Indo-European Reduplication:
Synchrony, Diachrony, and Theory

by

Sam Zukoff

M.A., University of Georgia (2012)
B.A., Georgetown University (2010)

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

© Sam Zukoff, MMXVII. All rights reserved.

Author . . . .
Department of Linguistics and Philosophy
September 8th, 2017

Certified by . . . .
Donca Steriade
Professor
Thesis Supervisor

Accepted by . . . .
David Pesetsky
Head, Department of Linguistics and Philosophy

<div align="center">

Indo-European Reduplication:
Synchrony, Diachrony, and Theory

by

Sam Zukoff

Submitted to the Department of Linguistics and Philosophy
on September 8th, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Linguistics

</div>

Abstract

The reduplicative systems of the ancient Indo-European languages are characterized by an unusual alternation in the shape of the reduplicant. The related languages Ancient Greek, Gothic, and Sanskrit share the property that root-initial consonant clusters exhibit different reduplicant shapes, depending on their featural composition. Moreover, even though the core featural distinction largely overlaps across the languages, the actual patterns which instantiate that distinction are themselves distinct across the languages. For roots beginning in stop-sonorant clusters (TRVX– roots), each of these languages agrees in displaying a prefixal CV reduplicant, where the consonant corresponds to the root-initial stop: TV-TRVX–. These three languages likewise agree that roots beginning in sibilant-stop clusters (STVX– roots) show some pattern other than the one exhibited by TRVX– roots. However, each of the three languages exhibits a distinct alternative pattern: V-STVX– in the case of Ancient Greek, STV-STVX– in the case of Gothic, TV-STVX– in the case of Sanskrit.

This dissertation provides an integrated synchronic and diachronic theoretical account of the morphophonological properties of verbal reduplication in the ancient Indo-European languages, with its central focus being to explain this core alternation between TRVX– roots and STVX– roots. Set within Base-Reduplicant Correspondence Theory, a framework for analyzing reduplication in Optimality Theory, the comprehensive synchronic analyses constructed in service of understanding this distinction and other interrelated distinctions allow us to probe complex theoretical questions regarding the constraints and constraint interactions involved in the determination of reduplicant shape.

This dissertation seeks not only to develop in depth, consistent accounts of both the productive and marginal/archaic morphophonological aspects of reduplication in the Indo-European languages, it aims to understand the origins of these patterns — from a historical and comparative perspective, and from the perspective of morphophonological learning and grammar change — and attempts to motivate the conditions for the onset, development, and retention of the changes that result in the systems observed in the attested languages. As such, these analyses constitute a valuable set of case studies on complex systemic change in phonological grammars.

Thesis Supervisor: Donca Steriade
Title: Professor

To perseverance, and to everyone who helped me along the way.

Acknowledgments

They say that the dissertation is not the end, it's just the beginning. Well, whatever this is, it began well before I ever conceived of writing a dissertation. This whole giant mess emerged out of a phonology assignment sophomore year of college at Georgetown. In Jim Gruber's phonology class, we had to come up with our own problem set. I happened to be taking my first semester of Ancient Greek at the time, and I had noticed that this whole reduplication thing was kind of weird. So I used that for my problem set, which I eventually decided to turn into my senior thesis (and no, I'm not going to let anyone see it), which morphed into my class paper for Anya Lunden's graduate phonology class at UGA, which morphed into my UGA Master's thesis (which I might let you see if you ask nicely), which prompted my independent study with Donca Steriade at MIT, which I turned into my first conference presentation at WECIEC, for which I wrote my first proceedings paper, which turned into my first journal article at *LI*, which forms the basis for the first content chapter of this dissertation, which is entirely an extension of the ideas contained therein. So many people have helped me get from there to here, and even to get there in the first place. I can't possibly thank them all here, let alone thank them properly. But here's my best attempt.

First I have to thank my committee, who have all been extremely supportive throughout the dissertation process, and through my whole time at MIT.

To Donca Steriade, a great advisor and my greatest champion. When I first arrived on campus for my accepted student visit, I stepped off the elevator and saw the department and said to myself *"Wow, this is different, and awesome!"*. Then the very first thing I did was meet with Donca, and I had the very same reaction. I made the department wine me and dine me for the rest of the open house, but I knew at that moment that I was going to MIT. Thank you, Donca, for always being willing to read and comment on everything I ever wrote, no matter how bad a shape it was in. And for always being open to my ideas, even (and especially) when you adamantly disagreed with them. For giving me the chance to argue for my ideas, and for be willing to be convinced when I did a good job at doing so. I can think of no greater compliment than, when after my second generals defense, you came to me and said *"This is the second time you came to me with something crazy... and it turned out to be true!"*. There are so many different possibilities that I can only guess about which was the first time she had in mind.

To Adam Albright, for always having the most positive and constructive feedback. I can't count the number of times I went into a meeting with you feeling really anxious and uncertain about how to proceed on a project and whether I would ever be able to make any headway on it, and came out with a game plan that I was confident about and believing that I could do it. And for, like clockwork, always asking the question *"But how would you learn this?"*.

And to Edward Flemming, for never being satisfied with anything less than the absolute clearest and strongest reasoning. I came out of every meeting with you with a clearer understanding of my own thoughts, and of the bigger picture.

And thank you to all the rest of the faculty and staff at MIT Linguistics, as well, especially David Pesetsky and Michael Kenstowicz. There were multiple times when you all went above and beyond to help me through difficult times, and your kindness and support made a huge difference.

I've also been extremely lucky to have support from giants of the field who weren't on my committee. There was no greater validation than when Craig Melchert read my Anatolian chapter and gave it the thumbs up. Thank you, Craig, for suggesting that I look into Anatolian in the first place, for being willing to read and comment on my chapter (and for doing so so quickly), and for, as always, providing the most helpful feedback. Though I didn't take full advantage of it, having Jay Jasanoff down the street has been an amazing resource for me. Thank you, Jay, for be willing to

I don't think I've ever had as much fun as I have with my friends from MIT. Thank you to Ted Levin and Coppe van Urk for always being willing to ball out of control, or, when circumstances dictated, to ball on a budget. Thank you to Ruth Brillman and Chris Davis for being the most supportive friends you could ever ask for, always inviting me over for a gourmet meal at the drop of a hat, and dropping everything to go grab a beer whenever I needed to talk. Michelle Yuan, Athulya Aravind, and Carrie Spadine insisted that they deserved special mention too. They're right. Thanks also to (in chronological order) Sam Steddy, Gretchen Kern, Aron Hirsch, Anthony Brohan, Chris O'Brien, Kenyon Branan, Ezer Rasin, Justin Colley, Colin Davis, Chris Baron, and all the other amazing friends I've had the honor of sharing the department with (and to Nicole Torres, who, at this point, is basically an honorary member of the department). Thanks to my linguist friends from near and far: Nico Baier, Laura Grestenberger, Jesse Lundquist, Chiara Bozzone, Andrew Bird, Jess DeLisi, Caley Smith, Hemanga Dutta, and too many others to count. And thanks also to my great friends from back in Summit, New Jersey: Dan Rufolo, David Richards, Scott Leighton, just to name a few. Thank you also to my Boston area cousins: Steve Zukoff, Adam and Beth Badik, Kenny Westerman and Brian Westerman.

Above all, however, I could not have made it to this point without the unwavering and unparalleled love and support from my family. To my parents, Mimi and Paul Zukoff, words can't describe how grateful I am for everything you've done for me, and everything you continue to do for me. I never would have made it through all the tough times without you in my corner. Thank you to my grandmother Irene Preiser and my late grandfather Aaron Preiser. You were always there to support me and facilitate my success. You know that I am who I am because I take after you. Thank you to my grandmother Ella Zukoff, my grandfather Marty Zukoff, and my uncle Ed Zukoff. No grandparents have ever shown their grandson as much love as you have always showed me. And thank you to my late brother Ben Zukoff. You have always been an inspiration to me, while you were and every day since.

The incomparable Ted Levin once imparted us with the following wisdom: "You can sweat the petty stuff, but don't pet the sweaty stuff." On this point I can quibble with only one thing: you shouldn't sweat the petty stuff either. To anyone who's still reading at this point, I must strongly recommend you stop now — it's all downhill from here. But if you insist, here it is.

Contents

Chapter 1

Introduction

1.1 Overview of the Dissertation

The reduplicative systems of the ancient Indo-European languages are characterized first and foremost by an unusual, almost unique, alternation in the shape of the reduplicant. The related languages Ancient Greek, Gothic, and Sanskrit share in common the property that root-initial consonant clusters exhibit different reduplicant shapes, depending on their featural composition. Moreover, even though the core featural distinction largely overlaps across the languages, the actual patterns which instantiate that distinction are themselves distinct across the languages.

For roots beginning in stop-sonorant clusters (TRVX– roots), each of these languages agrees in displaying a prefixal CV reduplicant, where the consonant corresponds to the root-initial stop: TV-TRVX–. These three languages likewise agree that roots beginning in sibilant-stop clusters (STVX– roots) show some pattern other than the one exhibited by TRVX– roots. However, each of the three languages exhibits a distinct alternative pattern: V-STVX– in the case of Ancient Greek, STV-STVX– in the case of Gothic, TV-STVX– in the case of Sanskrit.

This dissertation provides an integrated synchronic and diachronic theoretical account of the morphophonological properties of verbal reduplication in the ancient Indo-European languages, with its central focus being to explain this core alternation between TRVX– roots and STVX– roots — an alternation that runs throughout much of the language family, yet consistently differs in such fundamental respects. Set within Base-Reduplicant Correspondence Theory (McCarthy & Prince 1995, 1999, *et seq.*), a framework for analyzing reduplication in Optimality Theory (Prince & Smolensky 1993/2004), the comprehensive synchronic analyses constructed in service of understanding this distinction and other interrelated distinctions allow us to probe complex theoretical questions regarding the constraints and constraint interactions involved in the determination of reduplicant shape.

This dissertation seeks not only to develop in depth, consistent accounts of both the productive and marginal/archaic morphophonological aspects of reduplication in the Indo-European languages, it aims to understand the origins of these patterns from a historical and comparative perspective, and from the perspective of morphophonological learning and grammar change. In so doing, it develops fully articulated synchronic analyses of earlier stages of the languages, and attempts to motivate the conditions for the onset, development, and retention of the changes that result in the systems observed in the attested languages. As such, these analyses constitute a valuable set of case studies on complex systemic change in phonological grammars. Furthermore, having assembled a suite of comprehensive accounts of the reduplicative systems of the individual Indo-European languages,

coupled with internal reconstructions based on analysis of their archaic patterns, this dissertation provides new perspective on the comparative reconstruction of the reduplicative system of the common ancestor of the Indo-European languages — Proto-Indo-European — as a dynamic synchronic system.

1.2   Structure of the Dissertation

This dissertation is structured around four in depth case studies in the reduplicative systems of the individual ancient Indo-European languages: Ancient Greek (Chapter 2), the Anatolian languages Hittite and Luwian (Chapter 3), Gothic (Chapter 4), and Sanskrit (Chapter 5). These case studies culminate with the formalization of a new solution to the core TRVX– vs. STVX– distinction: the NO POORLY-CUED REPETITIONS constraint (*PCR). This approach derives the TRVX– vs. STVX– distinction — which actually only represents the endpoints of a larger, more diverse distributional pattern exhibited within these and other phonological systems (reduplicative and non-reduplicative alike) — through stringent contextual licensing conditions on repeated consonants, based on acoustic/auditory cues to consonantal contrast (Chapter 6). Bringing to bear the analytical and historical insights from the individual languages, and the formal mechanism of the *PCR constraint, the dissertation concludes with a new systemic reconstruction of reduplication in Proto-Indo-European (Chapter 7).

The case studies largely adhere to a consistent structure and explanatory trajectory. They begin with detailed theoretical analyses of the synchronic reduplicative systems (and the broader morphophonological systems in which they are embedded) as we have them in the attested languages. This analysis reveals some aspect of the system which has become, in one way or another, less than transparent, suggesting that this component of the system may be better understood by considering its context at a prior stage of the language. I identify this prior stage through internal and/or comparative reconstruction, and I systematically examine the properties that underlie the attested synchronic oddity, yielding a new synchronic analysis of the relevant aspects of the prior stage.

This raises the question of how the synchronic system of the prior stage, where the relevant component was transparently generated as part of the regular morphophonological system, came to develop into the attested system where this is no longer the case. To address these questions, I consider how principles of (morpho)phonological learning — focusing largely on versions of the Recursive Constraint Demotion (RCD; Tesar & Smolensky 1998, 2000; cf. Prince & Tesar 2004, Becker 2009) learning procedure — could have driven the change in the phonological and morphological grammar from one stage to the next, deterministically generating the innovative constraint rankings and morphophonological organizations.

I now outline the structure and main arguments of the individual chapters. Subsequently, in Section 1.3, I provide an overview of the approach to reduplicant shape that will be employed throughout the dissertation, including a limited factorial typology of the main constraints involved and a preview of the Indo-European reduplication data.

1.2.1   Ancient Greek

I begin the body of the dissertation in Chapter 2 with Ancient Greek. In addition to displaying the core Indo-European distinction between TRVX– roots and STVX– roots in reduplication — in the form of a phonologically predictable alternation (driven by *PCR) between "C_1-copying" for TRVX– roots (TV-TRVX–) and "non-copying" for STVX– roots (V-STVX–) — Ancient Greek

also shows a synchronically unpredictable distinction in its treatment of vowel-initial roots in the perfect tense, the main reduplicated morphological category of the language. While most vowel-initial roots (VCX– roots) form their perfect tense stem via lengthening of the root-initial vowel (V̄CX–), a small set of roots instead display a pattern referred to as "Attic Reduplication", whereby the root-initial VC sequence is copied and the root-initial vowel is lengthened (VC-V̄CX–).

After constructing an analysis that generates the productive behavior of the consonant-initial roots, I show that this reduplicative grammar is directly compatible with the productive vowel-lengthening pattern. Attic Reduplication, and also a set of consonant-initial roots which show reduplication in defiance of the normal pattern, can only be generated synchronically through appeal to lexical constraint indexation. The exceptional, lexically restricted behavior of these two types requires additional explanation. Such explanation can be achieved through consideration of their diachronic origins, and how they interact with the rest of the system over the time course of change.

I argue that Attic Reduplication can and should be traced back to an alternative reduplication strategy for laryngeal initial roots in Pre-Greek, triggered by laryngeal-related phonotactics. I construct a complete analysis of the Pre-Greek stage, and demonstrate that the properties of the precursor of the Attic Reduplication pattern are shaped by the interaction of the independently motivated reduplicative grammar and another laryngeal-related repairs, "laryngeal vocalization".

The subsequent loss of the laryngeals via sound change forced learners to reanalyze the pattern, such that Attic Reduplication came to be retained in Ancient Greek not directly by phonotactics, but by lexical constraint indexation. This situation is repeated by the consonant-initial roots that show exceptional copying in reduplication. These forms retain archaic reduplicative properties because of their connection to other morphologically-related reduplicated outputs in such a way that they are brought within the orbit of the lexically indexed constraint that is independently required for Attic Reduplication. The development and synchronic behavior of these types can be motivated and formalized using a procedure for constraint cloning under conditions of ranking inconsistency, based on the proposals of Becker (2009) and Pater (2009).

1.2.2   Anatolian

In Chapter 3, I examine and analyze the reduplicative systems of the Anatolian languages Hittite and Luwian, and I undertake a reconstruction of the reduplicative system of their proximate common ancestor — Proto-Anatolian. Hittite and Luwian fail to show the core Indo-European distinction between TRVX– roots and STVX– roots in their synchronic reduplicative systems: Hittite exhibits across-the-board cluster-copying (TRV-TRVX– and iSTV-STVX–), while Luwian lacks the distinction because it has eliminated STVX– roots from its lexicon via sound change (it retains C_1-copying for TRVX– roots: TV-TRVX–). Nevertheless, based on the languages' regular diachronic correspondences in the treatment of ST clusters, and considerations of parsimony in historical change, Proto-Anatolian is actually to be reconstructed as having had the core TRVX– vs. STVX– distinction — in the form of C_1-copying for TRVX– roots (TV-TRVX–) vs. cluster-copying for STVX– roots (STV-STVX–).

The synchronic activity of *PCR is essential in generating the TRVX– vs. STVX– distinction in Proto-Anatolian. However, this constraint plays no role in the synchronic grammar of its daughter languages Hittite or Luwian; in fact, the two languages bear an innovative pattern of reduplication for vowel-initial roots (VR-VRT–) that directly violates it. Therefore, *PCR must transition from a high-ranked, active constraint which can induce alternative reduplication strategies in Proto-Anatolian to a low-ranked, inactive constraint which can be violated in reduplication in the daughter languages. I will argue that independent phonological changes in the internal development of Hittite and Luwian had the incidental effect of eliminating the distinction between TRVX–

roots and STVX– roots in reduplication, and that this allowed learners to converge on a *PCR-free analysis, paving the way for the subsequent emergence of the *PCR-violating <u>VR</u>-VRT– pattern.

This represents an interesting case with respect to phonological learning, as it seems to be a diachronic counter-example to the Subset Principle (cf. Prince & Tesar 2004). That is, there is a point in the development of these languages where speakers evidently did not learn the most restrictive grammar, allowing for the later emergence of a more marked pattern. Given the particular constraint types that are involved, it might be possible to avoid the problem by using Biased Constraint Demotion (BCD; Prince & Tesar 2004) as the learning procedure. Alternatively, the problem may be solved by introducing a minor reformulation of RCD/BCD that favors the high ranking of maximally informative winner-preferring constraints. This approach yields a satisfactory step-wise account of the relevant diachronic developments, and it would support a view that, under the right conditions, learners will indeed be predicted to learn a non-subset grammar.

1.2.3 Gothic

In Chapter 4, I explore the analysis of the reduplicative system of Gothic. Unlike the other languages examined in detail in this dissertation, the reduplicated verbal forms of Gothic represent but one piece of a larger system of complex verbal morphology and morphophonology — the so-called "strong verb" system. The preterite stems of strong verbs are formed in a variety of ways, mostly involving vocalic alternations relative the base, but also including reduplication. What is especially noteworthy about this system is that the choice of stem-formation process is predictable based on the phonological properties of the verbal root; for example, reduplication is only found in roots having a long root vowel or a root vowel /a/ followed by two consonants. Therefore, the distribution of preterite stem formation patterns seems very clearly non-arbitrary from the synchronic point of view, and is thus in need of analysis.

I develop an analysis whereby the strong preterites select for a null underlying representation of the morpheme PRETERITE, and differentiation of stems is induced by a family of constraints that require overt exponence of morphosyntactic features, REALIZE MORPHEME (RM; Kurisu 2001), and thus phonological contrast between stems which are morphologically related in a particular way. The nature of the changes undergone to satisfy RM, of which reduplication is only one of many (and indeed a sort of "last resort" synchronically), falls out from the interaction between the phonological properties of individual roots and the ranking of markedness and faithfulness constraints. Furthermore, I show that the phonological grammar required to generate these patterns of stem formation are fully consistent with the phonological grammar required to account for the patterns of reduplicant shape, which are themselves subject the core TRVX– vs. STVX– distinction induced by *PCR.

1.2.4 Sanskrit

Chapter 5 examines certain aspects of the complex reduplicative system of Sanskrit. The perfect tense in Sanskrit demonstrates two distinct means of instantiating the core Indo-European TRVX– vs. STVX– distinction. For cluster-initial roots, this distinction is reflected in the difference between C_1-copying for TRVX– roots (<u>TV</u>-TRVX–) and C_2-copying for STVX– roots (<u>TV</u>-STVX–). On the other hand, in the inflectional categories of the perfect which normally call for deletion of the root vowel ("zero-grade ablaut"), collectively referred to as the "perfect weak stem", the TR vs. ST distinction plays out in a different way for roots of the shape C_1aC_2, which would be reduced to C_1C_2 clusters by root-vowel deletion. TaR roots show the expected root-vowel deletion and exhibit C_1-copying: i.e., $\sqrt{\text{TaR}} \rightarrow$ perfect weak stem <u>TV</u>-TR–. SaT roots, on the other hand, display neither

18

root-vowel deletion nor reduplication of any kind (at least overtly), instead showing a root allomorph in $C_1\bar{e}C_2$: i.e., $\sqrt{}$SaT $\rightarrow$ perfect weak stem SēT–.

In this chapter, I will develop an analysis of Sanskrit that simultaneously generates each of the distinct outcomes of the perfect, namely: (i) the C_1-copying pattern for TRVX– roots and TaR perfect weak stems, (ii) the C_2-copying pattern for STVX– roots, and (iii) the $C_1\bar{e}C_2$ pattern for SaT perfect weak stems. I will first present an analysis based on allomorph selection, where the underlying representation leading to the $C_1\bar{e}C_2$ form is selected in the phonological component just in case a licit reduplicative structure that satisfies *PCR and the Input-Reduplicant faithfulness constraint LINEARITY-IR (McCarthy & Prince 1995, 1999) cannot be obtained. I will then explore an analysis which derives the $C_1\bar{e}C_2$ pattern directly in the phonology, treating it as a *PCR-driven phonological repair of an overtly reduplicative structure; specifically, $//\underline{C_1}e\text{-}C_1C_2\text{-}// \rightarrow [\underline{C_1}e\text{:}C_2\text{-}]$, where C_1 represents the consonant of the reduplicant, the vowel represents the vowel of the reduplicant, and the length of the vowel derives from compensatory lengthening after the deletion of the root-initial consonant.

While the phonological solution encounters difficulties in accounting for the quality of the vowel synchronically in Sanskrit (suggesting that the allomorphy analysis is most appropriate for Sanskrit synchronically), the analysis extends seamlessly to a related $C_1\bar{e}C_2$ pattern found in the Germanic languages (namely, the Class V strong preterite plurals).[1] I adduce evidence for similar patterns in Old Irish and in Hittite, as well. Regardless of the synchronic analysis of the pattern in attested Sanskrit, this correspondence argues in favor of a phonological origin of the $C_1\bar{e}C_2$ type, even if they do not originate as a unitary formation in Proto-Indo-European.

1.2.5 The NO POORLY-CUED REPETITIONS Constraint (*PCR)

The analyses of the core Indo-European TRVX– vs. STVX– distinction developed in Chapters 2–5 are centered around the operation of a simplified version of the NO POORLY-CUED REPETITIONS constraint (*PCR), which prohibits sequences of repeated consonants ($C_\alpha V C_\alpha$) in immediate pre-obstruent position ($/\_C_{[-sonorant]}$). Chapter 6 expands the scope of this examination so as to include the full cluster-wise distribution of default vs. alternative pattern reduplication in the languages discussed. In each of these reduplicative systems, the TRVX– vs. STVX– distinction instantiates a broader distributional pattern regarding the treatment of different types of root/base-initial clusters; however, the languages diverge with respect to which clusters pattern with TRVX– in allowing the default C_1-copying pattern, and which clusters pattern with STVX– in resorting to the alternative pattern. For example, for stop-sibilant (TSVX–) roots, Ancient Greek exhibits the alternative pattern (grouping together with STVX–), while Sanskrit exhibits the default pattern (grouping together with TRVX–).

I argue that the full cluster-wise distributions of repetition avoidance effects in these systems are not to be explained in terms of traditional phonological features or abstract phonological properties like sonority, but rather in terms *acoustic/auditory cues* (cf. Steriade 1994, 1997, 1999, Flemming 1995/2002). Motivated by the empirical evidence, I propose the POORLY-CUED REPETITIONS HYPOTHESIS — previewed in (1) — which states that repetition imposes distinct burdens on the perceptual system with regards to the licensing of contrasts (specifically C∼Ø contrasts). Based upon this hypothesis, I formulate the final, precise version of the NO POORLY-CUED REPETITIONS constraint (*PCR) — previewed in (2) — which penalizes repetitions that would leave the repeated consonant without sufficient cues to its contrast with Ø. Specifically, I will show that *intensity rise*, alongside *transitions* and (perhaps) *stop release burst*, are central to licensing these

---

[1] Note, however, that the correspondence in vowel quality between the two patterns is illusory: Germanic $\bar{e}$ corresponds with Sanskrit $\bar{a}$.

contrasts under repetition. The different languages in effect select from among these cues which ones will be sufficient, either on their own or in combination, to license a consonant repetition.

(1) **THE POORLY-CUED REPETITIONS HYPOTHESIS**
There is some property of the perceptual system which degrades listeners' ability to apprehend the presence of a consonant (i.e. the contrast between that consonant and its absence) when that consonant is adjacent to an identical consonant.

    i.   This property diminishes the effectiveness of some or all acoustic/auditory cues to C∼Ø contrasts, such that some cues which are normally sufficient to license those C∼Ø contrasts (in otherwise equivalent positions) are no longer sufficient to license those contrasts under repetition.

    ii.   This property diminishes the effectiveness of different cues to different extents: the effectiveness of cues to acoustic events which are more difficult to anchor at a particular point in the speech stream and/or tend to extend across multiple segments are diminished to a greater degree than cues to acoustic events which are more reliably located at their correct position in the speech stream.

(2) **The NO POORLY-CUED REPETITIONS constraint (*PCR)**
Languages may set stricter conditions (in terms of cues) for the licensing of C∼Ø contrasts (i.e. the presence of C) when that C would be the second member of a transvocalic consonant repetition (i.e. C_α^2 in a $C_\alpha^1 V C_\alpha^2$ sequence) than in other contexts. Assign a violation mark * for each C_α^2 (i.e. each C∼Ø contrast where C is a C_α^2) which is not cued to the level required by the language-specific repetition licensing conditions.

The inclusion of this constraint in the grammar properly derives the patterns of repetition avoidance instantiated by the languages examined in this dissertation. In Chapter 6, I adduce additional empirical and analytical evidence in favor of the use of the *PCR constraint. I provide analyses of several additional reduplicative effects that require the use of *PCR, namely, infixal reduplication in Latin STVX– roots, infixal reduplication in Sanskrit vowel-initial desideratives, and a pattern equivalent to that of Gothic found in the non–Indo-European language Klamath (Barker 1964). I also discuss a few marginal patterns outside of reduplication which may be amenable to a *PCR-based analysis.

1.2.6   Reconstructing Proto-Indo-European Reduplication

Chapter 7 concludes the dissertation. Besides reviewing the main arguments advanced in the earlier chapters, this chapter lays out an updated reconstruction of the reduplicative system of Proto-Indo-European (PIE), focusing on the question of the core TRVX– vs. STVX– distinction. The more traditional view reconstructs for PIE the distribution found in Gothic (and now Proto-Anatolian): i.e., C_1-copying for TRVX– (TV-TRVX–) but cluster-copying for STVX– (STV-STVX–). This allows for the productive STVX– treatments of Ancient Greek (V-STVX–), Sanskrit (TV-STVX–), and Latin (S-TV-TVX–) to be seen as different kinds of reductions from the original type.

I argue for an alternative reconstruction for PIE, one which is present in the literature (e.g., Byrd 2010:100–105), though not the most commonly accepted view: in PIE, both TRVX– and STVX– roots, and indeed all types of root-initial clusters, exhibited C_1-copying; i.e., TV-TRVX– and SV-STVX–. The strongest evidence in favor of this alternative reconstruction comes from archaisms in Greek and Latin (Brugmann & Delbrück 1897–1916:40–41, Byrd 2010:103–104), and their agreement with Iranian (Byrd 2010:103), which all points to SV-STVX–.

Furthermore, now understanding the extent to which the precise behavior of *PCR (with respect to its repetition licensing conditions) differs across the Indo-European languages, it is less evident that the core Indo-European distinction between TRVX– and STVX– is an instance of inheritance. Rather, it seems just as if not more likely that it represents a parallel development driven by similar inherited conditions. If this is correct, then this presents a further argument that reconstructions of dynamic properties/systems like reduplication need to be based on fully articulated theoretical analyses, not just correspondences in surface forms.

1.3 Analysis of Reduplicant Shape in Indo-European

The core reduplication patterns of the Indo-European languages show a fairly remarkable distribution. Five constraints are required in order to capture the core facts of the pattern. The factorial typology of these five constraints, crossed with a potential distinction in the nature of the reduplicative vowel, predicts six possible copying systems. Five of the six are attested among the Indo-European languages; only one type is not found within the family (and indeed not attested at all).

This section lays out the analysis of these systems. It proceeds from the most straightforward types of cases – where there is consistent copying behavior across all types of bases, to those where a subset of base shapes (that is, those beginning with a particular sort of consonant cluster) displays a divergent pattern. It concludes with consideration of the restricted factorial typology of the constraints included in the analysis, and confirms the relatively strong fit to the typology within Indo-European itself.

1.3.1 Across-the-board Behavior

The most typical copying behavior in reduplicative systems is to simply copy a consistent amount of material from one edge of the base for all bases that allow reduplication. Such systems are formally the simplest, as the active constraints apply equally to all forms, regardless of the featural properties of the forms. Such patterns are attested among the Indo-European languages, though in each case there is some orthogonal complication which partially obscures its status as such a pattern. In this subsection, I detail the two types which we have in Indo-European: across-the-board cluster-copying, which is attested in Hittite; and across-the-board C_1-copying, which is attested in Old Irish (and reconstructed for other stages of several Indo-European languages, including Greek). In both cases, I will first schematize an idealized version of the pattern, and then present the data as we actually have it from these languages.

1.3.1.1 Across-the-board Cluster-copying: Hittite

The pattern which is perhaps formally simplest is one which copies the first vowel of the base and all material before it, i.e. the onset of the base. This pattern is schematized in (3). Subscripts in the "Red. Shape" column indicate which number segment of the base, counting from the left, each reduplicated segment corresponds to (via Base-Reduplicant correspondence; cf. McCarthy & Prince 1995, 1999).

(3) Across-the-board cluster-copying

	Base Type	Root		Reduplicated	Red. Shape
a.	Singleton	√*mako*	→	*ma-mako*	C_1V_2
b.	Stop-sonorant	√*prako*	→	*pra-prako*	$C_1C_2V_3$
c.	*s*-obstruent	√*stako*	→	*sta-stako*	$C_1C_2V_3$

There are several constraints which are directly relevant for calculating the shape of the reduplicant. First let us consider what constraint(s) might be violated by having reduplicants of these shapes. A CV reduplicant, especially for a CV-initial base, is virtually perfect in terms of violation profile, as it fully matches the content of the base, and also has the least marked syllable structure. On the other hand, the CCV reduplicants for cluster-initial bases, while perfectly matching their base, display a marked syllable structure, namely, a complex onset. In syllable-neutral terms, we can use the constraint *Cluster (*CC) to encode this markedness.[2]

(4) ***CC (≈ *COMPLEX or C/_V)**
Assign a violation mark * for every sequence of two consonants.

For a language with the across-the-board cluster-copying pattern, this constraint must be low-ranked enough to disallow the cluster from surfacing in the reduplicant. There must therefore be higher-ranked constraint(s) that advocate for copying the whole cluster. A commonly used constraint that could prefer cluster-copying in such a case is MAX-BR (McCarthy & Prince 1995), which advocates for copying everything in the base into the reduplicant, and thus will always prefer more copying to less. However, none of the Indo-European languages demonstrate any need for this constraint to be active in shaping their reduplication patterns, and the allowance of this constraint into the constraint inventory has well-known typological problems (for example, the "Kager-Hamilton Conundrum"; McCarthy & Prince 1999:258). For this reason, I will not be considering MAX-BR in this chapter.

Instead, I claim that there are two constraints, both of which are Base-Reduplicant faithfulness constraints, that collaborate to advocate for the across-the-board cluster-copying pattern: CONTIGUITY-BR and ANCHOR-L-BR. CONTIGUITY-BR (5a) requires contiguous copying from the base, and ANCHOR-L-BR (5b) requires copying that begins at the left edge of the base.

(5) BR-faithfulness constraints that promote cluster-copying

   a. **CONTIGUITY-BR**
   Assign one violation mark * for each pair of segments that are adjacent in the reduplicant but have non-adjacent correspondents in the base (i.e. no X_1X_3-$X_1X_2X_3$).

   b. **ANCHOR-L-BR**
   Assign a violation mark * if the segment at the left edge of the reduplicant does not stand in correspondence with the segment at the left edge of the base.[3]

---

[2] *CC falls under the notion of "size restrictor" constraints in the non-templatic analysis of reduplicant shape (see Spaelti 1997, Hendricks 1999, Riggle 2006, among others), which is the approach being advocated here. Therefore, in this introduction, and in the limited factorial typology presented below, I use *CC to stand in for *some size restrictor* with similar properties. In the actual analyses to be developed in this dissertation, in addition to *CC, alignment constraints (McCarthy & Prince 1993a) will frequently be used as size restrictors.

[3] The uses of ANCHOR-L-BR in this dissertation would largely also be compatible with Nelson's (2003) "LOCALITY" constraint. One case where LOCALITY will not work — and therefore we require ANCHOR-L-BR — is Ancient

The most obvious way of alleviating the *CC violation would be to copy only one member of the base-initial cluster. However, doing so would necessarily violate one of these two constraints. If only the first member of the cluster were copied along with the vowel, i.e. C_1V_3-$C_1C_2V_3C_4$... (6b), this would be discontiguous copying and violate CONTIGUITY-BR. If only the second member of the cluster were copied along with the vowel, i.e. C_2V_3-$C_1C_2V_3C_4$... (6c), this would mean that copying did not start from the left edge of the base and thus would violate ANCHOR-L-BR. Therefore, in order to generate across-the-board cluster-copying, we need the following ranking: CONTIGUITY-BR, ANCHOR-L-BR ≫ *CC. This is demonstrated in (6).

(6)   Generating across-the-board cluster-copying

/RED, prako/	CONTIGUITY-BR	ANCHOR-L-BR	*CC
a. ☞ <u>pra</u>-prako			**
b. <u>pa</u>-prako	*!		*
c. <u>ra</u>-prako		*!	*

Note that this pattern can *only* be derived if the vowel of the reduplicant stands in correspondence with the vowel of the base. If the base and reduplicant vowels did not correspond, then copying of just the base-initial consonant (6b) would not violate CONTIGUITY-BR, and candidate (6b) would thus harmonically bound desired candidate (6a). This is because, in such a case, only one segment is actually being copied (or, more precisely, standing in BR-correspondence), so there is no contiguity relationship between multiple reduplicant segments to be maintained. The status of the reduplicating vowel will be especially relevant for the analysis of the Ancient Greek pattern, and will be discussed further in that context (see Section 1.3.2.3).

