4.2.3, and 4.3.1). Because differences in the Historical Probabilities of unnatural alternations are considerably smaller than differences between natural-unnatural pairs, these confounds might influence the results more substantially. For example, some single sound changes might be counted as two independent events (see Section 4.3.1), the surveys of sound changes are potentially unrepresentative, and the temporal dimension is not modeled with BSC. Until more comprehensive surveys are available, however, the BSC technique makes, to my knowledge, the most accurate approximations of Historical Probabilities of alternations.

The inferential statements and predictions of the BSC method from the paragraph above can be compared to the observed typology. Kümmel (2007) reports that PND is attested once as a sound change — in Yaghnobi. As was argued in Chapter 3, PND in Yaghnobi in fact results from precisely the three sound changes in the Blurring Cycle (27) that I model with BSC. In other words, a combination of sound changes that lead to PND is attested once in

Kümmel's (2007) survey. While FV is also reported once in Kümmel (2007), i.e. an alleged *t surfaces as [d] word-finally in Latin, the reported case of FV targets only one series of stops and it is questionable if this sound change was indeed phonetically realized as final voicing (see Section 2.2.4). Kümmel's (2007) survey also reports a sound change that almost qualifies as IVD: South Italian *d is reported to change to [t] between two unstressed vowels. The scope of this sound change is, however, prosodically limited to unstressed vowels and its phonetic status is unconfirmed. PND as a diachronic development that results from the Blurring Process seems to be most securely attested in Kümmel's (2007) survey, which is in line with the predictions of the BSC technique.

Looking at the typology more broadly, BSC makes some further desirable predictions. Based on a survey in Chapter 2 that aims to collect all reported cases of PND, the Blurring Process that leads to PND is attested in 13 languages. In at least two languages the Blurring Process results in a productive synchronic alternation. IVD is reported in three languages as a result of the Blurring Process and once as a gradient phonotactic restriction (according to the survey in Chapter 2), but never as a productive synchronic alternation. Finally, FV is, to the my knowledge, never reported as a synchronic phonological alternation (for reasons why Lezgian and other cases are not analyzed as featuring FV, see Kiparsky 2006 and Section 2.2.4), and only in doubtful cases as a sound change. The typology thus suggests that PND is indeed the most frequent and FV the least frequent, just as predicted by BSC. For evaluation of further typological predictions of the BSC technique, see Sections 4.3.4 and 4.4.

As already mentioned, for the purposes of diachronic modeling I distinguish alternations from static phonotactics (Section 3.4). 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 less 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 processes as phonotactic restrictions. A further study is required to answer this question. A preliminary survey, however, reveals that FV as a static phonotactic restriction might not be as rare. Three languages might qualify to feature this process: Ho, some varieties of Spanish, and Tundra Nenets (see Section 3.4).

4.3.3 Prediction of attestedness

In addition to comparing Historical Probabilities of alternations and performing inference on the comparison, the BSC technique allows for prediction of attestedness in a given sample. We can estimate the difference between the Historical Probability of an alternation A_k and the probability of being attested once in a given sample (with N languages surveyed). If the estimated 95% BC_a confidence intervals fall below or above zero, we can claim that the Historical Probability of A_k is significantly higher or lower than the probability of being attested once in a given sample.

(31)

$$\Delta P_{\chi}(\mathrm{A}_k, \frac{1}{\mathrm{N}}) = P_{\chi}(\mathrm{A}_\mathrm{k}) - P(\frac{1}{\mathrm{N}})$$

The procedure for estimating the difference between the Historical Probability of an alternation A_k and the probability of being attested once in a sample is exactly the same as the procedure for estimating differences in the Historical Probabilities of two alternations (described in 4.3.2), except that the probability of being attested once in a sample is estimated from a sample of one success and $N-1$ failures. Estimates for each of the six natural/unnatural alternations are computed with functions *bsc2()* and *summary.bsc2()* and are summarized in Table 4.4. None of the three unnatural alternations differ significantly from the probability of being attested once in a sample of 294 languages. Historical Probabilities of all three natural alternations, on the other hand, are significantly higher than the probability of being attested once in the sample (Table 4.4).

Table 4.4: Estimated ΔP_{χ} (in %) for natural and unnatural alternations with 95% BC_a confidence intervals.

		95% BC_a CI		
A_k	**ΔP_{χ}**	**Lower**	**Upper**	
PNV	13.9	10.4	18.6	*
PND	−0.3	−1.7	0.05	
IVV	9.2	6.1	13.1	*
IVD	−0.3	−1.8	0.03	
FD	10.9	7.6	15.1	*
FV	−0.3	−2.0	0.01	

We can interpret these result as follows: because the Historical Probabilities of the natural alternations are significantly higher than the probability of being attested once in a sample of N languages, we expect the natural alternations to be attested more than once in a given sample of comparable size.

For the unnatural alternations, the interpretation requires further clarifications. Let us first compare the estimated 95% BC_a confidence intervals and the raw probability of being attested once in a sample. The 95% BC_a confidence intervals of all three unnatural alternations fall below the raw probability ($P(\frac{1}{294}) = 0.34\%$). The 95% BC_a CI of PND that reaches the highest probability is estimated at [0.02%, 0.11%], which is well below the raw probability of being attested once in a sample (0.34%). This observation would predict that none of the natural alternations would be attested in a given sample. The probability of being attested once in a sample, however, also bears a degree of uncertainty. Instead of comparing the 95% BC_a CIs of unnatural alternations with the raw probability, I estimate the 95% BC_a CI (based on BSC) of the difference between the probability of being attested once in a sample ($\mathrm{P}(\frac{1}{294})$) and the Historical Probability of an alternation A_k (as proposed in (31) above). As already mentioned, none of the Historical Probabilities of the unnatural alternations in question significantly differ from the probability of being attested once in a given sample (Table 4.4). We can interpret this result as follows: BSC predicts that unnatural PND, IVD, and FV might or might not be attested in a given sample.

Predicting attestedness with BSC reveals another important generalization about the derivation of phonological typology within the Channel Bias approach and about typological surveys in general. When estimating the probability of being attested once in a given sample with a non-parametric bootstrap, we sample randomly with replacement from a sample that includes only one success and many more failures. This means that many bootstrap replicates (random samples with replacement) will have probability zero. For example, in one bootstrapping draw with 10,000 replicates, 3,676 replicates (or 36.8%) yielded $\mathrm{P} = 0$ in estimating $\mathrm{P}(\frac{1}{294})$. For this reason, regardless of how low the Historical Probability of an alternation A_k is, it will likely not be significantly lower than the probability of being attested once in a given sample. This generalization persists even if we estimate the difference between the two probabilities with a smoothed bootstrap. Smoothed bootstrapping is a bootstrapping technique that samples from a "kernel density estimate of the distribution" instead of sampling from the observed distribution (Wolodzko 2017). For illustrative purposes, I estimate the difference between the Historical Probability of a hypothetical alternation that requires three sound changes to arise, each of which is attested only once in a given sample, and the probability of being attested once in that sample. Let us assume our hypothetical sample size is 294. I estimate the difference using a Gaussian kernel in the *kernelboot* package (Wolodzko 2017). The difference between the two probabilities is not significant: $\Delta P_\chi = -0.33\%$ with 95% quantiles crossing zero ($[-0.01\%, 0.27\%]$).

This observation points to a problem that typological generalizations face: virtually no alternation will have Historical Probability low enough that its P_χ will be significantly lower than the probability of being attested once in a given sample. In other words, we cannot claim with great confidence for any alternation that it will not be attested in a given sample. BSC thus either predicts that some alternations will be attested in a given sample (those with P_χ higher than $P(\frac{1}{N})$) or that some alternations might or might not be attested (those for which P_χ and $P(\frac{1}{N})$ do not differ significantly). No alternation is predicted to be unattested in a given sample by the BSC. This last generalization is in fact desirable: languages have productive synchronic alternations that require multiple sound changes, many of which are not common. One such example is Sardinian lateral sandhi (/l/ → [ʁ] / V__V; Scheer 2015) that requires at least five sound changes (l > *ɫ > w > ɡʷ > ɣʷ > ʁ / V__V; see Scheer 2015 and literature therein).[80]

4.3.4 Comparing P_χ to observed synchronic typology

Predictions of the BSC model can be evaluated by comparing Historical Probabilities with observed typology of synchronic alternations. Estimation of synchronic typological probabilities, however, faces many more difficulties and problematic assumptions than estimation of Historical Probabilities. 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

---

[80] The trajectory of the development of Sardinian sandhi is confirmed by related dialects that feature intermediate stages (Scheer 2015 and literature therein).

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, comparison between 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 my knowledge, better than alternative approaches (see Section 4.1). Table 4.5 compares Historical and observed synchronic probabilities. Historical Probabilities (P_χ) are estimated with the BSC technique as described above (see Section 4.3.1 and Table 4.3). The synchronic typology of natural processes is estimated based on surveys of alternations in Locke (1983) (reported in Hayes and Stivers 2000) and Gurevich (2004) (reported in Kaplan 2010). Because no systematic typologies of FD exist, this process is left out of the comparison. The synchronic typology of unnatural alternations is based on surveys in Chapter 2. The synchronic typology of unnatural processes is challenging to estimate, because it is difficult to confirm a productive synchronic status for an unnatural alternation compared to natural alternations, and because typological surveys of unnatural alternations are usually not performed in a systematic and controlled manner. The survey in Chapter 2, for example, are based on all sources and reports available to me. A reasonable estimate of the languages surveyed would be approximately 600.

PND has been confirmed as a fully productive synchronic alternation in two related languages and as a morphophonological alternation in a few others. For the purpose of comparison, only fully productive alternations are counted in the synchronic typology. Because Tswana and Shekgalagari are closely related, I count PND there as a single occurrence. IVD and FV are, to my knowledge, not attested as productive phonological alternations in any language, which is why I estimate their synchronic typological probability below $P(\frac{1}{600})$.[81] Confidence intervals for typological probabilities of synchronic alternations are estimated with a non-parametric bootstrap (according to the same procedure described in Section 4.2.2.1 for estimating Historical Probabilities of alternations requiring only one sound change) from the numbers of successes (languages in a sample with a synchronic alternation) and failures (languages in a sample without the synchronic alternation).

Table 4.5: 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

---

[81] If we counted two of the best candidates for FD and IVD, Lezgian and Sula, as featuring fully productive unnatural alternations (see Yu 2004, Bloyd 2017), the typological probabilities of FV and IVD would be estimated at $P(\frac{1}{600}) = 0.17\%$.

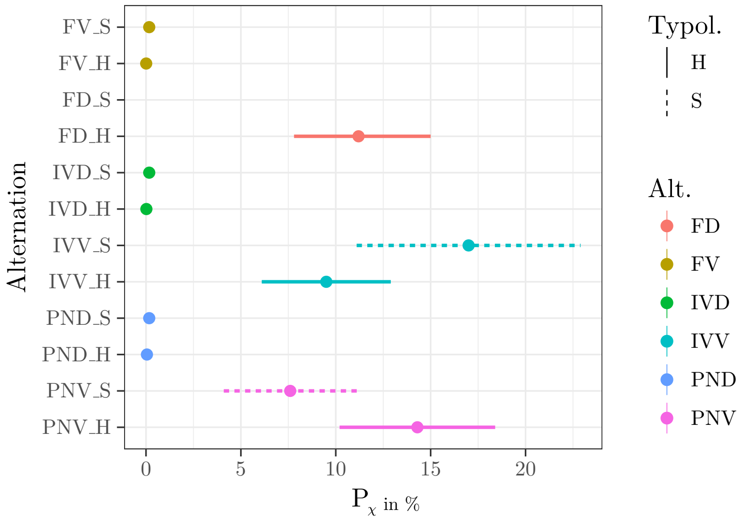


Figure 4.3: Observed Historical (H, solid line) and synchronic (S, dashed line) probabilities (in %) with 95% BC$_a$ CIs from Table 4.5 estimated with BSC.

Table 4.5 and the corresponding plot of estimated Historical and synchronic probabilities with 95% BC$_a$ CIs in Figure 4.3 show that BSC correctly predicts natural alternations to be significantly more frequent than the unnatural alternations. Historical Probabilities and observed synchronic typology also match to the degree that at least the 95% BC$_a$ confidence intervals of both Historical and synchronic typological probabilities always overlap for all five processes compared. BSC also correctly predicts that some unnatural alternations, such as PND, will be more frequent than others, such as FV, which is substantiated by the observed typology. To my knowledge, no other approach to typology (within AB or CB frameworks) makes such predictions, especially considering the fact that we estimate Historical Probabilities of unnatural alternations not directly from observed surface typology, but from the typology of natural sound changes that are independent of the unnatural result. Given the quality of current typological surveys, this is, to my knowledge, the most accurate comparison between predicted Historical and observed synchronic probabilities. The comparison thus shows that Historical Probabilities estimated with the BSC technique yield an accurate prediction of synchronic typological probabilities.

4.4 Implications

4.4.1 The AB-CB conflation problem

I argue that BSC derives several typological generalizations that have so far been considered problematic for both the CB and AB approaches. 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 or IVV, are non-existent. As was argued above, the BSC method derives these generalizations quantitatively: we predict FV and IVV to be very rare or possibly unattested in a given sample. On the other hand, the lack of evidence in favor of the Substantive Bias (see Moreton and Pater 2012a,b) has been used as a counterargument against the AB approach to typology. If alternations show no learnability differences, substantial typological differences in these alternations are not easily derivable within the AB approach.

The BSC method also offers a quantitative framework for disambiguating AB and CB influences on typology. If two typologically unequal alternations show no learnability differences, but have significantly different Historical Probabilities, we can explain differences in the observed typology between the two alternations within the CB approach. On the other hand, if two typologically unequal alternations have equal Historical Probabilities and show differences in learnability, we can explain typological differences within the AB approach. In the case of the unnatural alternations PND, IVD, and FV, the BSC technique suggests that CB primarily influences the typology (Section 4.3; for further disambiguation between AB and CB, see Section

4.4.2). Comparing the AB and CB influences according to this method in other alternations might yield different results.

Disambiguation and comparison of AB and CB influences on phonological typology face two challenges. First, further learnability experiments are needed to confirm the absence or presence of Substantive Bias, preferably for each alternation we are interested in. Opposing results have been reported from artificial grammar learning experiments on the existence of Substantive Bias, although rarely for the same alternation. It is entirely possible that Substantive Bias is present only in a subset of natural-unnatural alternation pairs. From an overview of the literature (Moreton and Pater 2012a,b) it appears that Substantive Bias has been more often reported for vocalic alternations than for consonantal alternations, but this observation requires further experimental evaluation. In any case, the AB influences on typology should ideally be estimated from artificial grammar learning experiments for every alternation in question.[82]

The second challenge facing the disambiguation of AB and CB influences is the fact that the outcomes and frequencies of sound change might be primarily influenced by learning biases themselves (cf. Kiparsky 1995, 2008). In other words, it is possible that the AB influences crucially affect the observed typology of sound changes (let us call this the *AB-CB conflation* problem). While this is not the position taken in many other works on sound change (Labov 1994) or this dissertation, the accuracy of the BSC model might be undermined by the argument stating that learnability crucially influences sound change typology. Argumentation against this

---

[82] Additionally, rather than deriving prior σ^2 from P-map-related metrics (in MaxEnt models of learning, Wilson 2006, White 2017), differences in learnability should be encoded directly from results of artificial grammar learning experiments. Relating differences in learnability from perceptual measures is problematic, especially because it has been recently shown in Greenwood (2016) that perceptual salience in experimental design can influence observed results. In other words, differences in learnability should ideally be estimated experimentally for every alternation from which typological generalizations are drawn.

position is beyond the scope of this dissertation; as will be argued below, however, the "AB-CB conflation" problem is controlled for at least in the case of unnatural alternations.

Unnatural alternations and MSCR play a crucial role in controlling for the "AB-CB conflation" problem. Even if the probabilities of individual sound changes are crucially influenced by learnability (and therefore by AB), the fact that for unmotivated and unnatural alternations to arise, at least two or three sound changes, respectively, are required to operate in a language (due to the MSCR) means that CB plays a crucial role in determining the synchronic typological probabilities of unmotivated and unnatural alternations. All else being equal, even if we assumed learnability is the only factor influencing probabilities of individual sound changes, the probability of a single sound change will necessarily be greater than the probability of a combination of three sound changes, and 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.

4.4.2 AB-CB complexity mismatch

The BSC technique proposed here identifies additional mismatches in predictions of the AB and CB approach, especially with respect to complexity of alternations and their typological attestedness. BSC not only predicts unnatural alternations will be rare (Sections 4.3 and 4.4.1), but also that, all else being equal, complex alternations will be less frequent than simple alternations. According to the "minimality principle", which states that sound change is a change in one feature in a given environment (see Section 2.1.2), 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. 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).

Natural and unnatural alternations that arise through the Blurring Process provide unique insight into the discussion on Complexity Bias and Substantive Bias as well as the discussion of the AB-CB conflation problem. There is a crucial mismatch in predictions between the AB and CB approach 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 4.6). 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, the 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 (MSCR). 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 complexity of the resulting alternation, but the third sound change decreases the complexity. Complexity Bias thus predicts that the alternations that arise from the first and the second sound change in the Blurring process will be increasingly rare, but predicts that the structurally simpler alternations resulting from the combination of all three sound changes will be more frequent than the alternation requiring only two sound changes. Let us call this prediction the *AB-CB complexity mismatch*.

We can estimate Historical Probabilities for each step in the Blurring Process that leads to unnatural alternations. Let us take as an example PND. The probability of the initial stage before the first sound change operates is calculated simply as $1-\mathrm{P}_{\chi}$, where P_{χ} is the Historical Probability of the first sound change. The Historical Probabilities of each alternation were estimated with BSC as described above.

Table 4.6: Mismatch in Historical Probabilities (P_{χ}) and probabilities predicted by complexity bias ($\mathrm{P}_{complex}$) for all four synchronic stages in a Blurring Cycle that involves three sound changes and results in a synchronic alternation PND. Historical Probabilities are estimated using BSC with 95% BC_a CI (Lo. and Up).

Sound change	**Alternation**	$\mathbf{P}_{\chi}$	**Lo.**	**Up.**	**Features**	$\mathbf{P}_{\mathbf{complex}}$	$\mathbf{P}_{\chi}$
	No alternation	67.0	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	↑	↓

The fact that the first two sound changes in the Blurring Process increase 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 operating in question are not crucially influenced by AB.

The mismatched predictions of BSC and Complexity Bias illustrated in Table 4.6 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, the 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 BSC 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, preliminary 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 BSC (and therefore CB), suggesting that Complexity Bias influences this distribution. The Historical Probability of Z → T / [+nas]__ is significantly higher than the Historical Probability of PND. The difference is estimated with BSC as $\Delta P_\chi(\mathrm{Z} \rightarrow \mathrm{T}\ /\ [+\mathrm{nas}]\_\_, \mathrm{PND}) = 1.1\%,\ [0.6\%, 1.7\%]$. The Historical Probability of the alternation Z → T / [+nas]__ that arises through two sound changes is thus predicted to be approximately 20 times more frequent than the Historical Probability of PND (see Table 4.6). 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, Nasioi, and Punu (Hyman 2001, Merrill 2014, 2016a,b, Santos 1996, Brown 2017).[83] Other languages are more difficult to classify because some of them seem to have 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.

---

[83] Punu is a language that undergoes a different development for PND from the one described in Chapter 3. For a discussion, see Hyman (2001).