As argued in Chapter 4, Hittite displays the across-the-board cluster-copying pattern schematized in (3). In obstruent-sonorant–initial bases (7a), the reduplicant copies the whole cluster. In STVX– bases (7b), the reduplicant also copies the whole cluster, but additionally epenthesizes [i] before the reduplicant cluster; prothesis to initial ST-clusters is a general process in the language, and thus does not have anything to do with reduplication. (There are also questions surrounding the length or laryngeal features of base-initial stops, as well as the nature of the reduplicative vowel; see Chapter 4 for discussion.)

(7)   Across-the-board cluster-copying in Hittite

   a.   **TRVX– bases → cluster-copying**

Root		Reduplicated stem	
√*par(a)i-*	'blow'	*parip(p)ar(a)i-*	[pri-pːr(a)i-]
√*ḥal(a)i-*	'kneel'	*ḥaliḥal(a)i-*	[χli-χl(a)i-]

   b.   **STVX– bases → (prothesis +) cluster-copying**

Root		Reduplicated stem	
√*stu-*	'become evident'	*išdušduške-*	[istu-stu-]

Greek (Chapter 2). The ANCHOR-L-BR vs. LOCALITY question will be addressed at relevant points throughout the dissertation.

1.3.1.2 Across-the-board C₁-copying: Old Irish (and elsewhere)

The other across-the-board reduplicative behavior which is attested among the Indo-European languages is across-the-board C_1-copying. In this pattern, all reduplicants, regardless of the shape of the base, surface as CV, where the consonant corresponds to the initial consonant of the base. This pattern, which is equivalent to candidate (b) in the tableau in (6), is schematized in (8). It is straightforward to see from (6) that this pattern can be derived simply by swapping the ranking of *CC and CONTIGUITY-BR. This is shown in (9) below.

(8)　Across-the-board C_1-copying

	Base Type	Root		Reduplicated	Red. Shape
a.	Singleton	√*mako*	→	*ma-mako*	C_1V_2
b.	Stop-sonorant	√*prako*	→	*pa-prako*	C_1V_3
c.	*s*-obstruent	√*stako*	→	*sa-stako*	C_1V_3

(9)　Generating across-the-board C_1-copying

/RED, prako/		ANCHOR-L-BR	*CC	CONTIGUITY-BR
a.	pra-prako		**!	
b. ☞	pa-prako		*	*
c.	ra-prako	*!	*	

In this pattern, it is more important to avoid creating an extra cluster than to avoid discontiguous copying. This pattern is generable regardless of the correspondence status of the reduplicative vowel, since CONTIGUITY-BR does not actively participate in determining the winner.

Across-the-board C_1-copying is attested in Old Irish, and is also reconstructible to Pre-Greek (see Chapter 2), and potentially other prior stages within the Indo-European family. Old Irish reduplicated preterites to cluster-initial bases are illustrated in (10). The root-initial stops in the TRVX–roots undergo lenition (spirantization), but this is not transferred to the reduplicant.

(10)　Old Irish reduplicated preterites (Thurneysen 1946 [1980]:424–428/§687–691)

　　a.　**TRVX– roots → C_1-copying**

Root		Reduplicated preterite	
√*-glenn-*	'learn'	*-geglann*	[-ge-ɣlənn]
√*-grenn-*	'persecute'	*-gegrann*	[-ge-ɣrənn]
√*brag-*	'bleat'	*bebrag-*	[be-vrəɣ-]
√*klad-*	'dig'	*cechlad-*	[ke-xləð-]

　　b.　**STVX– roots → C_1-copying**

Root		Reduplicated preterite	
√*skenn-*	'fly off'	*sescann-*	[se-skənn]

24

1.3.1.3 Across-the-board C_2-copying: Unattested

One will notice that there is one more pattern that can be generated by permuting the ranking of the constraints proposed thus far. If we had the ranking *CC, CONTIGUITY-BR $\gg$ ANCHOR-L-BR, and the reduplicative vowel stood in correspondence with the base vowel, we would generate a pattern where all cluster-initial roots copy their second member plus the base vowel. This pattern is illustrated in (11), and demonstrated in (12).

(11) Across-the-board C_2-copying

	Base Type	Root		Reduplicated	Red. Shape
a.	Singleton	$\sqrt{}$*mako*	$\rightarrow$	*ma-mako*	C_1V_2
b.	Stop-sonorant	$\sqrt{}$*prako*	$\rightarrow$	*ra-prako*	C_2V_3
c.	*s*-obstruent	$\sqrt{}$*stako*	$\rightarrow$	*ta-stako*	C_2V_3

(12) Generating across-the-board C_2-copying

/RED, prako/		CONTIGUITY-BR	*CC	ANCHOR-L-BR
a.	pra-prako		**!	
b.	pa-prako	*!	*	
c. ☞	ra-prako		*	*

This is the only system of the factorial typology not attested within Indo-European. To my knowledge, this system is also not attested anywhere outside of Indo-European either. Other than the fact that we are dealing with a fairly small number of languages that could be expected to display the conditions necessary for such a pattern, thus making it quite possible that this gap is an accidental one and not a significant one, there is no glaring reason why such a pattern should be unattested. This is a question for future consideration.

1.3.2 Cluster-Dependent Copying Patterns

In the patterns discussed thus far, all base-initial clusters behave identically. While formally simplest and perhaps typologically most common, this behavior is somewhat atypical of the Indo-European languages. In Gothic, Sanskrit, and Ancient Greek, different types of base-initial clusters trigger different copying patterns. In all of these languages, singleton-initial (CVX–) bases and TRVX– bases exhibit the C_1-copying pattern: CVX– $\rightarrow$ $\underline{C_1V}$-C_1VX–, TRVX– $\rightarrow$ $\underline{T_1V}$-T_1R_2VX–. However, for STVX– bases (and a subset of other initial cluster types, varying by language; see especially Chapter 6 for discussion and explanation), they all have some other copying pattern. Gothic shows cluster-copying: $\underline{S_1T_2V}$-S_1T_2VX– (Section 1.3.2.1). Sanskrit shows C_2-copying: $\underline{T_2V}$-S_1T_2VX– (Section 1.3.2.2). And Ancient Greek shows non-copying: $\underline{V}$-S_1T_2VX– (Section 1.3.2.3). One of the major claims of this dissertation is that these divergent copying behaviors are induced by a constraint which enforces restrictions on *consonant repetitions*, i.e. sequences of identical consonants separated only by a vowel ($C_\alpha VC_\alpha$), of particular types and in particular contexts.

In Chapter 6, I will develop a repetition avoidance analysis of these patterns based on the distribution and perception of acoustic/auditory cues to particular consonantal contrasts; I call this approach the NO POORLY-CUED REPETITIONS constraint (*PCR). However, simplifying significantly and putting aside for the time being investigation of these details, based purely upon the

empirical facts, it is justifiable to posit a constraint which militates against locally repeated conso-
nants in pre-obstruent position, as defined in (13). (In this chapter, for the purposes of maximal
clarity, I will refer to this constraint as $*C_\alpha VC_\alpha$ / $\_C_{[-sonorant]}$. In subsequent chapters, I will refer
to it as *PCR, in accordance with the full analysis that will be developed in Chapter 6.)

(13)   **$*C_\alpha VC_\alpha$ / $\_C_{[-sonorant]}$ [≈ NO POORLY-CUED REPETITIONS (*PCR)]**
       For each sequence of repeated identical consonants separated by a vowel ($C_\alpha VC_\alpha$), assign
       a violation * if that sequence immediately precedes an obstruent.

This analysis assumes that the C_1-copying pattern is the target pattern for these languages;
note that only a high ranking of CONTIGUITY-BR can lead to the other patterns discussed in
Section 1.3.1. This means that the language by default seeks to create reduplicated strings like
pa-prako and **sa-stako*. By admitting $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$ into the constraint set, we have a means
of treating these two strings differently. For TRVX– bases, C_1-copying places the consonant repe-
tition (*pap*) before a sonorant (*r*). This means C_1-copying can be carried out for TRVX– bases
without violating $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$. On the other hand, for STVX– bases, C_1-copying places the
consonant repetition (*sas*) before an obstruent (*t*). Performing C_1-copying to STVX– bases therefore
would violate $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$. When this constraint is ranked sufficiently high in the grammar,
it will prohibit C_1-copying from taking place (that is, from having a string *sast...* surface), and force
an alternative copying pattern to take over, just in case the base begins an obstruent-second cluster.
This is the case for Gothic, Sanskrit, and Ancient Greek, as will be demonstrated in the remainder
of this subsection.

1.3.2.1   TRVX– C_1-copying, STVX– Cluster-copying: Gothic

One response to $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$ is to copy the entire base-initial cluster. This has the effect
of having an additional consonant (namely, the root-second consonant) intervene between the
repeated base-initial consonant, disrupting the repetition: i.e. *C_1C_2V-C_1C_2V...* not **C_1V-C_1C_2V...*.
This pattern, where TRVX– bases show C_1-copying but STVX– bases show cluster copying, is
schematized in (14).

(14)   TRVX– C_1-copying, STVX– cluster-copying

	Base Type	Root		Reduplicated	Red. Shape
a.	Singleton	√*mako*	→	*ma-mako*	C_1V_2
b.	Stop-sonorant	√*prako*	→	*pa-prako*	C_1V_3
c.	*s*-obstruent	√*stako*	→	*sta-stako*	$C_1C_2V_3$

To generate C_1-copying in the basic case (i.e. TRVX–), we need the ranking ANCHOR-L-BR,
*CC ≫ CONTIGUITY-BR (cf. (9) above). This is shown in (15).[4] Then, in order to motivate diver-
sion from the C_1-copying pattern just in case the base begins in STVX–, $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$ must
dominate *CC, as shown in (16) below. In other words, it is generally preferable to avoid creating a
consonant cluster in the reduplicant, but this is tolerated if it allows a pre-obstruent repetition to be
avoided.

---

[4]   As mentioned before, the C_1-copying pattern is consistent with both types of reduplicative vowels, i.e. those that do
      correspond with the base vowel and those that do not. If the reduplicative vowel does not correspond to the base vowel,
      then the ranking of CONTIGUITY-BR is irrelevant.

(15)   Generating TRVX– C_1-copying

/RED, prako/		ANCHOR-L-BR	*$C_\alpha VC_\alpha$ / _$C_{[-son]}$	*CC	CONTIG-BR
a.	pra-prako			**!	
b. ☞	pa-prako			*	*
c.	ra-prako	*!		*	

(16)   Generating STVX– cluster-copying with TRVX– C_1-copying

/RED, stako/		ANCHOR-L-BR	*$C_\alpha VC_\alpha$ / _$C_{[-son]}$	*CC	CONTIG-BR
a. ☞	sta-stako			**	
b.	sa-stako		*!	*	*
c.	ta-stako	*!		*	

Gothic illustrates this pattern perfectly. (This same pattern is also reconstructible for Proto-Anatolian; see Chapter 4.) TRVX– bases follow the default C_1-copying pattern (17a), while STVX– bases display cluster-copying (17b).

(17)   Class VII preterites in Gothic (forms from Lambdin 2006:115)

a.   **TRVX– roots → C_1-copying preterites**

Root	Infinitive		Preterite		
'to weep'	*gretan*	[grēt-an]	*gaigrot*	[ge-grōt]	not **[gre-grōt]

b.   **STVX– roots → cluster-copying preterites**

Root	Infinitive		Preterite		
'to possess'	*staldan*	[stald-an]	*staistald*	[ste-stald]	not **[se-stald]
'to divide'	*skaidan*	[skaið-an]	*skaiskaiþ*	[ske-skaiθ]	not **[se-skaiθ]

1.3.2.2   TRVX– C_1-copying, STVX– C_2-copying: Sanskrit

Another *$C_\alpha VC_\alpha$ / _$C_{[-son]}$ avoidance strategy is to copy C_2 rather than C_1, as shown in (18). This is the cluster-dependent version of the unattested across-the-board C_2-copying pattern outlined in Section 1.3.1.3. Copying C_2 rather than C_1 avoids creating a repetition with the base-initial consonant altogether. There is a repetition of the base-second consonant, but, just as in the STVX– cluster-copying case, the repetition is interrupted by another consonant (here C_1): C_2V-$\boldsymbol{C_1}C_2V$... not *C_1V-C_1C_2V...

(18)   TRVX– C_1-copying, STVX– C_2-copying

	Base Type	Root		Reduplicated	Red. Shape
a.	Singleton	√*mako*	→	*ma-mako*	C_1V_2
b.	Stop-sonorant	√*prako*	→	*pa-prako*	C_1V_3
c.	*s*-obstruent	√*stako*	→	*ta-stako*	C_2V_3

Since this pattern shows the same C_1-copying behavior for TRVX– roots as the previous case, we can begin by importing the TRVX– C_1-copying ranking from (15): ANCHOR-L-BR, *CC ≫ CONTIGUITY-BR. (Again, CONTIGUITY-BR is relevant only if the reduplicative vowel corresponds to the base vowel, and the pattern is generable with either sort of reduplicative vowel.) The only difference from the STVX– cluster-copying pattern that is required to generate STVX– C_2-copying is to reverse the role of ANCHOR-L-BR and *CC. In the previous pattern, it was *CC that was dominated by $*C_\alpha VC_\alpha \mathbin{/} \_C_{[\text{-son}]}$, leading to cluster toleration under compulsion of $*C_\alpha VC_\alpha \mathbin{/} \_C_{[\text{-son}]}$ violation. Here, it is ANCHOR-L-BR that must be dominated by $*C_\alpha VC_\alpha \mathbin{/} \_C_{[\text{-son}]}$, such that mis-anchoring is allowed but only when compelled by $*C_\alpha VC_\alpha \mathbin{/} \_C_{[\text{-son}]}$. This is shown in (20).

(19) Generating TRVX– C_1-copying

/RED, prako/	$*C_\alpha VC_\alpha \mathbin{/} \_C_{[\text{-son}]}$	*CC	ANCHOR-L-BR	CONTIG-BR
a. pra-prako		**!		
b. ☞ pa-prako		*		*
c. ra-prako		*	*!	

(20) Generating STVX– C_2-copying with TRVX– C_1-copying

/RED, stako/	$*C_\alpha VC_\alpha \mathbin{/} \_C_{[\text{-son}]}$	*CC	ANCHOR-L-BR	CONTIG-BR
a. sta-stako		**!		
b. sa-stako	*!	*		*
c. ☞ ta-stako		*	*	

The TRVX– C_1-copying with STVX– C_2-copying pattern is instantiated in Sanskrit for cluster-initial roots, as illustrated in (21).[5]

(21) Perfects to cluster-initial roots in Sanskrit (forms from Whitney 1885 [1988])

a. **TRVX– roots → C_1-copying perfects**

	Root		Perfect Tense	
	√$b^h raj$-	'shine'	_ba_-$b^h r\bar{a}j$-_a_	not **_ra_-$b^h r\bar{a}j$-_a_
	√$prac^h$-	'ask'	_pa_-$pr\bar{a}c^h$-_a_	not **_ra_-$pr\bar{a}c^h$-_a_
	√dru-	'run'	_du_-$druv$-_ē_	not **_ru_-$druv$-_ē_
	√$tviṣ$-	'be stirred up'	_ti_-$tviṣ$-_ē_	not **_vi_-$tviṣ$-_ē_

b. **STVX– roots → C_2-copying perfects**

	Root		Perfect Tense	
	√$sparç$-	'touch'	_pa_-$spṛç$-_ē_	not **_sa_-$spṛç$-_ē_
	√$st^h\bar{a}$-	'stand'	_ta_-$st^h\bar{a}$-_u_	not **_sa_-$st^h\bar{a}$-_u_
	√$stamb^h$-	'prop'	_ta_-$stamb^h$-_a_	not **_sa_-$stamb^h$-_a_

---

[5] In Sanskrit, when base-initial ST-clusters arise through vowel deletion (i.e. "zero-grade ablaut"), a different repair — the "CēC pattern" — is initiated. See Chapter 5 for discussion and analysis.

1.3.2.3 TRVX– C$_1$-copying, STVX– Non-copying: Ancient Greek

When we take seriously the notion that it is possible for the reduplicative vowel to *not* stand in correspondence with the base vowel, we admit the possibility of one more type of pattern: TRVX– C$_1$-copying with STVX– non-copying, as schematized in (22). This pattern is attested in Ancient Greek, as shown in (23).

(22)   TRVX– C$_1$-copying, STVX– non-copying

	Base Type	Root		Reduplicated	Red. Shape
a.	Singleton	$\sqrt{mako}$	$\rightarrow$	*m-e-mako*	C$_1$-V
b.	Stop-sonorant	$\sqrt{prako}$	$\rightarrow$	*p-e-prako*	C$_1$-V
c.	*s*-obstruent	$\sqrt{stako}$	$\rightarrow$	*e-stako*	Ø-V

(23)   TRVX– C$_1$-copying, STVX– non-copying in Ancient Greek

a.   **TRVX– roots $\rightarrow$ C$_1$-copying perfects**

Root		Perfect Tense		
$\sqrt{kri}$-	'decide'	κέκριμαι	[k-e-kri-mai]	not **[e-kri-mai]
$\sqrt{pneu}$-	'breathe'	πέπνυμαι	[p-e-pnū-mai]	not **[e-pnū-mai]
$\sqrt{tla}$-	'suffer, dare'	τέτληκα	[t-e-tlɛ̄-k-a]	not **[e-tlɛ̄-k-a]

b.   **STVX– roots $\rightarrow$ Non-copying perfects**

Root		Perfect Tense		
$\sqrt{stel}$-	'prepare'	ἔσταλκα	[e-stal-k-a]	not **[s-e-stal-k-a]
$\sqrt{strat\text{-}eu}$-	'wage war'	ἔστρατευμαι	[e-strat-eu-mai]	not **[s-e-strat-eu-mai]

Before proceeding to the analysis of this sort of pattern, we must now carefully consider the possible properties of reduplicative vowels in these types of systems. The patterns of reduplicant vocalism in the Indo-European languages vacillate between two descriptive types: *copy* vocalism, where the reduplicative vowel is consistently identical (or partially identical) to the vowel of the root or base; or *fixed* vocalism, where the reduplicative vowel has a consistent value which does not co-vary with the vowel of the root or base.

Following Alderete et al. (1999), fixed vocalism (or, generally, fixed segmentism) comes in two types: *phonological* fixed segmentism and *morphological* fixed segmentism. Phonological fixed segmentism can be characterized as *copy + reduction* (see McCarthy & Prince 1995).[6] Within a Base-Reduplicant Correspondence Theory (McCarthy & Prince 1995, 1999) analysis of the present case, this means that the reduplicative vowel stands in BR-correspondence with the base vowel, but is subject to additional markedness pressures (via *the emergence of the unmarked*; McCarthy & Prince 1994) and surfaces as a less marked version of the base vowel.

The alternative type of fixed segmentism, the morphological type, is when the fixed segment is specified in the underlying representation, and thus it is not a "copy" of, or in correspondence

---

[6] Alderete et al. (1999) operate with an analysis in which the fixed segment is not copied but rather epenthetic, though they admit that, in the general case, the copy + reduction model is sufficient as well.

with, any element of the base. While it may be possible to assume in such cases that the fixed segment is literally part of the reduplicative morpheme underlyingly — for example, /REDe/ (cf. Zukoff 2014), which is meant to encode that there is an unspecified string which will be filled out through copying/correspondence which attaches to a subsequent /e/ vowel that is fully specified — this requires a substantivized conception of reduplicative morphemes which itself would require significant investigation and justification. (The specific conceptual issue that would arise from such a representation is whether the fixed segment counts as a part of the reduplicant proper for the purposes of calculating violations for constraints that make direct reference to the reduplicant, for example BR-correspondence constraints and realizational constraints.) A more conservative approach would take the morphological fixed segment to be a distinct affix from the reduplicative morpheme underlyingly — for example, /RED/ + /e/ (see Chapter 2 on Ancient Greek) — and for the two morphemes to be concatenated, along with the root and other suffixes, in the output. As such the fixed segment would definitively not be treated as part of the reduplicant proper.

These questions are significant for the patterns under discussion in this chapter because of the way in which they interact with BR-correspondence constraints. Consider the comparison of two possible $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$ avoidance strategies, assuming that the reduplicative vowel is fixed as [e]: the C_2-copying pattern $C_2e\text{-}C_1C_2VX–$ (cf. Section 1.3.2.2) vs. the non-copying pattern $e\text{-}C_1C_2VX–$ (cf. (22c)). Whether the fixed [e] is phonological or morphological has significant implications for the more detailed phonological and morphological representation of the output candidates. If the reduplicative vowel is a *phonologically* fixed segment, then it stands in correspondence with the root vowel (indicated with an additional correspondence index *3*), and is part of the reduplicant proper (indicated with underlining): $\underline{C_2e_3}\text{-}C_1C_2V_3X–$ and $\underline{e_3}\text{-}C_1C_2V_3X–$. In this case, both candidates have a reduplicant-initial segment that is standing in correspondence with a non–base-initial segment, and thus both definitively violate ANCHOR-L-BR. The non-copying candidate additionally incurs a violation of the constraint ONSET:

(24) **ONSET**
      Assign a violation mark * for each onsetless syllable.

Because of its ONSET violation, the non-copying candidate will be harmonically bounded by the C_2-copying candidate. This means that it is not possible (at least given the constraint set that is relevant for Indo-European reduplication) for a language to display the non-copying pattern if its reduplicative vowel stands in correspondence with base vowel, whether it is a (partial) copy or a fixed vowel.

However, if the reduplicative vowel is a *morphologically* fixed segment, then the detailed output representation is significantly different: $\underline{C_2}\text{-}e\text{-}C_1C_2V_3X–$ and $e\text{-}C_1C_2V_3X–$. In neither candidate does the reduplicative vowel stand in correspondence with a base vowel or belong to the reduplicant proper. In fact, in the non-copying candidate, there is no reduplicant at all in the output. Assuming that BR-correspondence constraints are only evaluated when material actually surfaces in the reduplicant, the non-copying candidate under this configuration no longer violates ANCHOR-L-BR, while the C_2-copying candidate still does. Therefore, with morphological fixed segmentism, non-copying is not harmonically bounded by C_2-copying, and can be selected under the ranking ANCHOR-L-BR ≫ ONSET.[7]

This thus permits the possibility of the TRVX– C_1-copying with STVX– non-copying pattern, exclusively to systems with a morphologically fixed reduplicative vowel. This is fully compatible with the analysis of Greek developed in Chapter 2, where the identification of the fixed reduplicative [e] as a morphologically fixed segment is confirmed by the behavior of vowel initial roots.

---

[7] See Chapter 2 for more explicit justification of this approach.

The complete ranking that generates this pattern is: $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$, ANCHOR-L-BR, $*CC \gg$ ONSET. The schematic analysis is provided in (25) and (26). (Note that CONTIGUITY-BR is necessarily irrelevant in this case, because the reduplicative vowel does not stand in correspondence with the base vowel; it is thus omitted.)

(25)   Generating TRVX– C_1-copying

/RED, e, prako/	$*C_\alpha VC_\alpha$ / $\_C_{[-son]}$	ANCHOR-L-BR	*CC	ONSET
a. pr-e-prako			**!	
b. ☞ p-e-prako			*	
c. r-e-prako		*!	*	
d. _-e-prako			*	*!

(26)   Generating STVX– C_2-copying with TRVX– C_1-copying

/RED, e, stako/	$*C_\alpha VC_\alpha$ / $\_C_{[-son]}$	ANCHOR-L-BR	*CC	ONSET
a. st-e-stako			**!	
b. s-e-stako	*!		*	
c. t-e-stako		*!	*	
d. ☞ _-e-stako			*	*

1.3.3   Factorial Typology

Holding constant that reduplication is at the left edge, and assuming that MAX-BR (the constraint mentioned above that advocates for maximal copying from the base) is effectively inactive (i.e. at the very bottom of the ranking if present in the grammar at all), the factorial typology of the five constraints employed in Sections 1.3.1–1.3.2 yields six possible reduplication systems, as confirmed by OTSoft (Hayes, Tesar, & Zuraw 2013). Five of these six systems are indeed attested within the Indo-European language family. This is briefly summarized in (27) on the following page. Each entry in the factorial typology is notated with the behavior of TRVX– roots and STVX– roots, the type(s) of vocalism which is compatible with the pattern, a language which displays the pattern, and one possible ranking that generates the pattern. (The "candidate" associated with each pattern refers to the corresponding candidate in the preceding tableaux.)

This demonstrates that the basic constraint types that will be employed in the analysis of reduplication in this dissertation lead to a good fit with the attested typology. This exercise of course did not include the full range of constraints that will be used throughout the dissertation, as the full details of the patterns examined are significantly more complex than presented here. Nonetheless, these constraints represent the core of all the analyses to be developed below.

(27) Factorial typology of constraints in Sections 1.3.1–1.3.2

i. *Across-the-board copying patterns*

a. **Across-the-board cluster-copying [C_1C_2V-C_1C_2VX–]**
   TRVX– behavior: Cluster-copying *pra-prako*
   STVX– behavior: Cluster-copying *sta-stako*
   Vocalism: Copy
   Language: **Hittite**   STVX– example: *istu-stu-*
   Ranking: ANCHOR-L-BR, CONTIG-BR, (ONSET) ≫ *CC, *$C_\alpha VC_\alpha$ / _$C_{[-son]}$

b. **Across-the-board C_1-copying [C_1V-C_1C_2VX–]**
   TRVX– behavior: C_1-copying *pa-prako*
   STVX– behavior: C_1-copying *sa-stako*
   Vocalism: Copy or Morphologically fixed
   Language: **Old Irish**   STVX– example: *se-skann*
   Ranking: ANCHOR-L-BR, ONSET, *CC ≫ *$C_\alpha VC_\alpha$ / _$C_{[-son]}$, CONTIG-BR

c. **Across-the-board C_2-copying [C_2V-C_1C_2VX–]**
   TRVX– behavior: C_2-copying *ra-prako*
   STVX– behavior: C_2-copying *ta-stako*
   Vocalism: Copy
   Language: **Unattested**   STVX– example: (hypothetical) *ta-sta-*
   Ranking: CONTIG-BR, (ONSET,) *CC ≫ ANCHOR-L-BR, *$C_\alpha VC_\alpha$ / _$C_{[-son]}$

ii. *Cluster-dependent copying patterns*

d. **TRVX– C_1-copying [T_1V-T_1R_2VX–], STVX– cluster-copying [S_1T_2V-S_1T_2VX–]**
   TRVX– behavior: C_1-copying *pa-prako*
   STVX– behavior: Cluster-copying *sta-stako*
   Vocalism: Copy or Morphologically fixed
   Language: **Gothic**   STVX– example: *ste-stald*
   Ranking: *$C_\alpha VC_\alpha$ / _$C_{[-son]}$, ANCHOR-L-BR, ONSET ≫ *CC ≫ CONTIG-BR

e. **TRVX– C_1-copying [T_1V-T_1R_2VX–], STVX– C_2-copying [T_2V-S_1T_2VX–]**
   TRVX– behavior: C_1-copying *pa-prako*
   STVX– behavior: C_2-copying *ta-stako*
   Vocalism: Copy or Morphologically fixed
   Language: **Sanskrit**   STVX– example: *ta-stamb^h-*
   Ranking: *$C_\alpha VC_\alpha$ / _$C_{[-son]}$, ONSET, *CC ≫ ANCHOR-L-BR ≫ CONTIG-BR

f. **TRVX– C_1-copying [T_1V-T_1R_2VX–], STVX– non-copying [V-S_1T_2VX–]**
   TRVX– behavior: C_1-copying *p-e-prako*
   STVX– behavior: Non-copying *_-e-stako*
   Vocalism: Morphologically fixed
   Language: **Ancient Greek** STVX– example: *e-stal-*
   Ranking: *$C_\alpha VC_\alpha$ / _$C_{[-son]}$, ANCHOR-L-BR, *CC ≫ ONSET, (CONTIG-BR)

Chapter 2

Greek

2.1 Introduction

In Chapter 1, I introduced the core reduplicative alternation exhibited by consonant-initial roots in Ancient Greek: stop-sonorant–initial roots (TRVX– roots) display the standard Indo-European C_1-copying pattern, but other roots beginning in other types of clusters, notably *s*-stop–initial roots (STVX– roots), display non-copying. This variation is phonologically predictable, driven by the NO POORLY-CUED REPETITIONS constraint (*PCR; $\approx$ *$C_\alpha VC_\alpha$ / _$C_{[\text{-sonorant}]}$). Ancient Greek also displays variation in the behavior of vowel-initial roots in the same category; however, this variation is *not* predictable based on phonological properties. While most vowel-initial roots show lengthening of the root-initial vowel, a small set of roots instead display a pattern referred to as "Attic Reduplication", whereby the root-initial *VC* sequence is copied and the root-initial vowel is lengthened.

In this chapter, I will show that, when fully fleshed out, the reduplicative grammar that is necessary to generate the patterns for consonant-initial roots is directly compatible with the productive vowel-lengthening pattern. Attic Reduplication, with respect to both its shape and its distribution, requires additional explanation. This comes in the form of careful consideration of diachrony. I will argue that the pattern arises as a response to phonotactic constraints on the laryngeal consonants in Pre-Greek, which force an alternative reduplication strategy, fully in line with the cluster-specific repetition avoidance strategies detailed in Chapter 1. The pattern itself is constrained by the normal reduplicative grammar and another laryngeal-related process, "laryngeal vocalization". The subsequent loss of the laryngeals forces reanalysis, such that Attic Reduplication comes to be retained in Ancient Greek not directly by phonotactics, but by lexical constraint indexation.

The account developed in this chapter yields three primary results. First, it provides a comprehensive analysis of the synchronic system of perfect-stem formation in Ancient Greek, integrating the minority pattern — Attic Reduplication — with the productive majority patterns. Secondly, it synthesizes previous, relatively informal proposals regarding the origin of the Attic Reduplication pattern into a full-fledged formal synchronic analysis, located at the Pre-Greek stage. And third, more generally, it addresses the problem of how to deal with residual morphophonological patterns within a language's morphological and phonological grammar. Minority patterns of the sort represented by Attic Reduplication are omnipresent cross-linguistically, yet analysts often overlook their value. This account not only demonstrates that such patterns can reveal significant insights about the larger systems in which they are embedded, but illustrates diachronic pathways by which

---

they arise and the diachronic tools which can be employed to yield a meaningful analysis of this kind.[1]

2.1.1 A Preview of the Data

In the Ancient Greek perfect tense, consonant-initial roots display a phonologically regular alternation between two stem-formation patterns, determined by the type of initial cluster. Roots with an initial singleton consonant or an initial stop-sonorant cluster show the overtly reduplicative pattern in (1a), namely, a prefixed copy of the root-initial consonant followed by a fixed vowel [e]. Roots with all other types of initial clusters, primarily obstruent-obstruent clusters, lack overt reduplicative copying and show just the prefixed [e], the "non-copying" pattern in (1b) (which we might think of as "covert" reduplication).[2]

(1) Distribution of stems in the perfect: consonant-initial roots
   a. C_1-copying
       Singleton roots: $\sqrt{}$CV- $\rightarrow$ Ce-CV-, e.g. $\sqrt{}$*dō-* 'give' $\rightarrow$ perfect *de-dō-*
       Stop-sonorant roots: $\sqrt{}$TRV- $\rightarrow$ Te-TRV-, e.g. $\sqrt{}$*kri-* 'judge' $\rightarrow$ perfect *ke-kri-*
   b. Non-copying
       Other cluster roots: $\sqrt{}$CCV- $\rightarrow$ e-CCV-, e.g. $\sqrt{}$*kten-* 'kill' $\rightarrow$ perfect *e-kton-*

Vowel-initial roots likewise show a dichotomy of perfect stem formation patterns. However, unlike among the consonant-initial roots, there is no clear (synchronic) phonological conditioning that regulates the variation — it simply varies by lexeme. Most vowel-initial roots form their perfect stem by lengthening the root-initial vowel, as shown below in (2a). However, a small set of roots, illustrated in (2b), instead displays copying of the root-initial *VC* sequence while simultaneously lengthening the root-initial vowel, a pattern referred to as Attic Reduplication (AR).[3]

(2) Distribution of stems in the perfect: vowel-initial roots
   a. Vowel lengthening: $\sqrt{}$VC- $\rightarrow$ V̄C-
       e.g. $\sqrt{}$*ag-*   'lead' $\rightarrow$ perfect *āg-*
             $\sqrt{}$*onoma-* 'name' $\rightarrow$ perfect *ōnoma-*
   b. Attic Reduplication: $\sqrt{}$VC- $\rightarrow$ VC-V̄C-
       e.g. $\sqrt{}$*ager-* 'gather' $\rightarrow$ perfect *ag-āger-*
             $\sqrt{}$*ol-*   'destroy' $\rightarrow$ perfect *ol-ōl-*

AR's distribution within the synchronic grammar is seemingly arbitrary; the roots which undergo AR have no discernible phonological characteristics that set them apart from roots which undergo the default pattern. Obviously, this requires explanation.

---

[1] In this chapter and throughout the dissertation, I will use the following notations: " > " indicates a diachronic development; " $\rightarrow$ " indicates a synchronic Input-Output mapping; " ** " indicates a form which never occurred; " * " indicates a reconstructed form.

[2] There is one principled exception that will be discussed in Section 2.2.3.

[3] For forms involving /a/, I use non-Attic-Ionic forms, such that its lengthened correspondent is [ā]. In the Attic-Ionic dialect group, [ā] has become [ɛ̄] (see, e.g., Sihler 1995:48–52), such that the relationship between short and long vowel is slightly less transparent. It may not be the case that all forms with [ā] are actually attested outside of Attic-Ionic (i.e. in Doric or other [ā] dialects), but all are at least attested in their [ɛ̄] forms in Attic-Ionic. When necessary, I will refer to [ā]-forms as belonging to "Common Greek."

2.1.2   Outline of the Chapter

This chapter will provide a comprehensive account of the synchronic reduplicative grammar of Ancient Greek, and the historical development of the Attic Reduplication pattern. In exploring the synchronic reduplicative system of attested Ancient Greek in Section 2.2, we will find that the grammar that generates the pattern displayed by consonant-initial roots also directly generates the productive vowel-lengthening pattern for vowel-initial roots. This reveals that it is indeed Attic Reduplication which requires further attention. I argue that copying in these forms is motivated synchronically by the operation of a lexically-indexed REDUP constraint (cf. Zuraw 2002). Additional evidence for the special activity of this constraint comes from a set of exceptions to the generalizations regarding cluster-type–dependent copying in (1), namely, C_1-copying perfects with associated reduplicated presents.

I will then answer the question of how the Attic Reduplication pattern came into being, in Section 2.3. Virtually all of the roots which display AR are reconstructed with an initial laryngeal consonant (cf. Winter 1950:368–369, Beekes 1969:113–126, and many others). With this in mind, I will propose that the historical source of Attic Reduplication (henceforth "Pre-AR") arose at a stage of the language in which the laryngeal consonants were still present ("Pre-Greek"), such that an AR form like Ancient Greek *ol-ɔ̄l-* derives historically from a Pre-AR form *h_3əl-e-h_3l-*. Pre-AR is a deviation from the normal reduplication pattern, restricted to laryngeal-initial roots, induced by the unique phonetic and phonological properties of the laryngeals. The exact nature of Pre-AR is determined in large part by the interaction of the default reduplicative grammar with another laryngeal-related phonological process known as "laryngeal vocalization". The distribution of default reduplication vs. the Pre-AR pattern in Pre-Greek is schematized in (3).

(3)   Default reduplication vs. Pre-AR in Pre-Greek (H = laryngeal consonant)
   a.   Default reduplication pre-forms: *$\underline{C^iV}$-C^iC^kVC- or *$\underline{C^iV}$-C^iC^k-
   b.   Attic Reduplication pre-forms: *$\underline{H^iVC^kV}$-H^iC^kVC- or *$\underline{H^iVC^kV}$-H^iC^k-

Having accounted for the origin of Attic Reduplication, in Section 2.4 I will explore the question of how the pattern could be retained as a minority pattern into attested Ancient Greek, and why it came to be represented in the synchronic grammar via a lexically-indexed REDUP constraint. Subsequent to the initial development of Pre-AR, the laryngeals were lost in Greek, and thus the phonotactics driving the pattern were no longer recoverable. The distinction between different copying patterns for the now vowel-initial roots led to inconsistency within the grammar, and thus required lexical indexation in order to be resolved. I demonstrate using a system for constraint indexation based on Becker (2009) that the evidence across multiple diachronic stages is consistent with a trajectory that results in the synchronic grammar posited in Section 2.2. Incorporating lexically-indexed REDUP into the grammar thus provides a principled way of generating the entire synchronic distribution of reduplicative forms in the Ancient Greek perfect.

Finally, I will examine the residue of AR forms which survive in the present and the aorist in Section 2.5. These forms follow from the general analysis proposed, but also have interesting implications regarding the underlying representation of reduplication in different morphological categories, and also the diachrony of the Greek vowel system.

2.2 Reduplication in Ancient Greek

This section presents the full synchronic analysis of the reduplicative system of the Ancient Greek perfect tense.[4] The three productive patterns of perfect-tense stem-formation ((1a), (1b), and (2a)) are generated from a single, consistent constraint ranking, without appeal to reduplicative templates, under an analysis where two morphemes — /RED/ and /e/ — compete for position at the left edge of the word. I begin in Section 2.2.1 by analyzing the two distinct patterns found among the consonant-initial roots, including a discussion of the underlying morphemic structure of the perfect (Section 2.2.1.3). Section 2.2.2 examines the behavior of vowel-initial roots, and shows that the analysis developed for the consonant-initial roots is directly compatible with the productive vowel-lengthening pattern. In order to resolve the inconsistency between the productive vowel-lengthening pattern and the minority Attic Reduplication pattern, I develop an analysis of AR based on constraint indexation (Section 2.2.2.2). This approach directly carries over to the one other area of exceptionality in the perfect, the cluster-initial roots with unexpected C_1-copying perfects that have associated reduplicated presents (Section 2.2.3). When projected back to the earlier stage of the language in which laryngeal consonants were still present, the default grammar developed in this section, adjusted only slightly and supplemented by an independently motivated phonotactic constraint, will generate the precursor of Attic Reduplication, and indeed the precursor of the cluster-initial roots with exceptional C_1-copying perfects.

2.2.1 Consonant-Initial Roots

2.2.1.1 Data and Generalizations

As discussed in Chapter 1, and summarized in (1) above, the Ancient Greek perfect shows two distinct stem-formation patterns for consonant-initial roots: C_1-copying and non-copying. These are exemplified further below in (4) and (5), respectively.[5] The distribution is determined by the composition of the root-initial string. If the root begins in a single consonant or a cluster comprised of *stop-sonorant*, it takes C_1-copying. All other consonant-initial roots show non-copying.[6]

The basic generalizations that we can draw from this data, which will serve as the basis for the analysis developed in this section, are as follows. When overt reduplication is successfully carried out in forms like those in (4), the string preposed to the root takes the shape *CV*. In such cases, C is always identical to the root-initial consonant. In the overt copying pattern in (4), and indeed also in those cases where consonant-copying fails to occur (as in (5)), *V* is always [e], regardless of the identity of the root vowel.

---

[4] For the facts of Ancient Greek reduplication, see, for example, Schwyzer (1939:646–650). For recent theoretical analyses, see Steriade (1982:195–208, 304–312, 1988:135–136), McIntyre (1992:85–136), Fleischhacker (2005), Keydana (2006:83–91), Halle (2008:333–336).

[5] The Ancient Greek data in this paper is drawn primarily from the survey of verbal forms conducted by van de Laar (2000). All generalizations comport with traditional descriptions, for example, Smyth (1920 [1984]), Schwyzer (1939), Steriade (1982), Sihler (1995).

[6] There is a systematic set of exceptions where roots with other cluster types unexpectedly show C_1-copying. These will be discussed in Section 2.2.3 below.

(4) C$_1$-copying reduplication

	Root		Present Tense		Perfect Tense	

a. Roots with initial singleton consonants (CVX–)

√dō-	'give'	δίδωμι	[<u>di</u>-dō-]	δέδωκα	[<u>de</u>-dō-]
√pemp-	'send'	πέμπω	[pemp-]	πέπεμπται	[<u>pe</u>-pemp-]
√lu-	'loosen'	λύω	[lu-]	λέλυκα	[<u>le</u>-lu-]

b. Roots with initial stop-sonorant clusters (TRVX–)

√kri(n)-	'decide'	κρίνω	[krī-n-]	κέκριμαι	[<u>ke</u>-kri-]
√pneu-	'breathe'	πνέω	[pne-]	πέπνῡμαι	[<u>pe</u>-pnū-]
√tla-	'suffer'	τλάω	[tla-]	τέτλᾱκα	[<u>te</u>-tlā-]

(5) Non-copying reduplication

	Root		Perfect Tense		

a. Roots with other initial clusters (STVX– and others)

√stel-	'prepare'	ἔσταλκα	[<u>e</u>-stal-]	not **[<u>se</u>-stal-]
√kten-	'kill'	ἔκτονα	[<u>e</u>-kton-]	not **[<u>ke</u>-kton-]
√pseud-	'lie'	ἔψευσμαι	[<u>e</u>-pseus-]	not **[<u>pe</u>-pseus-]

b. Roots with initial geminates (C:VX–)

| √rreu- | 'flow' | ἐρρύηκα | [<u>e</u>-rru-] | not **[<u>re</u>-rru-] |
| √sseu- | 'hasten' | ἔσσυμαι | [<u>e</u>-ssu-] | not **[<u>re</u>-ssu-] |

2.2.1.2 The Analysis of Reduplication in Base-Reduplicant Correspondence Theory

Before proceeding to the analysis, I must first introduce the foundational elements of the formal system I will be using to analyze reduplication. I adopt Base-Reduplicant Correspondence Theory (BRCT; McCarthy & Prince 1995, 1999) as my framework for analyzing reduplication.[7] In BRCT, the output string is divided into two substrings: (i) the Reduplicant (R) and (ii) the Base of reduplication (B). These two strings (and the segments comprising them) stand in Base-Reduplicant (BR) correspondence. Faithfulness constraints act on this correspondence relation to promote identity between the two strings.

As will be demonstrated throughout this chapter, BR-faithfulness itself actually plays a relatively small role in the analysis of reduplication in Ancient Greek. The one major exception is the BR-faithfulness constraint ANCHOR-L-BR (McCarthy & Prince 1995:123, 1999:295), as defined in (6). In effect, this constraint penalizes copying from non–root-initial position, as briefly illus-

---

[7] For recent alternative theoretical approaches to reduplication, consult Raimy (2000, 2009), Inkelas & Zoll (2005), Idsardi & Raimy (2008), Frampton (2009), Kiparsky (2010b), Saba Kirchner (2010, 2013), McCarthy, Kimper, & Mullin (2012).

trated in (7).[8] (On the composition of the input to this tableau, see the proceeding discussion.) In all tableaux, the output string constituting the reduplicant is underlined.

(6) **ANCHOR-L-BR**
    Assign one violation mark * if the segment at the left edge of the base does not stand in correspondence with the segment at the left edge of the reduplicant.

(7) C_1-copying reduplication: $\sqrt{kri\text{-}} \rightarrow \varkappa\acute{\varepsilon}\varkappa\rho\iota\mu\alpha\iota$ [k̲-é-kri-mai] 'he has (been) judged'

/RED, e, kri-/	ANCHOR-L-BR
a. ☞ k̲-e-kri-	
b. r̲-e-kri-	*!

Within standard BRCT, the way to represent the fact that a particular morpheme is a reduplicative morpheme is to record its underlying representation as /RED/. That is to say, for the specific case of Ancient Greek currently under discussion, the morphology selects /RED/ as the underlying representation for the morphosyntactic feature PERFECT. While the nature of this /RED/ morpheme is rarely made precise in work relating to BRCT, Zuraw (2002:403) makes its function explicit: "It is the presence of a RED morpheme in the input that requires morphological labelling of one part of the word as 'base' and another part as 'reduplicant'."

The labeling of output substrings as Base and Reduplicant is as an operation in GEN. For every surface string that is a possible output of GEN, there are multiple possible morphological labelings of that string; namely, there are candidates where particular substrings are respectively labeled as Base and Reduplicant (e.g. [{$_{red}$xy}{$_{base}$xyz}]), and string-identical candidates in which no substring has either of these labels (e.g. [x̄yxyz]). When an output candidate contains substrings labeled Base and Reduplicant, those substrings by definition stand in BR-correspondence. Set within Zuraw's (2002) larger proposal on "coupling" (which unites "aggressive reduplication" — a non-morphological type of surface string-internal correspondence — with more traditional morphological reduplication), there must be a constraint that is sensitive to whether or not an output candidate contains labeled Base and Reduplicant substrings (and, as a result, instantiates the BR-correspondence relation). Adapting Zuraw's (2002:403) "REDUP" constraint, which enforces the general coupling relation (that is, the division of the output into multiple corresponding substrings, irrespective of morphological labeling), I propose that the morphological labeling of output strings as Base and Reduplicant is regulated by the constraint REDUP(RED), defined in (8).

(8) **REDUP(RED)** (based on Zuraw's REDUP constraint)
    Assign a violation mark * if there is a RED morpheme in the input, but the output does not contain substrings labeled as Base and Reduplicant.[9]

---

[8] The function of ANCHOR-L-BR partially overlaps with the LOCALITY constraint proposed by Nelson (2003). However, it appears that directly replacing ANCHOR-L-BR with LOCALITY would not be compatible with the analysis developed in this chapter (though perhaps one could be developed, given a number of different assumptions), as it would lead to a ranking paradox between the C_1-copying pattern and the vowel-lengthening pattern. In any case, Nelson's (2003) system allows for both constraints to exist independently in the grammar.

[9] In Zukoff (2017c), I referred to the constraint that had this function as REALIZE MORPHEME(RED). This was not the most felicitous use of Kurisu's (2001) "REALIZE MORPHEME" proposal, since it went beyond the requirement that an input morpheme have an output exponent (which is the crux of Kurisu's proposal) to instead requiring that morpheme to have a specific type of surface exponent. I thus adopt the REDUP-based formulation developed here. Furthermore, in Chapter 4 I will employ REALIZE MORPHEME constraints in the analysis of the Germanic Strong Verb system, in the way Kurisu intended. Therefore, continued use of REALIZE MORPHEME(RED) in this context would be doubly inappropriate.

REDUP(RED) is a morphologically specific version of the general coupling constraint REDUP, as Base-Reduplicant labeling is a morphologically specific version of the general coupling labeling operation. The underlying /RED/ morpheme serves essentially as a null morpheme with a diacritic activating the morphologically restricted REDUP(RED) constraint.

The fact that this special labeling operation is morphologically restricted is important in one additional way. It must be the case that there is a restriction on which output substring(s) can be labeled as the Reduplicant. This relates to the notion of "Consistency of Exponence", which McCarthy & Prince (1993b:21) define as: "No changes in the exponence of a phonologically-specified morpheme are permitted". This is a property of GEN. While the reduplicant is phonologically *un*specified underlyingly in exactly the way that would allow for some freedom in the determination of its exponent, all non-reduplicative morphemes — which have phonologically-specified content underlyingly — are strictly bound by Consistency of Exponence. Their output content must be identified with their input content, and thus cannot be parsed as part of the string labeled as the Reduplicant. This means that the only output content that GEN can label as a Reduplicant is content that lacks an input correspondent. Therefore, any candidate which is labeled with a Reduplicant string must contain output content with no input correspondent.[10]

This clarifies a question that will arise in analyzing the non-copying pattern in Section 2.2.1.5 and the vowel-lengthening pattern in Section 2.2.2.1: in the derivation of a perfect-tense form, candidates where no overt copying takes place must not violate ANCHOR-L-BR, even though it is not the case that there is a reduplicant correspondent of the leftmost segment of the root. This follows directly from the preceding discussion. If a candidate lacks material that has no input correspondent, that candidate cannot be labeled with a Reduplicant in the output. If a candidate has no Reduplicant in the output, the Base-Reduplicant correspondence relation has failed to be instantiated, and thus all Base-Reduplicant faithfulness constraints (of which ANCHOR-L-BR is one) are vacuously satisfied. Since the input to such a derivation contains a /RED/ morpheme, an output candidate lacking overt copying violates REDUP(RED). However, REDUP(RED) is indeed a violable, ranked constraint in CON. Given a low-enough ranking of this constraint, such an output can be permitted to surface. This is the case for the non-copying and vowel-lengthening patterns in Ancient Greek.

2.2.1.3 Perfect Reduplication: One Morpheme or Two?

Returning to the stem-formation patterns of the consonant-initial roots laid out in Section 2.2.1.1, we can observe that the [e] vowel which precedes the root in the perfect does not co-vary with a segment in the base. *A priori*, cases where a fixed segment occurs in a reduplicative context admit of two analytical options (see Alderete et al. 1999; see also the discussion in Chapter 1): a phonological analysis or a morphological analysis. Under the phonological approach, the segment is taken to be copied from the base as part of the reduplicant, but markedness constraints induce phonological reduction (a case of *the emergence of the unmarked*; McCarthy & Prince 1994, 1995). As will be

---

[10] Strictly speaking, this means that the reduplicant label could also be applied to an epenthetic segment, since it lacks an input correspondent which would subject it to Consistency of Exponence. This perhaps is not an unwelcome result, for at least two reasons.

For one, Saba Kirchner (2010, 2013) develops a theory of reduplication where the underlying morpheme that leads to a surface reduplicant can, under the right conditions, surface with an epenthetic exponent rather than a strictly reduplicative one (see also Alderete et al. 1999 on phonological fixed segments in reduplication). Regardless of the specific analysis, this indicates a closer relationship between reduplication and epenthesis than might have been expected.

Second, some of the recent work on copy epenthesis (Kitto & de Lacy 1999, Stanton & Zukoff to appear) has argued that epenthetic vowels, regardless of whether they surface as copies of an adjacent vowel or as a default vowel, stand in correspondence with a neighboring segment. This also implies that reduplication and epenthesis bear some deeper structural similarity (see also Yu 2005, Stanton & Zukoff 2016a,b; cf. Kawahara 2007).

shown in Sections 2.2.1.4 and 2.2.1.5, such an analysis is unworkable for Ancient Greek, as it would lead to a ranking paradox.

Therefore, I will proceed with the alternative, morphological analysis. Rather than identifying the fixed segment as belonging to the reduplicant proper (i.e., arising via "copying"), we can view it as an independent morpheme, bound to co-occur with the reduplicative morpheme. This situation resembles, for example, that of *schm*-reduplication in English (Alderete et al. 1999:355–357; cf. Nevins & Vaux 2003).[11] Under this approach, a typical reduplicated form like *perfect* χέχριμαι [kékrimai] will be decomposed as in (9) below.

(9)   Morphological decomposition of the perfect

k̲-	e-	kri	-mai
REDUPLICANT	FIXED SEGMENT AFFIX	ROOT	INFLECTION

With the fixed [e] identified as an independent morpheme, two questions remain to be answered in order to complete an analysis of the consonant-initial roots: (i) how does the reduplicant come to take the shape of a single consonant in the pattern in (4); and (ii) how do we derive the $C \sim \emptyset$ alternation that distinguishes the C_1-copying pattern in (4) from the non-copying pattern in (5). These two questions are taken up immediately below.

2.2.1.4   The C_1-Copying Pattern

Since the non-copying pattern exists, the constraints which motivate having segments in the reduplicant must be violable in Ancient Greek. As will be shown in Sections 2.2.1.5 and 2.2.2.1, violation of these constraints can be compelled by higher-ranked phonotactic considerations. When these constraints are not in danger of being violated, the constraints which enforce copying are satisfied. This is the case for the roots with C_1-copying.

One of the two constraints that promote copying in the present case is the phonotactic constraint ONSET (cf. Itô 1989), defined in (10). ONSET promotes copying because failing to copy would cause the fixed /e/ to surface without a preceding consonant and thus head an onsetless syllable. While onsetless syllables are permitted in Ancient Greek, they are actively disfavored. This can be seen from a number of processes, including vowel contraction (/...V_1V_2.../ → [...$\bar{V}$...]), cross-word elision ("crasis"; /...V_1#V_2.../ → [...$\emptyset$#V_2...]), and "*nu* movable" (see Golston 2014). Therefore, ONSET will specifically militate for the presence of a consonant-final (and also consonant-initial) reduplicant, to accommodate the fixed /e/ morpheme.

(10)   **ONSET**
       Assign a violation mark * for each onsetless syllable in the output.

The other constraint that promotes copying has already been introduced: REDUP(RED) from (8) above, repeated below. This constraint demands that inputs containing an underlying /RED/ morpheme must contain an output string labeled as a reduplicant, which can only be instantiated through copying. Therefore, failure to copy will result in violation of this constraint.

(11)   **REDUP(RED)**
       Assign a violation mark * if there is a RED morpheme in the input, but the output does not contain substrings labeled as Base and Reduplicant.

---

[11] As with *schm*-reduplication and similar cases, it is unclear if these two morphs have distinct functions.

REDUP(RED) and ONSET thus prefer an overt reduplicant of the shape *#(C...)C-* (followed immediately by the fixed *e*), but make no further demands regarding reduplicant shape. McCarthy & Prince (1986 [1996], *et seq.*) argue that "reduplicative templates" must take the shape of "genuine units of prosody" (*syllable*, *foot*, *prosodic word*). In the Ancient Greek perfect, neither of the two overtly reduplicative patterns take on such a shape: the C_1-copying pattern currently under discussion is a single consonant; the Attic Reduplication pattern is a necessarily heterosyllabic *VC* sequence. Therefore, it seems unsuitable to pursue an analysis of reduplicant shape based on templates of any sort. Furthermore, the fact that such different reduplicant shapes result from roots of different shapes would make such an analysis difficult.

Instead, this section will develop an "a-templatic" analysis (cf. Spaelti 1997, Gafos 1998, Hendricks 1999, Riggle 2006, among many others). A-templatic accounts of minimal reduplication patterns such as these rely on the activity of a "size restrictor" constraint. A size restrictor constraint will in some way penalize the reduplicant for having excessive length (or indeed any length at all in most cases). When the size restrictor outranks MAX-BR (the constraint which advocates copying each segment of the base into the reduplicant; see McCarthy & Prince 1995), the minimal reduplicant shape emerges as optimal.

(12)   **MAX-BR**
         Assign a violation mark * for each segment of the base that does not have a correspondent in the reduplicant.

Following Hendricks (1999), I use an ALIGNMENT constraint (McCarthy & Prince 1993a, Prince & Smolensky 1993/2004) as the size restrictor. Given that the fixed [e] has been identified as a distinct morpheme, it can have an alignment constraint defined for it, as in (13).[12]

(13)   **ALIGN-/e/-L**
         Assign one violation * for every segment that intervenes between the left edge of the exponent of the fixed segment affix /e/ and left edge of the prosodic word.[13]

When ranked above MAX-BR, this constraint will induce the desired minimization effect, because increasing the length of the reduplicant will necessarily increase the number of violations of this constraint, as illustrated in (14).[14] Given that we do see copying in the general case, ALIGN-/e/-L must be ranked below ONSET and/or REDUP(RED), since failure to copy anything will satisfy ALIGN-/e/-L but violate ONSET and REDUP(RED). This is demonstrated in (15) below. (It will be shown in Section 2.2.2.1 that the vowel-lengthening pattern requires the ranking ONSET ≫ ALIGN-/e/-L ≫ REDUP(RED).)

---

[12] I employ gradient alignment constraints. McCarthy (2003) argues that alignment constraints (and indeed all OT constraints) should be defined categorically, not gradiently. However, Yu (2007:38–42) demonstrates that McCarthy's restriction to categorical alignment constraints does not actually avoid the typological overgeneration problem it seeks to solve. The facts here are compatible with a categorical analysis, in which the single gradient constraint is separated into two categorical constraints: one alignment constraint defined with reference to an intervening *segment*, and another defined with reference to an intervening *syllable*.

[13] In cases where the underlying /e/ morpheme coalesces with a root-initial vowel, this constraint is evaluated with respect to that coalesced vowel.

[14] To ensure that this constraint does not have the effect of placing the [e] to the left of the reduplicant, we may also need to include an alignment constraint on the reduplicant (ALIGN-RED-L), ranked above it. However, such an ordering would generally be disfavored anyway by higher-ranked ONSET.

(14) Minimizing the reduplicant: √*pemp-* → πέπεμπται [p-é-pemp-tai] 'he has (been) sent'[15]

/RED, e, pemp-/	ALIGN-/e/-L	MAX-BR
a. ☞ p-e-pemp-	*	***
b. pem-e-pemp-	**!*	*

(15) Ensuring consonant-copying: √*pemp-* → πέπεμπται [p-é-pemp-tai] 'he has (been) sent'

/RED, e, pemp-/	REDUP(RED)	ONSET	ALIGN-/e/-L
a. ☞ p-e-pemp-			*
b. _-e-pemp-	*!	*!	

ALIGN-/e/-L would also be capable of selecting the minimal C_1 reduplicant for roots with initial clusters. However, in accounting for the non-copying pattern in Section 2.2.1.5 below, we will see that the ranking of ONSET and/or REDUP(RED) over ALIGN-/e/-L would in that case prefer extending the reduplicant to include the whole cluster (i.e. *[kt-e-kton-] ≁ [_-e-kton-]). To avoid this outcome, we must supplement the ranking with a constraint against consonant clusters, *CLUSTER (*CC, defined in (16)), ranked above ONSET and REDUP(RED) (as shown in (17a) below).

(16) ***CC**
Assign one violation mark * for each consonant cluster.

We can view this as an *emergence of the unmarked* effect in reduplication; while consonant clusters are permitted generally, they are prevented from occurring in the reduplicant, even for roots beginning in consonant clusters. Therefore, MAX-IO and DEP-IO dominate *CC, but *CC dominates MAX-BR (see McCarthy & Prince 1994, 1995), as detailed in the ranking in (17b).[16] This ranking prefers the C_1-copying candidate (18a) to the cluster-copying candidate (18b) and the cluster-simplifying candidate (18c).

(17) **Rankings:**
    a.   *CC ≫ ONSET ≫ ALIGN-/e/-L ≫ REDUP(RED)   (see Section 2.2.2.1)
    b.   DEP-IO, MAX-IO ≫ *CC ≫ MAX-BR

These size-minimizing constraints show why reduplication cannot be larger than a single consonant; however, they do not specify which consonant should be copied into this position. The constraint which will enforce copying of root-C_1, as opposed to, for example, root-C_2 (as in a candidate like [r-e-kri-], (20b)) has already been mentioned: ANCHOR-L-BR. (Nothing yet fixes this constraint's relative ranking.) Candidates in which the /e/ is infixed, like [_-k-e-ri-] (20c) or

---

[15] We might also consider a candidate where the fixed /e/ intrudes on a multi-segmental reduplicant: *[p-e-m-pemp-]. This would do equally well on ALIGN-/e/-L as candidate (14a), but incur one fewer MAX-BR violations. This candidate has a thoroughly discontiguous reduplicant; however, given that copied segments from the base are also discontiguous, and in the very same way, it is unclear if this would actually violate CONTIGUITY-BR (the Base-Reduplicant version of CONTIGUITY-IO, as defined in (19) below). Regardless, this could be ruled out with an additional Alignment constraint: ALIGN-ROOT-L (see Section 2.5.3). If this constraint also dominates MAX-BR, then copying additional segments and placing them between the fixed /e/ and the root will not be superior to single-consonant copying.

[16] The constraint CONTIGUITY-IO ((19) below) must independently outrank *CC. It can recreate many of the cluster faithfulness effects of higher-ranked MAX-IO and DEP-IO (especially word-internally), but is not completely co-extensive (for example, with word-edge clusters).

(18)　C$_1$-copying reduplication: $\sqrt{kri\text{-}}$ → κέκριμαι [k̲-é-kri-mai] 'he has (been) judged'

/RED, e, kri-/	DEP-IO	MAX-IO	*CC	ALIGN-/e/-L	MAX-BR
a. ☞　k̲-e-kri-			*	*	**
b. 　k̲r-e-kri-			**!	**	*
c. 　k̲-e-ki-		*!		**	*
d. 　k̲er-e-k*e*ri-	*!			***	*

[k̲e-k-e-ri-] (20d), would alleviate the root's *CC violation, but these are ruled out if the constraint CONTIGUITY-IO (Kenstowicz 1994, McCarthy & Prince 1995), as defined in (19), dominates *CC.

(19)　**CONTIGUITY-IO**
　　　Assign one violation mark * for each pair of segments that are adjacent in the input but have non-adjacent correspondents in the output.

(20)　C$_1$-copying reduplication: $\sqrt{kri\text{-}}$ → κέκριμαι [k̲-é-kri-mai] 'he has (been) judged'

/RED, e, kri-/	ANCHOR-L-BR	CONTIGUITY-IO	*CC
a. ☞　k̲-e-kri-			*
b. 　r̲-e-kri-	*!		*
c. 　_-k-e-ri-		*!	
d. 　k̲e-k-e-ri-		*!	

　　This analysis successfully generates the basic C$_1$-copying reduplication pattern. The ranking of the constraints employed thus far is summarized in (21) below.[17]

(21)　**Ranking Summary**

MAX-IO　　DEP-IO　　CONTIGUITY-IO　　ANCHOR-L-BR

*CC　　　　ONSET　　　REDUP(RED)

ALIGN-/e/-L

MAX-BR
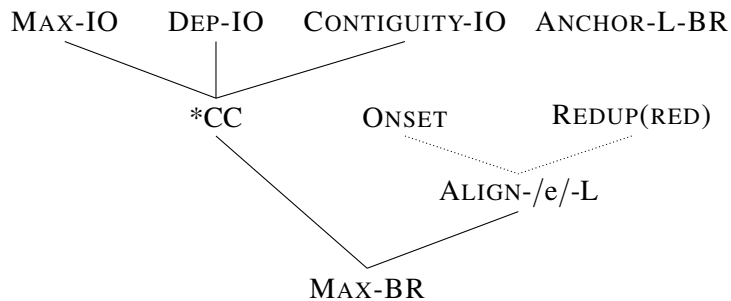
2.2.1.5　The Non-Copying Pattern

Roots that display the non-copying pattern are exemplified in (22) below (which repeats the forms from (5) above, plus additional data). These forms show that the C$_1$-copying pattern is blocked for cluster-initial roots *not* of the shape *stop-sonorant* — i.e., *fricative-stop* (22a), *stop-stop* (22b), and *stop-fricative* (22c) — and also for roots with initial geminates (22d). The behavior of the non–stop-sonorant cluster-initial roots is explained by the operation of the anti-repetition constraint

---

[17] We currently have evidence only that one of ONSET or REDUP(RED) dominates ALIGN-/e/-L. This is indicated with the dotted lines between these constraints.

*PCR (≈ *$C_\alpha VC_\alpha$ / _$C_{[-sonorant]}$), as proposed in Chapter 1. This constraint, which is a stand-in for a fuller analysis to be developed in Chapter 6, prohibits consonant repetitions (i.e. identical consonants across a vowel) in pre-obstruent position. The basic definition from Chapter 1 is repeated here in (23).

(22) Non-copying reduplication[18]

Root		Perfect Tense		

a. Roots with initial fricative-stop clusters (STVX–)

√stel-	'prepare'	ἔσταλκα	[e-stal-]	not **[s-e-stal-]
√sper-	'sow'	ἔσπαρμαι	[e-spar-]	not **[s-e-spar-]
√skep-	'view'	ἔσκεπται	[e-skep-]	not **[s-e-skep-]
√strateu-	'lead an army'	ἐστράτευμαι	[e-strateu-]	not **[s-e-strateu-]
√sbes-	'extinguish'	ἔσβεσμαι	[e-sbes-]	not **[s-e-sbes-]
√zdeug-	'yoke'	ἔζευγμαι	[e-zdeug-]	not **[z(d)-e-zdeug-]

b. Roots with initial stop-stop clusters (TTVX–)

√kten-	'kill'	ἔκτονα	[e-kton-]	not **[k-e-kton-]
√ktis-	'found'	ἔκτισμαι	[e-ktis-]	not **[k-e-ktis-]
√ptis-	'pound'	ἔπτισμαι	[e-ptis-]	not **[p-e-ptis-]
√pʰtʰer-	'destroy'	ἔφθαρμαι	[e-pʰtʰar-]	not **[p⁽ʰ⁾-e-pʰtʰar-]

c. Roots with initial stop-fricative clusters (TSVX–)

√pseud-	'lie'	ἔψευσμαι	[e-pseus-]	not **[p-e-pseus-]
√kses-	'shave'	ἔξεσμαι	[e-kses-]	not **[k-e-kses-]

d. Roots with initial geminates (C:X–)

√rreu-	'flow'	ἐρρύηκα	[e-rru-]	not **[r-e-rru-]
√sseu-	'hasten'	ἔσσυμαι	[e-ssu-]	not **[s-e-ssu-]

(23) **NO POORLY-CUED REPETITIONS (*PCR) [≈ *$C_\alpha VC_\alpha$ / _$C_{[-sonorant]}$]**
For each sequence of repeated identical consonants separated by a vowel ($C_\alpha VC_\alpha$), assign a violation * if that sequence immediately precedes an obstruent.

With respect to the analysis, *PCR penalizes C_1-copying candidates for roots with initial *consonant-obstruent* clusters, such as (24b). When ranked above ONSET and REDUP(RED), *PCR rules out the default pattern in favor of the non-copying candidate (24b). Besides non-copying, tableau (24) shows two additional ways of avoiding the problematic repetition: copying the entire root-initial cluster (candidate (24c)), and copying just root-C_2 (candidate (24d)). Since these are not the preferred solutions to the *PCR problem, this shows that *CC and ANCHOR-L-BR must outrank ONSET and REDUP(RED).

---

[18] There is also some variation between C_1-copying and non-copying for roots with initial *voiced stop* + *liquid* clusters. For discussion of these and a few other marginal facts of Greek, see Chapter 6.

44

(24)   Non-copying perfects: √*kten-* → ἔκτονα [é-kton-a] 'I have killed'

/RED, e, kton-/	*PCR	ANCHOR-L-BR	*CC	ONSET	REDUP(RED)
a. k̲-e-kton-	*!		*		
b. ☞ _-e-kton-			*	*	*
c. k̲t̲-e-kton-			**!		
d. t̲-e-kton-		*!	*		

In order to select the non-copying candidate (24b) over the C$_2$-copying candidate (24d), candidate (24d)'s ANCHOR-L-BR violation must be fatal. If it were the case that (24b) also suffered from an ANCHOR-L-BR violation, the evaluation would select (24d), since it avoids the ONSET and REDUP(RED) violations (in fact, candidate (24b) would be harmonically bounded by (24d)). Therefore, it is necessary that a candidate like (24b) *does not* violate ANCHOR-L-BR. This informs both the analysis of the fixed [e] and the abstract phonological representation of the non-copying form. If we had pursued a (copy + reduction) phonological fixed segmentism analysis of the [e] vowel, (24b) necessarily would violate ANCHOR-L-BR, since its leftmost reduplicant segment ([e]) would be in correspondence with a segment not at the left edge of the base (i.e. the root vowel). Therefore, the [e] must indeed be analyzed morphologically.

But the ANCHOR-L-BR question does remain even under the current morphological analysis, since it is conceivable that an empty reduplicant is still evaluated for ANCHOR-L-BR. Yet, the notion that this candidate violates REDUP(RED) requires that this candidate truly has no reduplicant in the output (see again Section 2.2.1.2). Without a reduplicant, there is nothing to instantiate the "R" in the BR (Base-Reduplicant) correspondence relation. This requires that BR-faithfulness constraints are vacuously satisfied when no reduplicative copying takes place, as the string(s) necessary to establish the correspondence relation is undefined. By this reasoning, (24a) vacuously satisfies ANCHOR-L-BR and is selected as the winner under the ranking shown in (24), schematized in the Hasse diagram in (25) below.