What is even more intriguing is the high frequency at which the third sound change in the Blurring Process, occlusion of voiced fricatives to stops, operates on synchronic systems that have 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 is $\mathrm{P}_{\chi} = 12.5\%,\ [7.9\%, 17.1\%]$. Of the languages in the survey in Chapters 2 and 3 that undergo the first two sound changes that lead to PND, 6 languages (out of 7, or approximately 85.7%)[84] feature occlusion of stops for at least one place of articulation or in at least in one position in the word. If we count only cases in which occlusion of fricatives targets more than two places of articulation, only Tswana, Shekgalagari, Makuwa, and Murik would count (two independent changes). 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 (and therefore CB) predicts for the occlusion of voiced fricatives in general.

We can compare the Historical Probability of the occlusion of fricatives based on the surveys of sound changes in Kümmel (2007), and the observed probability of occlusion of fricatives in those languages that have already undergone the first two sound changes in the Blurring Cycle that leads to PND. Because both of these estimations involve a single sound change and because the second sample is small (7 observations), I test the significance of the difference with Fisher's Exact Test. The number of languages with voiced fricatives (216 surveyed) that undergo occlusion of voiced fricatives is 27. As already mentioned, under the less conservative count, 6

---

[84] PND in Tswana, Shekgalagari, and Makhuwa are counted as 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.

out of 7 languages in the Blurring Cycle show occlusion for at least one place of articulation or at least for one context (word-initially in Nasioi). The difference between the two counts is statistically significant ($p < 0.0001$).

This suggests that the high occurrence of the third sound change in the Blurring Process (in the case of PND, the occlusion of fricatives), might be an influence of the Complexity Bias within the AB approach. While AB likely does not influence the probabilities of the first two sound changes in the Blurring Process, because they increase complexity and therefore lower the learnability, it is likely that the occurrence of the third sound change and the therefore lower probability of the more complex unmotivated alternation is influenced precisely by Complexity Bias. In Chapter 6, I present experimental evidence in favor of this conclusion.

This example shows that mismatches in predictions between AB and CB identified by the Blurring Process, MSCR, and BSC can shed new light on the discussion of AB vs. CB influences on typology. Instead of arguing in favor of or against one or the other approach, influences of AB and CB should be estimated quantitatively. Comparison of quantitative estimates can provide new infrmation on what aspects of typology are primarily influenced by AB and what aspects primarily by CB. This chapter proposes the BSC technique, a quantitative method for estimating CB influences. For every alternation, BSC offers a technique for estimating its Historical Probability, i.e. a CB contribution to the observed typology. Estimation of AB influences should involve further experimental data: ideally, we can test the learnability of every alternation and compare it to its Historical Probability estimated with BSC. Further and more precise estimations of AB influences with experimental data and application of the BSC method to novel unnatural alternations should yield further insights into the long-standing discussion in phonology.

Chapter 5

Theoretical implications

This chapter presents two theoretical implications of the framework proposed in Chapters 2, 3, and 4 above. The existence of unnatural categorical alternations, such as PND in Tswana (Coetzee and Pretorius 2010), poses a challenge for the current approaches to OT that restrict the universal constraint inventory CON. If only natural constraints are admitted to CON or if unnatural constraints such as *ND are prohibited in CON, we predict unnatural synchronic alternations to be underivable. Undergeneration of OT with restricted CON has been addressed in several previous works on categorical OT (Hyman 2001, Hayes 1999). Less is known, however, about implications of the restricted CON within weighted-constraint-approaches to phonology. Frameworks like Harmonic Grammar or MaxEnt grammar that operate with weighted constraints rather than with strict domination have a number of advantages. Among others, weighted-constraint approaches outperform classical OT in deriving gradient or variable processes. Most approaches to gradient processes so far, however, deal exclusively with natural processes (as per definition in (2)). This chapter argues that the existence of unnatural gradient processes bears implications for the weighted-constraint-approaches to phonology: I propose a thus far largely unnoticed prediction of weighted-constraint-approaches with restricted Con that states unnatural feature values in a given environment are always predicted to be less frequent than natural ones (the so-called Natural Gradience Bias). While this is in many ways a desirable prediction, it also faces a crucial problem — undergeneration, as suggested by the existence of unnatural gradient phonotactics (Chapter 2).

One possible solution to the undergeneration problem is to admit unnatural constraints in Con. Under this hypothesis, unnatural processes are derivable, but any predictive power of OT and OT-related approaches is lost. The second part of this chapter proposes a MaxEnt-based model of typology that operates with weighted constraints and has two desirable properties: it derives unnatural processes and at the same time encodes gradient typological observations.

5.1 Natural Gradience Bias

Gradient phenomena, and more specifically, gradient phonotactic restrictions are among the most prominent arguments in favor of weighted-constraint-based models like Harmonic Grammar (HG) or MaxEnt (Coetzee and Pater 2008, 2011, Pater 2008, Albright 2009). HG and MaxEnt solve several problems of categorical OT, most notably derivation of non-categorial phonological processes. Under the classical OT approach, a candidate either wins or loses. Under the HG approach, on the other hand, each candidate is assigned a Harmony score. Various proposals exist for how to transform Harmony scores into probability distribution over candidates (such as Noisy HG, for an overview, see Hayes 2017). Perhaps the most successful model, MaxEnt framework, assigns probability to each member in the candidate set that is directly derivable from the Harmony score (via (35) and (36)). Probability distribution over output candidates is desirable when modeling non-categorical processes — observed probabilities of different variants are directly modeled with MaxEnt output probabilities. At the same time, weighted-constraint models retain many desirable predictions of the classical OT. One problem, however, remains even under the HG/MaxEnt approach: derivation of unnatural processes. I will show that HG (or for this purpose, MaxEnt models) with restricted Con requires that, in any given context, the natural value will have a probability that is at least as high as that of the unnatural value.

The classic version of OT (Prince and Smolenky 1993/2004) restricts its universal constraint inventory Con with the assumption that only a subset of possible constraints is universal and thus encodes typological asymmetries in the grammar. In the absence of unnatural constraints, a mapping from a natural feature value in the input to a natural feature value in the output in a given environment is impossible, because the unnatural feature value violates both the Faithfulness and Markedness constraint and the unnatural candidate is harmonically bound. Tableaux in Table 5.1 illustrate harmonic bounding of the unnatural mapping /ND/ → [NT].

Table 5.1: Two tableaux illustrating harmonic bounding of mapping /ND/ → [NT].

/tampa/	*NT	Ident-IO
tampa	*!	
☞ tamba		*

/tamba/	*NT	Ident-IO
☞ tamba		
tampa	*	*

In HG, categorical typological asymmetries have also been tackled by restricting Con (Jesney 2016). However, less is known about implications of HG/MaxEnt approaches for modeling gradient typology. Anttila and Magri (2018) use the MaxEnt framework to discuss probabilistic implicational universals, but focus on natural rather than unnatural processes and disregard the possibility of unnatural mappings. Some of the predictions of classic OT are directly translatable into a weighted-constraint framework with two important differences: the predictions of weighted-constraint frameworks are gradient and harmonic bounding in its classical sense does not exist in weighted-constraint-frameworks. Unnatural mappings (e.g. /ND/ → [NT]) that are ruled out by classic OT are derivable within weighted-constraint-frameworks (see Tables 5.1 and 5.2). As I will argue below, however, harmonic bounding translates into a generalization that the natural feature value be more frequent in a given environment than the unnatural one (Natural Gradience Bias).

For example, if a natural Markedness constraint has a higher weight than the Faithfulness constraint, the output candidate with the natural feature value in a given environment has a higher probability than the candidate with the unnatural feature value, for both inputs with natural and unnatural feature values (/tampa/ and /tamba/) as illustrated in Table 5.2.

Table 5.2: Tableaux with weighted constraint and MaxEnt probabilities where the Markedness constraint weighted higher than the Faithfulness constraint.

/tampa/	*NT 2	IDENT-IO 1	$\mathscr{H}$	P
tampa	−1		−2	0.27
tamba		−1	−1	0.73

/tamba/	*NT 2	IDENT-IO 1	$\mathscr{H}$	P
tampa	−1	−1	−3	0.05
tamba			0	0.95

If Faithfulness constraint has a higher weight, on the other hand, the faithful candidate receives a higher probability. For an input with the natural feature in a given environment the mapping from the natural to unnatural feature value is predicted to be less frequent than the faithful mapping, regardless of weighting. (Table 5.3).

Table 5.3: A tableau with weighted constraint and MaxEnt probabilities with an input that features the natural feature value in a given environment.

/tamba/	*NT 1	IDENT-IO 2	$\mathscr{H}$	P
tampa	−1	−1	−3	0.05
tamba			0	0.95

For an input with the unnatural feature value in a given environment, this means that the unnatural feature value in the output will be more frequent than the natural feature value (Table 5.4). Cumulative output probabilities, however, still favor the natural feature value. Cumulative harmony scores (taken for both the natural and unnatural inputs) of [tampa] is −4, of [tamba] −2.

Table 5.4: A tableau with weighted constraint and MaxEnt probabilities where the Markedness constraint weighted lower than the Faithfulness constraint.

/tampa/	*NT 1	Ident-IO 2	$\mathscr{H}$	P
tampa	−1		−1	0.73
tamba		−1	−2	0.27

There is thus an additional aspect of the predictive power of HG under the restricted Con hypothesis that has gone largely unnoticed in the literature. If we restrict Con to only natural constraints, HG (or MaxEnt) will predict that, on a typological level, the natural feature value in a given environment will always be more frequent than the unnatural one (Beguš 2016).

Let us assume that the inputs have to be uniformly distributed (similar to what Jarosz 2006 and Hayes and Wilson 2008 assume, for example, for prior probability in phonotactic learning). This assumption can be extended to a typological level. If we interpret the Richness of the Base hypothesis (ROTB; Prince and Smolensky 1993/2004, Smolensky 1996) on a typological level, we can claim that a typological model has to derive surface probabilities regardless of the input probabilities, which is why the uniform probability distribution is the most neutral assumption. We can thus assume that inputs /NT/ and /ND/ have equal probabilities. Any other distribution of probabilities would be undesirable.

In a restricted version of Con, constraints Ident-IO(voice) and the natural Markedness constraint *NT are admitted, but crucially, the unnatural *ND is excluded. This restriction results in mapping /ND/ → [NT] to be predicted impossible as the candidate [NT] is harmonically bound. What are then implications of the restricted Con in HG/MaxEnt when we model gradient phenomena?

As already mentioned, the probabilities of inputs /NT/ and /ND/ are uniform. For modeling post-nasal (de)voicing, we thus assume both probabilities of voiced and voiceless input stops to be at chance level: if we limit the candidate set to those that manipulate [±voice], P(/NT/) and P(/ND/) equal 0.5. Restriction on Con and weighted constraints in combination yield the following implications: if the Faithfulness constraint ($\mathcal{F}$) Ident-IO(voi) is weighted (w) infinitely higher than the Markedness constraint ($\mathcal{M}$) *NT, we expect the probability of the outputs [NT] and [ND] to replicate the probability of the input under: P([NT]) and P([ND]) would equal 0.5. If, however, the markedness constraint is weighted less than infinitely lower or higher than the faithfulness constraint, the probability of the input [ND] will be greater than the probability of the input [NT].

(32) a. w(Ident-IO(voi))$\gg$ w(*NT): P([ND]) = P([NT]) = 0.5

b. w(*NT#)> w(Ident-IO(voi)): P([ND]) > P([NT])

If we allow only natural constraints into Con, we can only derive systems with gradient phonotactic restrictions in which the natural element in a given context is more frequent than the unnatural element. In other words, with restrictive Con and positive weights, no weighting exists that would yield a system in which the unnatural element is gradiently preferred: has greater probability than the natural one.[85]

---

[85] This assumption is confirmed by MaxEnt modeling in Hayes (2017).

(33) a. $w(\mathcal{F}) \gg w(\mathcal{M})$: $\mathrm{P(nat)} = \mathrm{P(unnat)} = 0.5$

b. $w(\mathcal{M}) > w(\mathcal{F})$: $\mathrm{P(nat)} > \mathrm{P(unnat)}$

From (33) it follows that either a full contrasts or a phonotactic restriction against unnatural feature values in favor of the natural feature values in a given environment are possible.

(34) *Natural Gradience Bias (NGB)*

Weighted-constraint frameworks (like HG and MaxEnt) with restricted Con predict that the probability of the natural feature value in a given environment is always equal or grater than the probability of the unnatural value in a given environment.

The NGB generalization correctly predicts the major typological trend with regard to gradient phonotactic restrictions: all cases reported previously (both as trends in the lexicon, e.g., Berkley 2000, Coetzee and Pater 2008, Anttila 2008, and as tacit phonotactic knowledge obtained from experiments, e.g. Albright 2009) indeed operate in the natural direction, where the natural element is preferred and more frequent than the unnatural one in a given environment. Moreover, the NGB assumption receives support from the modeling literature: Hayes (2017) has recently argued that in the MaxEnt framework with restricted Con "the bounded candidate can never receive a higher probability than the candidate that bounds it".[86]

However, the Tarma Quechua and Berawan systems of stop voicing presented in Chapter 2 suggest that weighted-constraint-models with restricted Con undergenerate, since the two languages require precisely the situation excluded by the Natural Gradience Bias: a higher frequency for the unnatural feature value in a certain context. In other words, no weighting of

---

[86] Hayes (2016) calls the generalization "stochastic harmonic bounding".

natural Markedness constraints can generate cases like Tarma Quechues and Berawan. This, in turn suggests, that Con must contain some unnatural Markedness constraints.

Before firm conclusions are drawn, however, the unnatural gradient phonotactic restrictions in the two languages would need to be confirmed experimentally. Unnatural categorical *alternations* have already been confirmed as being productive elsewhere: Coetzee and Pretorius (2010) show that post-nasal devoicing in Tswana extends to nonce-words. Chapter 2 presents evidence that unnatural gradient phonotactic restriction in Berawan dialects and especially in Tarma Quechua show signs of productivity, too. Experimental nonce-word tests would additionally reveal the degree of productivity and grammatical status of the two processes and therefore the ability of unnatural gradient phonotactic restriction to be productive in general.

As already mentioned, existence of unnatural categorical and gradient processes suggest that unnatural constraints should be admitted to Con. However, to simply relax Con and allow all possible Markedness constraints is not a desirable solution either. For example, Hayes and Wilson's (2008) phonotactic learner is able to derive unnatural phonotactics because they do not limit Con to natural constraints — the learner is only provided with feature values and constraint templates. Their model, however, does not encode typological rarity of unnatural processes (although, see Pater 2012, Staubs 2014, or Stanton 2016c how typological rarity could be derived with unrestricted Con). Ideally, the grammar would be able to derive unnatural patterns and encode their rarity at the same time. In the following, I propose a new typological model that is based on the Maximum Entropy framework (Goldwater and Johnson 2003) and that combines existing proposals for encoding the AB with a novel metric: prior Historical Weights. I argue that a typological model that combines both AB and CB influences outperforms the current split models. The main advantage of the proposed model is that it solves undergeneration by deriving the attested unnatural processes and at the same time encode gradient typological observations.

5.2 A typological model

Several proposals exist for modeling phonological learning and computationally encoding the fact that some processes are consistently underlearned. Wilson (2006) proposes a MaxEnt model (Goldwater and Johnson 2003) that differentiates prior variance (σ^2) in the regularization term of different constraints. The differing degrees of learnability of some alternations result in phonological typology according to the Analytic Bias, and the differentiating σ^2 can be used to model typology from learning biases. One of the objections against this approach is that it fails to derive typological generalizations for those processes that do not show learnability differences. A growing body of work argues that while structurally complex alternations are indeed underlearned compared to structurally simpler alternations (the so-called Complexity Bias), no such biases exist when structural complexity is controlled for, at least for a subset of natural-unnatural alternation pairs (the so-called Substantive Bias). There is indeed a substantial difference in the proportion of positive vs. negative results between studies that test Complexity and Substantive Bias (for a literature overview, see Moreton and Pater 2012a,b). While some studies do report differences in learnability when complexity is controlled for (e.g. Wilson 2006), many studies fail to find such differences (Pycha et al. 2003, Kuo 2009, Skoruppa and Peperkamp 2011, via Moreton and Pater 2012a,b). As already mentioned, the position of this dissertation is that evidence in favor or against AB should be evaluated for each alternation pair separately. Several studies specifically tested alternations that target feature [±voice] and found no differences between natural and unnatural processes Seidl et al. 2007, Do et al. 2016, Glewwe 2017, Glewwe et al. 2018). Absence of Substantive Bias evidence in alternations that target feature [*pm*voice] is especially problematic for the Analytic Bias approach to typology because there exist substantial differences in typology between natural and unnatural alternations targeting feature [±voice]. For example, 15 languages of 197 surveyed (Locke 1983; reported in Hayes and Stivers 2000) have PNV as a synchronic phonological alternation. 26 of 153 languages surveyed in Gurevich (2004; reported in Kaplan 2010) have IVV as a synchronic alternation (e.g. in Warndarang /p/, /t/, and /k/ surface as [b], [d], and [g] intervocalically; Kaplan 2010). To my knowledge, no systematic typological studies of final devoicing exist, but FD is one of the most frequent alternations in world's languages (Brockhaus 1995). As was already discussed in Section 4.3.2, the unnatural alternations PND, IVD, and FV are very rare with only PND being securely attested as a productive synchronic process (Coetzee and Pretorius 2010). The AB approach faces problems deriving this typological distribution if no differences in learnability are observed between natural and unnatural alternations.

Historical Probabilities of alternations estimated with the BSC technique offer a solution to this problem. By introducing a typological model that combines Historical Probabilities with the MaxEnt framework, we can maintain the MaxEnt approach to modeling phonological learning and differences in learnability of different alternations (Wilson 2006, Hayes and Wilson 2008) and at the same time derive typological generalizations for those processes for which no learnability biases exist.

In a MaxEnt model of phonological learning, the probability distribution over candidates is computed from the harmony of each candidate (Goldwater and Johnson 2003, Wilson 2006). Harmony (H) of an output(y)-input(x) mapping is derived from violations of constraints (C) and their respective weights (w). The model proposed here follows the assumption that weights are non-negative and constraint violations non-positive (e.g. Pater 2009, Boersma and Pater 2013, Magri 2015), although this assumption has no unique implications to the model.

(35)

$$H(y,x) = \sum_{i=1}^{m} w_i C_i(y,x)$$

The probability of an output y given an input x ($P(y|x)$) is computed as (Goldwater and Johnson 2003, Wilson 2006):

(36)

$$P(y|x) = \frac{e^{H(x,y)}}{\sum\limits_{y \in Y(x)} e^{H(x,y)}}.$$

I propose that we can extend the MaxEnt framework for modeling phonological learning to a model of *typological* probabilities. More specifically, I argue that we can model both AB and CB influences on typology in such a “typological” MaxEnt model.

Two different implementations of encoding learnability differences in MaxEnt models of phonological learning have been proposed: Wilson (2006) differentiates prior variance (σ^2) of constraints, whereas White (2017) differentiates prior means (μ) in the regularization term. In both proposals, one metric becomes redundant when the other is employed. For modeling AB, I adopt Wilson’s (2006) approach of differentiating prior variance (σ^2).