(25)   **Ranking Summary**

        *PCR   *CC   ANCHOR-L-BR
            \ | /
          {ONSET, REDUP(RED)}
                    |
             ALIGN-/e/-L
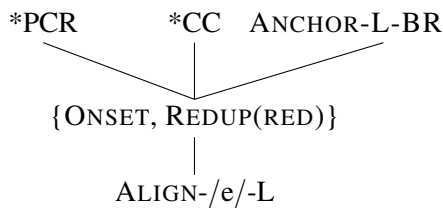
If we view geminates as long single consonants, rather than as sequences of two identical consonants, *PCR will not be able to explain the behavior of the geminate-initial roots in (22b). (If we did treat the geminates as sequences, then *PCR — at least in its final, most precise version, as laid out in Chapter 6 — would be sufficient.) Instead, the answer here lies in Base-Reduplicant faithfulness. A high-ranking constraint demanding identity for consonant length between base and reduplicant — IDENT[long]-C-BR[19] — would prevent copying a root-initial geminate as a reduplicant singleton. Initial geminates are disallowed, as evidenced by the initial degemination observed for these roots in isolation: for example, /sseu-/ → [seu-] (*#Cː ≫ IDENT[long]-C-IO;

---

[19] This constraint must be limited to consonant length, because base-reduplicant alternations in vowel length are present in Attic Reduplication forms.

defined in (26)). These two factors interact to make any sort of copying impossible for these roots, as demonstrated in (27) below.

(26)   Constraints relating to geminates

     a.   **IDENT[long]-C-BR**
        Assign one violation mark * for each pair of consonants in BR-correspondence which differ in length.

     b.   **IDENT[long]-C-IO**
        Assign one violation mark * for each pair of consonants in IO-correspondence which differ in length.

     c.   ***#C:**
        Assign one violation mark * for each word-initial geminate (i.e. long) consonant.

(27)   Non-copying to initial geminate roots: √*sseu* → ἔσσυμαι [é-ssu-mai] 'I have hastened'

/RED, e, ssu-/	*#C:	IDENT[long]-C-BR	IDENT[long]-C-IO	ONSET	REDUP(R)
a. ☞ _-e-ssu-				*	*
b. <u>ss</u>-e-ssu-	*!				
c. <u>s</u>-e-ssu-		*!			
d. <u>s</u>-e-su-			*!		

2.2.1.6   Local Summary

There is one exception to the non-copying pattern for non–stop-sonorant–initial roots. This will be detailed and analyzed in Section 2.2.3 below. These forms aside, we have now accounted for the behavior of all consonant-initial roots in the synchronic grammar of Ancient Greek.

2.2.2   Vowel-Initial Roots

2.2.2.1   Vowel-Lengthening Perfects

The productive pattern for perfect-stem formation for vowel-initial roots is lengthening of the root-initial vowel. Some examples of this pattern are given in (28) below. Beyond the fact that this is the most frequent behavior, an argument for the productivity of this pattern in Ancient Greek comes from the behavior of newly coined verbs such as denominative ἐμπολάω [empolá-ɔ̄] 'get by traffic; purchase' (Liddell & Scott 1889:255). This verb is derived from the noun ἐμπολή [empolɛ́] 'merchandise', which itself was formed from the prefix *en-* 'in' plus a verbal stem, perhaps that of πωλέω [pɔ̄lé-ɔ̄] 'sell' (cf. Chantraine 1980:344). ἐμπολάω makes a perfect ἠμπόληκα [ɛ̄mpólɛ̄-k-a] (← /RED, e, em(-)polɛ̄, k, a/), with lengthening of a vowel which etymologically belongs to a prefix.[20] Given that prefixes are normally excluded from reduplication (that is, preverbs normally surface to the left of the reduplicant), this is evidence of innovation, and thus productivity. Additionally, we find some roots with Attic Reduplication perfects that seem to have later developed vowel-lengthening perfects: for example, √*op* 'see' → (older) perfect ὄπωπα [óp-ɔ̄p-a], (later; perhaps only in the passive) perfect ὦπται [ɔ̂p-tai]. This indicates that the vowel-lengthening pattern is attracting older forms into its sphere, re-making them according to the productive pattern.

---

[20] Chantraine (1980:344) also provides a late perfect form ἐμπεπόληκα [em-p-e-pólɛ̄-k-a], where the prefix has been re-identified as such, and thus excluded from the root for the purposes of perfect reduplication.

(28)   Vowel-lengthening perfects (forms from Smyth 1920 [1984]:147, Schwyzer 1939:65)

Root		Present Tense		Perfect Tense	
√onoma-	'name, call'	ὀνομάζω	[onoma-]	ὠνόμακα	[ɔ̄noma-]
√ortʰo-	'set upright'	ὀρθόω	[ortʰo-]	ὤρθωκα	[ɔ̄rtʰɔ̄-]
√etʰel-	'wish'	ἐθέλω	[etʰel-]	ἠθέληκα	[ɛ̄tʰelɛ̄-]
√elpid-	'hope'	ἐλπίζω	[elpizd-]	ἤλπικα	[ɛ̄lpi-]
√aŋgel-	'announce'	ἀγγέλλω	[aŋgel-]	ἄγγελκα	[āŋgel-]
√ag-	'lead'	ἄγω	[ag-]	ἄγμαι	[āg-]

The grammar developed thus far is consistent with vowel-initial roots forming their perfects through vowel-lengthening. The length derives from the underlying mora contributed by the fixed segment affix /e/. The output long vowel is the result of coalescence of the root-initial vowel with the fixed /e/. This analysis requires that the constraint militating against coalescence — UNIFORMITY-IO (McCarthy & Prince 1995:123) — is not highly-ranked. There is independent evidence for this, as Ancient Greek has an extensive process of mora-preserving "vowel contraction" (see Smyth 1920 [1984]:20–21, de Haas 1988). To generate vowel coalescence/contraction, we can employ the ranking in (29), which is illustrated in (30). The activity of MAX-μ-IO selects candidate (30d) over (30c), but there is no evidence for its relative ranking.

(29)   **Ranking:** MAX-IO ≫ ONSET ≫ UNIFORMITY-IO[21]

(30)   Vowel contraction: /...Ce-o.../ → [...Cō...]

/...Ce$_1$-o$_2$.../	MAX-IO	ONSET	UNIFORMITY-IO	MAX-μ-IO
a. ...Ce$_1$-o$_2$...		*!		
b. ...Co$_2$...	*!			*
c. ...Co$_{1,2}$...			*	*!
d. ☞ ...Cō$_{1,2}$...			*	

Vowel-lengthening in the perfect, as well as the vowel-lengthening that occurs in the affixation of the past tense indicative "augment" prefix (see Smyth 1920 [1984]:145–146), which is also underlyingly /e/, results in coalescence outputs different than those generally found in vowel contraction. For vowel-lengthening perfects and augmented forms, coalescence of /e/ + /e,o/ generally produces *lax* long mid vowels [ɛ̄,ɔ̄] (orthographic < η, ω >);[22] yet, in vowel contraction, coalescence produces *tense* long mid vowels [ē,ō] (orthographic < ει, ου >). These distributions are straightforward when viewed from the diachronic perspective, as the lengthening pattern arises in a period of Greek prior to the first appearance of the tense long mid vowels.

Once the tense vowels become the normal result of contraction, the lax vowels of the perfect must be relegated to irregular morphophonology. This can be represented by a markedness constraint specific to the perfect that bans tense long mid vowels: *[ē,ō]$_{PERF}$ (see, e.g., Pater 2009 on constraint indexation). A higher-ranked IDENT constraint would protect *underlying* tense long mid vowels.

---

[21] The ranking of MAX-IO ≫ ONSET follows from transitivity relative to *CC (cf. (21) and (25)).

[22] There is some variation on this point, with some vowel-lengthening perfects and augmented forms attesting the contraction outputs [ē,ō].

Therefore *[ē,ō]$_{PERF}$ would only prohibit [ē,ō] from arising in the course of derivation, such as in perfect-tense vowel-lengthening.

When the vowel contraction facts are integrated with the evidence from consonant-initial reduplication, we derive the vowel-lengthening forms, subject to one adjustment to the ranking. In the preceding discussion, there was no way to disambiguate ONSET violations from REDUP(RED) violations. This is because properly-anchored copying always alleviated the ONSET violation that would be incurred by leaving the fixed [e] without an onset consonant. However, in the case of vowel-initial roots, properly-anchored copying itself induces a new ONSET violation, since the left-most copied element will be a vowel in word-initial position. Inspection of the ranking under these circumstances reveals that REDUP(RED) must in fact be ranked below ALIGN-/e/-L, while ONSET remains relatively highly-ranked. This is demonstrated in (31) below.

(31)   Vowel-lengthening perfect: $\sqrt{ag}$- → ἄγμαι [âg-mai] 'I have (been) led'

/RED, e$_1$, a$_2$g-/	ANCH-L-BR	ONSET	ALIGN-/e/-L	REDUP(R)	UNIF-IO
a. a̲.-e$_1$.-a$_2$g-		**!*	*		
b. ☞ _-ā$_{1,2}$g-		*		*	*
c. ag̲-ā$_{1,2}$g-		*	*!*		*
d. g̲-ā$_{1,2}$g-	*!		*		*

Due to the high-ranking of MAX-IO and MAX-μ-IO (omitted for reasons of space), candidates that delete a vowel, for example [g̲-e$_1$-g-], or have the root-initial vowel and the fixed /e/ coalesce as a short vowel, for example [_-a$_{1,2}$g-], are suboptimal. ONSET eliminates all candidates which display hiatus, here represented by candidate (31a). Since ANCHOR-L-BR ≫ ONSET, the word-initial ONSET violation cannot be avoided, as suboptimal candidate (31d) does by copying the root-second consonant. Only two candidates avoid hiatus and improper anchoring: the vowel-lengthening candidate (31b) [_-ā$_{1,2}$g-], and candidate (31c) [ag̲-ā$_{1,2}$g-], which is the potential output corresponding to the Attic Reduplication pattern.

Both candidates receive a single ONSET violation. Candidate (31b)'s violation is for the coalesced fixed /e/ + root /a/. Candidate (31c), on the other hand, repaired that particular ONSET violation by copying both the root-initial vowel and the root-second consonant, which serves as the onset for the coalesced vowel; however, the reduplicant-initial vowel now yields an ONSET violation of its own. The two candidates thus have equivalent violation profiles, but from different loci of violation. The choice comes down to the relative ranking of ALIGN-/e/-L and REDUP(RED). When the resolution of an ONSET violation is not at stake, the system prefers to leave the RED morpheme unrealized than to displace the /e/ from the left edge, selecting the vowel-lengthening candidate (31b). Nonetheless, the observation that the Attic Reduplication candidate survives this deep into the evaluation will serve as the starting point for an explanation of the Attic Reduplication pattern's survival in the language.

2.2.2.2 Attic Reduplication in Ancient Greek

The tableau in (31) above demonstrates the crucial rankings that select vowel-lengthening as the optimal strategy for forming perfects to vowel-initial roots in Ancient Greek. The choice between the optimal vowel-lengthening candidate (31b) and the suboptimal Attic Reduplication candidate (31c) comes down to the relative ranking of ALIGN-/e/-L and REDUP(RED): ALIGN-/e/-L ≫ REDUP(RED) selects vowel-lengthening; the reverse ranking — REDUP(RED) ≫ ALIGN-/e/-L — would have selected AR. That the ranking ALIGN-/e/-L ≫ REDUP(RED) does not generate

48

AR forms is further confirmed by the tableau in (32) below, where the AR root $\sqrt{ager}$ fails to be assigned an AR perfect.

(32)   AR forms in Greek: $\sqrt{ager\text{-}}$ → ἀγᾱ́γερμαι [ag-áger-mai] 'I have gathered together'

/RED, e_1, $a_2$ger-/	ANCH-L-BR	ONSET	ALIGN-/e/-L	REDUP(R)	UNIF-IO
a. <u>a</u>.-e_1.-$a_2$ger-		**!*	*		
b. 💣 _-$\bar{a}_{1,2}$ger-		*		*	*
c. ☺ <u>ag</u>-$\bar{a}_{1,2}$ger-		*	*!*		*
d. <u>g</u>-$\bar{a}_{1,2}$ger-	*!		*		*

While still superior to all other possibilities, the AR candidate (32c) loses to the vowel-lengthening candidate (32b), due to the ranking ALIGN-/e/-L ≫ REDUP(RED). The system thus prefers maximal left-edge alignment to overt realization of the reduplicative morpheme, in the general case. What is necessary to generate Attic Reduplication forms is a reversal of this preference, just in the case of roots which actually display AR. This can be accomplished using a lexically-indexed constraint (see Kraska-Szlenk 1997, 1999, Fukazawa 1999, Itô & Mester 1999, 2001, Pater 2000, 2009, Becker 2009) that favors overt realization of the reduplicative morpheme; that is, a lexically-indexed version of REDUP(RED) (cf. (8) above).

(33)   **REDUP(RED)**$_{lex}$
        If a root has the index *lex*, assign a violation mark * if there is a RED morpheme in the input, but the output does not contain substrings labeled as Base and Reduplicant.

When REDUP(RED)$_{lex}$ is ranked above ALIGN-/e/-L, and all and only the AR roots come with the lexical index *lex*, we derive the distinction between Attic Reduplication forms and vowel-lengthening forms within the synchronic grammar. In tableau (34), we generate copying for a root indexed with *lex*: $\sqrt{ager}_{lex}$ → *perfect* [ag-āger-]. In tableau (35), we generate just vowel-lengthening for a root not indexed with *lex*: $\sqrt{ag}$ → *perfect* [_-āg-].

(34)   REDUP(RED)$_{lex}$ **with** lexical indexation selects Attic Reduplication

/RED, e_1, $a_2^i g^k$er$_{\boxed{lex}}$-/	REDUP(RED)$_{lex}$	ONSET	ALIGN-/e/-L	REDUP(RED)
a. _-$\bar{a}_{1,2}$ger-	*!	*		*
b. ☞ $\underline{a^i g^k}$-$\bar{a}_{1,2}^i g^k$er-		*	**	

(35)   REDUP(RED)$_{lex}$ **without** lexical indexation selects vowel lengthening

/RED, e_1, $a_2^i g^k$-/	REDUP(RED)$_{lex}$	ONSET	ALIGN-/e/-L	REDUP(RED)
a. ☞ _-$\bar{a}_{1,2}$g-	not	*		*
b. $\underline{a^i g^k}$-$\bar{a}_{1,2}^i g^k$-	applicable	*	**!	

The crucial difference between these two derivations arises from the relationship between REDUP(RED)$_{lex}$ and ALIGN-/e/-L. When REDUP(RED)$_{lex}$ is not activated through the requisite indexation (as in (35)), there is nothing to differentiate the vowel-lengthening candidate (a) from the Attic Reduplication candidate (b) until ALIGN-/e/-L enters the evaluation. Since the AR candidate has extra copying, ALIGN-/e/-L selects vowel-lengthening. When REDUP(RED)$_{lex}$ is activated (as in (34)), ALIGN-/e/-L never gets to exert its force, because REDUP(RED)$_{lex}$ has already elim-

inated the vowel-lengthening candidate. This allows ONSET to adjudicate between the various copying candidates, ultimately selecting the Attic Reduplication output.

2.2.3 REDUP(RED)*lex*, Reduplicated Presents, and their Associated Perfects

Independent evidence for the activity of REDUP(RED)*lex* can be found elsewhere in the reduplicative system, namely, in the behavior of cluster-initial roots which are associated with reduplicated presents.

2.2.3.1 The Reduplicated Presents

In addition to having fully productive reduplication in the perfect, Ancient Greek also possesses a relatively small set of present stems which display reduplication (and also aorists; see Section 2.5 below). As illustrated in (36) below, these forms basically mirror the perfect, differing only in having a fixed [i] rather than a fixed [e] as the reduplicative vowel.

What is noteworthy, however, about the reduplicated presents relative to the perfect is the behavior of roots which begin in non–stop-sonorant clusters. As shown in (37), contrary to the productive pattern for the perfect, these roots display default C_1-copying rather than non-copying. Even more noteworthy are the perfect forms associated with these roots: these are perfects whose root allomorphs begin in non–stop-sonorant clusters, yet display default C_1-copying reduplication. That is to say, they contradict the productive pattern even though they are members of the productive category.[23]

(36)   Present reduplication (see Giannakis 1992)

Root		Present		Perfect	
√$d\bar{o}$-	'give'	δίδωμι	[<u>d</u>-í-d~~ō~~-mi]	δέδωκα	[<u>d</u>-é-d~~ō~~-k-a]
√$t^h\bar{e}$-	'place'	τίθημι	[<u>t</u>-í-t^h~~ē~~-mi]	τέθηκα	[<u>t</u>-é-t^h~~ē~~-k-a]
√p^hau-	'show'	πιφαύσκω	[<u>p</u>-i-p^haú-sk-~~ō~~]	not attested	
√$teuk^h$-	'prepare'	τιτύσκομαι	[<u>t</u>-i-tú-sko-mai]	τέτευχα	[<u>t</u>-é-teuk^h-a]
√$kl\bar{e}$-	'call'	κικλήσκω	[<u>k</u>-i-klê-sk-~~ō~~]	κέκληκα	[<u>k</u>-é-kl~~ē~~-k-a]

(37)   Present reduplication to non–stop-sonorant clusters

Root		Present		Perfect	
√pet-	'fall'	πίπτω	[<u>p</u>-í-pt-~~ō~~][24]	πέπτωκα	[<u>p</u>-é-pt~~ō~~-k-a]
√$st\bar{e}$-	'stand'	ἵστημι	[<u>h</u>-í-st~~ē~~-mi]	ἕστηκα	[<u>h</u>-é-st~~ē~~-k-a]
			($< *sist\bar{a}mi$)		($< *sest\bar{a}ka$)
√$mn\bar{e}$-	'remind'	μιμνήσκω	[<u>m</u>-i-mnê-sk-~~ō~~]	μέμνημαι	[<u>m</u>-é-mn~~ē~~-mai]

---

[23] While it does not follow from the formulation of *PCR used in this chapter, I will argue in Chapter 6 that repetitions are banned for nasal-nasal clusters; i.e., the sequence *mVmn* violates *PCR. This requires that *m-é-mnē-mai* result from the same mechanism which generates the other forms in (37).

[24] The ι of the reduplicant in this form is usually claimed to be long [ī], though the evidence is scarce (cf. Chantraine 1980:905–906). This length is usually taken to be analogical from ῥίπτω [r^hípt-~~ō~~] 'throw'.

What results is a striking gap: there are no C_1-copying perfects to roots beginning in *pt*, *st*, *mn*, and so on, that do not also have a reduplicated present in their individual verbal system.[25] (One further exception, κέκτημαι [k̲-é-ktɛ̄-mai], will be discussed below.[26]) This gap requires explanation. REDUP(RED)*lex* provides this explanation. If these roots are lexically indexed to **lex**, which is necessary to generate reduplication in the present, REDUP(RED)*lex* *predicts* that copying will also occur in the perfect, despite being dispreferred by the phonotactics.

2.2.3.2   Reduplicated Presents and Pre-Greek

First let us consider why the presents to these roots still retain the C_1-copying pattern. Since present reduplication is non-productive in Ancient Greek, it is possible to assume that the characteristics of the formation originate in an earlier period of the language, in a way similar to how Attic Reduplication will be accounted for in Section 2.3. The non-copying pattern rests upon the application of *PCR. However, we have no direct evidence that this constraint (or at least this version of the *PCR constraint) was active in Pre-Greek (see the discussion below, and also in Chapter 6). If *PCR was indeed not active in Pre-Greek, we predict across-the-board C_1-copying (except to laryngeal-initial roots; see Section 2.3). This generates Pre-Greek *pipt-* from (the zero-grade of) Pre-Greek root √*pet-*, as illustrated in (38). (The rankings are imported from those motivated for Pre-Greek in Section 2.3. All that is crucial is that *PCR is the lowest ranked among these constraints at the relevant stage.)

(38)   Copying to non–stop-sonorant roots in Pre-Greek: √*pet-* → *present* πίπτω [p̲-í-pt-ɔ̄]

/RED, i, pt, ɔ̄/	ONSET	DEP-IO	*CC	*PCR
a. ☞ p̲-i-pt-ɔ̄			*	*
b. p̲t-i-pt-ɔ̄			**!	
c. p̲ə-t-i-pt-ɔ̄		*!	*	
d. _-i-pt-ɔ̄	*!		*	

This solution entails that the non-copying pattern is an innovation induced by the change in sensitivity to the *PCR constraint. Prior to the higher ranking of *PCR, the non-copying perfects would have been normal C_1-copying perfects. This is supported by the existence of the perfects in (37). It is possibly also supported by the distribution of perfect forms built to the root √*kta*

[25] The forms associated with the root √*pet-* 'fall' appear to have contaminated forms of phonologically similar, and perhaps etymologically related, roots that have initial *p(e)t-*. Despite not having reduplicated presents of their own, the verbal systems associated with *petánnūmi* 'spread out', *pétomai* 'fly', and *ptḗssɔ̄* 'crouch' all attest perfects in *pept...*, alongside more expected perfects in *ept...* in the first two cases (see van de Laar 2000:246–248, 253, 259–260). There is also a form *pepterúgɔ̄mai*, from root *pterugízdɔ̄* 'flutter with wings', in a fragment of Sappho/Alcaeus.

   It appears as though the lexical idiosyncrasy which is proper to the root √*pet-* 'fall', due to the presence of reduplicated *píptɔ̄* in its verbal system, has come to be transferred to these other roots, such that they build C_1-copying perfects via REDUP(RED)*lex*. This state of affairs seems to have a comparandum among the Attic Reduplication forms. The root √*or-* 'incite' (< PIE √*h₃er-*) authentically builds an AR perfect stem *orɔ̄r-* (due to its laryngeal; see Section 2.3). The phonologically (nearly) identical root in Greek √*(h)or(a)-* 'see, watch' (< PIE √*(s)wer-*; see Chantraine 1980:813–815, Beekes & van Beek 2010:1095–1096) comes to have an AR perfect despite not having the shape to have acquired it in the normal way diachronically. It thus seems likely that the etymologically validated AR associated with √*or-* 'incite' has contaminated a similar root, just like √*pet-* 'fly' has done to other *pt* roots.

[26] There are perfect forms in *pepʰtʰan-* to the root √*pʰtʰan-* 'anticipate', but these are not attested until well after the Classical Period (Beekes & van Beek 2010:1568). In Classical and Pre-Classical Greek, this root shows the expected non-copying forms in *epʰtʰan-*.

'acquire'.[27] This root has two distinct stem-formation patterns in the perfect: the expected non-copying pattern *e-ktē̆-*, but also the unexpected C_1-copying pattern *k̲-e-ktē̆-*. The expected non-copying form has the expected perfect semantics 'have acquired'. The unexpected C_1-copying perfect, however, has unexpected behavior. First, it has present semantics, consistently meaning 'possess'. Second, it serves as the base of derivation for a future stem *kektē̄-s-* 'will possess' and other modal forms, which is not typical of perfect stems. These facts indicate that the stem *kektē̄-* became paradigmatically isolated at some point in its history, and was reinterpreted, either as an independent root or as a present stem. There is no reason why it should have, after becoming isolated, developed C_1-copying reduplication if it had previously shown non-copying. The only explanation is that the isolated stem *retained* C_1-copying (or at least the phonological string which it resulted in), and the paradigmatically regular stem changed according to the regular grammar to yield a non-copying stem. Therefore, *k̲-e-ktē̆-* must be an archaism, attesting to a pre-stage at which *kt* clusters copied C_1 just like *stop-sonorant* clusters, even in the perfect.

The reason why the reduplicated presents never get remodeled, as opposed to the perfects which do get remodeled (except when they are associated with a reduplicated present), must be due to differences in productivity between the two categories. For the perfect, reduplication is a productive marker of all forms, blocked on the surface in certain cases by phonotactics but always there "underlyingly", i.e. with a /RED/ morpheme in the underlying representation. In present tense stem-formation, reduplication is one of many derivational markers, and thus is never obligatory. That is to say, if faced with phonotactic problems in trying to produce a reduplicated present, speakers could opt to employ a different present-forming strategy (i.e., divert to a different morphological derivation). Present reduplication is completely unproductive by the time of Ancient Greek. This means that present reduplication, maybe even prior to the change in ranking of the *PCR constraint, must in some way be lexically restricted, possibly indexed to REDUP(RED)*lex*.

2.2.3.3 Generating the Unexpected C_1-Copying Presents and Perfects with REDUP(RED)*lex*

Before *PCR came to be active in the grammar, the reduplicated present forms could be productively generated as such once the proper morphemes were entered into the underlying representation, as was illustrated in (38) above. However, after *PCR becomes higher-ranked (above ONSET), the proper underlying form would fail to generate any copying. Armed with the mechanism of REDUP(RED)*lex*, which they independently had to deduce to account for the Attic Reduplication forms, speakers could avoid losing the reduplication here by assigning these roots to the lexical class of REDUP(RED)*lex*. If REDUP(RED)*lex* dominates the *PCR constraint active in Ancient Greek, then we can generate copying even to non–stop-sonorant cluster-initial roots. The tableau in (39) below illustrates how this generates *píptɔ̄* in the synchronic grammar of Ancient Greek.[28]

---

[27] I am indebted to Dieter Gunkel for bringing these distributional facts to my attention.

[28] Note that the ranking of ONSET relative to DEP-IO and *CC has changed between these two stages as well: it ranked above them in Pre-Greek, but ranks below them in Ancient Greek. This change is independently explainable: sound change eliminated various sorts of initial consonants, creating many new vowel-initial roots, which violate ONSET each time they surface as such. This will provide learners with significant evidence that ONSET is not highly ranked in their grammar.

(39)   Present reduplication in Ancient Greek: $\sqrt{pet}$- → present πίπτω [p̱-í-pt-ɔ̄]

/RED, i, pt͟[lex], ɔ̄/	REDUP(RED)lex	DEP-IO	*CC	*PCR	ONSET
a. ☞ p̱-i-pt-ɔ̄			*	*	
b. pt-i-pt-ɔ̄			**!		
c. pet-i-pt-ɔ̄		*!	*		
d. _-i-pt-ɔ̄	*!		*		*

What is most tantalizing about this solution is that it immediately provides an answer for the more surprising forms of this type, the perfect forms of these same roots that unexpectedly show C_1-copying as well. When these roots become indexed to REDUP(RED)*lex*, REDUP(RED)*lex* applies not only in the present, but also in the perfect, as demonstrated in (40). Therefore, the aberrant and idiosyncratic copying behavior of the present carries over to the perfect despite there being no category-internal reason for doing so.

(40)   REDUP(RED)*lex* in present-perfect pairs in Ancient Greek:

$\sqrt{p(e)t}$- → perfect πέπτωκα [p̱-é-ptɔ̄-k-a]; present πίπτω [p̱-í-pt-ɔ̄]

/RED, e, ptɔ̄͟[lex], k, a/	REDUP(RED)lex	*PCR	ONSET	REDUP(RED)
a. ☞ p̱-e-ptɔ̄-k-a		*		
b. _-e-ptɔ̄-k-a	*!		*	*

The crucial point here is the REDUP(RED)*lex* violation in the non-copying candidate (40b). This violation supersedes the *PCR violation of the C_1-copying candidate (40a). If the root were not indexed to REDUP(RED)*lex*, that *PCR violation would be fatal, as it is in the general case for roots with non–stop-sonorant clusters; exactly this is shown in (41) below (cf. (24) above). But due to REDUP(RED)*lex*, copying is required, and the C_1-copying candidate emerges. Thus, the grammar obeys the copying requirement at the expense of the phonotactics. This is the same sort of constraint interaction that led to the selection of the Attic Reduplication form in (34) and (35) above.[29]

(41)   Stop-sonorant roots without *lex* show non-copying perfects:

$\sqrt{kten}$- → ἔκτονα [é-kton-a] 'I have killed'

/RED, e, kton-/	REDUP(RED)lex	*PCR	ONSET	REDUP(RED)
a. ḵ-e-kton-	not	*!		
b. ☞ _-e-kton-	applicable		*	*

2.2.4   Interim Summary

This section has developed a grammar that generates the full distribution of stem-formation patterns in the Ancient Greek perfect tense, both productive and exceptional, both overtly reduplicative (as in

[29] There may be one more corner of the grammar that displays similar REDUP(RED)*lex* effects. Brent Vine (personal communication) has pointed out to me that there is a set of apparently reduplicated nouns built to *HeC roots which bear striking resemblance to Attic Reduplication verbal forms — in fact, they are built to many of the same roots which display Attic Reduplication in the perfect: for example, $\sqrt{ed}$ < *h₁ed 'eat' → ἐδωδή [edɔ̄dɛ́ː] 'food', $\sqrt{ag}$ < *h₂eǵ 'lead' → ἀγωγή [agɔ̄gɛ́ː] 'freight; movement; guidance' (see Vine 1998). As of yet, I cannot reconstruct the scenario by which these forms would have arisen, but the connection seems relevant.

the case of C_1-copying for consonant-initial roots and Attic Reduplication for vowel-initial roots), and non-reduplicative (as in the "non-copying" pattern for consonant-initial roots and the basic vowel-lengthening pattern for vowel-initial roots).

For consonant-initial roots, C_1-copying is the preferred pattern, applying to roots with an initial singleton consonant or an initial *stop-sonorant* cluster. This pattern is blocked for roots with other types of initial clusters by *PCR — a markedness constraint disfavoring consonant repetitions in certain environments, namely, in pre-obstruent position. It is also blocked for roots with initial geminates by constraints on consonant length. To avoid such violations, copying is eschewed altogether for these roots. The same strategy is ultimately employed for vowel-initial roots. Since, in such cases, it is impossible to completely alleviate ONSET violations without deletion or improper anchoring, the minimal reduplicant shape — i.e. null — is preferred, despite the violation of REDUP(RED).

In the special cases of Attic Reduplication and the unexpected C_1-copying perfects to cluster-initial roots with associated reduplicated presents, extra copying is enacted by a lexically-indexed version of REDUP(RED), which supersedes the constraints which normally induce non-copying: ALIGN-/e/-L in the case of vowel-initial roots, *PCR in the case of consonant-initial roots. Once that constraint is installed at the top of the ranking, the rest of the constraint ranking established for the basic cases is sufficient to generate the nature of the exceptional copying pattern.

The full ranking of the constraints posited in Section 2.2 is summarized in (42) below. Note that REDUP(RED)*lex* crucially must dominate both *PCR and ALIGN-/e/-L; the latter ranking also follows from transitivity through *PCR, as represented in (42).

(42)   **Total ranking for Ancient Greek reduplication**
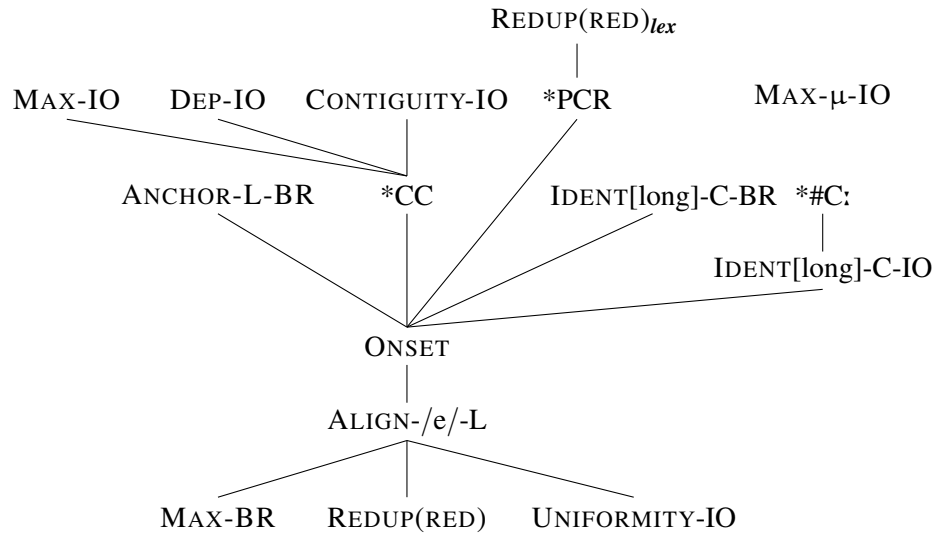
While this section has developed an analysis that can explain Attic Reduplication synchronically, we are still left with the question of why this divergent pattern should exist at all, and how could it persist in a grammar that generates a contradictory pattern? Section 2.3 will bring to bear insights from the diachrony of Greek and comparison with the other Indo-European languages to establish a phonologically-motivated origin for the pattern in a prior stage of the language. Section 2.4 will track the development from this prior stage into attested Ancient Greek, and show that the evidence at each stage is compatible with the retention of AR via constraint indexation.

2.3 Attic Reduplication

In investigating the productive reduplicative behavior of vowel-initial roots within the synchronic grammar of attested Ancient Greek, we saw that there is no obvious synchronic motivation for the presence of the Attic Reduplication pattern; it could only be explained by constraint indexation, which is, on its own, tantamount to pure stipulation. But given that it also has a very restricted distribution, the best explanation is that it is a retained archaism. This section shows that the origin of this archaic pattern can be generated directly in the phonology of an earlier stage of the language. It is thus to be viewed as retention of an original pattern in the face of a new productive pattern.

2.3.1 Attic Reduplication and the Laryngeals

Within the synchronic grammar of Ancient Greek, there are no obvious phonological properties which distinguish the vowel-initial roots that exhibit Attic Reduplication (AR) from the vowel-initial roots that exhibit vowel-lengthening. However, there is a clear distinction when we consider their etymologies. Virtually all of the roots which display AR can be reconstructed as having an initial laryngeal consonant in Proto-Indo-European (PIE) (see the reconstructions and evidence in, for example, Rix et al. 2001).[30] This connection between AR and the laryngeals has long been recognized in the Indo-European literature (Kuryłowicz 1927, Winter 1950:368–369, Beekes 1969:113–126, Suzuki 1994, Sihler 1995:489, Keydana 2006:90–91, 2012:107–108).[31] The table in (43) below shows example AR forms linked with their roots, both in Ancient Greek terms and in PIE terms, organized by which laryngeal each had in initial position in PIE. (See the Appendix — Section 2.7 — for the full list of AR perfects and their etymologies.)