When modeling CB influences on the typology, however, I introduce a diachronic metric into the MaxEnt model called Historical Weights. I propose that the prior Historical Weights of each Markedness-Faithfulness constraint pair for a given alternation can be directly derived from Historical Probabilities. The differences in Historical Weights between a Markedness ($\mathcal{M}$) and Faithfulness ($\mathcal{F}$) constraint for a given alternation A_k are calculated from Historical Probabilities according to:

(37)

$$\Delta w_{\chi}(\mathcal{M}, \mathcal{F}) = -\log\big(\frac{P_{\chi}(A_k)}{1 - P_{\chi}(A_k)}\big).$$

The distribution of prior weights and prior σ^2, where the first encodes CB (Historical Weights) and the latter AB is not based on any empirical facts, but on a more general conception of the two influences on typology. It would be equally possible to assume a model that would encode both AB and CB influences by differentiating prior μ. The CB approach would be encoded with prior Historical Weights and the AB with prior μ in the regularization term. White (2017) proposes a model which encodes learnability differences by differentiating prior μ. There is, however, an appealing conceptual reason for dividing AB and CB influences into prior σ^2 and Historical Weights. Differentiating prior μ, just like σ^2, means that some processes require more input data to be learned. The logic behind this generalization, however, is different between the two approaches. Prior μ encodes the "preferred weight" of constraints (White 2017). Penalization of the prior thus increases as a constraint moves away from its preferred weight (White 2017). This means that rate of learning is equal for all constraints and dependent primarily on the distance of the learned weight from the preferred weight. Differentiating prior σ^2, on the other hand, effectively means that learning of weights for some constraints occurs faster than for others. The latter model might be closer to reality, which is why it is chosen in this dissertation, but further empirical evidence is needed to distinguish between the two approaches.

Let us look at how the MaxEnt model of typology derives typological differences between PND and PNV. The two alternations have been tested experimentally and no differences in learnability have been observed (Seidl et al. 2007). Based on this experimental evidence and in the absence of positive results, we can assume that there is no difference in learnability between PND and PNV. This means prior variance of the markedness *NT and *ND constraints should be kept equal. When modeling phonological learning, equal prior variance results in a desirable generalization that both alternations are equally learnable. Because of equal prior variance, we can disregard influences of the AB on typology. Without the input from Historical Probabilities, the derivation would end here and we would incorrectly predict no differences in typology between PNV and PND. Under the new model that admits both AB and CB influences, we can maintain MaxEnt framework of phonological learning and at the same time derive the observed typological differences.

To encode the CB influences on typology, we calculate the differences in Historical Weights between *NT and Ident-IO(voi) and *ND and Ident-IO(voi) constraints according to (37). Historical Probabilities (P_{χ}) of each process, PND and PNV, are estimated as in Section 4.3.1 and repeated in (38).

(38) a. $P_{\chi}(\text{PNV}) = 0.143$

b. $P_{\chi}(\text{PND}) = 0.0005$

Differences in Historical Weights between corresponding Markedness and Faithfulness constraints are calculated from historical Probabilities according to (37).

(39) a. Δw_{χ}(*NT, Ident-IO(voi)) = 1.79

b. Δw_{χ}(*ND, Ident-IO(voi)) = 7.60

If we set prior w_{χ} of the Ident-IO(voi) constraint arbitrarily at 10, we derive the following Historical Weights for the two Markedness constraints *NT and *ND.

(40) a. w_{χ}(*NT) = 8.21

b. w_{χ}(*ND) = 2.40

Table 5.5 illustrates derivation of probabilities in a MaxEnt model of typology. Differences in Historical Weighs (w_{χ}) result in different harmonies ($\mathscr{H}$) of natural and unnatural candidates. Harmonies are derived from Historical Weights and constraint violations (41). The only difference between between the derivation of synchronic MaxEnt probabilities (35) and the typological probability (41) is that the latter derives Harmony score from Historical Weights (w_{χ}).

(41)

$$H(y, x) = \sum_{i=1}^{m} w_{\chi i} C_i(y, x).$$

Different Harmonies in turn results in different predicted typological probabilities (Predicted P that equal Historical Probabilities P_{χ}) via (36). Predicted typological probabilities can be compared to the actual observed synchronic typology (Observed P; see Section 4.3.4). To be sure, the framework in 5.5 only models typology and *not* phonological learning: speakers have no access to Historical Weights.

Table 5.5: A tableau illustrating Historical Weights.

/NT/	Ident-IO(voi) $w_\chi = 10$	*NT $w_\chi = 8.21$	$\mathscr{H}$	Predicted P	Observed P
a. [NT]		-1	-8.21	.857	.924
b. [ND]	-1		-10	.143	.076

/ND/	Ident-IO(voi) $w_\chi = 10$	*ND $w_\chi = 2.40$	$\mathscr{H}$	Predicted P	Observed P
a. [ND]		-1	-2.40	.9995	≈ .996
b. [NT]	-1		-10	.0005	≈ .003

The MaxEnt model of typology in Table 5.5 states that given input /NT/, the typological probability of mapping /NT/ → [ND] is 14.3%, while the typological probability of /NT/ → [NT] is 79.6%. For input /ND/, the probability of /ND/ → [NT] is 0.05%, the probability of /ND/ → [ND] is 99.9%. These predicted probabilities match closely the observed probabilities (derived from Locke 1983 and my own survey in Chapter 2; see also Chapter 4). The new model thus allows us to derive unnatural processes by including unnatural constraints in Con and at the same time derive observed gradient typology.

To be sure, a Faithfulness constraint, such as Ident-IO(voi) is in relationship with several Markedness constraints (e.g. *D, *NT, etc.). Each of these markedness constraints is in turn in a relationship with several other faithfulness constraints. It is quite likely that differences in weights for one set of relationship will differ from weights for another set of relationships, which would render the constraints impossible to order with respect to Historical Weights. The solution to this problem is to employ independently motivated positional faithfulness constraints (Beckman 1997, Lombardi 1999), e.g. Ident-IO([voice]/__#). Thus, each Faithfulness-Markedness relationship is independent of Faithfulness-Markedness relationships for other processes.

The current proposal makes no claims about how differences in prior variance (σ^2) of different constraints result in observed typology. One difficulty with deriving typology by differentiating

σ^2 is that the σ^2-metric effectively encodes that some alternations require more input data to be learned than others. In phonological acquisition, learners are provided with abundance of data — or at least with sufficiently enough data that they are able to reproduce inputs with a great degree of faithfulness, both on a phonetic and phonological level. While iterated learning models can amplify even minor learnability biases (Kirby and Sonderegger 2013, 2015, Staubs 2014), it is not immediately clear how gradient learnability preferences in phonology result in typology for processes for which abundant evidence exists in the primary linguistic data. The most promising line of explanation for deriving typology from learnability differences is to assume that learnability primarily influences those aspects of phonology for which input data are sparse. In such cases, learners can actually fail to learn a process or, alternatively, they favor reanalysis based on learnability difference, which can, over generations, results in observed typology. Learning stress patterns, for example, often involves sparse data that allows several competing analyses (e.g. Staubs 2014). The questions of the exact relationship between σ^2 and the observed typology is thus beyond the scope of this dissertation and is left for future work. The dissertation, does, however, outline a model called "catalysis" (Section 6.5) that explains how learnability differences (AB) might influence typology when processes do not involve restructuring or rule inversion, when input data is abundant, and when the change in phonological representation occurs in adult population and does not necessarily stem from phonological acquisition.

In sum, an advantage of introducing Historical Weights into a typological model within the MaxEnt framework is that such a model derives all attested processes and encodes differences in the observed typology that would be left unexplained if only the AB influences would be admitted to the model.

5.3 Too Many Solutions and future directions

In addition to solving the problem of undergeneration while retaining typological predictions, the model of typology proposed in this chapter bears the potential to solve the Too Many / Too Few Solutions (TMTFS) problem (Steriade 2001). For some markedness constraints, there exist several different repair strategies, whereas other markedness constraints seem to allow only one repair strategy (or rely on one strategy considerably more frequently than others). The new model of typology is able to encode such distributions: greater differences in Historical Weights between markedness constraints and different faithfulness constraints that correspond to these markedness constraints result in a strong preference for a single repair strategy, while at the same time allowing for other repair strategies (with considerably lower frequencies). Smaller differences in Historical Weights, on the other hand, mean that more repair strategies will be available and that they will tend to be more equiprobable.

Perhaps the two most famous cases illustrating the Too Many / Too Few Solution problem are repairs of the markedness constraints *NT and *D#. Pater (1999) shows that languages exhibit a variety of strategies for repairing the markedness constraint *NT, including deletion, nasalization, and voicing. On the other hand, in the majority of languages that feature alternation of word-final voiced stops, the repair strategy is that of devoicing. Steriade (2001) proposes a P-map explanation for this distribution, claiming that repair strategies are perceptually minimal. Because perceptual distance between voiced and voiceless stops is smallest in word-final position (smaller than the perceptual distance between voiced stops and any other segment), the only (or most common) repair strategy is final devoicing.

5.3.1 Nasalization as a repair?

Below, I argue that approaching the TMTFS problem by combining diachronic statistical modeling and experimental estimates of learnability biases might bring new solutions to the issue. Before presenting a potential solution within the proposed MaxEnt typological model using Historical Weights and prior σ^2, I present a diachronic treatment of a process that is directly linked to the TMTFS problem: final nasalization.

Final nasalization has long been believed impossible in synchronic phonological literature, rendering validity to Steriade's (2001) P-map hypothesis. Recently, however, a synchronic system that repairs *D# with nasalization has been reported in Noon (Merrill 2015) and in some Austronesian languages (Blust 2005, 2016).[87] Merrill (2015) argues that final nasalization in Noon arises through a combination of sound changes. Voiced stops in Pre-Noon are prenasalized. Word-finally, the oral part of prenasalized stops is lost and the result is word-final nasal stop. Intervocalically, on the other hand, prenasalized voiced stop develop to plain voiced stops. These sound changes result in a synchronic final nasalization of voiced stops, e.g. [tam] vs. [tab-in]. This development is proposed in Merrill (2015) and is substantiated with abundant evidence from neighboring dialects.

Final nasalization is never reported to operate as a sound change, except for some Austronesian languages in Blust (2005, 2016). Because there are no traces of prenasalized stops that could lead to final nasalization in these cases, Blust concludes that final nasalization had to operate as a single sound change.

---

[87] Final nasalization of voiced stops is a rare process, although not unnatural, but rather unmotivated because it does not operate against a universal phonetic tendency.

Four cases of final nasalization in Austronesian are reported in Blust (2016): in Kayan-Murik, Kalabakan Murut, Karo Batak, and Berawan dialects. An intriguing aspect that is immediately noticeable in this distribution is that out of four cases of final nasalization reported in Austronesian, two are independently argued to have undergone the Blurring Process: Murik (Section 3.2.6) and Berawan dialects (Section 3.5.2.1).

Spontaneous nasalization is generally very rare typologically, but the one environment in which spontaneous nasalization is commonly attested and phonetically motivated is precisely the position before voiceless fricatives (Ohala and Ohala 1992, Ohala and Busà 1995). In both, Murik and Berawan dialects, the Blurring Process explanation proposed in Sections 3.2.6 and 3.5.2.1 posits a stage in which voiced stops develop to voiced fricatives intervocalically (in Berawan dialects) or in non-post-nasal position (in Murik). There is no particular reason why we should assume fricativization of voiced stops in Berawan did not operate on post-vocalic stops as well. This means that the Blurring Process explanation assumes that voiced stops were realized as voiced fricatives in word-final position at some stage of the development of Murik and Berawan. Note that this assumption is completely independent of final nasalization. It is also likely that word-finally, voiced fricatives devoice, which yields exactly the condition in which spontaneous nasalization is well-motivated by the mechanism described in Ohala and Ohala (1992) and Ohala and Busà (1995).

That voiced fricatives underlie final nasalization is supported by both language-internal and dialectal data. While it is true that final voiced stops in proto-languages of the four languages with FN correspond to observed final nasals (as presented in Blust 2016), reflexes of voiced stops in other positions suggest that additional steps were involved in the development of FN.

Let us first take a look at the Kayan-Murik data. Blust (2016) reports two dialects to undergo FN in the Kayan-Murik group: Uma Bawang and Long Atip dialect of Kayan. As already mentioned, I have independently argued that PND in Murik proceeds through a change of voiced stops to voiced fricatives. This assumption is strengthened by both internal and dialectal data. Blust (2016) presents data from four Kayan-Murik dialects, two of which (Uma Bawang and Long Atip) have FN. Proto-Kayan-Murik (PKM) final *-b and *-d correspond to Uma Bawang and Long Atip [m] and [n], e.g. PKM *kələb > Uma Bawang [kələm]; PKM *anud > Uma Bawang [anun] (data from Blust 2016). Word-internally, however, voiced stops develop to voiced fricatives [v] and [r] (likely through and interstage *ð). There is strong evidence to suggest that Uma Bawang and Long Atip voiced stops first develop to voiced fricatives in all post-vocalic positions, including word-finally. In a closely related dialect, Uma Juman that lacks FN, voiced stops develop to voiced fricatives in all post-vocalic positions, including word-finally, e.g. PKM *kələb > Uma Juman [kələv]; PKM *anud > Uma Juman [anur] (data from Blust 2016). In other words, word final Uma Bawang and Long Atip nasals [-m] and [-n] correspond to Uma Juman [-v] and [-r]. Table 5.6 summarize these dialectal correspondences (from Blust 2016).

Table 5.6: Summary of developments in Kayan-Murik (from Blust 2016).

		#__	**V__V**	__#
Uma Bawang & Long Atip	labial	b	v	m
	alveolar	d	r	n
Uma Juman	labial	b	v	v
	alveolar	d	r	r

The second language with reported FN that shows independent evidence of fricativization of voiced stops is Berawan. That voiced stops first develop to voiced fricatives intervocalically in Berawan is strongly suggested by both the development of coronal series of stops for which lenition is still attested and by relative chronology (Section 3.5.2.1). For example, the voiced dental stop develops to [r] intervocalically that likely goes back to a fricative [ð]. Additionally, the development *b > [k] strongly suggests that the development proceeds through an intermediate stage with a voiceless fricative articulation [ɸ]. Otherwise, the sound changes that operated in Berawan cannot be ordered (see Section 3.5.2.1). Just like in Uma Bawang and Long Atip, we can assume that voiced stops develop to voiced fricatives in all post-vocalic positions in Berawan, including the word-final position.

In two of four reported languages with FN, the Blurring Process explanation reconstructs a stage with fricativization of voiced stops, completely independently of FN data. In one additional language, there exist strong internal evidence that FN arises through a stage with fricatives. Blust (2016) reports that Kalabakan Murut underwent FN and shows that Proto-Murutic final voiced stops *-b, *-d, and *-g develop to [-m], [-n], and [-ŋ]. What is more relevant for our approach, however, is that Proto-Murutic *b, *d, and *g yield [b], [r], and [h] intervocalically in morpheme boundaries when a vowel-initial prefix follows (Blust 2016). Table 5.7 summarizes the development.

Table 5.7: Summary of developments in Kalabakan Murut (based on Blust 2016).

Proto-Murutic	**Kalabakan Murut**		
	#__	**V__-V**	__#
*b	b	b	m
*d	d	r	n
*g	d	h	ŋ

Reflexes of Proto-Murutic *d and *g thus directly point to an interstage with a voiced fricative (*ð/[r] and *ɣ). The velar fricative additionally undergoes devoicing that is not limited to word-final position and debuccalization to [h].[88] The synchronic alternation in Kalabakan Murut is thus between final [-m], [-n], and [-ŋ] and non-final [b], [r], and [h], which again directly points to a stage with frication, e.g. [aŋ-ipa**ŋ**] vs. [ipa**h**-in] or [a-nako**n**] vs. [toko**r**-or] (data from Blust 2016).

That the deveopment of Kalabakan Murut FN proceeds through a stage with final voiced fricatives that devoice to voiceless fricatives is supported by dialectal data as well. In closely related dialects, voiced velar stop *g develops to a voiced velar fricative [ɣ] in Murut Paluan and Murut Timugon (*liʔog > [lijoɣ]) and further to [h] in Murut Tagol *liʔog > [lijoh]) (Lobel 2013).

In sum, three of four languages with FN reported in Blust (2016) show strong evidence for an interstage in which voiced stop develop to fricatives. In two out of four languages the evidence for a stage with fricativization is independent of FN data. In Kalabakan Murut, final nasals still alternate with a voiceless fricative ([h]) or [r] in the coronal and velar series, thus directly pointing to a stage with fricativization. In one language with reported FN, however, no direct evidence for an interstage with frication exists — in Karo Batak.

As already mentioned, spontaneous nasalization is a rare phenomenon, except in position before voiceless fricatives. In three of four Austronesian languages with reported FN, there exists strong evidence for a stage in which voiced stops surface as fricatives in word-final position, which is precisely the condition where spontaneous nasalization is well-motivated (Ohala and Ohala 1992, Ohala and Busà 1995). We can thus assume that fricatives devoice word-finally (if they are not already devoiced as in Berawan), after which final syllables become nasalized (VS > ṼS̃) due to the mechanism described in Ohala and Ohala (1992) and Ohala and Busà (1995).

---

[88] Voiceless velar fricative [x] is in free variation with voiceless glottal fricative [h] word-finally in Karo Batak according to Woollams (1996). It is thus likely that [h] develops from a devoiced velar fricative [x].

The last change that we need to assume for a complete FN is occlusion of nasalized fricatives to nasal stops. In sum, in four of five languages with FN (including Noon), evidence exist that FN has a complex diachronic history.

There are three possible explanations for FN in Karo Batak, the only language without direct evidence for a stage with frication. First, FN in Karo Batak can result from a single sound change, which would distinguishing Karo Batak from other languages discussed above. Second, FN in Karo Batak can result from the same combination of sound changes discussed above. Under this hypothesis, Karo Batak would undergo occlusion of fricatives in all places of articulation. For example, I reconstruct a stage with voiced fricatives for FN in Berawan, Kalabakan Murut, and Murik. In all these languages, fricatives then undergo occlusion. For example, Murik and Kalabakan Murut voiced labial fricative [β] that does no devoice and nasalize occludes back to [b]. In Murik, occlusion also occurs in the velar series. It is possible that Karo Batak voiced fricative undergo occlusion in all places of articulation, thus blurring the evidence for initial stage with frication. Finally, Karo Batak FN can also result from analogy. While it is true that Proto-Batak final voiced stops *-b, *-d, and *-g correspond to Karo Batak [-m], [-n], and [-ŋ], Blust (2016) also mentions that these nasalized final stops do not alternate. For example, *tərəb yields [tərəm], but when the stop appears in non-final position, it also surfaces as a nasal, e.g. [kə-tərəm-ən] (data from Blust 2016). This stands in contrasts with Kalabakan Murut that preserves synchronic alternation, e.g. [a-naku**m**] vs. [taku**b**-on]; [a-mila**n**] vs. [bila**r**-on]; or [aŋ-ipa**ŋ**] vs. [ipa**h**-in] (data from Blust 2016). While by itself, this piece of evidence does not necessarily mean nasalization has to be analogical, evidence from related Toba Batak suggest that nasalization in Karo Batak might arise analogically based on a sandhi phenomenon. In Toba Batak, Blust (2013) reports a sandhi phenomenon whereby word-final nasal and oral voiced stops merge to oral voiced stops before word-initial stops. If we assume a stage with the same sandhi phenomenon in Karo Batak, it is easy to explain how word-final nasal variant got generalized. If -N D- and -D D- both merge to -D D- in sandhi, the nasal articulation can get analogically generalized from this merger position.

It is difficult to asses which of the three possible scenarios actually gave rise to FN in Karo Batak, but given the evidence from related languages and a plausible scenario for analogy, it is likely that Karo Batak FN did not arise from a single sound change.