(43)   Some Attic Reduplication perfects and likely etymologies

Root (Greek < *PIE)		Present Tense		Perfect Tense	
#*h_1					
$\sqrt{eleut^h} < *h_1lewd^h$	'go, come'	n/a		ἐλήλυθα	[elēlut^h-]
$\sqrt{en(e)k} < *h_1neḱ̯}$	'bring'	n/a		ἐνήνοχα	[enēnok^h-]
#*h_2					
$\sqrt{ager} < *h_2ger$	'gather together'	ἀγείρω	[agēr-]	ἀγᾱ́γερμαι	[agāger-]
$\sqrt{ar} < *h_2er$	'join, fit together'	ἀραρίσκω	[arar-]	ἄρᾱρα	[arār-]
#*h_3					
$\sqrt{od} < *h_3ed$	'smell'	ὄζω	[ozd-]	ὄδωδα	[odōd-]
$\sqrt{or} < *h_3er$	'incite'	ὄρνυμι	[or-]	ὄρωρα	[orōr-]

---

[30] Only two of at least 20 AR roots, are definitively not laryngeal-initial, and both are structurally similar (or, in the case of $\sqrt{or}$- 'keep watch', identical) to roots which are historically laryngeal-initial.

[31] Cowgill (1965:153) takes the opposing view: "It seems also that the Attic Reduplication in Greek perfects must have started from roots which had a prothetic vowel of *non-laryngeal origin*" (my emphasis).

The laryngeals are a set of consonants reconstructed for PIE based on internal and comparative evidence (de Saussure 1879).[32] They are partially attested in the Anatolian languages, but have been lost in all other Indo-European branches. Their exact phonetic characteristics are unknown, but they are generally identified as having been fricatives with constriction in the rear of the vocal tract. The most commonly recognized reconstructed phonemic inventory of PIE contains three laryngeals (which will be represented in this dissertation as h_1, h_2, and h_3, respectively, H collectively), based in large part on the "triple reflex" in Greek. As represented in (44), in each of several environments where we can reconstruct a laryngeal, each of the three different (non-high) vowel qualities are found in Greek.

(44)   Laryngeal outcomes in Greek

	Laryngeal Contexts		
	Vocalization	Coloration	Coloration and Lengthening
	*H / {C,#}_C	*He / #_	*eH / _{C,#}
*h_1	e	e	$\bar{\varepsilon}$
*h_2	a	a	$\bar{a} \sim \bar{\varepsilon}$[33]
*h_3	o	o	$\bar{\mathfrak{o}}$

We find exactly this triple reflex in the Attic Reduplication forms. Of the approximately 20 vowel-initial roots which have AR perfects, none have an initial high-vowel; all begin in [e,a,o], the outcomes of laryngeals in word-initial position ("vocalization"/"coloration"). The long vowels of the second syllables of the AR forms are limited to [ɛ̄,ā,ɔ̄], the outcomes of tautosyllabic *-eH-sequences ("coloration and lengthening"). The vowels associated with the AR pattern are thus exactly those vowels associated with laryngeal reflexes. When coupled with comparative etymo-logical evidence for initial laryngeals in these roots, it is safe to assert a connection between Attic Reduplication and laryngeals.

Prior to Proto-Greek (the stage reconstructible based on comparison of the Greek dialects), the laryngeals were lost, leaving only indirect effects such as those listed above. In order to bring the laryngeals to bear on Attic Reduplication, therefore, the origin of the pattern must be local-ized in a stage of Greek prior to their loss. Since evidence of this stage comes from internal reconstruction of Common Greek or Proto-Greek, this stage will be identified as "Pre-Greek." I proceed under the conservative assumption that, in the absence of evidence to the contrary, the reduplicative grammar of Pre-Greek is minimally different from the directly observable grammar of Ancient Greek. (I do assume one difference: recall from Section 2.2.3 above that there is evidence that non-laryngeal cluster-initial roots in Pre-Greek did not display the non-copying pattern which is productive in Ancient Greek, but rather showed C_1-copying. This amounts to a lower ranking of *PCR, and different conditions for the ranking of ONSET and REDUP(RED).)

---

[32] For an introduction to laryngeal theory, consult, for example, Mayrhofer (1986:121–145), Fortson (2010:62–64, 81–83), Byrd (forthcoming). For further discussion of laryngeal behavior in Greek, see Cowgill (1965), Beekes (1969), Rix (1992:68–76), among others.

[33] Regarding the "coloration and lengthening" outcome of h_2, in Attic-Ionic, there is a sound change that changes /ā/ to /ɛ̄/. [ā] is attested in other dialects. See footnote 3.

2.3.2 Previous Approaches

With the connection between laryngeals and Attic Reduplication established, the null hypothesis would be that the Pre-Greek source of the AR pattern was generated by running the laryngeal-initial roots through the basic reduplicative grammar, as we have it still in Ancient Greek. Since the laryngeals were consonantal segments, the default C_1-copying reduplication pattern for consonantal roots would yield a preform of the shape H^i-e-$H^iC^k(VC)$-. For a root $\sqrt{}$*h_2ger 'gather together', this would predict the derivation in (45), which would yield Ancient Greek **$\bar{a}germai$. The actual attested form, which does display the AR pattern, is ἀγᾰ́γερμαι [agā́germai] (Attic-Ionic ἀγήγερμαι [agḗgermai]). This form is clearly incompatible with such a derivation.

(45)   If laryngeal roots reduplicated normally...
   a.   Pre-Greek input-output mapping: $\sqrt{}$*h_2ger → perfect *$\underline{h_2}$-e-h_2ger-mai
   b.   Diachrony: Pre-Greek *$h_2eh_2germai$ > Common Greek **$\bar{a}germai$

To fix this problem, most accounts have asserted that roots with initial *HC clusters exceptionally copied both elements to create a reduplicant of the shape *HCV- (Winter 1950:368–369, Beekes 1969:113–126, Rix 1992:204–205, Keydana 2006:90–91, 2012:107–108).[34] Once the forms are fixed in such a way, they would evolve correctly into Greek:

(46)   Copying root-initial HC
   a.   Pre-Greek input-output mapping: $\sqrt{}$*h_2ger → perfect *$\underline{h_2g}$-e-h_2ger-mai
   b.   Diachrony: Pre-Greek *$h_2geh_2germai$ (> *$\mathring{h}_2geh_2germai$) > Com. Greek $ag\bar{a}germai$

However, these accounts rarely consider what the motivation for such exceptional behavior (i.e., copying as $\underline{C_1C_2V}$-C_1C_2V- rather than $\underline{C_1V}$-C_1C_2V-), might have been, and simply pronounce it by stipulation. While some have tried to connect this cluster-copying for laryngeals to the behavior of *s-stop*–initial roots (Keydana 2012), it is demonstrably the case that such roots did indeed follow the normal C_1-copying pattern, at least among the reduplicated presents. As was pointed out already by Brugmann & Delbrück (1897–1916:40–41; via Byrd 2010:103–104), the exact correspondence between the archaic reduplicated present forms of the PIE root $\sqrt{}$*$steh_2$ 'stand' in Ancient Greek ἵστημι [hí-stɛ̄-mi] (< Proto-Greek *$\underline{si}$-$st\bar{a}$-mi) and Latin $sist\bar{o}$ ([s̲i̲-st-ō]), neither of which conform to the languages' productive patterns for reduplication, requires that we reconstruct this pattern for Proto-Indo-European, and thus Pre-Greek, as well.[35] Under the assumption that reduplication operated in the same way in both present and perfect at the periods in which both were productive, and thus that evidence from the present bears on the behavior of the perfect, we can infer that *STVX– roots copied C_1 (i.e. s) in Pre-Greek. This leaves *HC-initial roots as the only type not to follow the *CV* reduplication pattern.

But it is not necessary to stipulate that this one particular root shape should have copied in an exceptional way. Appealing to the process of "laryngeal vocalization," and considering the underlying motivation behind it, provides a recourse for deriving the divergent pattern directly through constraint interaction. Once markedness constraints targeting laryngeals are integrated

---

[34] Suzuki (1994) also asserts an exceptional copying pattern for laryngeal-initial roots, based on a rule of "laryngeal resyllabification." Under his account, *HC clusters employed single-consonant copy, but of C_2 rather than C_1, equivalent to what we find in Sanskrit $j\bar{a}g\acute{a}ra$ < *$\underline{g^w}e$-h_1g^wor-e, and also synchronically in Sanskrit STVX– roots. This generates a preform in *C_2V-HC_2VC-, but still requires "analogical" reintroduction of the initial vowel, perhaps through a sort of Base-Derivative faithfulness (though Suzuki does not use exactly those terms).

[35] See Chapter 7 for further discussion of these forms and their relevance for the reconstruction of the reduplication patterns of Proto-Indo-European.

into the reduplicative grammar, the precursor of Attic Reduplication ("Pre-AR") will emerge as the optimal resolution. This resolution yields a pre-form similar to that of the cluster-copying approaches, but with a phonological motivation for the exceptional behavior of laryngeal-initial roots. The proposed distribution of reduplicant shapes in Pre-Greek is shown in (47), slightly modified from (3) above.

(47)   Default reduplication vs. Pre-AR in Pre-Greek
    a.   Default reduplication pre-forms: $*\underline{C^i}$-e-C^iC^kVC- or $*\underline{C^i}$-e-C^iC^k-
    b.   Attic Reduplication pre-forms: $*\underline{H^i\vartheta C^k}$-e-$H^iC^k$VC- or $*\underline{H^i\vartheta C^k}$-e-$H^iC^k$-

2.3.3   Vowel Prothesis and Laryngeal Vocalization in Greek

In Ancient Greek, as well as in Armenian and Phrygian, reconstructed Proto-Indo-European word-initial *laryngeal + consonant* (*HC*) sequences ultimately surface as the sequence *VC* (see, e.g., Cowgill 1965, Clackson 1994). In Greek, the quality of the vowel corresponds to the quality of the laryngeal (cf. (44)); for example, Greek ἀνήρ [anḗr] 'man' < PIE *$h_2n\bar{e}r$ (cf. Skt *nar-*). This sound change is traditionally referred to as "vowel prothesis," and can be described with the following diachronic correspondence:

(48)   Vowel prothesis:
    PIE *#$\boxed{\text{H}}$CV > Ancient Greek #$\boxed{\text{V}}$CV (*H > V / #_CV)

Vowel prothesis, however, is really just a special case of the more general process of "laryngeal vocalization", whereby a reconstructed PIE laryngeal consonant displays a vocalic reflex in the daughter language. Laryngeal vocalization in Greek occurred when a laryngeal consonant would have occurred word-medially between consonants in a *-VCHCV-* sequence, as exemplified by the forms in (49). In terms of diachronic correspondence, the development can be stated as in (50).

(49)   Examples of laryngeal vocalization in Greek
    a.   PIE *$h_2en\boxed{h_1}$-*mos* > Ancient Greek ἄνεμος   [án$\boxed{\text{e}}$mos] 'breath'   (Rix 1992:71)
    b.   PIE *$\acute{g}en\boxed{h_1}$-*tōr*   > Ancient Greek γενέτωρ [gen$\boxed{\text{é}}$tōr] 'begetter' (Sihler 1995:99)

(50)   Laryngeal vocalization:
    PIE *C$\boxed{\text{H}}$C > Ancient Greek C$\boxed{\text{V}}$C   (*H > V / C_C)

The only difference in conditioning environment between vowel prothesis and traditional laryngeal vocalization is the preceding context: word-boundary in the first case and consonant in the second. The two contexts can be unified by the fact that the laryngeal is *not adjacent to a vowel* in either case.

Requiring adjacency to a vowel would be a means of ensuring that the laryngeal consonant has transitional cues. Given that the laryngeals were on their way towards complete loss (likely by way of a gradual lenition process), it is likely that they were relatively difficult to perceive at this stage. Maximizing what phonetic cues they had (by requiring an adjacent vowel) would have improved the laryngeal's perceptibility, both in terms of perceiving its presence and in terms of perceiving its contrastive place. The constraint demanding that laryngeals be adjacent to vowels, which was active in the grammar of Pre-Greek, is defined in (51).

(51)   **H//V**
    Assign one violation * for each laryngeal which is not adjacent to a vowel.

This constraint describes the conditioning environment for laryngeal vocalization, but not the change itself. I will be following the view in which laryngeal vocalization is not seen as direct vocalization of the consonantal segment, but rather as epenthesis of a vowel adjacent to the laryngeal (see Mayrhofer 1986:138, Byrd 2010, 2011). The alternative view involving direct laryngeal vocalization is not compatible with the analysis developed in this section, as it cannot make use of H//V, and requires an optimal output at the Pre-Greek stage which violates ONSET (a constraint which the analysis otherwise requires to be highly ranked). This means that our previous examples have the historical derivations in (52).[36]

(52)   Derivations of laryngeal vocalization in Greek
  a. Pre-Greek IO mapping: *$/h_2$nēr/ $\rightarrow$ *h_2ə$n\bar{e}r$
    Diachrony: Pre-Greek *h_2ə$n\bar{e}r$ > Ancient Greek ἀνήρ [anḗr]
  b. Pre-Greek IO mapping: *$/h_2$enh_1-mos/ $\rightarrow$ *$h_2$enh_1əmos/*$h_2$enə$h_1$mos
    Diachrony: Pre-Greek *$h_2$enh_1əmos/*$h_2$enə$h_1$mos > Ancient Greek ἄνεμος [ánemos]
  c. Pre-Greek IO mapping: *$/ǵen h_1$-tōr/ $\rightarrow$ *$ǵen h_1$ə$t\bar{o}r$/*$ǵen$ə$h_1 t\bar{o}r$
    Diachrony: Pre-Greek *$ǵen h_1$ə$t\bar{o}r$/*$ǵen$ə$h_1 t\bar{o}r$ > Ancient Greek γενέτωρ [genétōr]

The synchronic mappings in Pre-Greek are generated by the ranking in (53) below. The ranking ONSET ≫ CONTIGUITY-IO is responsible for cluster-internal epenthesis in word-initial position.

(53)   **Ranking (Pre-Greek)**

   MAX-IO H//V ONSET

     DEP-IO CONTIG-IO
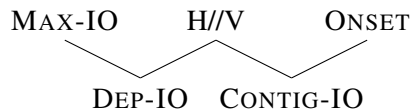
Tableau (54) below illustrates how this ranking selects the cluster-internal epenthesis candidate. In this and all subsequent tableaux in this section, the candidates (in the leftmost column) are followed by the form that such a candidate would evolve into in Common Greek. If a candidate would yield the attested outcome, it is accompanied by a " ✓ "; if it would yield an unattested outcome, it is marked by " ** ". (" >> " means "becomes, possibly via multiple sound changes.") I make the following two assumptions about the diachrony of laryngeals: (i) the synchronic process of laryngeal vocalization does not involve deletion of the laryngeal consonant; and (ii) the sound change that eliminates laryngeals occurs after laryngeal vocalization has already run its course, leaving behind the epenthetic vowel as part of the (underlying) phonological representation.

(54)   Laryngeal vocalization: PIE $\sqrt{}$*$h_2$ger- 'gather together' > Ancient Greek ἀγερ- [ager-]

/$h_2$ger-/		H//V	MAX-IO	DEP-IO	ONSET	CONTIG-IO
a.	$h_2$ger- >> **ger-	*!		*		
b.	ger >> **ger-		*!			
c.	$h_2$er >> **ar-		*!			*
d. ☞	h_2əger- >> ✓ager-			*		*
e.	ə$h_2$ger- >> ✓ager-			*	*!	

<hr/>

[36] I will not attempt to adjudicate the position of the epenthetic vowel relative to the laryngeal for all cases, as it is likely to vary depending on the specific phonotactics and morphological composition of any given string. However, consistency with the proposed analysis of Pre-AR (Section 2.3.4) requires cluster-internal (as opposed to cluster-preceding) epenthesis in cases of word-initial *HC clusters.

2.3.4 Generating (Pre-)Attic Reduplication in Pre-Greek

When the rankings just motivated for laryngeal vocalization are integrated with the grammar previously developed for reduplication in Ancient Greek, the grammar selects an output that will evolve into the Attic Reduplication pattern. The "Pre-AR" output that the Pre-Greek grammar will ultimately produce is [h₂əg-e-h₂ger-] (schematically [$\underline{H^i$əC^k}-e-H^{i}C^k(VC)-]). This form copies both members of the root-initial cluster, with an epenthetic vowel inserted between the copied segments in the reduplicant. This divergence from the normal C$_1$-copying pattern emerges as a repair for two high-ranking laryngeal-related markedness constraints in the system: H//V and a version of *PCR that specifically targets laryngeal repetitions.

2.3.4.1 Motivating the Pattern

Prior to the initiation of the Pre-AR pattern (i.e., in Proto-Indo-European and the subsequent period of independent development of Greek prior to the change about to be described), all root-shapes displayed C$_1$-copying, regardless of cluster type (see the discussion in Section 2.3.2, and evidence from archaisms in Ancient Greek in Section 2.2.3). This means that laryngeal-initial roots would have had a *CV* reduplicant (i.e. *HV*). This is likely reflected in the Vedic Sanskrit perfect stem [ānáś-] < PIE *$\underline{h_{1/2}e}$-$h_{1/2}nok$-,[37] which is directly cognate with the Old Irish preterite -*ánaic* [-ānək^j]. Since neither of these forms are synchronically regular, they are evidence for *HV*- reduplication to *HC* roots in PIE. Applying this pattern to our example root √*h_2ger*, these grammars select a candidate [$\underline{h_2}$-e-h₂ger-], which copies just C$_1$. Such a form would have evolved into Common Greek **$\bar{a}ger$-, which is clearly not the Attic Reduplication pattern.

What, then, changes such that [$\underline{h_2}$-e-h₂ger-] is no longer an acceptable output? This is another effect of *PCR (cf. (23) above; see Chapter 6 for further elaboration, especially regarding what types of repetitions this sort of constraint can target), specifically a version of the constraint which bans (only) laryngeal repetitions:

(55)   ***PCR-H (= *H$_\alpha$VH$_\alpha$ / _C)**
       Assign a violation * to any sequence of identical laryngeals separated by a vowel (H$_\alpha$VH$_\alpha$) which immediately precedes a consonant.

The presence and activity in the grammar of exactly this version of *PCR likely correlates with the factors that led to laryngeal vocalization. Given that laryngeals required epenthesis of an adjacent prop vowel to license their presence (via newly high ranked H//V), likely as a means of maximizing their phonetic cues, it is reasonable that they would be specially targeted in the repetition context, as well, if indeed repetition avoidance is sensitive to phonetic cues and perceptibility. Conversely, the lack of epenthesis of this sort for other consonants correlates with tolerance of their repetition.

When *PCR-H comes to be active in the grammar (i.e., is promoted to a position in the ranking high enough to induce repairs), the default C$_1$-copying pattern is prevented from surfacing, since the previously optimal C$_1$-copying form [$\underline{h_2}$-e-h₂ger-] violates this constraint. The new form which will ultimately be chosen as optimal, [h₂əg-e-h₂ger-] (> Common Greek [agāger-]), satisfies *PCR-H, as illustrated in (56) below.

---

[37] For the etymology with *h_1*, see Cowgill (1965:151); for *h_2*, see Kümmel (2000:289).

(56)   Ruling out C$_1$-copying reduplication

/RED, e, h$_2$ger-/				*PCR-H
a.		h$_2$-e-h$_2$ger-	> **āger-	*!
b.	☞	h$_2$əg-e-h$_2$ger-	> ✓agāger-	

With *PCR-H blocking the selection of the C$_1$-copying candidate, an alternative copying pattern must take over. The characteristics of this alternative pattern, i.e. (56b) — which, note, is distinct from the alternative pattern triggered by *PCR in Ancient Greek (namely, non-copying) — are determined by the relative ranking of the remaining constraints, which has in large part already been determined.

2.3.4.2 The Alternative Pattern

There are a number of ways in which the *PCR-H problem could potentially be avoided. Viable repairs are provided in the table in (57) below. The Pre-AR form is the *cluster-copy + reduplicant-internal epenthesis* candidate (a) [h$_2$əg-e-h$_2$ger-], which violates DEP-IO and CONTIGUITY-BR (the Base-Reduplicant version of CONTIGUITY-IO, as defined in (19) above); it also has additional violations of ALIGN-/e/-L relative to the default pattern.[38]

(57)   Potential repairs and their associated constraints

	Repair	Candidate output	Constraint(s) violated
a.	Red-internal epenthesis:	[h$_2$əg-e-h$_2$ger-]	DEP-IO & CONTIG-BR
b.	Root-internal epenthesis:	[h$_2$-e-h$_2$əger-]	DEP-IO & CONTIG-IO
c.	Infixation with copying:	[h$_2$e-h$_2$-e-ger-]	CONTIG-IO
d.	Infixation without copying:	[_-h$_2$-e-ger-]	CONTIG-IO & REDUP(RED)
e.	Unfilled onset:	[_-e-h$_2$ger-]	ONSET & REDUP(RED)
f.	Cluster-copying:	[h$_2$g-e-h$_2$ger-]	H//V
g.	Deletion of root-C$_1$:	[g-e-ger-]	MAX-IO
h.	Deletion of root-C$_2$:	[h$_2$-e-h$_2$er-]	MAX-IO & CONTIG-IO
i.	Improper anchoring:	[g-e-h$_2$ger-]	ANCHOR-L-BR

These repairs coincide with operations modulated by the constraints introduced previously in order to account for the basic reduplication pattern and for laryngeal vocalization, respectively. These rankings are repeated below in (58). Note that I import the Ancient Greek ranking that was constructed based solely on the analysis of the basic pattern for stop-sonorant–initial roots (codified in (21) in Section 2.2.1.4), not the one further developed from the evidence of non-copying, vowel-lengthening, and the exceptional patterns (i.e. not the one shown in (42)). This is because, according to the analysis being proposed here, these patterns were not present in the language at the stage in which Pre-AR developed. Therefore, ranking arguments based on that evidence should not be attributed to this stage. Note especially that this means that the ranking of REDUP(RED) is under-determined; it was only because of evidence from the vowel-lengthening pattern that we could definitively rank it below ALIGN-/e/-L. This multiplicity of possible rankings for REDUP(RED)

---

[38] Notice that this pattern is equivalent to the avoidance strategy for Gothic STVX– roots, namely simple cluster-copying (see Chapter 4), but with reduplicant-internal epenthesis in addition.

will be important in understanding the development of the system after the inception of the Pre-AR pattern, as will be discussed in Section 2.4 below.

(58)   a.   **Ranking for C$_1$-copying reduplication** (Section 2.2.1.4, ex. (21))

MAX-IO   DEP-IO   CONTIGUITY-IO   ANCHOR-L-BR

*CC   ONSET   REDUP(RED)

ALIGN-/e/-L

MAX-BR
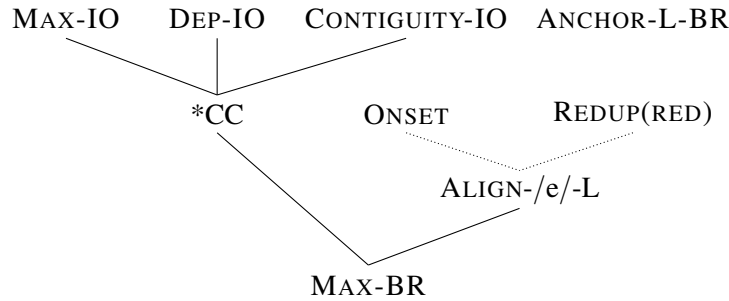
   b.   **Ranking for laryngeal vocalization** (Section 2.3.3, ex. (53))

MAX-IO   H//V   ONSET

DEP-IO   CONTIG-IO
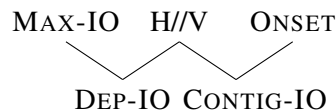
When we compare these rankings, we find that there are no ranking contradictions. The two rankings can therefore be reconciled without changing the results of either process independently. The result of integrating the two rankings without asserting any additional rankings which do not follow from transitivity — other than the addition of undominated *PCR-H — is shown in the Hasse diagram in (59) below. Per the preceding discussion, REDUP(RED) is now omitted, as there is neither evidence nor need for its ranking at this stage.

(59)   **Integrated ranking**

MAX-IO   H//V   ONSET   ANCHOR-L-BR   *PCR-H

DEP-IO   CONTIG-IO

*CC   ALIGN-/e/-L

MAX-BR
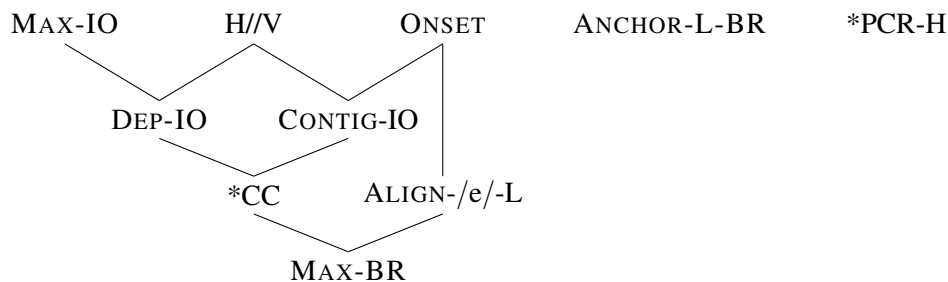
The critical rankings contained in (59) successfully eliminate a majority of the candidates listed in (57) (assuming the undominated constraints ANCHOR-L-BR and *PCR-H cannot become dominated by otherwise dominated constraints). Among the candidates not eliminated is our presumed Pre-AR candidate in (57a). However, a unique winner cannot be determined with these critical rankings alone. To select the Pre-AR candidate, all that is necessary is to fix CONTIGUITY-IO above DEP-IO, ALIGN-/e/-L, and CONTIGUITY-BR. These additional rankings are shown in the Hasse diagram in (60). The tableau in (61) demonstrates that these rankings properly select the desired Pre-AR candidate from among the remaining candidates.

(60) **Reconciled total ranking for Pre-AR**

ANCHOR-L-BR *PCR-H H//V   ONSET

MAX-IO   CONTIG-IO

DEP-IO

*CC   ALIGN-/e/-L CONTIG-BR

MAX-BR
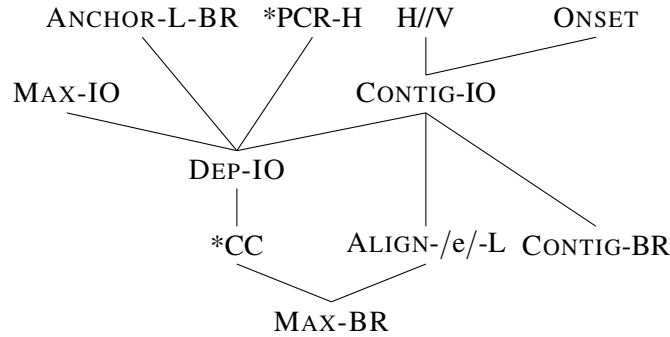
(61) The alternative repair: cluster-copying + reduplicant internal epenthesis

/RED, e, h$_2$ger-/		CNTG-IO	DEP-IO	ALIGN-/e/-L	CNTG-BR
a. ☞ h$_2$əg-e-h$_2$ger- ¦ > ✓agāger-			*	***	*
b. h$_2$-e-h$_2$əger- ¦ > **āger-		*!	*	*	
c. h$_2$e-h$_2$-e-ger- ¦ > **āger-		*!		**	
d. _-h$_2$-e-ger- ¦ > **ager-		*!		*	

Finally, the generalizations captured by the ranking in (60) are summarized in (62). A full summary tableau of the candidates from (57) is shown in (63) below.

(62) Generalizations and ranking arguments

a. **Laryngeals must be adjacent to a vowel**:
H//V is undominated

b. **H//V violations are repaired by epenthesis**:
H//V, MAX-IO ≫ DEP-IO

c. **Consonant-initial roots reduplicate with C$_1$-copying**:
ANCHOR-L-BR is active; ONSET ≫ ALIGN-/e/-L ≫ MAX-BR

d. **This is interrupted for H-initial roots due to a dispreference for H repetitions**:
PCR-H is active

e. ***PCR-H violations are avoided by epenthesis + extra copying**:
*PCR-H, MAX-IO, ANCHOR-L-BR, ONSET ≫ DEP-IO, ALIGN-/e/-L

f. **Reduplicant-internal epenthesis is preferred to root-internal epenthesis**:
CONTIG-IO ≫ CONTIG-BR, ALIGN-/e/-L

(63) Generating Pre-AR

/RED, e, h₂ger-/			*PCR-H	ANCHOR-L-BR	MAX-IO	H//V	ONSET	CONTIG-IO	DEP-IO	ALIGN-/e/-L	CONTIG-BR
a. ☞	h_2əg-e-$h_2$ger-	> √agāger-							*	***	*
b.	$\underline{h_2}$-e-h_2əger-	> **āger-						*!	*	*	
c.	$\underline{h_2}$e-h_2-e-ger-	> **āger-						*!		**	
d.	_-h_2-e-ger-	> **ager-						*!		*	
e.	_-e-$h_2$ger-	> **āger-					*!				
f.	$\underline{h_2}$g-e-$h_2$ger-	> **gāger-				*!				**	
g.	$\underline{g}$-e-ger-	> **geger-			*!					*	
h.	$\underline{h_2}$-e-$h_2$er-	> **ār-			*!			*		*	
i.	$\underline{g}$-e-$h_2$ger-	> **gāger-		*!						*	
j.	$\underline{h_2}$-e-$h_2$ger-	> **āger-	*!							*	

2.3.5 Attic Reduplication for *HeC roots

While the solution proposed above derives the Pre-AR pattern for roots of the shape *HCeC without problem, there is a complication that arises for *HeC roots with respect to the operation of ablaut.[39] The expected ablaut grade for the perfect active singular is the "*o-*grade." Therefore, for an *HeC root like √*h₁ed 'eat' (> Ancient Greek √ed 'eat'), the root allomorph which should be entered into the derivation (for the perfect active singular) is /h₁od/. Since the normal pattern for reduplication is C₁-copying, the default candidate for this allomorph would be [$\underline{h_1}$-e-h₁od-]. In this output, the laryngeal is intervocalic, and thus not in violation of *PCR-H or H//V. Therefore, there would be nothing to rule out this candidate, and it should be chosen as the winner. A Pre-Greek form *$\underline{h_1}$-e-h₁od- would yield Ancient Greek **ɔ̄d-, which is not the attested perfect stem for this root; instead we have a perfect stem edɛ̄d- which shows Attic Reduplication. While the [ɔ̄C-] outcome is not attested for this particular root, it is likely seen in the lexicalized perfect stem ἄνωγα [án-**ɔ̄g**-a] 'I command', which is typically identified as belonging to the PIE root √*h₂eǵ 'say' (Rix et al. 2001:256; the *an* sequence is the preverb *an(a)-* 'up, on, upon'): Pre-Greek *an(a)-$\underline{h_2}$-e-h₂og- > Ancient Greek *anɔ̄g-*. Thus, the system developed to account for reduplication in Pre-Greek does generate attested outcomes for *o*-grade perfects of Pre-Greek *HeC roots, just not AR outcomes. This means that the Pre-AR pattern in *HeC roots cannot come from the *o*-grade.

In order to generate Pre-AR to the *HeC roots that display it, we must instead start with a formation which takes "zero-grade" ablaut. For √*h₁ed, the best and oldest attested perfect form is the participle ἐδηδώς [edɛ̄d-ɔ́s]. Since the participle is indeed a zero-grade formation, the input would be: /RED, e, h₁d, wɔ́s/. Plugging in the default C₁-copying candidate, we do encounter our *PCR-H violation: [$\underline{\textbf{h}_1}$-**e-h₁d**-wɔ́s]. This leads us down the same road as with the *HCeC roots, ultimately choosing the candidate [h₁əd-e-h₁d-wɔ́s], which directly yields the attested AR form ἐδηδώς [edɛ̄d-ɔ́s]. Therefore, while AR should not arise on *o*-grade (or indeed *e*-grade) formations for *HeC roots, it should arise in zero-grade formations, which include all categories in the perfect other than the active singular.

This predicts that, for a time, *HeC roots would have had normal C₁-copying reduplication in *e/o*-grade categories, as is reflected in *anɔ̄g-* (< *an(a)-$\underline{h_2}$-e-h₂og-), but Pre-AR reduplication in

---

[39] For an introduction to Indo-European ablaut, see, for example, Mayrhofer (1986), Fortson (2010).

zero-grade categories, as is reflected in *edɛ̄d-* (< *$h_1ǝd$-e-h_1d-). As ablaut distinctions collapsed, and as the transparency of the relationship between the two reduplicative allomorphs was eroded by the loss of the laryngeals, speakers could have easily generalized one or the other of the stem forms throughout the perfect paradigm. Nevertheless, the co-existence of the two types at the very earliest stage after the loss of the laryngeals will lay the groundwork for the ultimate fate of Attic Reduplication.

2.3.6   Interim Conclusions

In this section, I showed how the phonological properties of the laryngeals, likely deriving from the weakness of their phonetic cues, had significant effects on Pre-Greek. In the general case, laryngeals required epenthesis of a prop vowel when not otherwise vowel-adjacent (i.e. laryngeal vocalization, driven by H//V). In reduplication, the desire to avoid the local repetition of laryngeals in pre-consonantal position (*PCR-H) made it impossible for laryngeal-initial roots to reduplicate according to the default C_1-copying pattern of the language. This led to the precursor of the Attic Reduplication pattern of attested Ancient Greek. The constraint ranking needed to generate this pattern, which ultimately selects cluster-copying and reduplicant-internal epenthesis as the optimal alternative reduplication pattern, is in large part independently motivated by the default C_1-copying reduplication pattern and laryngeal vocalization. The independent activity of these two parts of the grammar may, in a certain sense, have predestined this particular resolution of the laryngeal markedness problem. The Pre-AR pattern, generated productively and transparently in Pre-Greek, is maintained in Ancient Greek as Attic Reduplication. In the following section, we will consider how this pattern came to persist into Ancient Greek despite the loss of its original conditioning factors.

2.4   The Diachrony of Attic Reduplication

The synchronic analysis of (Pre-)Attic Reduplication presented above hinges crucially on the presence of laryngeals in the phonemic/phonetic inventory. However, the AR forms clearly survive beyond the period at which laryngeals are lost from the inventory, or else we would have no trace of the pattern. If the perfects built to these roots had been generated according to the productive pattern throughout their history, they would ultimately have fallen together with the outcomes of other vowel-initial roots[40] — and thus come to display vowel-lengthening (see again Section 2.2.2.1) — rather than retaining AR. The retention of AR thus requires special explanation.

The system developed in Section 2.2.2.2 — whereby Attic Reduplication forms are derived in the synchronic grammar of Ancient Greek via a higher-ranked lexically-indexed version of REDUP(RED), as shown in the tableaux in (64) and (65) below (repeated from (34) and (35)) — is sufficient to capture the synchronic distribution of stem-formation patterns for vowel-initial roots in Ancient Greek. However, it raises an obvious question: how did Greek speakers arrive at this system?

---

[40] Note that there were no (or extremely few) vowel-initial roots in Proto-Indo-European (cf. Rix et al. 2001). They come about in Greek (and elsewhere) due primarily to loss of certain consonants in initial position, namely the laryngeals, glides, and *s*. Chronologically, the laryngeals are lost first. Therefore, there is no pre-existing pattern for vowel-initial roots in the perfect when they first come into being due to the loss of the laryngeals. See more below on the development of vowel-initial roots and vowel-initial perfect forms.

(64) REDUP(RED)$_{lex}$ **with** lexical indexation selects Attic Reduplication (cf. (34))

/RED, e$_1$, a$_2^i$g^{i}er$_{\boxed{lex}}$-/	REDUP(RED)$_{lex}$	ONSET	ALIGN-/e/-L	REDUP(RED)
a. $\quad$ _-ā$_{1,2}$ger-	*!	*		*
b. ☞ $\quad$ $\underline{\text{a}^i\text{g}^k}$-ā$_{1,2}^i$g^{k}er-		*	**	

(65) REDUP(RED)$_{lex}$ **without** lexical indexation selects vowel lengthening (cf. (35))

/RED, e$_1$, a$_2^i$g^i-/	REDUP(RED)$_{lex}$	ONSET	ALIGN-/e/-L	REDUP(RED)
a. ☞ $\quad$ _-ā$_{1,2}$g-	not	*		*
b. $\quad$ $\underline{\text{a}^i\text{g}^k}$-ā$_{1,2}^i$g^k-	applicable	*	**!	

I will now go step-by-step through the diachronic stages of development, from the first emergence of (Pre-)AR in the period of Pre-Greek still containing laryngeals down to the language as we have it in attested Ancient Greek. I will show that the evidence at each stage is compatible with a trajectory that ends with Attic Reduplication forms being synchronically derived through a lexically-indexed REDUP(RED) constraint. I will also show that the development of Attic Reduplication vis-à-vis REDUP(RED) is mirrored by the later development of the class of exceptional C$_1$-copying perfects with reduplicated presents (cf. Section 2.2.3).

The procedure for making these claims is based on the system developed by Becker (2009), which uses Recursive Constraint Demotion, Inconsistency Detection, and Constraint Cloning as a deterministic mechanism for deriving lexically-indexed constraints in the grammar. I diverge from Becker in adopting the more traditional view of indexed constraints as specific vs. general, at least for cases like the ones under discussion.

2.4.1 Compositionality in Greek Reduplication

Before proceeding to the explanation of *how* Attic Reduplication was retained as a phonologically generable pattern, it is necessary first to show that Attic Reduplication should in fact be treated as a phonologically generable pattern throughout its history. That is to say, a possibly simpler alternative account of the persistence of Attic Reduplication would be to claim that the AR forms are retained as *non-compositional listed allomorphs* to particular roots. However, there is clear evidence for compositionality in reduplication in Greek in general, and thus compositionality for the AR forms. The best such evidence comes from the treatment of reconstructed root-initial labiovelar consonants in the unproductive reduplicated present.

While Ancient Greek productively/obligatorily displays reduplication only in the perfect tense, it does show remnants of reduplicative processes in its two other tense-stems: the present and the aorist.[41] In Proto-Indo-European and Pre-Greek, we can reconstruct reduplication of a form virtually identical to that of the perfect (*Ci-* in the present, *Ce-* in the aorist; see Section 2.5) that was used as an optional derivational process of stem-formation in these two tense categories. By the time of attested Ancient Greek, it appears that new forms could not be generated in this way; but many relics remain (particularly in the present). The unproductiveness of present reduplication gives us a window into the nature of reduplication in the system, vis-à-vis its interaction with sound change.

Proto-Indo-European contained a series of consonants which are reconstructed as labiovelar stops (*k^w, *g^w, *g^{wh}). These sounds are retained as such (modulo devoicing of *g^{wh}) in Mycenaean, the earliest attested dialect of Greek. Subsequently, they undergo a series of conditioned partial

---

[41] See van de Laar (2000) for a catalog of Greek verbal forms; see Giannakis (1992) for a study of the reduplicated presents in Greek, and Beckwith (1996) for a study of the reduplicated aorists.

mergers with the other stop series (see, e.g., Schwyzer 1939:293–296, Rix 1992, Sihler 1995), and have completely merged with the other stops by the period of Common Greek. The laryngeals have already been lost by the Mycenaean period. Therefore, any process relating to the conditioned outcomes of the labiovelars necessarily post-dates any processes affecting the laryngeals.

Of interest here are two particular outcomes of the labiovelars: labiovelars generally became coronals before a front vowel ($[\breve{e},\breve{i}]$), but, for the most part, they became labials elsewhere. When a root-initial labiovelar entered into reduplication, the possibility arose that the copied consonant and the root-initial consonant might surface in contexts which would condition different outcomes. Specifically, the reduplicated consonant would be in the coronalizing context (because of the fixed [e] in perfect and aorist reduplication and the fixed [i] in present reduplication) even if the root-initial consonant was in the default (i.e. labializing) context. If such a form surfaces in Greek with a coronal in the reduplicant but a labial in the root, we would know that the form was "frozen" prior to the application of the labiovelar sound changes. If, on the other hand, the reduplicant consonant matches the outcome in the root, we can surmise that the form was generated compositionally subsequent to the application of the sound change. Since we know that reduplication is fully productive in the perfect, perfect forms would all be expected to display the latter behavior (and they do). This question will therefore only be probative when asked about reduplication in the present or the aorist.

The root $\sqrt{}*g^wer(h_3)$ 'eat' gives us exactly the desired test case. This root has a reduplicated present, which takes the form βιβρώσκω [b̠-i-brɔ́-sk-ɔ̄]; it is not **διβρώσκω **[d̠-i-brɔ́-sk-ɔ̄], which would be the outcome predicted by regular sound change (as if from $*\underline{g^w}\text{-}i\text{-}g^wr(e)h_3\text{-}sk\text{-}\bar{o}$).[42] The fact that such forms do not show the outcomes of regular sound change demonstrates that they were subject to compositional production past the stage at which the labiovelars changed (Schwyzer 1939:649). That is to say, if they had come to be stored non-compositionally, Base-Reduplicant identity would not have protected the copied consonant from undergoing the expected sound change. If the unproductive reduplicated presents were being generated compositionally at this stage, it seems extremely likely that all perfect forms — including Attic Reduplication perfects — were being generated compositionally as well, since reduplication was fully productive in the perfect tense well beyond that point. This strongly indicates that AR forms are being generated compositionally past the point at which the laryngeals are lost. This should lead us to eschew the non-compositional analysis, and explore an analysis in which the AR pattern is continuously generated in the phonology.

2.4.2 The Diachrony of Laryngeal Loss and the Reflexes of the Perfect

The grammar developed in Section 2.3 is sufficient to generate the output distribution of perfect forms only until such time as the laryngeals are lost via sound change. Obviously, once the laryngeal segments are no longer present in the language, phonotactic constraints governing their distribution will no longer be able to explain the pattern. Therefore, if the forms are to be retained in their post–laryngeal-loss reflexes, it must be the case that a different grammatical mechanism is responsible for doing so.

To understand what this would have to look like, let us first consider what the reflexes of Pre-AR–stage laryngeal-initial perfect forms would be immediately after laryngeal loss. To do this, we must first know the nature of the processes, both synchronic and diachronic, affecting the laryngeals. We already encountered the reflexes of laryngeals in (44) above. It is actually more

---

[42] The same argument can be made with the reduplicated aorist πέφνον [pe-pʰn-on] 'he slayed'. This form is built from the PIE root $*\sqrt{}g^{wh}en$, which normally yields Ancient Greek $\sqrt{}t^hen$ (present θείνω [tʰēn-ɔ̄]). The regular outcome of a reduplicated aorist to the root in PIE terms, i.e. $*\underline{g^{w(h)}}e\text{-}g^{wh}n\text{-}on$, would have been **$\underline{te}\text{-}p^hn\text{-}on$, contrary to fact.

appropriate to think of these outcomes as the reflexes of the (non-high, non-back) *vowels* that were adjacent to laryngeals, rather than direct reflexes of the laryngeals themselves. This is illustrated in (66) below, which reflects the outcomes of laryngeal-adjacent vowels in Pre-Greek,[43] now with the understanding that the "vocalization" outcome from (44) is simply the coloration of an epenthetic vowel (cf. Section 2.3.3).

(66)   Laryngeal-adjacent vowel reflexes in Pre-Greek

	Coloration	Lengthening
	*ə / H_ , _H	*e / _H{C,#}
	*e / H_ , _HV	
$*h_1$	e	ē
$*h_2$	a	ā
$*h_3$	o	ō

The processes that interact to generate this distribution, some of which are best thought of as synchronic processes, others as diachronic processes (i.e. sound changes), are spelled out in (67) below. Coloration (67a) was likely a synchronic allophonic process, applying perhaps already in PIE. This process does not affect the shape of words (since it only changes vowel quality), but will be helpful for tracking later changes in the derivations below. It is after this process is in effect that laryngeals are lost via sound change (67b). The results of loss can be split into two types, based on the position of the laryngeal. If the laryngeal was in coda position (i.e. V_{C,#}), and thus moraic (67b.i), laryngeal loss results in compensatory lengthening of the preceding vowel (as long as it was not an epenthetic vowel).[44] On the other hand, when the laryngeal was in onset position (i.e. _V), and thus non-moraic (67b.ii), it was lost without a trace (other than any coloration effects it had already induced). When this onset laryngeal was indeed intervocalic (V_V), its loss would have created a hiatus. Throughout the history of Greek, including at this stage, hiatus is resolved through (mora-preserving) coalescence (67c), resulting in a long vowel.

(67)   Laryngeal-related processes
   a.   **Synchronic:** Coloration: $\{e,ə\} \rightarrow V_i$ // H_i
         i. $\{e,ə\} \rightarrow e$ // h_1
         ii. $\{e,ə\} \rightarrow a$ // h_2
         iii. $\{e,ə\} \rightarrow o$ // h_3
   b.   **Diachronic:** Laryngeal Loss
         i.   In coda: $H_\mu > \mu$ (VH > V̄ / __{C,#} ; V ≠ ə)
         ii.   In onset: $H_\emptyset > \emptyset$
   c.   **Synchronic:** Coalescence: (*VHV >) V.V $\rightarrow$ V̄

Understanding the interaction between these processes, we can now derive the reflexes of the laryngeal-initial perfect forms.[45] *HCeC roots uniformly displayed Pre-AR, as their reduplicative bases always began in *HC clusters regardless of the ablaut category (68a). *HeC roots, on the

---

[43] The long mid vowels of this stage later become lax vowels. At this stage, there is no tense/lax contrast.

[44] It is also possible that this was a synchronic, recoverable process prior to full laryngeal loss. This makes no difference for the points relevant to this section.

[45] I will assume that no significant remodeling takes place while the sound changes are in progress. This is likely an oversimplification, but hopefully an innocuous one.

other hand, must have displayed two patterns, depending on the ablaut category (see Section 2.3.5): in zero-grade categories, they had a base-initial *HC cluster ($\sqrt{HeC} \to$ //HC//), and thus displayed Pre-AR (68b); in *o*-grade categories (and hypothetical *e*-grade categories), their root-initial laryngeal was pre-vocalic in the reduplicative base ($\sqrt{HeC} \to$ //HoC//), and thus they supported default C_1-copying (68c). This produces a crucial difference at the Pre-AR stage between *o*-grade *HeC bases on the one hand, and all other laryngeal-initial forms on the other. This distinction at the Pre-AR stage is maintained when these forms are affected by the sound changes in (67). This is demonstrated in (68) below.

(68)   Diachrony of laryngeal-initial perfects

Pre-AR Stage examples	a. H∂C-e-HCeC- [h₂∂g-e-h₂ger-]	b. H∂C-e-H<Ø>C- [h₁∂d-e-h₁d-]	c. H-e-H{e,o}C- [h₂-e-h₂og-]
(i) Coloration	*h₂agah₂ger-*	*h₁edeh₁d-*	*h₂ah₂og-*
(ii.a) Lengthening	*h₂agāger-*	*h₁edēd-*	—
(ii.b) Loss	*agāger-*	*edēd-*	*a.og-*
(iii) Coalescence	—	—	*ōg-*
Ancient Greek	[agāger-]	[edēd-]	[ōg-]

Two things are especially noteworthy. First, both types end up obtaining a long vowel at the juncture between the fixed /e/ and the root, though from different diachronic sources: the *HCeC roots and the zero-grade *HeC bases get the long vowel from laryngeal lengthening; the *e-/o*-grade *HeC bases get the long vowel from vowel coalescence following laryngeal loss. Nonetheless, given that coalescence was a "persistent" process, the long vowels from the two sources probably arose virtually simultaneously.

Secondly, and most significantly, the laryngeal sound changes create a categorical distinction between the two types with respect to whether or not any material can be transparently attributed to a reduplicant. The former [H∂C] reduplicated string of the first type retains phonological substance as *VC* after the loss of the laryngeals. In the *e-/o*-grade *HeC bases, however, all that previously inhabited the reduplicant was a laryngeal. This laryngeal is lost without any trace via the general laryngeal loss sound change (67b.ii).[46] Therefore, after these sound changes took effect, there was effectively *no reduplicant* in the output of former *e-/o*-grade *HeC bases. This is the first corner of the language where a perfect lacked an overt reduplicant.[47] Without the laryngeal-related phonotactics as a guide, these two patterns, *VC*-*V̄C(VC)*- (Attic Reduplication) and Ø-*V̄C*- (vowel-lengthening), cannot be generated by a single grammar consisting of fully general constraints.

2.4.3   Inconsistency Detection and Constraint Cloning for Attic Reduplication

Consider now the learning situation that obtained when learners were faced with the distribution produced by the sound changes discussed above. For vowel-initial roots, they encounter two distinct, irreconcilable types of perfects: Attic Reduplication perfects of the shape *VC*-*V̄C(VC)*- (e.g. *agāger-*) and vowel-lengthening perfects of the shape Ø-*V̄C*- (e.g. *ōg-*). While, at a previous

---

[46] The reduplicant-initial laryngeal would not even have left any unique coloration effects. First of all, any coloration effects which might have obtained on the fixed /e/ were equally well attributable to the root-initial laryngeal. And secondly, in *o*-grade forms, the [o] quality of the base vowel is not subject to coloration effects at all, and is dominant in coalescence, such that potential coloration effects on the fixed /e/ would have been overridden.

[47] Other than the archaic non-reduplicated perfect οἶδα [óìd-a] 'know' < *woid-.

69

stage, some aspects of this distinction might have been derivable from ablaut differences (e.g., the vowel-lengthening perfects should have had a preponderance of initial [ō]'s, since they would have authentically only originated in *o*-grade forms), the morphological signification of ablaut is by this point highly degraded, in no small part due to laryngeal-related changes obscuring the core $e \sim o \sim \varnothing$ (*root $\sim$ perfect active singular $\sim$ perfect elsewhere*) ablaut pattern.

We can demonstrate the problem most clearly using the comparative tableau format introduced by Prince (2002), where the violation profiles of *winner~loser* pairs across different derivations are compared directly: constraints which prefer the winner over the loser for a given pair are marked with **W**, constraints which prefer the loser are marked with **L**, and constraints on which the winner and loser tie are marked with **e** (for *even*). This format is conducive for using Recursive Constraint Demotion (RCD; Tesar 1995, Tesar & Smolensky 1998, 2000, *et seq.*) to attempt to find a consistent ranking over multiple distinct input-output mappings; convergence is possible only when there is a constraint ranking which allows each *winner~loser* pair to have at least one W outranking all L's. RCD finds this constraint ranking by (i) collecting all constraints which prefer no losers and installing them at the top of the ranking, (ii) removing from the tabulation all mappings for which the just-installed constraint(s) assign a **W**, and (iii) repeating until all mappings are accounted for.

For the distribution presented to learners at the stage immediately after laryngeal-loss, there are four types of mappings observable in the perfect. All consonant-initial roots/bases still map to C_1-copying perfects (the change to non-copying happens only later; see below). I split them into two types: (i) C(R)VX, i.e. roots/bases beginning in a consonant followed by a sonorant (either a vowel or a sonorant consonant); and (ii) STVX, i.e. roots/bases beginning in a sequence of obstruents (covering ST, TS, and TT, and perhaps other non-obstruent clusters, and also initial geminates). The two types differ in their violation profile with respect to the constraint *PCR; this constraint is omitted from the set of tableaux in this subsection, as it is not yet active, but will be included and significant in the following stage (see Section 2.4.4). The two remaining types are the two mappings for vowel-initial roots: (i) Attic Reduplication (*VC-V̄C(VC)-*), which will be notated as the *ager* class; and (ii) vowel-lengthening (*∅-V̄C-*), which will be notated as the *og* class.

Each *winner~loser* pair consists of the (properly-anchored) copying candidate (C_1-copying for C-initial roots, VC-copying — i.e. AR — for V-initial roots) and the non-copying candidate (basic non-copying for C-initial roots, vowel-lengthening for V-initial roots). The three constraints which are relevant at this stage are REDUP(RED), ALIGN-/e/-L, and ONSET. The first two of these will be abbreviated RDP and ALIGN, respectively.

Now we can begin reasoning about the rankings. The table in (69) provides the violation profiles for each of the four mappings.

(69)   Violation profiles of post–laryngeal-loss mappings

	RDP	ALIGN	ONSET
/C(R)VX/ → [C̲-e-C(R)VX] ≻ [∅̲-e-CVX]	W	L	W
/STVX/ → [S̲-e-STVX] ≻ [∅̲-e-STVX]	W	L	W
ager → [ag̲-āger] ≻ [∅̲-āger]	W	L	e
og → [∅̲-ōg] ≻ [og̲-ōg]	L	W	e

When RCD is run over these mappings, it is able to install ONSET in the first stratum, since, for each mapping, it either prefers the winner or treats the winner and loser equally; it never prefers a loser. Ranking ONSET high is sufficient to account for both of the consonant-initial patterns (indicated with gray rows in (70)). However, it does not account for either of the vowel-initial

patterns. In both cases, both winner and loser have one ONSET violation each, and are thus equivalent with respect to ONSET (and thus their *winner~loser* pairs both have **e** in the ONSET column). With ONSET installed, as shown in (70), the two remaining constraints conflict with respect to the remaining mappings: REDUP(RED) prefers the winner for *ager* but the loser for *og*; whereas ALIGN-/e/-L prefers the winner for *og* but the loser for *ager*. This situation crashes traditional RCD, because there is no consistent ranking that can generate all the *winner~loser* pairs.

(70)   RCD installs ONSET and then crashes

		ONSET	RDP	ALIGN
/C(R)VX/ → [C̲-e-C(R)VX] ≻ [Ø̲-e-C(R)VX]		W	W	L
/STVX/ → [S̲-e-STVX] ≻ [Ø̲-e-STVX]		W	W	L
ager → [ag̲-āger] ≻ [Ø̲-āger]		e	W	L
og → [Ø̲-ōg] ≻ [og̲-ōg]		e	L	W

There is, however, a way to rescue an RCD crash: constraint cloning. As developed by Pater (2009), Becker (2009), *et seq.*, the way to overcome inconsistency of this sort is to introduce a copy of one of the conflicting constraints, and index this copy to a subset of the mappings, namely, those for which it prefers the winner (at least among the mappings not already accounted for by already installed constraints). Given this approach, we can explain the inconsistent distribution above by cloning either RDP or ALIGN. I opt for RDP, though this choice is somewhat arbitrary.[48] As shown in (71), introducing and installing a cloned version of RDP which is indexed (minimally) to the *ager* mapping allows RCD to progress to completion. Once the cloned RDP constraint accounts for the *ager* mapping, ALIGN can be installed to account for the *og* mapping. All mappings are now accounted for, and consistency is achieved.

(71)   Clone RDP; Install Align; RCD completes

		ONSET	RDP$_{ager}$	ALIGN	RDP$_{og}$
/C(R)VX/ → [C̲-e-C(R)VX] ≻ [Ø̲-e-C(R)VX]		W	?	L	?
/STVX/ → [S̲-e-STVX] ≻ [Ø̲-e-STVX]		W	?	L	?
ager → [ag̲-āger] ≻ [Ø̲-āger]		e	W	L	?
og → [Ø̲-ōg] ≻ [og̲-ōg]		e	e	W	L

One significant question remains: which mappings are indexed to which RDP constraints? We know that the higher-ranked RDP constraint must evaluate *ager* and not evaluate *og*, and that the lower-ranked RDP constraint must evaluate *og*. *A priori*, we do not know what the violation profile of the consonant-initial roots are with respect to the two constraints, since these constraints play no active role in selecting their winners; hence the question marks in the first two rows of (71). But furthermore, there is also a question about the behavior of the lower-ranked RDP constraint with respect to the *ager* mapping, as indicated by the last question mark in (71).

---

[48] There are several reasons one might prefer RDP in this situation. First, phenomenologically, the thing which appears to be at issue here is whether you copy or not. This is exactly what REDUP(RED) governs. Therefore, indexing this constraint seems to get directly to the heart of the issue. Second, in Zukoff (2017a,b), I have argued that the ranking and function of alignment constraints is not entirely arbitrary (namely, they are the phonological avatars of morphosyntactic instructions for linearization). If they are non-arbitrary, it might be preferable (or required) for speakers not to posit indexed versions of them.

If the cloning process simply yields two distinct constraints with non-overlapping sets of indexed mappings, then *ager* will not be evaluated by the lower-ranked RDP constraint (and thus the cell would receive an **e**), as this applies *only* to the *og* mapping (i.e., all the roots or bases that display vowel-lengthening perfects). Becker (2009) argues that this interpretation is necessary in order to properly identify the members of "lexical trends", i.e. variation over the lexicon that is categorical by lexeme, such that statistics about the lexical trend can be brought to bear on the production of a novel form. This is a desideratum given the phenomenon of "frequency matching": speakers frequently re-create the statistical distribution of lexical trends in *wug*-test experiments (see Zuraw 2000, Albright & Hayes 2003, Hayes & Londe 2006, Becker 2009, and many others).

On the other hand, *contra* Becker (2009), we could assume that cloning results not in two constraints with distinct indexed sets of applicable mappings, but rather (as in most of the previous literature on this subject) in a *specific* constraint (the high-ranked one) and a *general* constraint (the low-ranked one). This approach would have the benefit of immediately knowing how to treat the violation profiles of the already-accounted for mappings: they are subject to the general constraint. This distinction could be important for an actual language learner, who is receiving evidence serially, not all at once. If the learner receives new evidence that invalidates an early-installed constraint, and thus triggers a reset of RCD, the learner will potentially need to know how the existing mappings fare on both versions of the cloned constraints, as certain mappings which formerly were explained by the early-installed constraint can be no longer. If the cloned constraints are simply specific and general, and mappings can be diverted to the specific constraint only when inconsistency is detected, the learner will know to run RCD first under the assumption that each mapping that has not already received an index is evaluated only by the general constraint. If and when this fails, the learner can try to run RCD again with this assumption reversed, namely that it actually should be indexed to the special version. Only if and when this fails would cloning of an additional constraint be triggered.

Beyond this general rationale, one reason to think that the current case should be analyzed as a specific-general relationship is that it is unclear if the *ager* mapping is a "lexical trend" that speakers of the language actually internalized as such. Becker (2009) and others distinguish among statistical patterns present in the lexicon between those that speakers *do* replicate in the experimental setting and those that they *don't*. If speakers fail to replicate the pattern, Becker and others reason, they have not indexed across the lexicon according to the difference instantiated by the lexical statistics. That is to say, the distinction in mappings based on that apparent lexical difference is not encoded by speakers as an indexation to distinct constraints. Obviously, we cannot perform *wug*-tests on a dead language, so we cannot know for sure whether speakers internalized the distinction between the two mappings for vowel-initial roots as a "lexical trend." The best available proxy for this, though, would seem to be the status of innovation with respect to the two mappings.

In the history of Greek through at least the Classical Period, there seems to be at most two examples of roots which could not have organically developed Attic Reduplication from an original laryngeal. One is the root $\sqrt{eme}$ 'vomit', which etymologically has an initial *w (PIE *$\sqrt{wemh_1}$-); it makes an AR perfect ἐμήμεχα [em-ēme-k-a]. There is no direct explanation for this, and thus it must be viewed as an extension of the AR pattern. The second is $\sqrt{or}$ 'keep watch', which certainly had a pre-vocalic *w, and possibly initial *sw (i.e. < *(s)wer-) if indeed related to the present ὁράω [horá-ō] (cf. Chantraine 1980:813–815); it makes an AR perfect ὄρωρα [ór-ōr-a]. Both the root and the perfect stem are identical to a separate root $\sqrt{or}$ 'incite' (present ὄρνυμι [ór-nū-mi]) → perfect ὄρωρα [ór-ōr-a]. It would be reasonable to assume that the inauthentic AR perfect for $\sqrt{or}$ 'keep watch' arises because of direct confusion with $\sqrt{or}$ 'incite', not because of meaningful extension of the AR pattern. Thus, there are at most two examples, and perhaps just one example, of extension of the AR pattern beyond its original scope.

If the two types of vowel-initial perfect mappings existed as a lexical trend in the mind of speakers, we should expect there to be some number of innovative AR forms attested, given that there were a great many new vowel-initial roots that entered the language throughout its history which required the coinage of a new perfect form. Even if the statistical power of the vowel-lengthening pattern greatly exceeded that of the AR pattern (and it seems likely to be the case that it did *not* do so in the very early stages at least), we should expect at least a smattering of innovative AR perfects to be attested, even if only as infrequent variants or "dialectal" forms. With the exception of ἐμήμεχα and possibly ὄρωρα 'keep watch', no such cases are (to my knowledge) attested. Given this striking lack of extension, it seems that we do not want to invoke a system whereby the AR forms are encoded as a lexical trend; rather, we want a system where they are encoded simply as a closed class of archaic exceptions. The specific-general constraint cloning approach captures this, while Becker's (2009) dual list approach does not.

I thus adopt the specific-general approach for the present case. The tableau in (72) reflects the violation profiles yielded by this approach.[49]

(72)   Violation profile and ranking with specific vs. general REDUP(RED)

		ONSET	RDP$_{spec}$	ALIGN	RDP$_{gen}$
/C(R)VX/ → [C̲-e-C(R)VX] ≻ [Ø̲-e-C(R)VX]		W	e	L	W
/STVX/ → [S̲-e-STVX] ≻ [Ø̲-e-STVX]		W	e	L	W
ager → [ag̲-āger] ≻ [Ø̲-āger]		e	W	L	W
og → [Ø̲-ōg] ≻ [og̲-ōg]		e	e	W	L

Regardless of which manner of constraint cloning we employ, it is clear that a single cloned — i.e. lexically-indexed — constraint is sufficient to capture the distribution at this stage. This comports with the analytical approach developed in Section 2.2 based purely on the synchronic state of Ancient Greek.

2.4.4 Deriving the Behavior of REDUP(RED)$_{lex}$ After the Rise of *PCR

In Section 2.4.3, I derived the emergence of the constraint REDUP(RED)$_{lex}$ (i.e. a specific version of REDUP(RED) indexed to a subset of roots in the language) through constraint cloning. This becomes slightly more complicated when we introduce non-copying as the general pattern for obstruent-obstruent–initial roots, with a special over-copying pattern for the subset of obstruent-obstruent roots with a reduplicated present (see Section 2.2.3). This is because we now have two, partially overlapping inconsistencies, rather than just the one which was present at the prior stage. Nevertheless, given several additional assumptions, the same approach that derived the constraint indexation for the Attic Reduplication forms is capable of deriving the same result for the exceptional C$_1$-copying perfects.

The violation profile in (73) reflects two changes relative to the violation profile of the earlier stage given in (69) above. First, the *winner∼loser* pair for STVX is reversed, now selecting the non-copying pattern over the C$_1$-copying pattern. This mapping's violation profile is thus the reverse of that shown in the earlier discussion. Second, there is one additional mapping: the *pipto* type, which represents the exceptional C$_1$-copying perfects detailed in Section 2.2.3. This mapping's violation profile is equivalent to the former STVX mapping, as it still shows C$_1$-copying despite

---

[49] Note that, if RCD is restarted when cloning occurs, there would be no ranking difference between ONSET and RDP$_{spec}$.

having an obstruent-obstruent–initial root. Additionally, *PCR has been added to the constraint set, as this will now be crucial in generating non-copying for the STVX mapping.

(73)   Violation profiles of mappings after change of ST to non-copying

	*PCR	RDP	ALIGN	ONSET
/C(R)VX/ → [C̲-e-C(R)VX] ≻ [Ø̲-e-C(R)VX]	e	W	L	W
/STVX/ → [Ø̲-e-STVX] ≻ [S̲-e-STVX]	W	L	W	L
pipto → [p̲-e-pto] ≻ [Ø̲-e-pto]	L	W	L	W
ager → [a̲g-āger] ≻ [Ø̲-āger]	e	W	L	e
og → [Ø̲-ōg] ≻ [o̲g-ōg]	e	L	W	e

Unlike at the earlier stage, ONSET cannot be installed on the initial round of RCD, as it is now loser-preferring for the STVX mapping. In fact, no constraints are uniquely winner-preferring, and thus we immediately have inconsistency. According to Becker's (2009:149–155) proposal, when there are multiple constraints which could be cloned to resolve inconsistency, the constraint cloning algorithm selects the one with the fewest combined **W**'s and **L**'s; this is a means of separating out distinct lexical trends. This would result in cloning of *PCR, and ultimately lead to the following ranking with two separate cloned constraints, capturing the two separate lexical trends (i.e. among the obstruent-obstruent roots and the vowel-initial roots):

(74)   **Constraint ranking derived from cloning *PCR:**
        *PCR*STVX* ≫ ONSET ≫ RDP*ager* ≫ ALIGN ≫ *PCR*pipto*, RDP*og*

This again raises the questions discussed in Section 2.4.3 regarding the appropriateness of encoding the lexical trends in these cases. Membership in the *pipto* class is completely non-arbitrary from a synchronic perspective (albeit almost completely arbitrary from a diachronic perspective): the only roots that comprise this class are roots which make reduplicated presents. This is a discrete morphological fact, not an arbitrary phonological distribution. While the class is small enough that we probably do not expect diachronic extension, it is definitively the case that there is no diachronic extension of this pattern. These two notions together make it again reasonable to assume that this pattern should not be encoded in the grammar as a lexical trend in the sense of Becker (2009).

Thus, rather than following Becker's proposal for the choice of constraint to clone, let's assume that the mechanism for determining which constraint to clone allows for RDP to be selected for cloning. If the goal is to minimize the number of constraints cloned, then this will be the only possible choice, as it is the only constraint whose cloning can explain the *ager* class.

(75)   Clone RDP (specific-general version)

	RDP*spec*	*PCR	RDP	ALIGN	ONSET
/C(R)VX/→ [C̲-e-C(R)VX] ≻ [Ø̲-e-C(R)VX]	?	e	W	L	W
/STVX/ → [Ø̲-e-STVX] ≻ [S̲-e-STVX]	–	W	L	W	L
pipto → [p̲-e-pto] ≻ [Ø̲-e-pto]	W	L	W	L	W
ager → [a̲g-āger] ≻ [Ø̲-āger]	W	e	W	L	e
og → [Ø̲-ōg] ≻ [o̲g-ōg]	–	e	L	W	e

The cloned RDP*spec* constraint *must* be indexed to both the *pipto* mapping and the *ager* mapping (hence the **W**'s in their columns, and the shading in their rows), and the STVX mapping and the

og mapping *must not* be indexed to RDP$_{spec}$ (indicated by the " – " in their columns). There is a question, though, about whether the basic C(R)VX mapping should be indexed to RDP$_{spec}$. In terms of the basic algorithm, we might expect the answer to be yes, as the clone is supposed to collect all mappings that have a **W** in its column (Becker 2009; cf. Pater 2009). However, given that we are operating with a specific-general version of cloning, it might instead be reasonable to assume that there is a bias against indexation to the special constraint; that is, a bias only to index to the special constraint if inconsistency results otherwise. Put another way, this system would assume a general bias *against* indexation, whereas Becker advocates for a system where essentially everything is indexed.

Once we make the assumption that C(R)VX is not indexed to RDP$_{spec}$, RCD can run to completion, yielding the ranking shown in (76).[50]

(76)   RCD assuming C(R)VX is not indexed to RDP$_{spec}$

	RDP$_{spec}$	*PCR	ONSET	ALIGN	RDP
/C(R)VX/→ [C̲-e-C(R)VX] ≻ [Ø̲-e-C(R)VX]	e	e	W	L	W
/STVX/ → [Ø̲-e-STVX] ≻ [S̲-e-STVX]	e	W	L	W	L
pipto → [p̲-e-pto] ≻ [Ø̲-e-pto]	W	L	W	L	W
ager → [ag̲-āger] ≻ [Ø̲-āger]	W	e	e	L	W
og → [Ø̲-ōg] ≻ [og̲-ōg]	e	e	e	W	L

2.4.5   Local Summary and Discussion

The set of assumptions laid out in this section allows us to arrive at the full synchronic analysis of Ancient Greek posited in Section 2.2 by means of morphophonological learning at each stage of the development of the language. Constraint indexation is crucial to resolving the inconsistency brought about by diachronic factors. By adopting a specific-general approach to constraint indexation and cloning (*contra* Becker 2009), we are able link two types of exceptional behavior within the system: Attic Reduplication and the exceptional C$_1$-copying perfects associated with reduplicated presents.