5.3.2 Potential solution

The proposed model of typology allows us to encode that some repairs are less frequent than others and even specify whether the relative rarity is due to AB or CB. Based on BSC, we can estimate the Historical Probability of final nasalization based on the two trajectories it can arise from (in Noon and in Austronesian). To encode relative frequencies of different repair strategies, we can differentiate prior Historical Weights of the Markedness constraint *D# and two Faithfulness constraints, Ident-IO(voice) and Ident-IO(nasal). To be sure, Faithfulness constraints always have higher weights than the Markedness constraint because absence of an alternation is, to my knowledge, always typologically more frequent than presence of an alternation. While the exact estimation of Historical Probabilities and therefore Historical Weights for the two processes is lacking, we expect the difference between *D# and the Faithfulness constraint that preserves feature [±voice] to be much smaller than the difference between *D# and the Faithfulness constraint preserving feature [±nasal]. This would result in the prediction that mapping from /D/ → [T] be more frequent than mapping /D/ → [N] in word-final position.

Such reasoning can be extended to other constraints and other processes with the TMTFS. To avoid a problem discussed in Section 5.2 where different sets of Markedness-Faithfulness relations require different weights, independently motivated positional faithfulness constraints (Beckman 1997, Lombardi 1999) should be employed here as well (in line with Steriade 2001).

While Historical Weights encode the Channel Bias, we can at the same time encode that one mapping is also preferred by learning. Experiments that would test learnability of the two processes, final devoicing and final nasalization, would shed light on this discussion. To my knowledge, no such studies exist. A related experiment in Do and Albright (2016), however, suggests that cases with Too Many Solutions problem do show learnability differences. Further experimental work combined with the BSC technique would shed light on the distinction between the AB and CB influences when modeling different repair strategies.

In sum, the typological model proposed in this chapter allows us to encode that some repair strategies are relatively equally frequent, other are disproportionately more or less frequent and that a part of these differences is due to AB and part due to CB influences. Initial investigation into the TMTFS problem thus shows that Historical Weights (or CB) affect the typology of repair strategies (TMTFS) as well, as they arise from complex diachronic scenarios that involve multiple sound change. The relationship between Historical Weights (CB) and prior variance (AB) when deriving the typology of repair strategies should be explored further.

Page intentionally left blank.

Chapter 6

Learning the Blurring Process

One of the major problems in the discussion of AB and CB influences on typology is that empirical evidence often supports both hypotheses equally well. Additionally, it is often difficult to isolate AB influences from CB influences and vice versa (see the so-called AB-CB Conflation Problem in Chapter 4 above). One of the advantages of the new historical model and the BSC technique proposed above is that it allows us to identify mismatches in predictions between the AB and CB approaches and test our hypotheses on those mismatched predictions. In other words, we can control for the CB influences when testing the AB factor and vice versa.

For example, several studies test learnability differences between the natural PNV, IVV, and FD and the unnatural PND, IVD, and FV. None of the studies, to my knowledge, report positive results: learning of natural and unnatural processes does not differ significantly in any of these studies (Seidl et al. 2007, Do et al. 2016, Glewwe 2017, Glewwe et al. 2018). Based on these experimental results, we can claim that when testing the CB influences on these processes, the AB factor is controlled for because we expect no differences in typology due to learning. The historical model proposed in this dissertation predicts unnatural processes targeting the feature [±voice] to be significantly more frequent than unnatural processes (Chapter 4). Figure 4.3 shows that Historical Probabilities match the independently observed synchronic typology well, which suggests the CB model influences the typology of natural and unnatural processes targeting feature [±voice] when AB influences are controlled for.

On the other hand, the BSC technique allows us to identify and design experiments that control for the CB influences when testing the AB factor. Unnatural processes are crucial in this respect, as they always arise through the Blurring Process (Chapter 3). The AB-CB Complexity Mismatch identified by the BSC technique (Section 4.4.2) is one such case on which we can directly test the AB influences while controlling for the CB factor. In the rest of this chapter, I present an experiment that tested the prediction of the AB approach identified in Section 4.4.2 that the lower frequency of complex alternations is influenced by AB and that AB increases the frequency of operation of the last sound change in the Blurring Process. This prediction is borne out by the experimental results: subjects are more likely to generalize a simple unnatural alternation compared to a complex alternation when presented with ambiguous data. Based on these results, I also outline a mechanism called “catalysis” that explains how Analytic Bias influences phonological typology. Deriving typology within AB is challenged by the fact that learners are able to reproduce the input grammar to a great degree of faithfulness. The catalysis explains at least one aspect of AB’s influence on typology through a combination of phonetic variation and learnability preferences. The advantage of this approach is that it can accommodate both the fact that input data in learning is abundant and that change often occurs in adult populations.

6.1 Experiment

As already mentioned, studies that test influences of AB with artificial grammar learning experiments so far fail to control for the CB influences. Testing learnability differences for processes with mismatched predictions between the AB and CB approaches allow us to control for diachronic factors when testing learnability. To test predictions of the proposed framework from Chapter 4, I conducted an artificial grammar learning experiment that tests the learnability of a complex process and unnatural process. Many artificial grammar learning experiments test learnability of complex ∼ simple natural or natural ∼ unnatural. To my knowledge, no experiment thus far tests learnability of complex ∼ unnatural processes.

The experiment tested learnability of two alternations: one that arises after two (Stage 3) and one after three sound changes (Stage 4) in the Blurring Process that leads to post-nasal devoicing. The first alternation is a complex alternation where voiced fricatives in the elsewhere condition alternate with voiceless stops in post-nasal position. While this alternation is complex in that it manipulates two features ([±voice] and [±continuant]), it can also be viewed as "unnatural" (*not* according to the definition in (2)) in that stops surface as voiceless in post-nasal position. This complex alternation is, however, easily derivable by very common and natural constraints if we assume that devoiced stops are voiced fricatives in the underlying representation. Post nasally, devoicing result from two natural markedness constraints — a constraint against post-nasal fricatives (*[+nasal][+cont]) and a global constraint against voiced stops [*D].

Table 6.1: Tableaux deriving the complex alternation tested in the experiment.

	/NZ/	*[+nas][+cont]	*D	Ident-IO(voi)	Ident-IO(cont)
	NZ	*!			
	NS	*!		*	
	ND		*!		*
☞	NT			*	*

	/Z/	*[+nas][+cont]	*D	Ident-IO(voi)	Ident-IO(cont)
☞	Z				
	D				*!

The second alternation tested in the experiment is indeed unnatural (according to the definition in (2)), but simple — post-nasal devoicing where voiced stops in the elsewhere position alternate with voiceless stops post-nasally.

As already mentioned, the first, complex alternation involves two features, [±voice] and [±continuant], whereas the second alternation, PND, involves only the feature [±voice]. Complexity is in this experiment, as in many others, conflated with perceptual distance. Voiced and voiceless stops are perceptually closer than voiceless stops and voiced fricatives. For the purpose of this dissertation, I will assume complexity to be the causing factor of learnability differences (following the body of research summarized in Moreton and Pater 2012a,b). This assumption, however, bears no crucial implications and as will be discussed in Section 6.4, the results presented here are equally relevant under the perceptual hypothesis as well.

The experiment was designed so that subjects had to learn both the complex and unnatural alternation in a made-up language called Martian. The experiment included an explicit task, vowel harmony in feature [+front] and the implicit task, learning of complex vs. unnatural process. Subjects were trained on forming plural nouns from singular nouns in Martian. The plural prefixes were [ɔn-] (before coronals) and [ɔm-] (before labials). If the vowel of the noun following the prefix is [+front] ([ɛ, i]), the prefixes are [ɛn-] (before coronals) and [ɛm-] (before labials). In other words, the [+front] feature in the first vowel triggers vowel harmony in the prefix. Subjects were instructed to pay attention to how plural nouns are formed in this language and were told that final task will involve forming plural nouns.

In addition to this explicit task, the training phase also involved data for the implicit task: learning of the complex vs. unnatural alternations. Word-initial voiced stops and fricatives alternate with post-nasal voiceless stops in plural forms with prefix-final nasal. Word-initial voiceless stops and fricatives remain unchanged in the post-nasal prefixed forms. The alternations in the implicit task are summarized in Tables 6.2 and 6.3.

Table 6.2: Alternations in voiceless and voiced stops and fricatives according to position (implicit task).

	[−voice]		[+voice]	
	[−cont]	[+cont]	[−cont]	[+cont]
#__	T	S	D	Z
N__	T	S	T	T

Table 6.3: Examples of Martian words for the implicit task.

	[−voice]				[+voice]			
	[−cont]		[+cont]		[−cont]		[+cont]	
	[LAB]	[COR]	[LAB]	[COR]	[LAB]	[COR]	[LAB]	[COR]
#__	pɔrɔ	taru	fura	sanu	balu	dɔru	vɔna	zɔlɛ
N__	ɔmpɔrɔ	ɔntaru	ɔmfura	ɔnsanu	ɔmpalu	ɔntɔru	ɔmpɔna	ɔntɔlɛ

In the test phase, subjects are given ambiguous stimuli in the plural with devoiced noun-initial stop that can go back to either a voiced stop or a voiced fricative. The subjects have to choose between two singular forms: one with a voiced fricative and one with a voiced stop word-initially. Figure 6.1 exemplifies this task. Preference for one option would suggest that one alternation is easier to learn than the other.

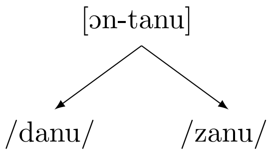


Figure 6.1: An example of an ambiguous stimulus in the test phase.

6.2 Experimental design

6.2.1 Stimuli

Stimuli were read by a phonetically trained native speaker of American English who was unaware of the task of the experiment. The speaker was paid $20 per hour for participation. Recordings were made in a sound-attenuated booth using Shure Beta 53 omnidirectional condenser head-mounted microphone and Sound Devices USB Pre2 pre-amplifier with Audacity recording software, sampled at 44.1 kHz (16-bit). To facilitate constant rate and intensity across recorded words, the stimuli were recorded in a carrier phrase "Say ______". The speaker was instructed to make a longer pause between the carrier word "Say" and the Martian made-up word and was asked to prevoice voiced stops. The speaker was asked to check if any word sounded familiar or English-like and reported no such words in the final inventory of words.

Stimuli were presented in a randomized order to the speaker on sheets of paper, but singular-plural pairs of the same noun were presented together. After the recording, every word-initial voiced stop was manually inspected for voicing. The speaker failed to prevoice voiced stop in some stimuli. Those were replaced by novel recordings, so that all word-initial voiced stops were prevoiced in the final inventory of recorded stimuli. The recorded stimuli were also RMS-equalized in Praat using .praat script *rms equalize* written by Gabriel J. L. Beckers (Beckers

2002). Recordings were converted to .mp3 format from .wav format using Audacity's LAME mp3 encoder.

6.2.2 Subjects

Subjects were recruited using Amazon Mechanical Turk in three recruitments of 80 participants (the last recruitment was stopped after 43 participants reported completion). Of altogether 203 participants, 198 finished the experiment in full and filled a final demographic questionnaire. Participants were paid \$3.30, \$2.00, and \$2.50 for their participation (according to the three recruitment groups described above). Average time to completion across all three groups was approximately 22 min (28 min, 19 min, 20 min). A US IP-address was a requirement for all three recruitments. The first recruitment included no further requirements, the second required "HIT Approval Rate (%) for all Requesters' HITs greater than 99" and "Number of HITs Approved greater than 50". The third recruitment required "HIT Approval Rate (%) for all Requesters' HITs greater than 99" and "Number of HITs Approved greater than 500".

6.2.3 Procedure

The experiment was conducted using the Experigen experimental software (Becker and Levine 2014), adjusted by Adam Albright and additionally adjusted by me for the experimental design in this dissertation. Pictures of Martian creatures in the experiments are taken from van de Vijver and Baer-Henney (2014) with their permission. Pairings of the pictures and stimuli were randomized for every subject.

Subjects were asked to go to a quiet room and put on headphones. To prompt subjects to do so, they were asked to transcribe a nonce-word in Martian. After this initial task, subjects had to agree to informed consent if they wished to proceed. Each speaker was assigned a unique identifier number that they used to identify themselves in the Amazon Mechanical Turk interface. The following page included spelling instructions with recorded examples. The last page before the training phase included instructions on how to proceed with the experiment. Speakers were told that plural nouns in Martian are formed by either [ɔn-] or [ɛn-] and were given two preliminary examples. Subjects were encouraged to repeat words they would hear and were told that they would "be asked some questions about Martian singular and plural forms" at the end of the experiment.[89]

The training phase consisted of 58 singular-plural pairs of words. All nouns were stressed on the first syllable of the root, which means that the first syllable is stressed in the singular (without the prefix) and the second syllable in the plural (with the prefix). Twelve (out of 58) words were fillers. The twelve filler words were of the shape $C_xV_yC_aV_b$ or $C_xV_yC_a$ and were created from initial (C_x) [j], [l], and [r] that did not change in the plural. Four words were created for each initial consonant ([j], [l], and [r]), so that each group contained two mono- and two disyllabic words. Each group also contained two front first vowels (V_y) and two non-front vowels. In other words, there was an equal number of mono- and bisyllabic words and equal number of front and back vowels that triggered frontness harmony in prefixes. C_a were taken from the following inventory: [s, r, l, n, m]. V_b included vowels from the following inventory: [ɔ, u, i]. Table B.1 lists all filler words used in the experiment.

---

[89] Subjects were given the following additional instructions (adjusted from Adam Albright's version of Experigen):

- In order to make sure that you are paying attention to how Martian words sound, we will periodically ask you to type the word that you have just heard.
- Don't despair if you don't see a pattern to the use of the en- and on-! Just listen to the words, and by the end, you will probably have an opinion about which one "sounds right".
- Don't try to memorize every word that you hear. You will not be asked to remember the meanings of the words.

Fourteen (out of 58) words were created with initial voiceless stops and fricatives. Because the English inventory lacks the voiced velar fricative, the experiment involves only labials and coronals. Six words started with initial voiceless stops [p] and [t] (3 each). Two of three words for each voiceless-stop-initial word contained a front V_y that triggered frontness harmony. One word for each group contained a non-front V_y. All six words were of the shape $C_xV_yC_aV_b$. Eight words were created with initial voiceless fricatives [f] and [s], four for each group. All eight words were of the shape C_xV_yCV too, and two words per group involved a front V_y. Two words per group involved a non-front V_y. C_a included consonant from the following inventory: [r, l, n, m]. V_b included vowels from the following inventory: [i, a, ɔ, u]. Table B.2 lists all nonce words with voiceless fricatives or stops in initial position used in the training phase of the experiment.

The remaining 32 words of the training phase were created with initial voiced stops (16 words) and voiced fricatives (16 words). Of the 16 voiced-stop and voiced-fricative-initial words, half (8) involved a labial and half a coronal place of articulation. Each group contained four words that triggered frontness harmony and four words that did not. The design was thus as balanced as possible in order to minimize undesired effects that are not controlled by our experiment. All words in this group were of the shape $C_xV_yC_aV_b$ as well. In order to control for influences of the sequence -$V_yC_aV_b$ following the initial consonant that is of interest to the experiment, design was balanced for vowel and consonant identity in these positions across the stop vs. fricative groups. Groups of words with labial and coronal voiced stops had the same inventory of vowels and consonants in the three positions V_y, C_a, and V_b. Each of the two groups of words (labial vs. coronal) had approximately the same inventory and number of vowels in V_y and V_a and the same inventory and number of consonants in C_a across the stop vs. fricative groups. To avoid too many minimal pairs, the order of vowels in V_y, C_a, and V_b was mixed across groups. For example, labial voiced-stop-initial words had the following inventory of V_x = [i, i, ɛ, ɛ, ɔ, a, u, u]; C_a = [l, l, l, r, r, r, n, m]; V_b = [i, ɛ, ɛ, ɔ, ɔ, a, a, u]. Words with initial voiced fricatives had exact same inventories, but in different combinations to avoid too many minimal pairs.[90] Some minimal pairs were included in the training phase, e.g. /tɛlɔ/ and /dɛlɔ/, /dilɔ/ and /bilɔ/, /bunɛ/ and /zunɛ/, /dulɛ/ and /vulɛ/, /dɛma/ and /bɛma/ or /taru/, /zaru/, and /varu/; /bila/, /zila/, and /vila/. C_a included consonant from the following inventory: [r, l, n, m]. V_b included vowels from the following inventory: [a, ɛ, ɔ, i, u]. Table B.3 lists all nonce words with voiced fricatives or stops in initial position used in the training phase of the experiment.

During the training phase, the stimuli (as described above and summarized in Appendix B) were presented in randomized order (randomized for each subject). A unique picture of a Martian creature was associated with each stimulus (also randomized for each subject). Generally, no orthography was given with the training stimuli. To prompt subjects into focusing on the experiment, however, ten words in the training phase were chosen randomly, different for each speakers. Those ten words were presented orthographically and auditorily and subjects had to enter a transcription of the plural form they heard (the singular form was presented both in audio form and printed on the screen, whereas the plural form only appeared in audio form). This orthographic task was never chosen for words that started with voiced stops or fricatives in order to minimize the effect of orthography and the effect of increased memorization due to this task.

---

[90] In the coronal series, C_a of the stop-initial group consisted of one [l] more and one [n] less than the fricative-initial group. In the labial series, V_b of the stop-initial group consisted of one [a] more and one [ɔ] less than the fricative-initial group.

After the training period, subjects were told they would hear some Martian words that they have not heard before and they would be asked what the most likely way to say those words in Martian would be.

The test phase consisted of two tasks, the explicit and implicit task. The order of the explicit and implicit tasks was randomized for each subject. In the explicit task, the experiment tested whether or not the subjects learned the vowel harmony in plural prefixes. Six words in the singular were created for this task, three that trigger frontness harmony and three that do not: [rema], [liro], [leni] vs. [lonu], [ruro], and [lona]. All six words included initial [l] or [r]. The second consonant (C_a) are either [m], [n], or [r]. Subjects were presented with the six words in the singular (both orthographically and in a recording) with a picture of a single Martian creature. After they heard the word, subjects were shown four pictures of the same creature and were given two orthographic stimuli: with the correct and incorrect harmonies. The order of the two options was randomized.

For the implicit task, subjects were first presented with four pictures of the same Martian creature accompanied by a recording of the test word in plural without any orthographic stimuli. The button that played the recording was embedded in a written sentence "These are ______". After they played the sound, subjects were presented with a single picture of the same Martian creature. The subjects played the first stimulus presented in a sentence "Is this a ______,". After the first stimulus played, they pressed the second stimulus embedded in "or a ______?". After they heard both stimuli, subjects were asked "Which one is it?"[91] and were given two choices ("1st" and "2nd"). Stop-initial and fricative-initial singular forms were equally assigned

---

[91] The question was followed by "(you can replay all three words)."

the first or second position, so that the position of the answer would not influence the results. None of the stimuli in the implicit task were presented orthographically.

The stimuli for the implicit task consisted of eight plural forms of the shape prefix-$C_x V_y C_a V_b$. The Noun-initial consonant (C_x) in all eight stimuli was a voiceless ("devoiced") labial or coronal stop (four nouns each). Four forms included a noun with a front first vowel (V_y) and four with a non-front first vowel. In other words, half of the stimuli were formed with the prefix [ɛn/m-] and half with [ɔn/m-]. C_a consisted of [m, n, r, l, j, w]. Test words were created such that there were no minimal pairs with the training words. Vowels V_y and V_b were composed from [a, ɛ, ɔ, i]. A sample spectrogram of a stimulus [ɛmpɛjɔ] and the two possible responses [bɛjɔ] and [vɛjɔ] are given in Figures 6.2 and 6.3.[92]

---

[92] Note that the low energy during the closure of [p] in [ɛmpɛjɔ] (Figure 6.2) is not voicing, but a low-frequency noise due to wide frequency response of the Shure Beta 53 microphone (20–20,000 Hz). If we filter out frequencies 0–50 Hz with 20 Hz smoothing (stop Hann band in Praat; Boersma and Weenink 2015), we get no energy in the voicing bar during closure.

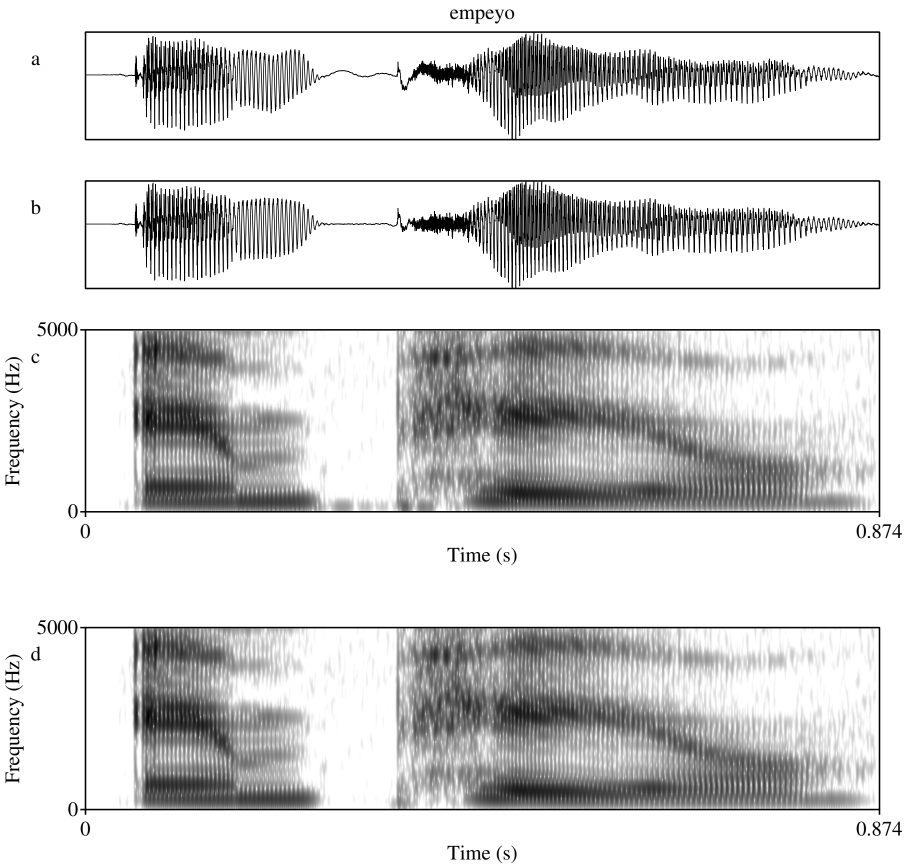


Figure 6.2: A waveform and a spectrogram (0–5000 Hz) of a sample stimulus [ɛmpɛjɔ] after the stimuli have been converted into a .mp3 format where a and c represent unfiltered stimulus and b and c represent filtered stimulus with 0–50 Hz stop Hann band with 20 Hz smoothing.

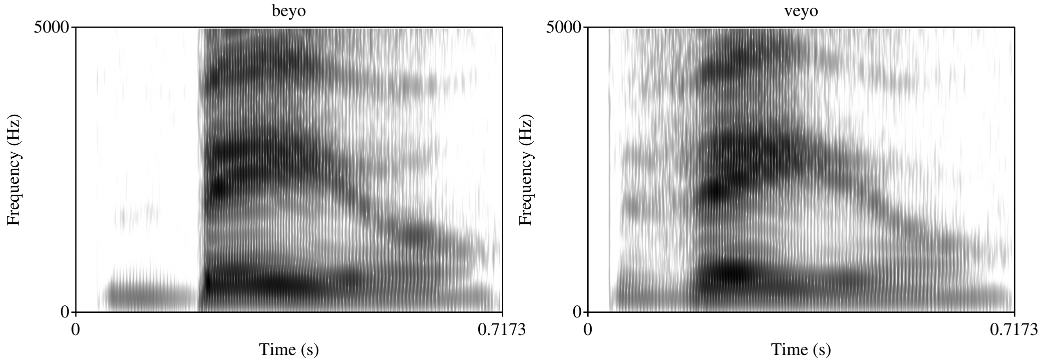
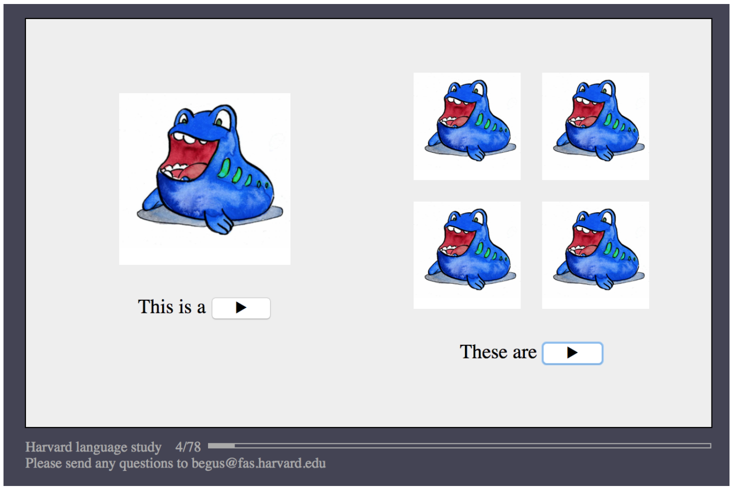


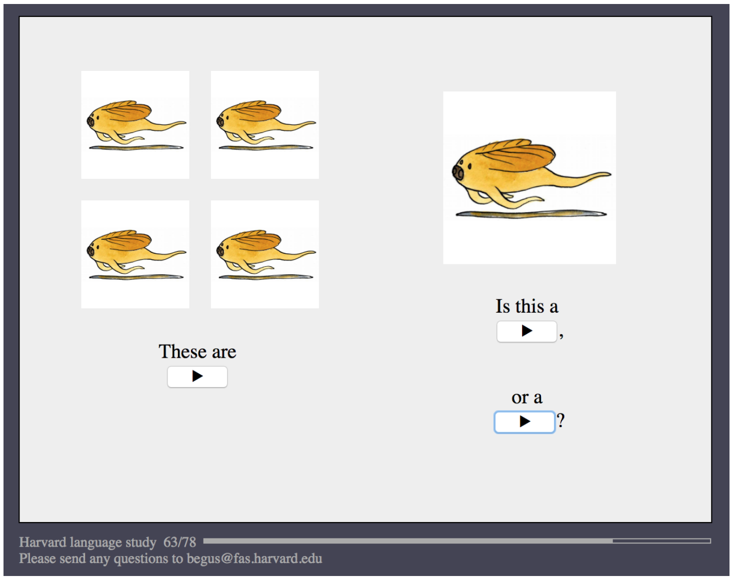
Figure 6.3: Spectrograms (0–5000 Hz) of two response stimuli [bɛjɔ] (left) and [vɛjɔ] (right).

After the test phase, subjects were asked to fill out a short demographic questionnaire. They were asked about their year of birth, gender, their origin, where other people think they are from (based on their speech), whether they are native speakers of English, whether they have taken any linguistics classes, and how did they choose their answers (based on a rule, intuition, or guesses).



Harvard language study 4/78
Please send any questions to begus@fas.harvard.edu

Figure 6.4: An example of the Experigen interface as used in the experiment for the training phase.



Harvard language study 63/78
Please send any questions to begus@fas.harvard.edu

Figure 6.5: An example of the Experigen interface as used in the experiment for the test phase.

6.3 Results

The explicit task (vowel harmony in prefixes) was used to eliminate subjects that did not learn alternations and potentially did not pay attention to the task. For this reason, I only analyze subjects who made no more than one mistake on the explicit task where they had to choose between options with correct and incorrect harmonies. In other words, only speakers that chose the correct vowel harmony 5- or 6-times out of 6 total nouns were analyzed. Additionally, only self-reported English native speakers who have taken no linguistic classes were analyzed in our study. There were 49 such subjects (24.7% of total 198 who completed the questionnaire). All 49 subjects also had a unique Amazon Mechanical Turk WorkerID.[93]

As already mentioned, speakers had to choose between singular noun forms from ambiguous plural data for our implicit task. In other words, plural nouns with nasal-final prefix featured devoiced initial labial or coronal stops. Subjects had to choose between a voiced fricative-initial (complex alternation, not unnatural; see Table 6.1) and stop-initial (simple unnatural alternation) singular nouns.

Each of the 49 subjects analyzed in this experiment had to choose between the two options for eight nouns. Altogether, the experiment thus includes 392 responses. The raw data shows a preference towards the response with voiced stops (simple unnatural). Of altogether 392 responses, subjects chose the stop-initial noun in 225 (or 57.4%) cases and fricative-initial noun in 167 (or 42.6%) cases. The same preference in the raw data holds within the two places of articulation (labial and coronal). Of 196 responses that involve labial-initial nouns, subjects

---

[93] One of the subjects included in this study had the same IP address as another subject that was excluded for independent reasons: they self-described as a non-native speaker. The two subjects had unique WorkerIDs and responded very differently to the demographic questionnaire, which is why I keep the subject that self-reported as a native speaker in our analysis.

chose the stop-initial variant in 108 cases (or 55.1%). For coronal-initial nouns, the number of stop-initial responses is 117 (or 59.7%). Table 6.4 summarizes the raw data.

Table 6.4: Counts of responses.

	Unnatural	**Complex**	**% Unnatural**
Labial	108	88	55.1
Coronal	117	79	59.7
Total	225	167	57.4

Raw by-subject responses show a preference for unnatural response, both for subjects that categorically choose either the unnatural or complex alternation and for subjects that show gradient preference for one or the other. Figure 6.6 plots the number of subject for each response score, where 0 means the subject choose all complex options and 8 means the subject all unnatural options (of altogether 8 test words). Altogether five subjects show a categorical unnatural response (they chose all 8 unnatural responses). On the other hand, only one subject categorically choses the complex response. There are also more subjects that gradiently favor the unnatural response compared to the complex response (Figure 6.6).

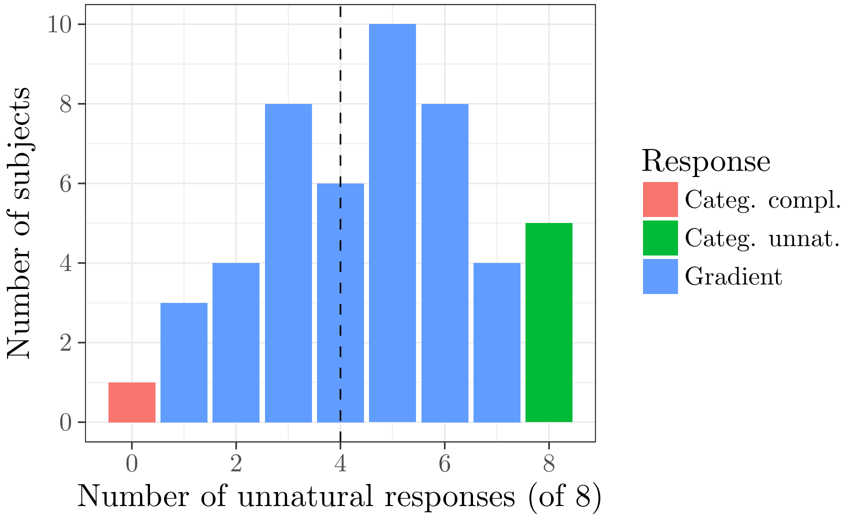


Figure 6.6: Counts of number of subjects and their response score, where 0 means all subject's responses were categorically complex (Categ. compl., red) and 8 means all subject's responses were categorically unnatural (Categ. unnat., green) with gradient responses colored in blue.

To test significance of the preference for stop-initial nouns (and therefore for simple unnatural alternations), the data were fitted to a mixed effects logistic regression model using the *glmer()* function in the *lme4* package (Bates et al. 2015). The final model was selected by a step-wise backward model selection technique. The full model includes presence or absence of the stop-initial response (unnatural) as a response variable and two fixed effects: place of articulation (Place) and frontness/backness (Frontness) of the first vowel/prefix with an interaction between the two. The random effects structure of the full model included a crossed random intercepts for Subject and Item with by-subject and by-item random slopes for Place and Frontness and their interactions. The full model fails to converge. Random slopes (both for interactions and predictors) and random intercepts are removed from the model, each at a time, based on the Akaike Information Criterion (AIC). The random effects structure of the final model selected by step-wise backward model selection technique includes only a random intercept for Subject. Adding a random intercept for Item is not justified by AIC (516.4 with vs. 515.4 without the random intercept). The fixed effect structure of the final model was also obtained with backward model selection technique, based on Likelihood Ratio Test (LRT) and AIC. The interaction Place:Frontness is not significant ($\chi^2 = 0.11, \text{df} = 1,\ p = 0.74$). Both Place and Frontness as main predictors fail to significantly improve the fit ($\chi^2 = 1.02, \text{df} = 1,\ p = 0.31$ for Place and $\chi^2 = 0.31, \text{df} = 1,\ p = 0.58$ for Frontness), which is why they are excluded from the final model. The final model thus includes only the intercept and a random intercept for Subject. The estimate of the intercept is $\beta = 0.32, = 2.02,\ p < 0.05$. While it is not justified by AIC probably due to a small sample size, one might argue that the random intercept for Item should be included in our model to control for the potential undesired effect of different stimuli. If we introduce a random intercept for Item in the model (unjustified by AIC), the estimates are $\beta = 0.37, z = 1.8, p = 0.07$, where p is slightly higher than 0.05. In any case, the preference for stop-response is weak. Depending on the choice of the model, it is either a week significant effect or a clear trend. In other words, there exist a weak preference (note the small effect size) for stop-initial responses (i.e. simple unnatural) compared to fricative-initial responses (complex). Figure 6.7 illustrates this preference (for the purpose of demonstrating the effect of place of articulation, Figure 6.7 shows a model without the random intercept for Item that includes Place as a predictor).

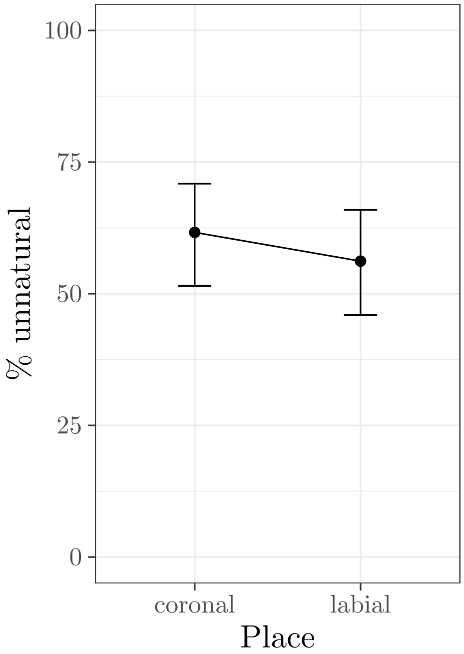


Figure 6.7: Estimates in % unnatural for the model with Place as a predictor (with 95% CIs).

6.4 Discussion

The experiment presented above tests which alternation subjects generalize more in an artificial grammar learning setting: a simple unnatural post-nasal devoicing or a complex alternation between voiceless stops that surface post-nasally and voiced fricatives that surface elsewhere. Subjects were exposed to the equal number of complex and simple unnatural alternations during the training phase. The training phase includes non-alternating voiceless stop- and fricative-initial nouns in addition to the alternating voiced stop- and fricative-initial nouns. Such experimental design prompts subjects into analyzing the two alternations as a simple alternation that involves only feature [±voice] and a complex alternation that involves two features: [±voice] and [±continuant].

When subjects are presented with ambiguous plural data with noun-inital devoiced stops that can go back either to a voiced stop or a voiced fricative, a preference for voiced stop emerges. This preference is weak (but significant under the model justified by AIC) and does not differ significantly across the two places of articulation.[94] These results suggest that it is easier to learn simple alternations compared to complex alternations, even if the simple alternation is an unnatural process operating in the exact opposite direction to universal phonetic tendencies. To my knowledge, this is the first report of an experiment that tests learnability of complex vs. simple alternations in which the simple alternation is unnatural.

These results also provide some evidence against the possibility that the unnatural processes such as PND are derived as opaque interactions. An unnatural process such as PND or IVD can be synchronically analyzed either as a simple process with an unnatural markedness constraint *NT or *VTV ranked above the faithfulness constraint or as an opaque interaction with only natural constraints that get reranked in different strata under the Stratal OT approach (Kiparsky 2000, 2015). Prickett (2017) proposes a potential analysis of PND in Stratal OT. The first stratum in the process is equivalent to the analysis of the complex alternation in our experiment (cf. Table 6.1 and 6.5; based on Prickett 2017). In the second stratum, the winning output

---

[94] It is possible that Place is not a significant predictor because of lower power of the experiment.

candidate for a voiced fricative in the input is a stop due to reordering of the *Z and Ident-IO(cont) constraints.

Table 6.5: Tableaux deriving the complex alternation tested in the experiment.

Stratum 1

/NZ/	*[+nas][+cont]	*D	Ident-IO(voi)	Ident-IO(cont)	*Z
NZ	*!				*
NS	*!		*		
ND		*!		*	
☞ NT			*	*	

/Z/	*[+nas][+cont]	*D	Ident-IO(voi)	Ident-IO(cont)	*Z
☞ Z					
D				*!	

Stratum 2

/Z/	*Z	Ident-IO(cont)
Z	*!	
☞ D		*

Our experiment tests two alternations: the complex alternation that is derivable from Stratum 1 (Table 6.5) and the simple unnatural alternation that requires the two strata (Stratum 1 and 2) under the assumption that unnatural processes are opaque and derivable with natural constraints only. Computational models (Jarosz 2016, Nazarov and Pater 2017) as well as experimental work (Prickett 2018) suggest that opaque processes are more difficult to learn than non-opaque processes. It is likewise reasonable to assume that a process that requires constraint reranking in the second stratum will be more difficult to learn simply because constraint reranking introduces an extra step that might require additional cognitive effort. If our subject were able to analyze the simple alternation as opaque interaction involving constraint reranking, we would expect the unnatural process to be more difficult to learn than the complex one that does not require constraint reranking in Stratum 2. Because this is not the case, we can conclude that speakers do not analyze PND as an opaque process with constraint reranking, at least not in the experimental design presented in this dissertation.

I have so far assumed that the difference in learnability stems from a different degree of complexity between the two tested processes. It is, however, also possible that the preference for learnability of the unnatural process is not caused by complexity of the competing alternation, but by perceptual distance in line with the P-map hypothesis (Steriade 2001). In other words, it is possible that the complex alternation is preferred because perceptual distance between T and D is smaller than between T and Z. The results of the experiment presented here remain relevant also under the perceptual approach: subjects prefer an alternation that incurs smaller perceptual distance, even if such an alternation operates in the unnatural direction. While this question does not have a direct implication for my framework, it is a matter worthy of further pursuit.