An important point that must be made here is that the learning procedure just discussed does not constitute a theory of sound change. In the two case studies here — learning indexation and ranking after the loss of laryngeals, and learning indexation and ranking after the change to non-copying (i.e. after the promotion of *PCR) — the learning procedure takes as its starting point the set of outputs produced by a sound change which forces a reorganization of the morphophonological grammar. The role of this learning procedure is to (try to) establish a consistent ranking over that set of inputs. An explicit theory for the learning of phonotactic grammars resulting from sound changes is still required, and the learning procedure employed here must ultimately be made consistent with that theory of phonotactic learning. I must defer further exploration of this subject to future work. (See Bowers 2015 for recent work on related issues.)

---

[50] If we made the reverse assumption, RCD would still converge, but with a different ranking:

   (i)   **Ranking:** RDP$_{spec}$ ≫ *PCR, ALIGN ≫ RDP, ONSET

   The reason for the higher ranking of ALIGN is that, when C(R)VX *is* indexed to RDP$_{spec}$, that **W** covers C(R)VX's **L** for ALIGN. Thus RCD can install it on the next iteration.

2.5 Attic Reduplicated Presents and Aorists

While reduplication remains productive in Ancient Greek only in the perfect, there were parallel reduplication patterns in both the present (as discussed in Section 2.2.3) and the aorist at earlier stages of the language. Among these, there are a few forms which do appear to reflect the Attic Reduplication pattern. In the present, lack of data obscures the situation, though there is at least one form which (if it is rightly to be analyzed as AR) is completely consistent with what would be predicted based on the perfect. In the aorist, on the other hand, there is sufficient evidence to establish the nature of the pattern. This pattern differs from the perfect only in the length of its second vowel, yet this has significant implications for the analysis of the underlying representation of all three patterns, and for the diachrony of the vowel system in Greek.

2.5.1 Attic Reduplicated Presents

Two present forms, ὀνίνημι [onínɛ̄-mi] and ὀπῑπεύω [opīp-eú-ɔ̄], given in (77), look as if they could have come about through Attic Reduplication.[51] While Chantraine (1980:808) entertains a redupli-cated origin for at least ὀπῑπεύω, Beekes & van Beek (2010:1083–1084,1090–1091) propose non-reduplicated etymologies for both (a nasal-infix present in the first case, a prefix *opi-* in the second), though both analyses are quite tentative.

(77)   Attic Reduplicated presents (forms from Cowgill 1965:§2, Giannakis 1992)

Root		Attic Reduplicated Present
$\sqrt{on\bar{\varepsilon}} < *h_3neh_2$ 'help'	ὀνίνημι	[oninɛ̄-mi]
$\sqrt{op} < *h_3ek^w$ 'see'	ὀπῑπεύω	[opīp-eu-ɔ̄] ('watch, spy')

Assuming that these forms are to be connected with Attic Reduplication, an obvious problem is the disagreement in the length of the second vowel in the two forms: ὀπῑπεύω has long [ī] but ὀνίνημι has short [i]. The analysis developed above for Attic Reduplication in the perfect predicts the long [ī] of ὀπῑπεύω, correcting for the difference in vowel quality between present and perfect reduplication; as seen in Section 2.2.3, present reduplication has a fixed /i/ vowel, rather than the fixed /e/ of the perfect. Positing an underlying representation for present reduplication with /RED, i/, the AR derivation proposed for the perfect follows directly: *[HəC-i-HC...] > [VCīC...]. The reason we see [ī] in the place of the perfect's [ɛ̄,ā,ɔ̄] alternation is that laryngeals had no coloration effect on high vowels, but still did lengthen them in the same environments (namely, pre-consonantally and word-finally; see (44) above). Given that the analysis predicts the long [ī], the short [i] of ὀνίνημι [onínɛ̄mi] presents a problem. With such a small data set, and with such uncertainty around both potential forms, it is impossible to say if one or the other (or either) should rightly be considered the

---

[51] One more word might also belong on this list: ἀτιτάλλω [atitállɔ̄] 'rear, raise'. The word is of uncertain origin (Chantraine 1980:131–132, Giannakis 1992:317–320, Beekes & van Beek 2010:161) and without obvious cognates, but, with respect to its form, seems to resemble the forms in (77). If it were from a root *$h_2t(e)l$ or *$h_2t(a)l$ (which looks distinctly un-Indo-European), it could have a derivation equivalent to that of ὀνίνημι [onínɛ̄mi]. Chantraine cites a view whereby the initial [a] is actually the "alpha-privative" negative marker. If it was synchronically analyzable as such, the root would be /tal/ rather than /atal/. This would then form a normal reduplicated present in [ti-tal-], to which the negative marker would be added externally, to yield [a-ti-tal-]. Given the uncertainty surrounding this form, it should not be taken as significant evidence for the current discussion, one way or the other.

"regular" outcome, but it is noteworthy that one of the options is predicted by the account provided for the perfect.[52]

2.5.2 Attic Reduplicated Aorists

The inventory of Attic Reduplicated aorists is slightly larger (cf. Beckwith 1996), and thus we will be able to determine the regular pattern with somewhat more certainty. However, accounting for the pattern will require several reasonable though not externally motivated assumptions. (All aorist forms in this section are given in the aorist infinitive, even if that particular form may not be attested, so as to avoid complications relating to vowel length stemming from the affixation of the past tense indicative augment prefix.)

(78)   Attic Reduplicated aorists (forms from Beckwith 1996, van de Laar 2000:393–397)

Root			AR Aorist		Perfect	
√or	< *h₃er	'incite'	ὀρορεῖν	[or-or-]	ὄρωρα	[orɔ̄r-]
√en(e)k	< *h₁neḱ	'bring'	ἐνεγκεῖν	[en-eŋk-]	ἐνήνοχα	[enɛ̄nokʰ-]
√ag	< *h₂eǵ	'lead'	ἀγαγεῖν	[ag-ag-]	ἆγμαι	[āg-]
√akʰ	< *h₂egʰ	'be troubled'	ἀκαχεῖν	[ak-akʰ-]	ἀκαχῆμαι	[akakʰɛ̄-][53]
√alek	< *h₂lek	'ward off'	ἀλαλκεῖν	[al-alk-]	not attested	
√apʰ	< *h₂ebʰ (?)	'cheat, beguile'	ἀπαφεῖν	[ap-apʰ-]	not attested	
√ar	< *h₂er	'join, fit together'	ἀραρεῖν	[ar-ar-]	ἄρᾱρα	[ar-ār-]

These Attic Reduplicated aorist forms are equivalent to what we observe in the perfect in all respects but one: their second vowel is short rather than long. This difference in vowel length can be explained by positing a difference in the underlying representation of the reduplicative morpheme(s). In the perfect and (potentially) in the present, the long vowel in the output of the AR forms is ultimately due to a sequence of *full vowel + laryngeal* at the Pre-Greek stage. That is to say, both categories have a morphologically fixed vowel in the underlying representation in addition to the reduplicative morpheme proper (/RED, e/ in the perfect, /RED, i/ in the present). When the laryngeals are lost, it is the fixed vowel which is compensatorily lengthened to yield the surface long vowel in attested Ancient Greek. A way to avoid a long vowel, therefore, would be to remove the full vowel from the underlying specification. If the reduplicated aorist lacks a fixed segment morpheme and instead simply contains /RED/ underlyingly, we obtain the desired results, pending certain assumptions.

First and foremost, we must assume that the reduplicative vowel in the aorist arises through epenthesis — the details of this will be motivated below. Second, we must assume that the sequence of *epenthetic* vowel + laryngeal yields a short vowel in Greek. Byrd (2011) argues for exactly this

---

[52] There is a further problem with ὀπῑπεύω [opīpeúɔ̄]. The second root consonant is originally the labiovelar *kʷ. The unconditioned development for this class of sounds into Greek is indeed labial stops (see Section 2.4.1). However, there is a conditioned development (in most dialects) to coronal stops before front vowels: *kʷ > t / _{i,e} (see, e.g., Sihler 1995). This would predict that both the root and reduplicated consonant should develop to [t], not [p] (i.e. *okʷīkʷ-eu-ɔ̄ > **otīt-eu-ɔ̄). If the [-eu-] stem-forming element is a later innovation (as suggested by Beekes & van Beek 2010:1091), however, it may be the case that the root-final /*kʷ/ would not have changed to [t], but rather reflect the default reflex [p]. This would permit an analysis whereby the *kʷ > t change under-applies in the reduplicant due to BR-identity effects.

[53] The (lack of) length in the second vowel of this form is unexpected. Note that the following syllable has a long vowel. An explanation might therefore lie in a dispreference for adjacent long vowels. I do not consider this further here.

77

state of affairs. Under such an account, the epenthetic vowel, which is (eventually) colored to [e,a,o] adjacent to the corresponding laryngeal, is non-moraic. Therefore, when it acquires an extra mora through compensatory lengthening, it will have a total of one mora, rather than the two moras that a full vowel would have. This requires that, when one of these epenthetic vowels is "phonologized" (i.e. first interpreted as belonging to the underlying representation) in an open syllable, as in cases of vowel prothesis (Pre-Greek /h₂ger-/ → [h₂əøger-] > Greek /a_μger-/; cf. Section 2.3.3), it acquires a mora by virtue of its position.

Given these assumptions, the analysis of Attic Reduplication in the perfect predicts double epenthesis in reduplicated aorists of laryngeal-initial roots, as follows:

(79)   Generating AR in the aorist: A. Greek $\sqrt{ag}$ (< *h₂eǵ) 'lead' : aorist inf. ἀγαγεῖν [agagên]
        ⇒ Pre-Greek /RED, h₂g, ein/ → [h₂əg-e-h₂g-ein]

/RED, h₂g, ein/			*PCR-H	CONTIG-IO	DEP-IO
a.		h₂ə-h₂g-ein ⋮ >> **agein	*!		*
b.	☞	h₂əgə-h₂g-ein ⋮ >> ✓ agagein			**
c.		h₂ə-h₂əg-ein ⋮ >> **āgein		*!	*(*)

Modulo the second vowel, winning candidate (79b) is identical to the AR pre-forms generated for perfect and present. The difference is the weight of that second vowel. Since we are assuming epenthetic vowels (laryngeally colored or not) to be mora-less, it will yield a short vowel in opposition to the long vowel produced by the other two categories' morphologically fixed vowels. This is summarized for the three different types of AR forms in (80). (Subscript Ø is meant to indicate that the vowel has no mora inherently.)

(80)   Moras and vowel length in AR forms
        a. **Perfect:**

            Pre-Greek [h₂əø.g-e_μ-h₂_μ.ge_μr-] > Proto-Greek [a_μga_μμge_μr-] > Attic-Ionic [agēger-]
        b. **Present:**

            Pre-Greek [h₃əø.kʷ-i_μ-h₃_μ.kʷ-] > Proto-Greek [o_μkʷi_μμkʷ-] > Attic-Ionic [opīp-]
        c. **Aorist:**

            Pre-Greek [h₂əø.gəø-h₂_μ.g-]   > Proto-Greek [a_μga_μg-]   > Attic-Ionic [agag-]

Once we arrive at the Greek stage, where these roots now begin with a vowel, deriving the AR aorists is completely straightforward. Simply importing the total ranking for Greek perfect reduplication from (42) above, we derive the right result, as shown in (81).

(81)   Generating AR aorists in *Greek*

/RED, ag-/			ANCHOR-L-BR	ONSET	REDUP(RED)	ALIGN-ROOT-L
a.	☞	ag-ag-		*		**
b.		g-ag-	*!			*
c.		Ø-ag-		*	*!	

ANCHOR-L-BR still rules out any mis-anchored candidates (81b). But now, since the aorist lacks anything equivalent to the fixed /e/ morpheme of the perfect, nothing triggers minimizing

the reduplicant for vowel-initial roots. For vowel-initial roots in the perfect (in the general case), the ranking ALIGN-/e/-L ≫ REDUP(RED) led to vowel-lengthening (via non-copying). However, in the aorist, ALIGN-/e/-L is not applicable, and thus REDUP(RED) is the constraint which selects between vowel-initial candidates. It prefers the AR candidate (81a) to the non-copying/vowel-lengthening candidate (81c), because of (81a)'s overt reduplicant. In this tableau, another alignment constraint, ALIGN-ROOT-L, is included. This constraint applies in all cases, but up until now never would have been particularly relevant. Here though, we see that it must in fact be ranked below REDUP(RED), as it otherwise could have had the same minimizing effect as ALIGN-/e/-L in the perfect. This interaction thus allows us to synchronically derive AR in the aorist even without resorting to REDUP(RED)$_{lex}$.[54]

2.5.3 Consonant-initial Reduplicated Aorists

This approach to the Attic Reduplicated aorists is centered on the notion that, at the Pre-Greek stage at least, the reduplicative vowel of the aorist was epenthetic. Given the appearance of the consonant-initial reduplicated aorists, as exemplified in (82), this might give us pause.

(82)   Consonant-initial reduplicated aorists (forms from Beckwith 1996, van de Laar 2000)

Root		Reduplicated Aorist	
√$keut^h$	'hide'	κεκυθεῖν	[ke-kut^h-éin]
√lat^h	'escape notice; forget'	λελαθεῖν	[le-lat^h-éin]
√$peit^h$	'persuade; trust'	πεπιθεῖν	[pe-pit^h-éin]
√tem	'find'	τετμεῖν	[te-tm-éin]

These non-AR reduplicated aorists show the same fixed surface [e] vowel as the perfect. However, if the preceding analysis of the AR aorists is correct, this fixed [e] must arise in some way other than that of the perfect. Without recourse to morphological fixed segmentism, the only way to derive the fixed vowel quality will be to employ a version of *phonological fixed segmentism* (Alderete et al. 1999). The epenthesis-based proposal of Alderete et al., whereby the phonologically fixed segment arises via epenthesis, will be ideal for generating the pattern at the Pre-Greek stage, if we assume that the [e] we observe in Greek can be derived from a Pre-Greek (non-laryngeal-adjacent) schwa (see below).

To use this approach for the Greek stage, this means that [e] must be the epenthetic vowel synchronically in Greek. There does not seem to be any other evidence for or against this claim within Greek. For convenience, I will assume a markedness-based approach to the determination of epenthetic vowel quality. (See Steriade 2009 for a faithfulness-based approach to epenthetic vowel quality; that approach is also compatible with the current problem.) To generate [e] as the epenthetic vowel, then, we need markedness constraints against the vowel features (or feature combinations) that are *not* found in [e] (e.g., [+low], [+long], [+round], [+high], etc.) to outrank those features found in [e] (i.e., [-low], [-long], [-round], [-high]).[55] I will abbreviate this with a constraint *[¬e], which penalizes all features not present in [e]; this constraint dominates *[e].

---

[54] It must be noted that a few of the AR aorists appear to violate *PCR, namely, a⟨lalk⟩ein and (if the difference in place between the nasals does not obviate *PCR) e⟨neŋk⟩ein. If they do, indexation of these roots to REDUP(RED)$_{lex}$ would be necessary and sufficient to license the *PCR violation.

[55] Sommerstein (1973), de Haas (1988), McIntyre (1992), and others take [e] to be the "unmarked" vowel in Greek due to the nature of its feature specifications (i.e. the minus value for each distinctive feature).

Both constraints are in turn dominated by IDENT-V-IO, such that they have no effects on underlying vowels. This is then an instance of the emergence of the unmarked (McCarthy & Prince 1994, 1995). The ranking IDENT-V-IO ≫ *[¬e], when coupled with IDENT-V-BR (all defined in (83)), will be sufficient to rule out any candidates that don't epenthesize their reduplicative vowel if all three constraints rank above DEP-IO.[56]

(83)   a.   **IDENT-V-BR**

Assign one violation mark * for each pair of vowels in BR-correspondence which differ in any vowel feature.

    b.   **IDENT-V-IO**

Assign one violation mark * for each pair of consonants in IO-correspondence which differ in any vowel feature.

    c.   ***[¬e]**: Assign one violation mark * for each surface vowel which is not [e].

The tableau in (84) illustrates the derivation in attested Ancient Greek of a CV-initial reduplicated aorist, where the V is /a/. Reduction of an underlying vowel to [e] (candidate (84d)), which would allow for base-reduplicant identity and maximal unmarkedness, is not permitted because of the high ranking of input-output faithfulness (IDENT-V-IO). Copying the root's [a] into the reduplicant (candidate (84c)) is disallowed by *[¬e], since the markedness of the reduplicant [a] is not protected by input-output faithfulness. Imperfect copying of the base vowel as [e] (candidate (84b)) is blocked by base-reduplicant faithfulness (IDENT-V-BR). The only remaining possibility, then, is candidate (84a), where the reduplicant vowel is epenthetic and thus not in correspondence with the base vowel (and thus not subject to base-reduplicant faithfulness). This candidate violates DEP-IO, and has increased violations of MAX-BR (since there is more material from the base that has not been copied), but these violations are tolerated under compulsion.

(84)   Generating an epenthetic reduplicative vowel in consonant-initial aorists in *Ancient Greek*

/RED, latʰ-/	IDENT-V-IO	*[¬e]	IDENT-V-BR	DEP-IO	MAX-BR
a. ☞ l<u>e</u>-latʰ-		*		*	**
b. l<u>e</u>-latʰ-		*	*!		*
c. l<u>a</u>-latʰ-		**!			*
d. l<u>e</u>-letʰ-	*!				*

The same exact approach is sufficient for the Pre-Greek stage, modulo the identity of the epenthetic vowel. As demonstrated above, the way we get the vowel length to work out right in AR aorists is by positing double epenthesis of a weightless vowel, i.e. [ə]. For consistency, we thus need to assume that it is this same vowel which is epenthesized here. With this in mind, the relevant markedness constraint at this stage is *[¬ə]. This is demonstrated in (85) below.

---

[56] This would actually select copying if the base vowel was [e], though no such cases are attested. In any event, this would be surface identical to the epenthetic candidate in Greek.

(85)   Generating an epenthetic reduplicative vowel in consonant-initial aorists in *Pre-Greek*

/RED, latʰ-/	IDENT-V-IO	*[¬ə]	IDENT-V-BR	DEP-IO	MAX-BR
a. ☞ lə-latʰ-		*		*	**
b. lə-latʰ-		*	*!		*
c. la-latʰ-		**!			*
d. lə-lətʰ-	*!				*

The one caveat surrounding this approach is that we must posit what is essentially an *ad hoc* sound change: Pre-Greek *ə > Greek *e* when not adjacent to a laryngeal. Since this outcome is distinct from the outcome of PIE "schwa secundum" (i.e. non–laryngeal-adjacent schwa), namely [i] (cf. Mayrhofer 1986), it would need to be the case that (the precise phonological details of) these forms, and perhaps also the AR aorists, post-date the phonologization of schwa secundum in the development into Greek.[57]

Given, though, that this vowel is being generated by the synchronic grammar at this and every subsequent stage, since it is involved in the process of reduplication, we could alternatively conceive of the change in terms of the grammar, rather than strictly as sound change (insofar as there may be some sort of difference). This line of thought would say that the change can be explained by a change in ranking of the constraints on vowel features, especially *[¬ə], and indeed *[ə]. At the Pre-Greek stage, *[¬ə] is the highest ranked of all such constraints, though still below IDENT-V-IO. By attested Ancient Greek, schwa has been completely eliminated from the inventory. This means that *[ə] now dominates IDENT-V-IO. (This is necessary in order to prevent potential underlying schwas from surfacing, per Richness of the Base; Prince & Smolensky 1993/2004.) This mitigates any effects of *[¬ə] (regardless of its ranking, as long as it is below *[ə]). As long as *[¬e] was (or became) the next highest ranked of the vowel markedness constraints, [e] would become the new epenthetic vowel. A rationale for the loss of schwa could come from the effects of laryngeal coloration. The laryngeals allophonically turned epenthetic would-be schwas into non-high peripheral vowels (cf. (66)). These would have been the vast majority of schwas in the language, thus removing most of the support for speakers positing schwas in their inventory.

2.5.4   Local Summary

This section has showed that the reduplicated aorists, both of the normal type and of the AR type, as well as AR presents (such as they are), can be generated in Ancient Greek by the synchronic reduplicative grammar developed for the perfect in Section 2.2. The AR presents and aorists have essentially the same source as the AR perfects, namely laryngeal-initial bases in Pre-Greek. The differences between them, namely the quality and quantity of their second vowel, result from differences in the underlying specification of their reduplicative vowel: /e/ in the perfect, /i/ in the present, and nothing in aorist. In order to make this claim consistent with the normal consonant-initial aorists, whose reduplication is surface identical to reduplicated perfects, it must be the case that their reduplicative vowel, both in Pre-Greek and Ancient Greek, arises through epenthesis. While there is little external motivation for these claims, they are nonetheless fully consistent with the rest of the reduplicative grammar, as shown in (86) below, which integrates the new vowel-related constraints employed in this section with the total ranking developed in Section 2.2 (cf. (42)).

---

[57] Thanks to Andrew Byrd for bringing this issue to my attention.

(86) **Updated total ranking for Ancient Greek reduplication**

IDENT-V-IO

IDENT-V-BR *[¬e]   REDUP(RED)*lex*

MAX-IO DEP-IO CONTIGUITY-IO *PCR   MAX-μ-IO

ANCHOR-L-BR *CC   IDENT[long]-C-BR *#Cː

IDENT[long]-C-IO

ONSET

ALIGN-/e/-L

MAX-BR REDUP(RED) UNIFORMITY-IO

ALIGN-ROOT-L
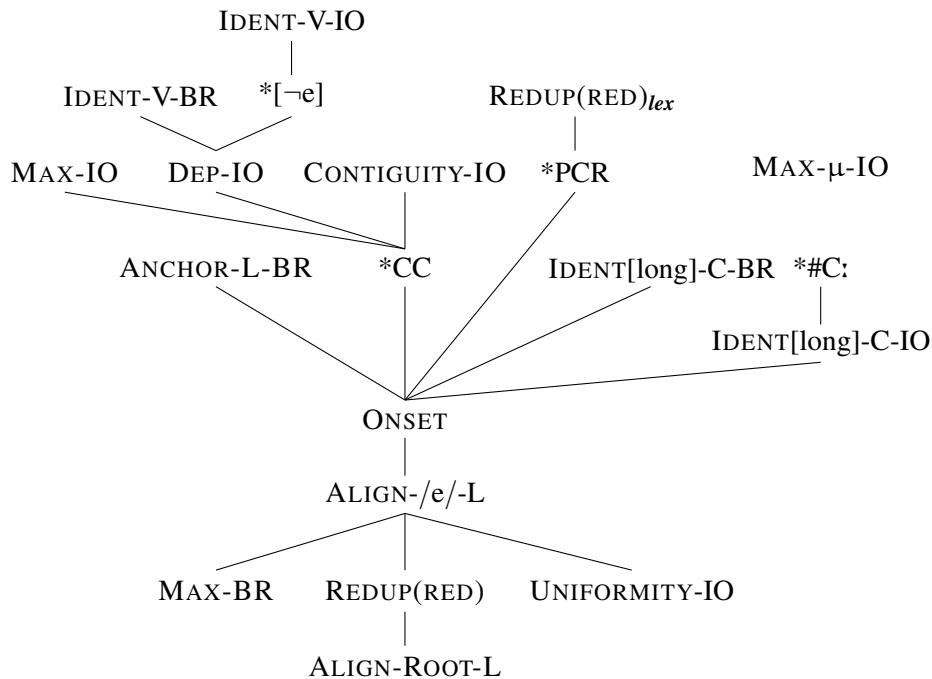
2.6 Conclusions

This chapter has provided a comprehensive account of the synchronic reduplicative system of Ancient Greek, and the historical development of the Attic Reduplication pattern. I demonstrated that the synchronic reduplicative system of Ancient Greek simultaneously generates the productive patterns displayed by consonant-initial roots and the productive vowel-lengthening pattern for vowel-initial roots. The Attic Reduplication pattern, and the previously unrecognized sub-regularity of the unexpectedly copying cluster-initial perfects associated with reduplicated presents, had to be accounted for through lexical indexation of a special constraint requiring copying in cases where the phonotactics or alignment would otherwise block it. The existence of the AR pattern alongside the productive vowel-lengthening pattern, and the way in which AR has to be encoded in the synchronic grammar, raised questions about its diachronic origin.

Based on the clear etymological connection between Attic Reduplication and the laryngeals, I argued that laryngeal-specific phonotactics operative in Pre-Greek spawned the precursor of Attic Reduplication (Pre-AR). Pre-AR was then shown to be consistent with the interaction between another laryngeal-specific phonotactic repair (laryngeal vocalization) and the general reduplicative grammar as still evidenced in attested Ancient Greek.

In an attempt to retain the pattern as faithfully as possible subsequent to the loss of the laryngeals (and thus the loss of the pattern's conditioning factors), speakers innovated a new constraint system based on lexical indexation. This was a natural result of the learning process given the inconsistency posed by the data regarding vowel-initial perfects at the stage immediately following the loss of the laryngeals. This same system can be used to derive indexation for the exceptionally copying perfects to roots with reduplicated presents. This demonstrates that both patterns are not simply frozen, archaic forms which have arbitrarily persisted in the language, but rather synchronically generable minority patterns which are subject to the normal demands of the grammar.

The same can be said of the reduplicated aorists, of both the normal consonant-initial type and the Attic Reduplication type. Given the proper analysis of epenthesis and the vowel system generally, aorist reduplication falls out from the same principles and rankings that generate the reduplicative distribution in the perfect, where reduplication remains fully productive.

This chapter illustrates how synchrony and diachrony can be used in tandem to help explain systematic irregularities. By constructing the synchronic grammar of Ancient Greek, we formalized the exceptionality of the Attic Reduplication pattern. Through consideration of historical reconstruction, we generated a clear hypothesis about why the irregularity should exist, namely the behavior of laryngeals. This suggested the possibility of integrating other known phonological processes of a similar time depth and scope, namely laryngeal vocalization, into a new synchronic account of the phenomenon at a distinct diachronic stage. In turn, consideration of how the output of this stage interacted with subsequent diachronic change allowed us to connect the exceptional behavior of Attic Reduplication roots to other, very different root types with similar exceptional behavior (the reduplicated presents and their exceptionally copying associated perfects), which was hitherto completely without principled explanation.

2.7 Appendix: Attic Reduplication Perfects

Table (87) lists the set of attested AR perfects, coupled with their root etymologies and their forms reconstructed for Pre-Greek based on the analysis developed in this chapter. Forms are drawn primarily from van de Laar (2000:59–320); see also Beekes (1969:116–120).

(87)   Attic Reduplication perfects

Root (Greek < *PIE)			Present Stem	Perfect Stem	Pre-Greek Reconstructed Perfect Stem[58]
#*h₁					
√eger	< *h₁ger	'wake'	egēr-	eg(r)ēgor-[59]	*h₁əg-e-h₁gor-
√ela(u)	< *h₁elh₂	'drive'	ela(u)-	elēla-	*h₁əl-e-h₁la-
√eleutʰ	< *h₁lewdʰ	'go, come'	—	elēl(o)utʰ-	*h₁əl-e-h₁l(o)udʰ-
√en(e)k	< *h₁nek	'bring'	—	enēnokʰ-	*h₁ən-e-h₁nokʰ-
√ered	< *h₁reyd	'cause to lean'	ereid-	erēreis-	*h₁ər-e-h₁reid/s-
√ereip	< *h₁reyp	'throw down'	ereip-	erērip-[60]	*h₁ər-e-h₁rip-
√eme	< *wemh₁	'vomit'	eme-	emēme-	—
#*h₂					
√ager	< *h₂ger	'gather'	agēr-	agāger-	*h₂əg-e-h₂ger-
√ako(u)	< *h₂kow(s)	'hear'	akou-	akāko-	*h₂ək-e-h₂kow(s)-
√ale	< *h₂elh₁	'grind'	ale-	alāle-s-	*h₂əl-e-h₂le-s-
√ar	< *h₂er	'join'	arar-	arār-	*h₂ər-e-h₂r-
√aro	< *h₂erh₃	'plow'	aro-	arāro-	*h₂ər-e-h₂r-o-
#*h₃					
√od	< *h₃ed	'smell'	ozd-	odɔ̄d-	*h₃əd-e-h₃d-
√ol	< *h₃elh₁	'destroy'	ol-	olɔ̄l-	*h₃əl-e-h₃l-
√om	< *h₃emh₃	'swear'	om-	omɔ̄mo-	*h₃əm-e-h₃m-o-
√op	< *h₃ekʷ	'see'	—	opɔ̄p-	*h₃əkʷ-e-h₃kʷ-
√or	< *h₃er	'incite'	or-	orɔ̄r(e)-	*h₃ər-e-h₃r(-e)-
√oreg	< *h₃reg	'stretch'	oreg-	orɔ̄reg-	*h₃ər-e-h₃reg-
√orug	< *h₃ru-gʰ	'dig'	orus-	orɔ̄rug-	*h₃ər-e-h₃ru-gʰ-
√or	< *(s)wer	'keep watch'	or-o-	orɔ̄r-	—

---

[58] Stem-final material may be anachronistic.

[59] The [r] in the reduplicant is secondary. Brent Vine (personal communication) has suggested that it is the result of hyper-corrective *r*-insertion, along the lines of the phenomenon discussed in Vine (2011).

[60] Beside this there is also *ererim-* with short [e] for long [ɛ̄].

Chapter 3

Anatolian

3.1 Introduction

The compilation and philological treatment of the reduplicated verbal forms of the Anatolian languages recently completed by Dempsey (2015) allows for the first time for a comprehensive reckoning of the phonological behavior of partial reduplication in Anatolian. I will argue here, following Yates & Zukoff (2016a,b) (see also Dempsey 2015), that the reduplicative distributions in Hittite and Luwian — the two Anatolian languages well-enough attested to make significant generalizations about, and Proto-Anatolian — their reconstructed common ancestor, are those shown in (1) below.[1]

(1) Anatolian partial reduplication patterns by base shape

Base Shape	a. Hittite	b. Luwian	c. Proto-Anatolian
CVX–	<u>CV</u>-CVX–	<u>CV</u>-CVX–	*<u>CV</u>-CVX–
TRVX–	<u>TRV</u>-TRVX–	<u>TV</u>-TRVX–	*<u>TV</u>-TRVX–
STVX–	*i*<u>STV</u>-STVX–	(<u>TV</u>-STVX–)	*<u>STV</u>-STVX–
VCX–	<u>VC</u>-VCX–	<u>VC</u>-VCX–	*does not exist yet*

Hittite (1a), unlike any of the other attested Indo-European languages, shows cluster-copying for all types of initial clusters (*TRVX–* → <u>*TRV*</u>*-TRVX–*, *STVX–* → *i*<u>*STV*</u>*-STVX–*), *modulo* the normal phonological treatment of word-initial *ST* clusters (namely, prothesis of [i]). Luwian (1b), on the other hand, shows the more expected Indo-European C_1-copying pattern for *obstruent-sonorant* bases (*TRVX–* → <u>*TV*</u>*-TRVX–*); it appears to attest C_2-copying for STVX– bases, but, as I will demonstrate below (Section 3.6.1), this is not a synchronically generated pattern, but rather a frozen relic which reflects the effects of regular sound change. Based on the diachronic treatment of Proto-Anatolian initial *ST* clusters (Section 3.6.1), consideration of parsimony in reconstruction with respect to TRVX– bases (Section 3.6.2), and both theory internal and diachronic considerations regarding the relative chronology of *h_1 loss and the advent of vowel-initial roots (Section 3.6.3),

---

[1] This chapter employs the following abbreviations for segment type: *C* = any consonant, *T* = obstruent, *R* = sonorant consonant, *S* = [s], *V* = vowel, *X* = optional string of additional segments.

I argue for the reconstruction of Proto-Anatolian provided in (1c). This reconstruction exhibits the typical Indo-European asymmetry in the treatment of bases with different types of initial clusters: C_1-copying for TRVX– bases, but cluster-copying for STVX– bases. This asymmetry is identical to Gothic (see Chapter 4). Just as in Gothic, this asymmetric behavior can be explained by the operation of the constraint *PCR, whose definition is repeated in (2).

(2)   **NO POORLY-CUED REPETITIONS (*PCR) [** $\approx$ ***$C_\alpha V C_\alpha$ / _C[-sonorant]]**
    For each sequence of repeated identical consonants separated by a vowel ($C_\alpha V C_\alpha$), assign a violation * if that sequence immediately precedes an obstruent.

While *PCR is thus essential in generating the cluster-dependent reduplicative patterns of *Proto*-Anatolian, it plays no role in the synchronic grammar of its daughter languages Hittite or Luwian, in which there is no evidence that different cluster types behave differently with respect to reduplicant shape. Furthermore, the *VC*-*VCX*– reduplication pattern of vowel-initial roots in each language directly *violates* *PCR. Thus, *PCR must be *inactive* in both languages; that is to say, *PCR must be situated at the very lowest stratum of the ranking in both languages. This obviously raises the following question: how did Anatolian go from a system where *PCR was active to one where it is completely inactive? In this chapter, I will argue that independent phonological changes in both Hittite and Luwian eliminated the distinction between *TRVX*– and *STVX*– roots in reduplication, and this caused learners to converge on a *PCR-free analysis.

The structure of this chapter is as follows. After a brief discussion in Section 3.2 of Hittite orthography and its impact on the interpretation of certain forms, I lay out the attested partial reduplication forms in Hittite and Luwian in Section 3.3. Section 3.4 details the synchronic analysis of the reduplicative system of Hittite. Section 3.5 briefly demonstrates that the same constraints and considerations employed for the analysis of Hittite can be deployed to analyze the minimally different reduplicative system of Luwian. In Section 3.6, I develop the arguments for my proposed reconstruction of the Proto-Anatolian reduplicative system, and then in Section 3.7 show again that the constraints and considerations deployed for Hittite and Luwian easily account for Proto-Anatolian, as well. Lastly, Section 3.8 develops a complete diachronic account of the transitions between Proto-Anatolian and Hittite and Luwian, respectively. Specifically, I argue that a minor reformulation of the Recursive Constraint Demotion algorithm (RCD; Tesar 1995, Tesar & Smolensky 1998, 2000), which favors the high ranking of maximally informative winner-preferring constraints, yields a satisfactory step-wise account of the relevant diachronic developments.