It is possible that the results of the experiment do not show learnability differences, but a more general dispreference against voiced fricatives (an anti-fricative bias). In English and many other languages, this anti-fricative bias is also conflated with frequency effects: voiced stops (less marked) are more frequent than corresponding voiced fricatives (more marked). It is thus possible that the observed preference for the simple unnatural alternation emerges not due to learnability, but because subjects simply favor voiced stops compared to voiced fricatives. This preference can either be due to voiced stops being less marked or more frequent in the system (Mines et al. 1978, although in Hayden 1950 /v/ is more frequent than /b/). This would mean that there exist no differences in learning of complex vs. simple unnatural alternations and that the preference for the latter results from a general anti-fricative bias. Even if this is the case, the results of the experiments are still relevant for our model. In this case, the application of the third sound change in the Blurring Process would be explained as a general avoidance of voiced fricatives: this dispreference could in time result in the third sound change of the Blurring Process. It is difficult to explain under this assumption, however, why fricatives occlude to stops significantly more frequently when this change simplifies an alternation (compared to cases in which it does not, see Section 4.4.2).[95]

These results combined with the probabilistic model of typology within the CB approach provide grounds for disambiguating AB and CB influences on typology. The main advantage of the framework presented in this dissertation is its ability to control for AB and CB influences. As already discussed in Section 4.4.2 and 4.4.1, the CB predicts decreased probabilities for every additional sound change that is required for an alternation to arise. Also, there exists a significant difference in frequencies of the last sound change (Z > D) in the Blurring Process that leads to PND and the same sound change operating independently. In other words, CB predicts the last sound change to be significantly less frequent than it actually is, estimated from our typological survey of PND. We can thus conclude that CB alone fails to explain the relative high frequency of the last sound change in the Blurring Process.

My experiment confirms what was predicted using the BSC technique: the last sound change in the Blurring Process is likely *catalyzed* by AB and its higher frequency cannot be attributed to CB. In other words, I experimentally support the claim that avoidance of complex alternations is primarily due to AB in these types of cases (when CB influences are controlled for).

---

[95] 95. The only case in which the experimental results are not relevant is if there were no learnability differences between the complex and simple unnatural alternation and if subjects were simply frequency-matching: they would prefer voiced stops not due to learnability or fricative bias, but because voiced stops are more frequent in English than voiced fricatives. It is unclear, however, what would explain the significantly higher rate of application of the third sound change in the Blurring Process (compared to the same change operating independently).

On the other hand, combining statistical modeling of diachronic developments with experimental results also allow us to control for AB influences when testing the predictions of CB. As already mentioned, experiments that tested learning of natural and unnatural patterns of PNV vs. PND, IVV vs. IVD, and FV vs. FD fail to yield positive results (Seidl et al. 2007, Do et al. 2016, Glewwe 2017; Glewwe et al. 2018). In other words, the AB approach would predict no typological differences between the natural and unnatural pairs of alternations. Typological surveys of these processes show substantial differences in frequencies between the natural and unnatural alternations. I argued in Chapter 4 that CB estimated with the BSC technique predicts these differences with a relatively high accuracy (Table 4.3).

We can therefore conclude that the rarity of unnatural alternations compared to natural ones is most likely influenced by CB and not by AB. While similar conclusions were suggested by surveys of experimental work (Moreton and Pater 2012a,b), most of these studies fail to control for one influence when testing the other.

The proposed framework has implications for our initial discussion (in Chapter 1) on the relationship between phonological GRAMMAR and observed typology. A combination of experimental and statistical modeling suggest that dispreference against unnatural processes (as defined in (2)) is not part of the universal phonological GRAMMAR and should therefore not be encoded in the grammar design. Dispreference against complex alternations, on the other hand *is* part of the universal GRAMMAR.

In Chapter 5, I propose a model of typology that has the ability to encode precisely these relations. We encode that dispreference against unnatural processes is not part of the grammar by admitting unnatural constraints into the CON and by not differentiating prior variance of the unnatural and natural constraint pairs. At the same time, we can encode that unnatural processes are typologically rare due to CB by differentiating prior Historical Weights.

6.5 Catalysis

Results of the present experiment have another implication: they suggest a potential answer to the question of how exactly the Analytic Bias influences the observed typology. This question has long been a central topic of historical and synchronic discussions in phonology. One of the objections against the AB approach to typology has been that more or less any alternation can be learned given enough exposure (Moreton and Pater 2012a,b). Human infants with no speech disabilities are able to acquire language and reproduce it with more or less perfect faithfulness to the input if they are exposed to primary linguistic data by the time of the critical period (cf. Dodd et al. 2003). It is thus not trivial to show how learnability preferences result in observed typology.

Morley (2015) and Stanton (2016c) propose models that explain how learnability and accidental gaps in phonological data interact to result in the observed typology. Both models, however, involve phonological "restructuring" (Kiparsky 1965) or rule inversion, rather than a gradual sound change based on phonetic precursors that results from phonologization. Less is known about how AB influences typology without rule inversion. Moreton (2008) proposes a model claiming that learnability discriminates between phonetic precursors that do get phonologized and those that do not. This proposal faces a challenge because it is difficult to distinguish the effect of learnability from the effect of perception (Yu 2011). Finally, numerous models exist that simulate diachronic development computationally, especially within the Exemplar Theory and related approaches (Pierrehumbert 2001, Wedel 2006, Kirby and Sonderegger 2013, 2015).

While these approaches primarily focus on modeling initiation and propagation of sound change based on noisy transmission and phonetic biases, it should in principle be possible to introduce learning biases in the speaker-learner loop. The problem with such an approach, however, is that articulatory or perceptual phonetic forces often align perfectly with learnability preferences (see also Chapter 4 above). It is thus difficult to find empirical evidence that would warrant introducing the learning bias into the model when there is empirical support in favor of the phonetic biases.

Historical and experimental data presented here point to a potential mechanism of how Analytic Bias might directly influence phonological typology. I have argued in Chapter 3 that unnatural processes arise through a specific combination of sound changes called the Blurring Process. After the first two sound changes in the Blurring Process operate, the resulting synchronic alternation is complex and involves voiced fricatives surfacing in the elsewhere position and voiceless stop in the post-nasal position. Example (42) illustrates this alternation.

(42) $\text{Z} \rightarrow \text{T} \ / \ \text{N}\_\_$

At this point, articulatory factors most likely cause phonetic variation, where voiced fricatives are occasionally produced with occlusion. As already, discussed in Section 3.3, fricatives require a greater level of articulatory precision. Deviation form articulatory targets with a higher level of precision can result in occlusion of fricatives (Ladefoged and Maddieson 1996:137). Articulatory forces thus cause phonetic variation between voiced fricatives and voiced stops. The variation is illustrated in (43).

(43) $\text{D} \sim \text{Z} \rightarrow \text{T} \ / \ \text{N}\_\_$

In the initial stages, this variation that cross-linguistically arises from automatic articulatory factors is expected to be highly skewed towards the faithful, in our case, fricative articulation. However, the learnability preference for simple (albeit unnatural) alternation favors the variant with closure. Precisely such preference is confirmed by my experimental design in this chapter. Despite the preference for unnatural alternations vs. complex being weak, over time this preference can result in reversal of the skewed variation. When speakers are faced with ambiguous surface forms, they generalize the unnatural variant with closure more than the complex variant with frication. Over time, this higher rate of generalization due to learnability (Analytic) bias can results in catalysis of the sound change that occludes fricatives and consequently in avoidance of complex alternations.

In fact, because Channel Bias alone cannot explain the high frequency of operation of the last sound change in the Blurring Process, I argue that this gradual learning preference for stop articulation operating on gradient phonetic variation is precisely what catalyzes occlusion of voiced fricatives in the development of PND. That catalysis of the last sound change in the Blurring Process proceeds through phonetic variation is also strongly suggested by Tarma Quechua data. Lexical variation between the fricative and stop articulations is still attested and phonetically confirmed in TQ (see Section 2.4.2). It is thus possible that TQ shows a midpoint in precisely the variation that I suggest is the condition for catalysis.

One might argue that the last sound change, Z > D, is simply a resolution of a higher degree of markedness of voiced fricatives (compared to voiced stops; see a discussion on anti-fricative bias potentially affecting experimental results in Section 6.4 above). That the last sound change in the Blurring Process is catalyzed by learning biases and not simply a regular markedness resolution is strongly suggested by the fact that sound change Z > D in the Blurring Process where it resolves a complex alternation operates significantly more frequently than independently in cases in which it does not resolve complexity ($p < 0.0001$, see Section 4.4.2). In other words, if catalysis is simply a markedness resolution, we would expect equal rates of application in cases in which it does not and in cases in which it does resolve complexity.

The advantage of catalysis is that it can explain how AB influences observed typology when controlling for the CB influences. Additionally, catalysis explains how AB influences the typology outside of the scope of language acquisition. We know that at least a subset of sound changes occurs in adult population (Labov 1994, Bybee 2001, Pierrehumbert 2003, among others; see Wedel 2006). The experiment presented in this dissertation is especially relevant to this discussion because it tests learnability of unnatural and complex processes in an adult population.

Page intentionally left blank.

Chapter 7

Conclusions

This dissertation proposes a framework for estimating and disambiguating influences of Analytic Bias and Channel Bias on phonological typology.

First, I propose a new division of alternations with respect to naturalness: processes traditionally labeled as "unnatural" are subdivided into unmotivated and unnatural. Unnatural processes are those that operate against universal phonetic tendencies. I then present a typological survey of unnatural categorical and gradient processes that target the feature [±voice]. Based on a statistical and phonetic analysis of Tarma Quechua and Berawan dialects, I argue that not only categorical, but also gradient unnatural processes are attested and show signs of productivity. I then focus on one of the most prominent cases of unnatural alternations and sound changes, PND. By collecting all known cases of this process, I show that common patterns emerge, yielding the conclusion that PND is not only in Tswana and Shekgalagari (Hyman 2001), but in all thirteen cases the result of three natural sound changes. This conclusion allows us to maintain the long-held position that sound change cannot operate against a universal phonetic tendency. I provide new crucial evidence from Sogdian and Yaghnobi that historically confirms the reconstructed development. On this basis, I present a new diachronic model, the Blurring Process, which can serve as a strategy for explaining unnatural phenomena historically. The Blurring Process states that unnatural processes always arise through a combination of three sound changes where the first sound change creates complementary distribution, the second sound change targets a changed or unchanged subset in the complementary distribution, and finally, the third sound change blurs the original complementary distribution. I argue that the Blurring Process approach is superior to the hypercorrection approach and that the Blurring Process outperforms other approaches in explaining not only PND, but also other categorical and gradient unnatural processes (such as IVD). The Blurring Process provides grounds to define a Minimal Sound Change Requirement (MSCR) and consequently a new model of typology within the CB approach. I argue that minimally three sound changes are required for an unnatural process to arise and minimally two for an unmotivated process.

To estimate the Channel Bias influences on typology, I introduce the notion of Historical Probabilities of Alternations building on the Blurring Process and MSCR. I demonstrate that we can estimate the Channel Bias contribution to phonological typology for any given alternation by identifying the number of sound changes required for that alternation (MSCR) and by estimating the respective probabilities of those sound changes. In Chapter 4, I propose a method for estimating Historical Probabilities that I label Bootstrapping Sound Changes. First, I give a detailed description of the statistical model and discuss its assumptions, properties of the sample, and implementation. The dissertation also includes functions in the R Statistical Software language for performing the BSC analysis (Appendix A.1). Several applications of the BSC technique are presented. For any synchronic alternation, both attested and unattested, the BSC technique estimates its Historical Probability from the number of sound changes the alternation requires and their respective probabilities. In other words, the BSC technique quantifies predictions of the Channel Bias approach to typology that can be compared to the actual observed synchronic typology. BSC also allows for inferential statistical tests comparing the Historical Probabilities of any two alternations. The BSC technique additionally predicts (un)attestedness of an alternation in a given sample. Alternations are either predicted to be attested more than once in a given sample or we predict an alternation to be possibly attested or unattested.

I argue that comparing Historical Probabilities with the observed typology yields new insights for the discussion on AB vs. CB influences on typology. First, I show that CB estimated by BSC predicts the observed typology with relatively high accuracy. This is especially true for differences between natural and unnatural alternations, which pose a problem for the AB approach to typology. To my knowledge, BSC makes the most accurate typological predictions currently possible within the CB approach. For example, no other proposals known to me predict that unnatural alternations will be substantially less frequent than natural alternations and at the same time predicts some unnatural alternations will be significantly less frequent than other unnatural alternations, a situation that is substantiated by the observed typology. Finally, BSC identifies crucial mismatches in predictions between the AB and CB approaches that provide new information for disambiguating AB vs. CB influences on typology. Both AB and CB predict that complex alternations will be less frequent, but within CB this prediction can be violated in the case of unnatural vs. unmotivated alternations.

The predictions of the proposed model have theoretical implications. First, I argue that the existence of unnatural processes poses a challenge to the assumption that the universal constraint inventory (Con) excludes unnatural constraints. In Chapter 5, I show that weighted-constraint frameworks with restricted Con predict that unnatural feature values in a given environment are always more frequent than the natural values — the so-called Natural Gradience Bias. While this prediction is valuable in many respects, I show that at least Tarma Quechua and Berawan dialects violate it. In other words, the data presented in Chapter 2 suggest that both categorical OT and weighted-constraint frameworks with restricted Con undergenerate as they fail to derive the attested categorial and gradient unnatural processes.

To address the undergeneration problem, I propose a formal model of typology that derives unnatural processes and at the same time keeps typological predictions of constraint-based frameworks (in Section 5.2). I adopt Wilson's (2006) MaxEnt approach to encoding learnability differences by differentiating prior variance of constraints. To encode that some processes are typologically rare due to CB, on the other hand, I introduce a diachronic metric: Historical Weights of constraints that are directly derivable from Historical Probabilities. This model derives all attested patterns (solves undergeneration), keeps typological predictions, and encodes that some typological restriction are due to learnability differences (AB) and others due to historical factors (CB). I also outline a potential solution for the TMTFS problem within this framework.

One of the main advantages of the proposed framework of combining diachronic statistical modeling with experimental work is the ability to disambiguate the AB and CB factors. In other words, instead of arguing in favor or against one or the other approach, phonological research should attempt to disambiguate the two influences and quantitatively estimate the relative contribution of one or the other to the observed typology. The main challenge that the attempts to disambiguate the two factors face is the fact that empirical evidence often supports both approaches equally well: existing studies testing predictions of AB and CB fail to control for the influence of one approach when testing the other. The BSC technique combined with experimental work allows us to do precisely this: I identify mismatches in predictions between the AB and CB approaches, which allows me to test the AB influences while controlling for the CB factor and vice versa.

In the case of natural-unnatural alternation pairs that target feature [±voice], we can assume that the AB influences are controlled for, as no differences have been observed in several studies that tested learnability of such pairs. I argue that the CB approach estimated by the BSC technique derives the observed typology with a relatively high accuracy. This allows me to conclude that typological rarity of unnatural processes that target feature [±voice] is likely not part of universal phonological GRAMMAR, but rather emergent from the transmission of language.

To control for the CB influences, on the other hand, I test learnability differences of those processes which CB predicts would operate with frequencies lower than observed. For example, I show that the last sound change in the Blurring Process that turns a complex alternation into an unnatural one operates with higher frequency than predicted by CB. This suggest that dispreference against complex processes is likely influenced by AB, even when we control for CB.

Chapter 6 presents an experiment that tests learnability of a complex and a simple unnatural alternation (Stages 3 and 4 of the Blurring Process that leads to PND). To my knowledge, this is the first artificial grammar learning experiment that tests learnability of a simple vs. complex alternation where the simple alternation is unnatural, as well as the first experiment that explicitly controls for the CB factor. The results suggest that the simple alternation is easier to learn than the complex alternation despite the simple alternation being unnatural. Experimental data thus confirms what has been identified with the BSC technique: avoidance of complex alternations is primarily influenced by AB and should therefore be encoded as part of universal phonological GRAMMAR. Additionally, I outline a model for explaining how AB influences might result in observed typology that I call catalysis.

This dissertation aims to shift the discussion in phonological typology from arguing in favor or against the Analytic Bias and Channel Bias approach to estimating both influences quantitatively by proposing a framework that combines statistical modeling of diachronic developments with experimental work. The dissertation focuses on the feature [±voice]. Application of the BSC technique on new alternations combined with new experimental data has the potential to yield further information for this long-standing discussion in phonology. Applying this method to different alternations and different features might yield different results. In any case, further information on what aspects of observed typology are primarily emergent from diachronic factors and what aspects truly reflect universal phonological GRAMMAR are desirable and might help us better understand the phonological aspect of language faculty.

The framework outlined in this dissertation calls for a number of future research directions. First, further unnatural alternations have been reported in the literature. Historical and experimental treatment of these processes have the potential to yield further insights into the discussion of the different influences on phonological typology. Especially desired would be fieldwork experiments on productivity of unnatural gradient phonotactics in languages such as Tarma Quechua and Berawan dialects.

As already mentioned, applying the BSC technique and combining it with experimental work on other features might yield different results from the ones obtained in this dissertation, which is limited to feature [±voice]. Regardless of the content of the results, applying the techniques to other features and processes should give us a better understanding of which types of processes are primarily influenced by learning and which types are emergent from phonetic forces in transmission of language. The more information of this kind we have, the better our understanding of what kinds of processes are truly favored by the universal phonological GRAMMAR. Once a better typology of processes that are influenced by learning emerges, we can test the correlation between these processes with perceptual and articulatory mechanisms. Finally, the new information should allow us to design more accurate computational models of learning in phonology that would exclude historical aspects and model only processes that are part of universal phonological GRAMMAR.

Page intentionally left blank.

Appendix A

BSC

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 (22), (23), (24), and (25). 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.

```
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))
  } else {
      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 (22), (23), (24), (25), and (26). 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 $\Delta \mathrm{P}_\chi$ 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.

```r
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)

#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.

```
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.

```
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")
}
```

A.5 plot.bsc()

The function *plot.bsc()* takes the output of *bsc()* as input and plots the distribution of bootstrap replicates with the observed Historical Probability of the process (solid line) and 95% BC_a CI (dashed line), calculated with the *boot.ci()* function from the *boot* package. The plotting is based on the *ggplot2* package (Wickham 2009). An optional argument *Alternation* allows for the change of the name of the alternation in the legend.

```
plot.bsc <- function (bsc.alt, Alternation = c("Alternation")) {
  library(ggplot2)
  bsc.ci.alt <- boot.ci(bsc.alt, type = "bca")
```

```r
  bsc.alt.df <- data.frame(bsc.alt$t)
  bsc.alt.df$name <- Alternation
  names(bsc.alt.df) <- c("boot", "Alternation")
  boot.plot <- ggplot(bsc.alt.df, aes(boot, fill = Alternation)) + geom_density(alpha = 0.5)  +
    geom_vline(xintercept = bsc.alt$t0, colour="red",linetype = "solid") +
    geom_vline(xintercept = bsc.ci.alt$bca[4],
               colour="red", linetype = "dashed") +
    geom_vline(xintercept = bsc.ci.alt$bca[5],
               colour="red", linetype = "dashed") +
    theme_bw() + xlab("Px in %") + ylab("")
  return(boot.plot)
}

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

pnd <- bsc(pnd.counts, pnd.surveyed)
plot.bsc(pnd, alternation = "PND")
```

A.6 plot.bsc2()

The function *plot.bsc2()* takes the output of *bsc()* as its input (two alternations) and plots the distribution of bootstrap replicates with the observed Historical Probability of the process (solid line) and 95% BC$_a$ CI (dashed line), calculated with the *boot.ci()* function from the *boot* package for each alternation. The plotting is based on the *ggplot2* package (Wickham 2009). An optional argument *Alternation* allows for the change of the name of the two alternations in the legend. Note that *plot.bsc2()* does not plot bootstrap replicates of the difference between two Historical Probabilities, but rather bootstrap replicates of Historical Probabilities of each of the two alternations. To plot the bootstrap replicates of the difference between two Historical Probabilities, apply *plot.bsc()* to the output of *bsc2*.

```
plot.bsc2 <- function (bsc.alt1, bsc.alt2, Alternation = c("Alternation␣1","
    Alternation␣2")) {
  library(ggplot2)
  bsc.ci.alt1 <- boot.ci(bsc.alt1, type = "bca")
  bsc.ci.alt2 <- boot.ci(bsc.alt2, type = "bca")
  bsc.alt1.df <- data.frame(bsc.alt1$t)
  bsc.alt2.df <- data.frame(bsc.alt2$t)
  bsc.alt1.df$name <- Alternation[1]
  bsc.alt2.df$name <- Alternation[2]
  names(bsc.alt1.df) <- c("boot", "Alternation")
  names(bsc.alt2.df) <- c("boot", "Alternation")
  bsc.alt.df<-rbind(bsc.alt1.df, bsc.alt2.df)
  boot.plot <- ggplot(bsc.alt.df, aes(boot, fill = Alternation)) +
    geom_density(alpha = 0.5)  +
    geom_vline(xintercept = bsc.ci.alt1$bca[4],
               colour = "red", linetype = "dashed") +
    geom_vline(xintercept = bsc.ci.alt1$bca[5],
               colour = "red", linetype = "dashed") +
    geom_vline(xintercept = bsc.alt1$t0,
               colour = "red", linetype = "solid") +
    geom_vline(xintercept = bsc.ci.alt2$bca[4],
               colour = "turquoise4", linetype = "dashed") +
    geom_vline(xintercept = bsc.ci.alt2$bca[5],
               colour = "turquoise4", linetype = "dashed") +
```

```
    geom_vline(xintercept = bsc.alt2$t0,
                colour = "turquoise4", linetype = "solid") +
    theme_bw() + xlab("Px␣in␣%") + ylab("")
  return(boot.plot)
}

#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)

plot.bsc2(pndfv)
```

Appendix B

Experimental stimuli

Table B.1: Filler words.

		Fillers	
#__	**Harm.**	**Sg.**	**Pl.**
[l]	[+fr]	len	enlen
		lino	enlino
	[−fr]	lor	onlor
		luru	onluru
[r]	[+fr]	rel	enrel
		rinu	enrinu
	[−fr]	ras	onras
		rolo	onrolo
[j]	[+fr]	yim	enyim
		yeni	enyeni
	[−fr]	yam	onyam
		yalu	onyalu

Table B.2: Nouns with initial voiceless sts and fricatives in the training phase.

Place	**Voiceless** **#__**	**Harm.**	**Sg.**	**Pl.**
Labial	[−cont]	[+fr]	pina	empina
			pimi	empimi
		[−fr]	poro	omporo
	[+cont]	[+fr]	fini	emfini
			fima	emfima
		[−fr]	fura	omfura
			folo	omfolo
Coronal	[−cont]	[+fr]	telo	entelo
			tina	entina
		[−fr]	taru	ontaru
	[+cont]	[+fr]	seno	enseno
			sila	ensila
		[−fr]	soro	onsoro
			sanu	onsanu

Table B.3: Nouns with initial voiced stops and fricatives in the training phase.

Place	**#__**	**Voiced Harm.**	**Sg.**	**Pl.**
Labial	[−cont]	[+fr]	bila	empila
			bera	empera
			bilo	empilo
			bema	empema
		[−fr]	bula	ompula
			balu	ompalu
			bora	ompora
			bune	ompune
	[+cont]	[+fr]	vila	empila
			vemo	empemo
			vira	empira
			vela	empela
		[−fr]	vulo	ompulo
			varu	omparu
			vona	ompona
			vule	ompule
Coronal	[−cont]	[+fr]	dilo	entilo
			diri	entiri
			delo	entelo
			dema	entema
		[−fr]	dule	ontule
			doru	ontoru
			dale	ontale
			duna	ontuna
	[+cont]	[+fr]	zila	entila
			zira	entira
			zemo	entemo
			zeni	enteni
		[−fr]	zulo	ontulo
			zaru	ontaru
			zole	ontole
			zune	ontune

Table B.4: Nouns in the test phase.

		Test		
Place	**Harm.**	**Pl.**	**Sg. 1**	**Sg. 2**
Labial	[+fr]	empene	bene	vene
		empeyo	beyo	veyo
	[−fr]	ompaya	baya	vaya
		ompara	bara	vara
Coronal	[+fr]	entele	dele	zele
		entiwo	diwo	ziwo
	[−fr]	ontami	dami	zami
		ontawo	dawo	zawo

References

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

Adelaar, Willem F. H., and Pieter C. Muysken 2004. *The Languages of the Andes.* Cambridge: Cambridge University Press.

Albright, Adam. 2009. Feature-based generalization as a source of gradient acceptability. *Phonology* 26(1): 9-41.

Ali, Latif, Ray Daniloff, and Robert Hammarberg. 1979. Intrusive stops in nasal-fricative clusters: An aerodynamic and acoustic investigation. *Phonetica* 36: 85-97.

Alwan, Abeer, Jintao Jiang, and Willa Chen. 2011. Perception of place of articulation for plosives and fricatives in noise. *Speech Communication* 53(2): 195-209.

Anderson, Stephen R. 1981. Why phonology isn't "natural". *Linguistic Inquiry* 12(4): 493-539.

Anttila, Arto. 2008. Gradient phonotactics and the Complexity Hypothesis. *Natural Language and Linguistic Theory* 26(4): 695-729.

Anttila, Arto and Giorgio Magri. 2018. Does MaxEnt Overgenerate? Implicational Universals in Maximum Entropy Grammar. In Proceedings of the 2017 Annual Meeting on Phonology. 1–12.

Archangeli, Diana and Douglas Pulleyblank. 1994. *Grounded Phonology.* Cambridge, MA: MIT Press.

Bach, Emmon and Robert T. Harms 1972. How do languages get crazy rules? *Linguistic Change and Generative Theory*, edited by Robert Stockwell and Ronald Macaulay. 1-21. Bloomington: Indiana University Press.

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

Barry, Martin and Ryad Teifour. 1999. Temporal patterns in Syrian Arabic voicing assimilation. *14th International Congress of Phonetic Sciences*, vol. 3, San Francisco, CA, 2429-2432.

Bartholomae, Christian. 1961. *Altiranisches wörterbuch.* Berlin: Walter de Gruyter.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1-48.

Becker, Michael, Andrew Nevins and Jonathan Levine. 2012. Asymmetries in generalizing alternations to and from initial syllables. *Language* 88(2): 231-268.

Becker, Michael and Jonathan Levine 2014. Experigen – an online experiment platform. Available at `http://becker.phonologist.org/experigen`.

Becker, Michael, Lauren Eby Clemens, and Andrew Nevins. 2017. Generalization of French and Portuguese Plural Alternations and Initial Syllable Protection. *Natural Language and Linguistic Theory* 35(2): 299-345.

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.

Beckers, Gabriel J. L. 2002. *rms equalize.* Praat script, available on March 31, 2018 at `http://www.gbeckers.nl/pages/praat_scripts/rms_equalize.praat_script`.

Beckman, Jill N. 1997. Positional Faithfulness, Positional Neutralisation and Shona Vowel Harmony. *Phonology* 14(1): 1–46.

Beddor, Patrice S. 2009. A coarticulatory path to sound change. *Language* 85(4): 785-821.

Beguš, Gašper. 2016. A Natural Origin for Unnatural Gradient Phonotactics. Talk at the 13th Old World Con- ference in Phonology, Workshop on variation in phonology. Budapest, Hungary. January 13–16, 2016.

Beguš, Gašper. 2018. Segmental phonetics and phonology. To appear in *The Oxford Handbook of Languages of the Caucasus*, edited by Maria Polinsky.

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.

Berkley, Deborah. 2000. Gradient Obligatory Contour Principle Effects. Ph.D. dissertation, Northwestern University.

Berkovits, Rochele. 1993. Utterance-final lengthening and the duration of final-stop closures. *Journal of Phonetics* 21: 479-489.

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.

Blevins, Juliette and Andrew Garrett. 1998. The origins of consonant-vowel metathesis. *Language* 74(3): 508-556.

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. 1974. A Murik vocabulary, with a note on the linguistic position of Murik. *The Sarawak Museum Journal* 22(43): 153-189.

Blust, Robert. 1992. The Long Consonants of Long Terawan. *Bijdragen tot de Taal-, Land- en Volkenkunde* 148 (3-4): 409-427.

Blust, Robert. 2002. Kiput Historical Phonology. *Oceanic Linguistics* 42(2): 384-438.

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

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

Blust, Robert. 2016. Austronesian against the world: where the P-map ends. Talk and handout presented at the 42nd meeting of Berkeley Linguistic Society. Feb. 5-6, 2016. Available on March 20, 2018 at: `www.researchgate.net/publication/294732706`.

Blust, Robert. 2017. Odd conditions: Context-sensitive sound change in unexpected contexts. *Journal of Historical Linguistics* 7(3): 322-371.

Boersma, Paul and David Weenink. 2016. *Praat: doing phonetics by computer* [Computer program]. Version 5.4.6, retrieved 21 Februray 2015 from www.praat.org.

Boersma, Paul and Joe Pater. 2013. Convergence Properties of a Gradual Learning Algorithm for Harmonic Grammar. In *Harmonic Grammar and Harmonic Serialism*, edited by John McCarthy and Joe Pater. 389-434. London: Equinox Press.

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.

Brockhaus, Wiebke. 1995. *Final Obstruent Devoicing in the Phonology of German.* Tübingen: Niemeyer.

Browman, Catherine P., and Louis Goldstein. 1992. *Articulatory phonology: An overview.* Phonetica 49: 155-180.

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.

Burkard, Monja and Kristina Dziallas. 2018. Final /d/ in the varieties of Madrid, Barcelona and Seville: Regional and stylistic variation. In *Proceedings of the Conference on Phonetics & Phonology in German-speaking Countries* 13, edited by Malte Belz, Christine Mooshammer, Susanne Fuchs, Stefanie Jannedy, Oksana Rasskazova, and Marzena Żygis. Berlin: Humboldt-Universität zu Berlin.

Burkhardt, Jürgen M. *The reconstruction of the phonology of Proto-Berawan.* Ph.D. Dissertation, Johann-Wolfgang-Goethe-Universität zu Frankfurt am Main.

Busà, M. Grazia. 2007. Coarticulatory nasalization and phonological developments: data from Italian and English nasal-fricative sequences. In *Experimental Approaches to Phonology*, edited by Maria-Josep Solé, Patrice Speeter Beddor, and Manjari Ohala. 155-174. Oxford: Oxford University Press.

Bybee, Joan. 2001. *Phonology and Language Use.*Cambridge: Cambridge University Press.

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.

Campbell, Lyle. 2006. Areal linguistics: A closer scrutiny. In *Linguistic Areas: Convergence in Historical and Typological Perspective*, edited by Yaron Matras, April McMahon, and Nigel Vincent. 1-31. London: Palgrave Macmillan.

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

Carpenter, Angela C. 2005. Acquisition of a natural vs. an unnatural stress system. In *Papers from the 29th Boston University Conference on Language Development (BUCLD 29)*, edited by A. Brugos, M. R. Clark-Cotton, and S. Han. 134-143. Somerville, MA: Cascadilla Press.

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.

Carter, Richard T. Jr. 1974. Teton Dakota Phonology. Ph.D. Dissertation, University of Manitoba.

Catford, John C. 1972. Labialization in Caucasian languages, with special reference to Abkhaz. In *Proceedings of the Seventh International Congress of Phonetic Sciences*, edited by André Rigault and René Charbonneau. 679-682. The Hague: Mouton.

Catford, John C. 1974. 'Natural' sound changes: Some questions of directionality and diachronic phonetics. 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.

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.

Chomsky, Noam and Morris Halle. 1968. The Sound Pattern of English. New York: Harper & Row.

Coetzee, Andries W. and Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory*, 26(2), 289-337.

Coetzee, Andries W. and Joe Pater. 2011. The place of variation in phonological theory. In *The Handbook of Phonological Theory.* Second Edition, edited by John Goldsmith, Jason Riggle and Alan C. L. Yu. 401-434. Oxford: Blackwell.

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.

Coetzee, Andries W., Susan Lin, and Rigardt Pretorius. 2007. Post-nasal devoicing in Tswana. In *Proceedings of the 16th International Congress of Phonetic Sciences*, edited by Jřgen Trouvain and William J. Barry. 861-864. Saarbrücken.

Cohn, Abigail C. 2006. Is there gradient phonology? In *Gradience in Grammar: Generative Perspectives*, edited by G. Fanselow, C. Féry, and M. Schlesewsky. 25-44. Oxford: Oxford University Press.

Creider, Chet A. 1968. *A morphological sketch of Tarma Quechua.* MA Thesis, University of Minnesota.

Davidson, Lisa. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54: 35-50.

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

de Lacy, Paul. 2002. The formal expression of markedness. Ph.D. Dissertation, University of Massachusetts, Amherst.

de Lacy, Paul 2004. Markedness conflation in Optimality Theory. *Phonology* 21(2): 145-199.

de Lacy, Paul 2006. *Markedness: Reduction and Preservation in Phonology.* 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.

Dixon, Robert M. W. 1977. *A grammar of Yidin.* Cambridge: Cambridge University Press.

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.

Do, Youngah and Adam Albright. 2016. Three biases for learning phonological alternations. Ms., MIT.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents.* Berlin: Foris Publications.

Dodd, Barbara, Alison Holm, Hua Zhu, and Sharon Crosbie. 2003. Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics* 17: 617-643.

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.

Donegan, Patricia J. and David Stampe. 2002. South-East Asian features in the Munda languages: Evidence for the analytic-to-synthetic drift of Munda. In *Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics.* 111-120.

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.

Fallon, Paul D. 1995. Synchronic and Diachronic Typology: The Case of Ejective Voicing. *Proceedings of the Twenty-First Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Historical Issues in Sociolinguistics/Social Issues in Historical Linguistics.* 105-116.

Fenwick, Rohan S. H. 2011. *A Grammar of Ubykh.* Munich: Lincom Europa.

Ferry, Marie-Paule. 1991. *Thesaurus tenda: dictionnaire ethnolinguistique de langues sénégalo-guinnéennes (bassari, bedik, konyagi).* Langues et cultures africaines, 13-15. Paris: Société des Etudes Linguistiques et Anthropologiques de France.

Fourakis, Marios and Robert Port. 1986. Stop epenthesis in English. *Journal of Phonetics* 14(2): 197-221.

Frisch, Stefan, Janet Pierrehumbert and Michael Broe. 2004. Similarity Avoidance and the OCP. *Natural Language and Linguistic Theory* 22(1): 179-228.

Fruehwald, Josef. 2016. The early influence of phonology on a phonetic change, *Language* 92(2): 376-410.

Gajdarov, R. I., A. G. Gjulmagomedov, U. A. Mejlanova, and B. B. Talibov. 2009. *Sovremennyj lezginskij jazyk*. Maxačkala: IJaLI DNC RAN.

Gallagher, Gillian. 2016. Distinguishing phonotactic restrictions and accidental gaps. Talk at the 2016 Annual Meeting on Phonology, University of Southern California. October 21 - 23, 2016.

Garcia, Guilherme. D. 2017. Weight gradience and stress in Portuguese. *Phonology* 34(1): 41-79.

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.

Goddard, Ives. 2007. Phonetically unmotivated sound change. In *Verba Docenti: Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues and Friends*, edited by Alan J. Nussbaum. 115-130. Ann Arbor: Beech Stave Press.

Goldwater, Sharon and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson and Östen Dahl (eds.), *Proceedings of the Workshop on Variation within Optimality Theory*. 111-120. Stockholm University.

Gouskova, Maria, Elizabeth Zsiga, and One Tlale Boyer. 2011. Grounded constraints and the consonants of Setswana. *Lingua* 121(15): 2120-2152.

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.

Guion, Susan G. 1996. Velar palatalization: Coarticulation, perception, and sound change. Ph.D. Dissertation, University of Texas at Austin.

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

Haggard, Mark. 1978. The devoicing of voiced fricatives. *Journal of Phonetics* 6(2): 95-102.

Hale, Mark and Charles Reiss. 2008. *Phonological Enterprise.* Oxford: Oxford University Press.

Hall, T. Alan. 2007. Segmental features. In *The Cambridge Handbook of Phonology*, edited by Paul de Lacy. 311-333. Cambridge: Cambridge University Press.

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.

Harris, Alice C. 2005. The challenge of typologically unusual structures. In *Morphology and Linguistic Typology: On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4), Catania, 21-23 September 2003*, edited by G. Booij, E. Guevara, A. Ralli, S. Sgroi, and S. Scalise. 277-284. Available on July 7, 2015 at: `morbo.lingue.unibo.it/mmm/`.

Harris, Alice C. 2008. On the explanation of typologically unusual structures. In *Language Universals and Language Change*, edited by Jeff Good. 54-76. Oxford: Oxford University Press.

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.

Hayden, Rebecca E. 1950. The relative frequency of phonemes in General-American English. *Word* 6(3): 217-223.

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. 2016. Varieties of Noisy Harmonic Grammar. Handout at the 2016 Annual Meeting on Phonology, University of Southern California. October 21 - 23, 2016.

Hayes, Bruce. 2017. Varieties of Noisy Harmonic Grammar. *Proceedings of the 2016 Annual Meeting on Phonology.* 1-17.

Hayes, Bruce and Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39(3): 379-440.

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`

Hellberg, Staffan. 1978. Unnatural phonology. *Journal of Linguistics* 14(2): 157-177.

Hock, Hans H. 1991. *Principles of Historical Linguistics.* Berlin: De Gruyter Mouton.

Hoffmann, Karl, and Bernhard Forssman. 2004. *Avestische Laut- Und Flexionslehre.* Innsbruck: Institut für Sprachwissenschaft.

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.

Hurd, Conrad, and Phyllis Hurd. 1970. Nasioi verbs. *Oceanic Linguistics* 9: 37-78.

Hyman, Larry M. 1972. Nasals and nasalization in Kwa. *Studies in African Linguistics* 4: 167-206.

Hyman, Larry M. 1975. *Phonology: Theory and Analysis.* New York: Holt, Rinehart & Winston.

Hyman, Larry M. 1976. Phonologization. In *Linguistic studies presented to Joseph H. Greenberg*, edited by A. Juilland. 407-418. Saratoga, CA: Anna Libri.

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.

Itô, Junko, and Armin Mester. 2002. One phonology or many? Issues in stratal faithfulness theory. In The Phonological Society of Japan (ed.), *Phonological Studies 5.* 121-126. Tokyo: Kaitakusha.

Itô, Junko, and Armin Mester. 2003. *Japanese Morphophonemics: Markedness and Word Structure.* Cambridge, MA: MIT Press.

Iverson, Gregory. 1983. Voice alternations in Lac Simon Algonquian. *Journal of Linguistics* 19: 161-164.

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.

Jansen, Wouter. 2004. Laryngeal contrast and phonetic voicing: A Laboratory Phonology approach to English, Hungarian, and Dutch. Ph.D. Thesis, University of Groningen.

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.

Jarosz, Gaja. 2006. Rich Lexicons and Restrictive Grammars — Maximum Likelihood Learning in Optimality Theory. PhD dissertation, Johns Hopkins University. (ROA 884).

Jarosz, Gaja. 2016. Learning opaque and transparent interactions in Harmonic Serialism. In *Proceedings of the Annual Meetings on Phonology*, edited by Gunnar Ólafur Hansson, Ashley Farris-Trimble, Kevin McMullin, and Douglas Pulleyblank. 1-12.

Jesney, Karen. 2016. Positional constraints in Optimality Theory and Harmonic Grammar. In John McCarthy and Joe Pater (eds.), *Harmonic Grammar and Harmonic Serialism.* 176-220. Sheffield: Equinox.

Kaplan, Abby. 2008. Perceptual, Articulatory, and Systemic Influences on Lenition. Ms., University of California, Santa Cruz.

Kaplan, Abby. 2010. Phonology shaped by phonetics: The case of intervocalic lenition. Ph.D. Dissertation, University of California, Santa Cruz.

Katzir Cozier, Franz. 2008. The Role of Perception in Phonotactic Constraints: Evidence from Trinidad English. Ph.D. Dissertation, Massachusetts Institute of Technology.

Kenstowicz, Michael J. and Charles W. Kisseberth. 1977. *Topics in Phonological Theory.* New York: Academic Press.

Kent, Raymond D. and Kenneth L. Moll. 1969. Vocal tract characteristics of the stop consonants. *Journal of the Acoustical Society of America* 46(6): 1549-1555.

Kent, Raymond D., Patrick J. Carney, and Larry R. Severeid. 1974. Velar movement and timing: Evaluation of a model for binary control. *Journal of Speech Language and Hearing Research* 17(3): 470-488.

Keyser, Samuel J. and Kenneth N. Stevens. 2001. Enhancement revisited. In *Ken Hale: A Life in Language*, edited by Michael Kenstowicz. 271-291. Cambridge, MA: MIT.

Keyser, Samuel J. and Kenneth N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82(1): 33-63.

Kingston, John and Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70(3): 419-454.

Kingston, John. 2008. Lenition. *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, edited by Laura Colantoni and Jeffrey Steele. 1-31. Somerville, MA: Cascadilla Proceedings Project.

Kiparsky, Paul. 1965. Phonological Change. Ph.D. Thesis, MIT.

Kiparsky, Paul. 1971. Historical linguistics. In *A Survey of Linguistic Science*, edited by William O. Dingwall. 576-642. College Park: University of Maryland Linguistics Program.

Kiparsky, Paul. 1973. Abstractness, opacity, and global rules. In *Three Dimensions of Linguistic Theory*, edited by Osamu Fujimura. 57-86. Tokyo: TEC.

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. 2000. Opacity and cyclicity. *The Linguistic Review* 17(2-4): 351-366.

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.

Kiparsky, Paul. 2015. Stratal OT: A synopsis and FAQs. In *Capturing Phonological Shades Within and Across Languages*, edited by Yuchau E. Hsiao and Lian-Hee Wee. 2-44. Cambridge Scholars Publishing.

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.

Kirchner, Robert. 2001. *An Effort Based Approach to Consonant Lenition.* Outstanding Dissertations in Linguistics. New York, NY: Routledge.

Kosmidis, Ioannis. 2013. *brglm: Bias reduction in binomial-response Generalized Linear Models.* `ucl.ac.uk/~ucakiko/software.html`

Kümmel, Martin. 2006. Einführung ins Ostmitteliranische (Introduction to Eastern Middle Iranian). Available on March 30, 2018 at `https://www.academia.edu/30130317`.

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.

Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages.* Oxford: Blackwell.

Leben, William. 1973. Suprasegmental Phonology. PhD thesis, MIT.

Legendre, Géraldine, Antonella Sorace, and Paul Smolensky. 2006. The Optimality Theory—Harmonic Grammar connection. In Smolensky, Paul, and Géraldine Legendre. (eds.), *The harmonic mind: From neural computation to Optimality-Theoretic grammar*. Cambridge, MA: MIT Press.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the World*. Eighteenth edition. Dallas, TX: SIL International. Online version: www.ethnologue.com.

Liljencrants, Johan and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48(4): 839-862.

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.

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, Susan Guion, Susan Hura, Seung-Jae Moon, and Raquel Willerman. 1995. Is sound change adaptive? *Rivista di Linguistica* 7: 5-36.

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.

Lobel, Jason W. 2013. Southwest Sabah revisited. *Oceanic Linguistics* 52(1): 36-68.

Locke, John. 1983. *Phonological Acquisition and Change*. New York: Academic Press.

Lofqvist, Anders. 2006. Position variability in stop and fricative production. *Journal of Acoustical Society of America* 119: 3393.

Lombardi, Linda. 1999. Positional Faithfulness and Voicing Assimilation in Optimality Theory. *Natural Language & Linguistic Theory* 17(2): 267–302.

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

Magri, Giorgio. 2015. How to keep the HG weights non-negative: the truncated perceptron reweighing rule. *Journal of Language Modelling* 3(2): 345–375.

Makalela, Leketi. 2009. Harmonizing South African Sotho language varieties: Lessons from reading proficiency assessment. *International Multilingual Research Journal* 3(2): 120-133.

Martinet, André. 1950. Some problems of Italic consonantism. *Word* 6(1): 26-41.

McCarthy, John. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17: 207-263.

McCarthy, John. 2008. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26: 499-546.

Merrill, John. 2014. A historical account of the Fula and Sereer consonant mutation and noun class systems. Ms., University of California, Berkeley.

Merrill, John. 2015. Nasalization as a repair for voiced obstruent codas in Noon. *LSA Annual Meeting Extended Abstracts* 6:14. 1–5.

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.

Miller, George A. and Patricia Nicely. 1955. An Analysis of Perceptual Confusions among some English Consonants. *Journal of the Acoustical Society of America* 27(2): 338-352.

Mills, Frederick R. 1975. *Proto South Sulawesi and Proto Austronesian Phonology.* Ph.D. Dissertation, University of Michigan, Ann Arbor.

Mines, M. Ardussi, Barbara F. Hanson, and June E. Shoup. 1978. Frequency of cccurrence of phonemes in conversational English. *Language and Speech* 21(3): 221-241.

Möbius, Bernd. 2004. Corpus-based investigations on the phonetics of consonant voicing. *Folia Linguistica* 38(1-2): 5-26.

Moll, Kenneth L. and Raymond G. Daniloff. 1971. Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America* 50(2B): 678-684.

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.

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. 2011. From phonetics to phonology: learning epenthesis. In *Proceedings of the 47th Meeting of the Chicago Linguistics Society.* 137-151. Chicago, IL: University of Chicago.

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

Morley, Rebecca. L. 2014. Implications of an exemplar-theoretic model of phoneme genesis: A velar palatalization case study. *Language and Speech* 57(1): 3-41.

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.

Myers, Scott. 2010. Regressive voicing assimilation: Production and perception studies. *Journal of the International Phonetic Association* 40(2): 163-179.

Nazarov, Aleksei. 2008. *Stop voicing in Tarma Quechua.* Ms., Leiden University.

Nazarov, Aleksei, and Joe Pater. 2017. Learning opacity in Stratal Maximum Entropy Grammar. *Phonology* 34(2): 299-324.

Noorduyn, Jacobus. 2012/1955. The Bugis language. In *Bugis and Makasar: two short grammars*, edited and translated by Campbell Macknight. Based on Noorduyn's 1955 work. 33-55. Canberra: Karuda Press

Novák, L'ubomír. 2010. *Jaghnóbsko-český slovník s přehledem jaghnóbské grammatiky.* Prague: Univerzita Karlova v Praze, Filozofická fakulta.

Novák, L'ubomír. 2013. Problem of archaism and innovation in the Eastern Iranian languages. Ph.D. Dissertation, Charles University in Prague.

Novák, L'ubomír. 2014. Historical phonology of Yaghnōbīand Sogdian. Ms, DOI: 10.13140/RG.2.2.19663.28326.

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.1989. Sound change is drawn from a pool of synchronic variation. In *Language Change: Contributions to the study of its causes*, edited by L. E. Breivik and E. H. Jahr. 173-198. Berlin: Mouton de Gruyter.

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. 1997a. Aerodynamics of phonology. *Proceedings of 4th Seoul International Conference on Linguistics [SICOL]*, 11-15 Aug 1997. 92-97.

Ohala, John J. 1997b. Emergent stops. In *Proceedings of the 4th Seoul International Conference on Linguistics (SICOL)* 84-91. Seoul: Linguistic Society of Korea.

Ohala, John J. 2006. Speech aerodynamics. In *Encyclopedia of Language & Linguistics.* Second Edition, edited by Keith Brown. 684-689. Oxford: Elsevier.

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.

Ohala, John J. and Carol J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In *Speech communication papers*, edited by J. J. Wolf and D. H. Klatt. 89-92. New York: Acoustical Society of America.

Ohala, John. J., and M. Grazia Busà. 1995. Nasal loss before voiceless fricatives: a perceptually based sound change. Special issue on The Phonetic basis of Sound Change, edited by Carol A. Fowler. *Rivista di Linguistica* 7:125-144.

Ohala, John J. and Manjari Ohala. 1992. Phonetic universals and Hindi segment duration. In Ohala, J. J., Nearey, T., Derwing, B., Hodge, M., and Wiebe, G. (eds.) *Proceedings, International Conference on Spoken Language Processing*, Banff, 12-16 Oct 1992. Edmonton: University of Alberta. 831-834.

Ohala, John J. and Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In *Nasals, Nasalization, and the Velum*, edited by Marie K. Huffman and Rena A. Krakow. 225-249. San Diego, CA: Academic Press.

Onishi, Masayuki. 2011. *A grammar of Motuna.* Munich: Lincom Europa.

Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In *The Prosody-Morphology Interface*, edited by René Kager, Harry van der Hulst, and Wim Zonneveld. 310-343 . Cambridge: Cambridge University Press.

Pater, Joe. 2008. Gradient phonotactics in Harmonic Grammar and Optimality Theory. Ms., University of Massachusetts, Amherst.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33:999–1035.

Pater, Joe. 2012. Emergent systemic simplicity (and complexity). In J. Loughran and A. McKillen (eds.), *Proceedings from Phonology in the 21st Century: In Honour of Glyne Piggott. McGill Working Papers in Linguistics* 22(1): 1-19.

Pater, Joe and Paul Boersma. Convergence Properties of a Gradual Learning Algorithm for Harmonic Grammar. Manuscript. Ms., available at: `http://works.bepress.com/joe_pater/5/`.

Paul, Daniel, Elisabeth Abbess, Katja Müller, Calvin Tiessen, and Gabriela Tiessen. 2010. The ethnolinguistic vitality of Yaghnobi. In *SIL Electronic Survey Report 2010-017, May 2010.* SIL International. Available on August 23, 2017 at `http://www-01.sil.org/silesr/2010/silesr2010-017.pdf`.

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

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In *Frequency Effects and the Emergence of Lexical Structure*, edited by Joan L. Bybee and Paul J. Hopper. 137-157. Amsterdam: John Benjamins.

Pierrehumbert, Janet. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115–154.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt, and Michael Becker. 2010. Harmonic Grammar with Linear Programming: From linear systems to linguistic typology. *Phonology* 27: 77-117.

Prickett, Brandon. 2017. Post-nasal devoicing as opacity: A problem for natural constraints. In *Proceedings of the 35th West Coast Conference on Formal Linguistics*, edited by W. Bennett et al., 331-340. Somerville, MA: Cascadilla Proceedings Project.

Prickett, Brandon. 2018. Learning biases in opaque interactions. Poster at the 2018 meeting of *PhoNE.*

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.

Pucilowski, Anna. 2013. Topics in Ho Morphophonology and Morphosyntax. Ph.D. Dissertation, University of Oregon.

Puente Baldoceda, Blas. 1977. *Fonología del quechua tarmeño.* Lima: Universidad Nacional Mayor de San Marcos, Centro de Investigación de Lingüística Aplicada.

Pullum, Geoffrey K. 1976. The Duke of York gambit. *Journal of Linguistics* 12: 83-102.

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.

Quilis, Antonio. 1993. *Tratado De Fonología Y Fonética Españolas.* Biblioteca Románica Hispánica. III. Manuales 74. Madrid: Gredos.

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/.

Rankin, Robert L. 2001. On Dakotan pre-boundary and cluster phonology. Unpublished paper, presented at the Siouan and Caddoan Languages at the University of Chicago.

Rankin, Robert L., Richard T. Carter, and A. Wesley Jones. 1998. Proto-Siouan phonology and grammar. In *Papers from the 1997 Mid-America Linguistics Conference*, edited by Xingzhong Li, Luis López, and Tom Stroik. 366-375. Columbia: University of Missouri Press.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood, and Iren Hartmann (eds.). 2015. *Comparative Siouan Dictionary.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Available on May 5, 2018 at `csd.clld.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.

Recasens, Daniel. 2012. The phonetic implementation of underlying and epenthetic stops in word final clusters in Valencian Catalan. *Journal of the International Phonetic Association* 42(1): 65-90.

Redford, Melissa A. and Randy L. Diehl 1999. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America* 106(3): 1555-1565

Rensch, Calvin R., Carolyn M. Rensch, Jonas Noeb, and Robert S. Ridu. 2006. *The Bidayuh language: yesterday, today, and tomorrow.* Kuching, Sarawak: Dayak Bidayuh National Association.

Riggs, Stephen Return. 1893. *Dakota Grammar, Texts, and Ethnography.* Washington: Government Printing Office.

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.

Rothenberg, Martin. 1968. *The Breath-Stream Dynamics of Simple-Released-Plosive Production.* Basel: S. Karger.

Ruhlen, Merritt. 1975. *A Guide to Languages of the World.* Stanford, CA: Language Universals Project.

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

Salminen, Tapani. 1997. *Tundra Nenets Inflection.* Suomalais-ugrilaisen Seuran toimituksia = Mémoires de la Société finno-ougrienne 227. Helsinki.

Samuels, Bridget. 2009. *The Structure of Phonological Theory.* Ph.D. Dissertation, Harvard University.

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

Scheer, Tobias. How diachronic is synchronic grammar? Crazy rules, regularity, and naturalness. In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons. Oxford: Oxford University Press.

Shaw, Patricia A. 1980. *Theoretical Issues in Dakota Phonology and Morphology.* New York: Garland Publishing.

Shiraishi, Hidetoshi. 2006. Topics in Nivkh phonology. Ph.D. dissertation, Rijksuniversiteit, Groningen, Netherlands.

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.

Sims-Williams, Nicholas. 1987. Sogdian. In *Compendium Linguarum Iranicarum*, edited by Rüdiger Schmitt. 173-192. Wiesbaden: Ludwig Reichert Verlag.

Sirk, Ülo. (1983). *The Buginese Language.* Moscow: Nauka.

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

Smith, Caroline. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25(4): 471-500.

Smolensky, Paul. 1996. *The initial state and 'richness of the base' in Optimality Theory.* Rutgers Optimality Archive, 293.

Snoeren, Natalie, Pierre Hallé, and Juan Segui. 2006. A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34(2), 241-268.

Solé, Maria-Josep. 2012. Natural and unnatural patterns of sound change? In *The Initiation of Sound Change: Perception, Production, and Social Factors*, edited by Maria-Josep Solé and Daniel Recasens. 123-145. Amsterdam: John Benjamins.

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.

Stadler, Kevin, Richard A. Blythe, Kenny Smith, and Simon Kirby. 2016. Momentum in language change. *Language Dynamics and Change* 6: 171-198.

Stampe, David. 1973. *A Dissertation on Natural Phonology.* Ph.D. Dissertation, University of Chicago.

Stanton, Juliet. 2016a. Predicting distributional restrictions on prenasalized stops. *Natural Language & Linguistic Theory* 34(3): 1089-1133.

Stanton, Juliet. 2016b. Effects of allophonic vowel nasalization on NC clusters: A contrast-based analysis. In *Proceedings of the 46th Meeting of the North East Linguistic Society.* Volume 3, edited by Christopher Hammerly and Brandon Prickett. 193-206. Amherst, MA: GLSA.

Stanton, Juliet. 2016c. Learnability shapes typology: the case of the midpoint pathology. *Language* 92(4): 753-791.

Staubs, Robert. 2014. Computational modeling of learning biases in stress typology. Ph.D. dissertation, University of Massachusetts Amherst.

Stausland Johnsen, Sverre. 2012. A diachronic account of phonological unnaturalness. *Phonology* 29: 505-531.

Steriade, Donca. 1993. Closure, release, and nasal contours. In *Phonetics and Phonology.* Volume 5: Nasals, Nasalization, and the Velum, edited by Marie K. Huffman and Rena A. Krakow. 401-470. San Diego: Academic Press.

Steriade, Donca. 1997. *Phonetics in phonology: the case of laryngeal neutralization.* Ms., University of California, Los Angeles.

Steriade, Donca. 2001. *The Phonology of Perceptibility Effects: The P-map and its Consequences for Constraint Organization.* Ms., UCLA.

Stevens, Kenneth, Sheila Blumstein, Laura Glicksman, Martha Burton, and Kathleen Ku-rowski. 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91(5): 2979-3000.

Strycharczuk, Patrycja. 2012. Phonetics-phonology interactions in pre-sonorant voicing. Ph.D. Thesis, University of Manchester.

Strycharczuk, Patrycja and E. Simon. 2013. *Natural Language and Linguistic Theory* 31(2): 563-588.

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

Valls, David. 2014. *A grammar sketch of the Bugis language.* Ms. Available online on August 31, 2017 at `https://www.academia.edu/23400897/A_grammar_sketch_of_the_Bugis_language`.

van de Vijver, Ruben and Baer-Henney, Dinah 2014. Developing biases. *Frontiers in Language Sciences* 5: 634.

Vennemann, Theo. 1988. Preference Laws for Syllable Structure. Berlin: Mouton de Gruyter.

Wang, William S-Y. 1968. Vowel features, paired variables, and the English vowel shift. *Language* 44(4): 695-708.

Warner, Natasha and Andrea Weber. 2001. Perception of epenthetic stops. *Journal of Phonetics* 29(1): 53-87.

Wedel, Andrew. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23: 247–274.

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.

Whatmough, Joshua. 1937. The development of Indo-European labiovelars. *Acta Jutlandica* 9: 45-56.

Wheeler, Max W. (2005). The Phonology of Catalan. New York: Oxford University Press.

White, James. 2013. Bias in phonological learning: Evidence from saltation. Ph.D. dissertation, UCLA.

White, James. 2014. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130(1): 96-115.

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.

Wickham, Hadley. 2009. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer.

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

Wilson, Colin and Marieke Obdeyn. 2009. Simplifying subsidiary theory: statistical evidence from Arabic, Muna, Shona, and Wargamay. Ms., Johns Hopkins University.

Wolodzko, Tymoteusz. 2017. *kernelboot: Smoothed Bootstrap and Random Generation from Kernel Densities.* R package version 0.1.4. `CRAN.R-project.org/package=kernelboot`

Woollams, Geoff. 1996. *A Grammar of Karo Batak, Sumatra.* Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

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

Xromov, Al'bert. 1987. Jagnobskij jazyk. In *Osnovy iranskogo jazykoznanija.* Novoiranskie jazyki: vostočnaja gruppa. 644–701. Moscow: Nauka.

Yoo, Isaiah WonHo and Barbara Blankenship. 2003. Duration of epenthetic [t] in polysyllabic American English words. *Journal of the International Phonetic Association* 33(2): 153-164.

Yoshida, Yutaka. 2016. Sogdian language: i. Description. *Encyclopdia Iranica*, online edition, available at `www.iranicaonline.org/articles/sogdian-language-01` (accessed on 11 November 2016).

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.

Zsiga, Elizabeth, Maria Gouskova, and One Tlale. 2006. On the status of voiced stops in Tswana: Against *ND. In *Proceedings of the 36th Annual Meeting of the North East Linguistic Society*, edited by Christopher Davis, Amy Rose Deal, and Youri Zabbal. 721-734. Amherst, MA: GLSA.

Zuraw, Kie 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language and Linguistic Theory* 28(2): 417-472.