3.2 Hittite Orthography, Epenthesis, and Cluster Phonotactics

This chapter will in large part focus on the reduplicative behavior of roots and bases which begin in consonant clusters. However, in Hittite, the interpretation and understanding of consonant clusters — with respect to their orthography and their phonology, and, specifically, the relationship between the two — is an extremely vexed question. In this short section, I will briefly outline the nature of the writing system, and discuss how this clouds our understanding of Hittite cluster phonotactics, as well as the interpretation of particular reduplicated forms.

The Hittite orthographic system is a version of the Cuneiform syllabary originally borrowed from the Akkadians (which was itself borrowed originally from the Sumerians). As such, it is not terribly well-suited to writing a language with Hittite's phonological structure, resulting in a great many complications relating to the phonological interpretation of orthographic forms (see Hoffner

& Melchert 2008:§1.1–1.44 for general discussion). The specific problem that is most relevant to the issues taken up in this chapter is the writing of (inherited and synchronic) consonant clusters.

Hittite orthography possessed four types of syllabic signs (Hoffner & Melchert 2008:11): V, CV, VC, and CVC. If the language were to contain word-edge bi-consonantal clusters and/or word-medial tri-consonantal clusters — which Melchert (1994:110–111) and many others show convincingly that it must, and which undoubtedly were present in Proto-Indo-European — then this orthographic system would be unable to represent them directly with the signs available to it. In many types of inherited consonant sequences (i.e., those consonant sequences guaranteed by extra-Anatolian comparative evidence to have existed at some point in the language's history), we find "non-etymological" vowels (Kavitskaya 2001) appearing in Hittite orthography. *A priori*, given the insufficiencies of the writing system, any particular non-etymological vowel could be assigned either of two interpretations: (i) "empty" vowels (Kavitskaya 2001, Hoffner & Melchert 2008:12–13), which appear in the orthography as a conventional means of writing phonological consonant sequences; or (ii) phonologically real epenthetic vowels which have been inserted for phonotactic reasons. (See Melchert 1994:108–109, 2016, Kavitskaya 2001, Yates 2014, 2016, among many others, for work on the phonotactic conditioning of epenthesis in Hittite.)

Kavitskaya's (2001) investigation of orthographic practice relative to different types of inherited consonant clusters makes it clear that both types of non-etymological vowels exist, and that they can be teased apart (though in many cases this is quite a difficult task). All cases of inherited word-initial *obstruent*+{*liquid/nasal*} (abbreviated TR) clusters[2] (except *tr*; cf. Melchert 1994, 2016) are written with either (i) an initial *Ca* sign, as shown in (3) below, or, less frequently, (ii) vacillation between an initial *Ca* sign and an initial *Ci* and/or *Ce* sign, as shown in (4) below.[3] (The same orthographic distribution holds of inherited word-internal sequences involving *C*+*obstruent*+*sonorant* and *sonorant*+*obstruent*+*C* sequences.) Kavitskaya (2001:276) further asserts that "there are no contemporary alternations between *Ce* and *Ci* without the existing alternates written with the help of *Ca* (Melchert, p.c.), which provides additional evidence that these vowels are merely orthographic."

(3)   Consistent writing of *<Ca->* in inherited initial TR clusters

    a.   ***<pa*-*ra*-*a*-*i>*** 'blow' (3SG.PRES.ACT)
        = [praː-i]   ← /pra-/   < *pra-*   (Kavitskaya 2001:274, ex. 9a)

    b.   ***<ka*-*ra*-*a*-*wa*-*ar>*** 'horn'
        = [kraː-war]   ← /kra-/   < *kra-*   (Kavitskaya 2001:274, ex. 9b)

    c.   *<ša*-*ma*-*an*-*ku*-*úr*-*wa*-*an*-*te*-*eš>* 'bearded' (NOM.PL)
        = [smankurwantes] ← /smankur-/ < *smo(n)ḱ-*   (Kloekhorst 2008:1029)

(4)   Vacillation between *<Ca->* and *<Ci->* (Kavitskaya 2001:275, exx. 12, 13)

    a.   ***<ga*-*ra*-*pV->*** ∼ ***<gi*-*ra*-*pV->*** 'devour'   = [grapV-]

    b.   ***<ga*-*ri*-*it>***   ∼ ***<gi*-*ri*-*it>***   'flood'   = [grit]

    c.   *<ša*-*me*-*en->* ∼ *<še*-*me*-*en->* 'cause to disappear' = [smen-]

In contrast to inherited initial TR clusters, inherited initial *s*+*obstruent* (ST) clusters show a different, consistent treatment: [STV-] sequences are written as ***<iS*-*TV->***.[4] This consistent behavior,

---

[2] On the writing of *obstruent*+*glide* with an empty homorganic vowel, see Kavitskaya (2001:292–294).

[3] The treatment of initial *sm-* appears to vary between faithful retention (as in (4c)) and internal epenthesis with [u]; cf. Melchert (2016:188–189).

[4] Initial *sp-* is a partial exception: it appears to attest three different outcomes (Melchert 2016). First, initial *sp-* sometimes shows the expected prothesis with *i* [isp-]. Second, initial *sp-* sometimes shows retention [sp-], as indicated

which is mirrored in word-internal position (Kavitskaya 2001), clearly indicates that this orthographic practice encodes something different than the empty *Ca* signs for TR clusters: namely, phonologically real epenthesis (see also Yates 2014, 2016). (The operation of word-initial epenthesis, i.e. prothesis, before ST clusters will be important for the analysis of Hittite's reduplicative behavior for STVX– bases; see Section 3.4.3 below). The difference in consistency between the two types of orthographic non-etymological vowels also strongly indicates that the *a* (and its vacillations) which breaks TR clusters is not to be analyzed as epenthetic, but rather purely orthographic.

We thus see that the orthographic sequence $<T_1a$-R_2V_3-$>$ has the potential to encode a phonological sequence $[T_1R_2V_3]$. Yet this writing also uncontroversially has the potential to represent actual $[T_1aR_2V_3]$, with phonologically real [a], as such sequences also freely exist in the language, and would naturally be written as such. When we have strong etymological evidence in favor of positing the presence or absence of that *a* vowel, we can say with some degree of certainty which phonological sequence the orthography is representing. However, this type of reasoning is not applicable in assessing reduplicated forms with such sequences, because "etymological" evidence of the relevant sort does not exist; that is to say, even if we had strong comparative evidence regarding the expected shape of the reduplicant, we cannot be sure that it still holds of the synchronic system of Hittite, in which reduplication is productive and thus potentially divergent from previous stages.

This problem does indeed come to fruition. In Section 3.3.3 and throughout, I will claim that Hittite reduplicates the entire cluster when the base begins in a TR cluster. The reduplicant is word-initial, so it can only be written *<Ta-RV->*. As just discussed, such an orthographic sequence could, *a priori*, represent either $[T_1aR_2V_3]$ or $[T_1R_2V_3]$. For example, the form which I interpret as [χli-χlai] 'kneel' is spelled *<ḫa-li-iḫ-la-i>* (Dempsey 2015:319), and thus has the potential to represent [χ<u>a</u>li-χlai] instead. This possibility is particularly problematic because, beside the partial reduplication pattern which is the subject of this chapter, Hittite also possesses the so-called "intensive" reduplication pattern, which frequently takes the shape [<u>CaC</u>-i-CaC-] (e.g. *wariwar* 'burn up'). The non-cluster interpretation of *<ḫa-li-iḫ-la-i>* would strongly resemble the intensive type. Thus, beyond casting doubt on the phonological composition of the reduplicant in a written form like *<ḫa-li-iḫ-la-i>*, the orthography is further casting doubt on which broad reduplicative type such a form represents to begin with. We are thus faced with a very serious ambiguity for the purposes of the current enterprise, and so we must proceed with caution as we consider these forms. Nevertheless, I will proceed under the assumption that we do have evidence for this type of cluster-copying reduplication in Hittite.

3.3   Hittite and Luwian Reduplication Data

This section attempts to provide a near exhaustive list of forms from the Anatolian languages displaying partial reduplication, assembled from Dempsey (2015).[5] The Anatolian languages seem to attest two distinct types of reduplicated verbal forms. Dempsey (2015:331) concludes that partial reduplication, the type we will be concerned with here, generally indicates imperfective/pluractional morphosemantics. On the other hand, Anatolian also exhibits a type indicating "expressive" or "intensive" morphosemantics (Dempsey 2015:332). These forms more closely resemble total redu-

---

by orthographic vacillation like that in (4). For example, the word for 'pin' attests the spellings *<še-pí-ik-ku-uš-ta->*, *<ši-pí-ik-ku-uš-ta->*, and *<ša-pí-ik-ku-uš-ta->*); this must be interpreted as [spikusta-]. Lastly, initial **sp-* sometimes shows internal epenthesis with *u* [sup-]. Melchert (2016) connects this variation to similar treatment of initial **sm-*.

[5]   A few "un-paired" stems, i.e. likely reduplicated forms without independently identifiable roots, have been omitted (consult especially Dempsey 2015:Ch. 4). Such forms do not generally seem to display any divergent characteristics, with the exception of several forms which will be referenced at appropriate points in the proceeding discussion.

plication of the root; for example, Hittite *wariwar-* [war-i-war-] "burn up". I will not examine the properties of this type here. However, it must be noted that, given the prototypical monosyllabicity of roots in Anatolian (as in Indo-European generally), and the ambiguities of the writing system discussed above, in a few cases (particularly for TRVX– bases; see Section 3.3.3 immediately below), it is not entirely clear whether a reduplicated form ought to be identified as the partial reduplication type or the total reduplication type.

With respect to the phonological properties of partial reduplication in Anatolian, just as in the other languages examined in this dissertation, there are four relevant categories to consider based on the phonological shape of the base:[6] single-consonant–initial bases *CVX–* (Section 3.3.1), vowel-initial bases *VCX–* (Section 3.3.2), obstruent-sonorant–initial bases *TRVX–* (Section 3.3.3), and *s*-obstruent–initial bases *STVX–* (Section 3.3.4).[7] All data discussed here, and most of the data discussed throughout this chapter, is drawn from the compilation of reduplicated verbal forms in Dempsey (2015). The data for the relatively more poorly attested base types (*VCX–*, *TRVX–*, and *STVX–*) are annotated with the source of their entries in Dempsey (2015) (abbreviated D) and, where appropriate, Kloekhorst (2008) (abbreviated K). (Language name abbreviations are as follows: Hitt. = Hittite, CLuw. = Cuneiform Luwian, HLuw. = Hieroglyphic Luwian, Lyc. = Lycian.)

3.3.1 CVX– Data

In both Hittite and Luwian (and also Lycian, though the evidence is quite limited), whenever the base begins in a C_1V sequence, the reduplicant takes the shape C_1V. The table in (5) on the following page lists all such reduplicated forms that stand beside an independently occurring verbal base in the language ("paired stems"). The table in (6) on the following page shows other attested reduplicated forms whose presumed verbal base is attested elsewhere in Anatolian. As might be expected, CVX– bases constitute the vast majority of the available data on partial reduplication in Anatolian.

It can be seen from this data that there are several different sub-patterns for the reduplicative vowel. While most forms show identical copy vocalism (e.g. Hitt. *kikkiš–*, *mammalt–*; virtually all of the Luwian forms), there seem to be at least a few forms with fixed [e] (e.g. Hitt. *wewakk–*) or fixed [i] (e.g. Hitt. *wiw(a)i(ške–)*, CLuw. *lilūwa–*). Given that [e] and [i] are the most common fixed vocalisms found in the other Indo-European languages (especially Greek; see Chapter 2), it is likely that these represent archaisms, and that copy vocalism should be identified as the productive pattern for Hittite and Luwian. In any event, as will be shown in the analysis of TRVX– bases in Section 3.4.2 below, the reduplicative vowel in Hittite *must* be standing in correspondence with the base vowel (at least for cluster-initial bases), whether or not it is an identical copy. The discussion in this chapter will be focused on reduplicant shape, so I leave fuller investigation of reduplicant vocalism in Anatolian as a question for future inquiry. It can also be observed that some root-initial consonants show gemination[8] in reduplication. This distribution does not appear to admit to any obvious synchronic explanation, so I leave this as a question for future inquiry, as well.

---

[6] Throughout this chapter, I will consistently speak of the *base* (by which I mean the base of reduplication within a surface reduplicated form), rather than the *root*, as cluster-initial reduplicative bases, which are of primary interest here, often result from syncope of a root vowel.

[7] There is, additionally, one isolated case of a *TTVX–* base and one isolated case of a *RTVX–* base in Hittite. It is doubtful that these forms should be treated as productively generated forms within synchronic Hittite, so discussion of these forms will be postponed until Section 3.8.7.

[8] There is undoubtedly a contrast between two stop series in Anatolian in word-medial position (Melchert 1994:14); however, there is not a consensus on what precise phonological/phonetic feature(s) actually distinguished between the two series (Hoffner & Melchert 2008:35). I follow one common practice in transcribing this distinction as one of single vs. geminate stops. Other proposals include a distinction in voicing or aspiration, or some sort of "fortis" vs. "lenis" distinction. This question is orthogonal to those taken up in this chapter.

(5) Reduplication with CVX– bases (paired stems)

	Gloss	Base	Reduplicated stem	
Hitt.	'happen'	*kiš–*	*kikkiš–*	[<u>ki</u>-kːis-]
	'cut'	*kuwarške–*	*kuwakuwarške–*	[<u>kʷa</u>-kʷar-]
	'bend'	*lak–*	*lelakk–*	[<u>le</u>-lakː-]
	'chant'	*mald–*	*mammalt–*	[<u>ma</u>-malt-]
	'fall'	*mau(šš)–*	*mummiye–*	[<u>mu</u>-mː-]
	'shoot'	*šiye–*	*šišiye–*	[<u>si</u>-si-]
	'place'	*d(a)i–*	*titti–*	[<u>ti</u>-tːi-]
	'step'	*tiya–*	*titti–*	[<u>ti</u>-tːi-]
	'cry out'	*wai–*	*wiw(a)i(ške–)*	[<u>wi</u>-w(a)i-]
	'wipe'	*warš–*	*wawarš–*	[<u>wa</u>-warsː-]
	'demand'	*wēk–*	*wewakk–*	[<u>we</u>-wakː-]
CLuw.	'run'	*ḫuiya–*	*ḫuihuiya–*	[<u>χʷi</u>-χʷi-]
	'take'	*la–*	*lala–*	[<u>la</u>-la-]
	'pour'	*lūwa–*	*lilūwa–*	[<u>li</u>-luː-]
	'give'	*pī(ya)–*	*pipišša–*	[<u>pi</u>-pi-]
	'break'	*malḫu–/*	*mammalḫu–/*	[<u>ma</u>-mːalχʷ-] /
		malwa–	*mammalwa–*	[<u>ma</u>-mːalwa-]
	'strike'	*dūp(a)i–*	*dūdupa–*	[<u>tu</u>-tupa-]
HLuw.	'exalt'	*sarla–*	*sasarla–*	[<u>sa</u>-sarla-]
	'release'	*sa–*	*sasa–*	[<u>sa</u>-sa-]
	'fill'	*su(wa)–*	*susu–*	[<u>su</u>-su-]
	'stand'	*ta–*	*tata–*	[<u>ta</u>-ta-]
Lyc.	'give'	*pije–*	*pibije–*	[<u>pi</u>-βi-]

(6) Paired CVX– stems reconstructible for PA by Anatolian-internal comparison

Hitt. *lipp–* 'lick'   : Luw(o-Hitt). *lilipa(i)–*

Hitt. *pašš–* 'swallow' : Luw(o-Hitt.) *papašša–*

Hitt. *nai–* 'turn'   : CLuw. *nana–*

Hitt. *tar–* 'say'   : CLuw. *tatariya–* 'curse' /

                           HLuw. *ta-ta-ra/i-ya–* [tatar(i)ya–] 'id.'

3.3.2 VCX– Data

For each of the remaining base types, the data is much more scarce. Vowel-initial bases in both Hittite and Luwian show a _VC_-_VCX_– reduplicative pattern, as shown by the forms in (7).

(7)   Reduplication with VCX– bases

	Gloss	Base	Reduplicated stem		
Hitt.	'mount'	_ark–_	_ararkiške–_	[<u>ar</u>-ark-]	(D:58–60, 260)
	'seat'	_ēs–_	_asāš–_	[<u>as</u>-aːs-][9]	(D:61–65, 282–284; K:218–219)
CLuw.	'wash'	_īlḫa–_	_ililḫa–_	[<u>il</u>-ilχa-]	(D:218–219, 263)

3.3.3 TRVX– Data

For _obstruent + sonorant_–initial bases (TRVX–), Hittite has a cluster-copying pattern _TRV_-_TRVX_–, while Luwian exhibits the standard Indo-European C_1-copying pattern _TV_-_TRVX_–. The attested forms are shown in (8) below. If analyzable as reduplicated forms, the Lycian verbs _pabra-_ and _pabla-_ (both of unknown meaning; cf. Dempsey 2015:273–274), would be consistent with the Luwian and Indo-European pattern of _TV_-_TRVX_– reduplication.

(8)   Reduplication with TRVX– bases

	Gloss	Base	Reduplicated stem		
Hitt.	'blow'	_par(a)i–_	_parip(p)ar(a)i–_	[<u>pri</u>-pːr(a)i-]	(D:121–126, 275–276; K:631–632)
	'kneel'	_ḫal(a)i–_	_ḫaliḫal(a)i–_	[χ<u>li</u>-χl(a)i-]	(D:66–71, 319–320; K:273–274)
CLuw.	'carry off'	_par(a)–_	_papra–_	[<u>pa</u>-pra-]	(D:230, 272–273)

 Due to orthographic limitations (see Section 3.2 above), the Hittite forms are ambiguous: they can reflect either partial reduplication or "intensive" reduplication (e.g. Hitt. _wariwar_ 'burn up', from √_war_ 'burn'). Hittite orthography did not contain any CC(V) signs, and thus had no way to directly represent word-initial clusters. Hittite therefore wrote word-initial $[C_1C_2V_3]$ sequences as $<C_1V$-C_2V_3->. The choice between "intensive" reduplication and partial reduplication thus comes down to whether the first _a_ is a phonologically real vowel, pointing towards the intensive type, or rather just a dummy vowel present for orthographic purposes, pointing towards the partial reduplication type. Dempsey (2015:275–276, 319–320) argues for "intensive" reduplication in 'kneel' but partial reduplication in 'blow', while Kloekhorst (2008:273–274, 631–632) argues for partial reduplication in both. As long as at least one of the Hittite forms is properly partial reduplication, then the generalization still holds. This is what I will assume in this chapter.[10]

---

[9] This form is included for completeness, as it seems to synchronically mirror the other vowel-initial forms. However, it seems quite likely that it has a somewhat different history than the others; see Dempsey (2015:282–284) and references therein.

[10] Following Dempsey (2015:329), I assume that the verbal stems _tatraḫḫ–_ 'incite' and _papraḫḫ–/papre(šš)–_ 'make/be(come) impure' are deadjectival in Hittite, rather than synchronic partial verbal reduplication to unattested verbal bases. These forms thus do not constitute evidence of the synchronic reduplicative system of Hittite. Nevertheless, their adjectival bases _tatrant–_ 'sharp-edged; prone to goring' and _paprant–_ 'impure' may preserve traces of an earlier _TV_-_TRVX_– C_1-copying pattern; see Section 3.6.2 below.

3.3.4 STVX– Data

For inherited *s* + *obstruent*–initial roots (PIE/Proto-Anatolian *STVX*–), Hittite and Luwian again diverge, as shown in (9). Hittite shows copying of the full cluster (as in *TRVX*– bases), plus a prothetic [i]. The initial [i] must be epenthetic, and outside of the reduplicant proper; if the root were underlyingly vowel-initial, we would expect the copy pattern for *VCX*– roots, yielding ****_is-istu_**–, contrary to fact.[11] Luwian synchronically lacks *STVX*– bases; the Luwian reduplicated forms in (9) are relics of the Proto-Anatolian *STV-STVX*– pattern, with deletion of **s* in Pre-Luwian according to regular sound change (cf. Section 3.6.1 below).

(9)   Reduplication with STVX– bases

	Gloss	Root/Base	Reduplicated stem
Hitt.	'become evident'	√*stu* ([istu–])	*išdušdu-ške*– [i<u>stu</u>-stu-] (D:71–73, 263; K:419)
CLuw.	'become evident'	PA *stu*–	*dušdu-ma/i*– [<u>tu</u>-stu-] (K:419–420)
	'bind'	PA *sh₂(o)i*–	*ḫišḫi(ya)*– [χ<u>i</u>-sχi-] (D:212–213, 261–262)

There are two additional Hittite forms that deserve to be mentioned here: (i) Hittite *sisd*– 'prosper' (Dempsey 2015:204–205, 301–302), and (ii) Hittite *šišḫa*– 'decide, appoint' (Dempsey 2015:206–207, 303–304). If we were to view these as productively generated reduplicated forms within synchronic Hittite, they would clearly run counter to the generalization presented above, showing C_1-copying rather than cluster-copying. Dempsey (2015:301–304), however, argues on both semantic and formal grounds that neither of these should be viewed as having been synchronically analyzed by speakers as reduplicated. I follow Dempsey on this point. If these forms are rightly to be analyzed as reflecting reduplicated forms of some earlier stage, they could be seen either as equivalent to Hittite *titḫa*– and *lilḫuwai*– in reflecting a late syncope process (see Section 3.8.7), or as actually bearing witness to the situation in Proto-Indo-European (see Chapter 7).

3.4 Synchronic Analysis of Hittite Copying Patterns

With the data established, we can now move on to the analysis. The reduplicative patterns are all generable using essentially the same set of constraints that have been used to explain reduplication in the other languages examined in this dissertation thus far. I will start with the analysis of Hittite.

3.4.1 CVX– Bases in Hittite

First let us consider the most basic type, the <u>CV</u>- copying pattern to CVX– bases. Most of the interesting analytical points arise only in the other types of bases, but this pattern can serve to illustrate one noteworthy point, namely that post-nuclear segments (e.g. C_2 and V_2 in a $C_1V_1C_2V_2$– base) are not generally copied; the exception will be with vowel-initial roots, where exactly one post-nuclear consonant does get copied — see Section 3.4.4 below. This fact can be derived using a size minimizer constraint (see Spaelti 1997, Hendricks 1999, among others) that prefers smaller reduplicants. In the analysis of reduplication in Ancient Greek in Chapter 2, we required two separate size minimizer constraints. One was *CC (defined in (10a)), a markedness constraint disfavoring consonant clusters. The other was ALIGN-/e/-L, an alignment constraint dictating that the

---

[11] Kloekhorst (2008:419) makes a similar observation.

independent reduplicative vowel morpheme /e/ be as far to the left as possible. Since, as alluded to above, Hittite requires that the reduplicative vowel be identified as a (potentially imperfect) copy of the base vowel (see Section 3.4.2 below), rather than an independent morpheme, if we are to use an alignment constraint as a size minimizer, it must be instead an alignment constraint on the root — namely, ALIGN-ROOT-L, as defined in (10b).

(10)   Size minimizer constraints

        a.   ***CC**
             Assign one violation mark * for each consonant cluster.

        b.   **ALIGN-ROOT-L**
             Assign one violation mark * for each segment which intervenes between the left edge of the word and the left edge of the root

For roots/bases of the shape *CVC(C)–*, these two constraints will have the equivalent effect, as shown in (12) for the root √*wars-* 'wipe'. Copying one post-nuclear consonant (candidate (12b)) or both post-nuclear consonants (candidate (12c)) adds gratuitous violations both of *CC — since there are now more consonant clusters than necessary, and of ALIGN-ROOT-L — since there are now more segments preceding the root than necessary.[12] *CC and ALIGN-ROOT-L thus advocate for a smaller reduplicant. At least one of these constraints must outrank the constraint which prefers fuller copying, MAX-BR, as defined in (11).

(11)   **MAX-BR**
        Assign one violation mark * for each segment in the base without a correspondent in the reduplicant.

(12)   CVX– bases: *warš-* 'wipe' → *wa-warš-*

/RED, wars-/	*CC	ALIGN-ROOT-L	MAX-BR
a. ☞ wa-wars-	*	**	**
b. war-wars-	**!	***!	*
c. wars-wars-	**!*	***!*	

In Ancient Greek, due to more complex interactions dictated by a wider variety of reduplicative patterns, both *CC and an alignment constraint were required, and were situated at different spots in the ranking. In Hittite, the relevant complications are not present, and one size minimizer will be sufficient. When we turn to *CVCV–* bases in (13),[13] we find that only ALIGN-ROOT-L is sufficient to fully rule out over-copying. When copying an additional *CV* sequence is possible (candidate (13c)), this will not create additional clusters, and thus cannot be ruled out by *CC. It can, however, be ruled out by ALIGN-ROOT-L, since that constraint is not sensitive to segment type. This demonstrates that ALIGN-ROOT-L must dominate MAX-BR, as given in (14).

Since this ranking is sufficient to rule out all types of over-copying, we have no evidence for the ranking of *CC. *CC will continue to play no distinct role in the analysis, so it will be omitted from further discussion.

---

[12] Both constraints would be fully satisfied if there were no material in the reduplicant. I assume that copying is necessitated by a constraint like REALIZE MORPHEME(RED), as employed for Greek in Chapter 2.

[13] There do not appear to be any actual reduplicating roots in Hittite that are polysyllabic; however, the CVCV shape does very clearly present itself when we include suffixes, either derivational or inflectional, in the potential domain of copying. In any case, we must account for the behavior of such a root shape given richness of the base.

(13)   CVCV– bases: (hypothetical) *wari-* → *wa-wari-*

/RED, wari-/		*CC	ALIGN-ROOT-L	MAX-BR
a. ☞ wa-wari-			**	**
b. war-wari-		*!	***!	*
c. wari-wari-			***!*	

(14)   **Hittite Ranking:** ALIGN-ROOT-L ≫ MAX-BR

3.4.2   TRVX– Bases in Hittite

Bases with initial *obstruent + sonorant* (TR) clusters copy both base-initial consonants, plus the reduplicative vowel; for example, *prai-* 'blow' → *pri-prai-*. The constraint which is crucial in preferring the cluster-copying candidate *pri-prai-* (18a) over a cluster-reducing CV reduplicant candidate *pi-prai-* (18b) is CONTIGUITY-BR, defined in (15). In order to select the cluster-copying candidate, CONTIGUITY-BR must dominate ALIGN-ROOT-L, as cluster-copying introduces additional segments intervening between the root and the left edge of the word.

(15)   **CONTIGUITY-BR**
         Assign one violation mark * if two segments which are contiguous in the base have correspondents in the reduplicant that are not contiguous.

To make use of this constraint for the current example, we must assume that the reduplicative vowel corresponds either to the entire diphthong of the base [ai], or the first base vowel [a], such that the base correspondent of the reduplicative vowel is contiguous with the base-second consonant. Alternatively, we could assume that the reduplicant is calculated relative to the weak stem, which regularly exhibits zero-grade ablaut for this type of verb (Craig Melchert, personal communication), i.e. *pri-pri-*. Either way, this problem does not arise with monophthongal nuclei, as in the *STVX–* example below. Nevertheless, since this constraint is crucial in the analysis, we must assume that in all cases, whether or not the vowel of the reduplicant is identical to the base vowel, that the vowel of the reduplicant stands in BR-correspondence with the vowel of the base.

A third candidate considered here is *ri-prai-* (18c), one which copies just the second root consonant rather than the first. This avoids creating a reduplicant cluster, and thus improves satisfaction of ALIGN-ROOT-L, without violating CONTIGUITY-BR. However, it does so at the expense of another constraint, ANCHOR-L-BR (defined in (16)), which is undominated in Anatolian. With the ranking shown in the (17) below, we select the cluster-copying candidate, as demonstrated in (18).

(16)   **ANCHOR-L-BR**
         Assign one violation mark * if the leftmost segment of the reduplicant does not correspond to the leftmost segment of the base.

(17)   **Hittite Ranking:** ANCHOR-L-BR, CONTIGUITY-BR ≫ ALIGN-ROOT-L

(18)　　TRVX– bases: *prai-* 'blow' → *pri-prai–*

/RED, prai-/	ANCHOR-L-BR	CONTIGUITY-BR	ALIGN-ROOT-L
a. ☞ <u>pri</u>-prai-			***
b. <u>pi</u>-prai-		*!	**
c. <u>ri</u>-prai-	*!		**

This pattern is quite noteworthy from the Indo-European perspective. It appears to be the only Indo-European language that attests reduplication to TR-initial bases but does not maintain the C_1-copying pattern for them. (Latin lacks reduplicated forms to TR-initial bases entirely; see Cser 2009.) While the motivations for this change remain largely mysterious (though see Section 3.6.2 for some speculation on the matter), the ramifications of this change — vis-à-vis constraint re-ranking and *PCR — will have serious implications for the reduplication patterns of other base types, namely, vowel-initial bases.

3.4.3　STVX– Bases in Hittite

Bases beginning in *s-obstruent* clusters (*#ST*) comprise a special case of cluster-initial bases, though not a distinct copying pattern in the same way as Proto-Anatolian or the other languages encountered previously. In Hittite, they display the same cluster-copying pattern as do TRVX– bases, but they additionally display prothesis of [i] to the (word-initial) reduplicant cluster: simplex *istu–* 'become evident' → reduplicated *istu-stu–*. This complication follows directly from the independent process of prothesis to initial ST clusters. Prothesis must still be synchronically active, as this root does not behave like vowel-initial roots, i.e. ****is-istu–*.[14] Therefore, such surface forms must have input-output mappings like /stu–/ → [istu–], where the root is stored with an initial ST-cluster, and the prothetic vowel arises in the course of the derivation.

　　To generate prothesis, the constraint militating against initial ST clusters (*#ST, (19a)) must outrank the constraint militating against epenthesis (DEPV-IO, (19b)). This is illustrated in (21). DEPV-IO must also be dominated by other faithfulness constraints whose violation could repair an initial ST cluster, for example MAXC-IO (19c), which penalizes consonant deletion.[15]

(19)　　Constraints involved in epenthesis
　　a.　***#ST**
　　　　Assign one violation mark * for each word-initial ST cluster.[16]
　　b.　**DEPV-IO**
　　　　Assign one violation mark * for each output vowel without an input correspondent.
　　c.　**MAXC-IO**
　　　　Assign one violation mark * for each input consonant without an output correspondent.

(20)　　**Hittite Ranking:** *#ST, MAXC-IO ≫ DEPV-IO ≫ *CC

---

[14] Additional evidence for the synchronic status of prothesis in Hittite comes from stress assignment; see Yates (2017).

[15] Both DEPV-IO and MAXC-IO probably dominate *CC, as clusters are generally permitted, though this matter is complicated by the independently necessary activity of CONTIGUITY-IO.

[16] It is probably appropriate to view this as a more general phonotactic restriction that requires [s] to surface adjacent to a vowel, as medial sequences of *CsC* (when *s* is not part of the affricate [t͡s]) are also actively avoided (see Kavitskaya 2001, Yates 2014, 2016). This extension would have no effect on the analysis presented here, so I employ the more limited formulation here.

(21) Epenthesis to ST roots: /stu-/ → [istu-]

/stu-/		*#ST	MAXC-IO	DEPV-IO
a.	stu-	*!		
b. ☞	istu-			*
c.	su-/tu-		*!	

Epenthesis could solve the *#ST problem in one of two ways: (i) prothesis (i.e., external epenthesis), as is attested; or (ii) internal epenthesis, as in an unattested output like **situ. To select prothesis over internal epenthesis, we need to consider the ranking of *#ST and two additional constraints: CONTIGUITY-IO and ONSET.[17]

(22) Constraints for epenthesis site

    a. **CONTIGUITY-IO**
    Assign one violation mark * for each pair of segments which are adjacent in the input that have non-adjacent correspondents in the output.

    b. **ONSET**
    Assign one violation mark * for each onsetless syllable.

CONTIGUITY-IO prefers external epenthesis, as internal epenthesis would disrupt adjacency relationships. ONSET, on the other hand, prefers internal epenthesis, because external epenthesis would create a word-initial (and thus syllable-initial) vowel. Furthermore, ALIGN-ROOT-L actually prefers internal epenthesis, as well, since external epenthesis introduces a new segment between the root and the left edge of the word. Given that external epenthesis is selected, we know that CONTIGUITY-IO must dominate both ONSET and ALIGN-ROOT-L, and also that *#ST dominates ONSET and ALIGN-ROOT-L, or else epenthesis would not be harmonically improving. This ranking is shown in (23) and illustrated in (24) below.

(23) **Hittite Ranking:** *#ST, CONTIGUITY-IO ≫ ONSET, ALIGN-ROOT-L

(24) Epenthesis site in ST roots: /stu-/ → [istu-]

/stu-/		*#ST	CONTIGUITY-IO	ONSET	ALIGN-ROOT-L
a.	stu-	*!			
b. ☞	istu-			*	*
c.	situ-		*!		

Since ST roots are indeed stored underlyingly with an initial ST cluster, these roots can act just like other cluster-initial roots (i.e. TRVX–), subject to the additional condition of prothesis, which applies now to the reduplicant cluster rather than the base cluster. This results from combining the constraints and rankings already established independently for TRVX– roots in reduplication (cf. (18)), and STVX– roots in isolation (cf. (21) and (24)). This is demonstrated in (25).

---

[17] As argued by Yun (2016) and others, the choice of cluster-external epenthesis for ST clusters is typologically fully expected, and may better be viewed as deriving from perceptual conditions rather than the more abstract phonological conditions employed here. Adopting Yun's (2016) approach in full would not have any effect on the current analysis, so I employ this more basic strategy, which will be more familiar to readers.

(25)   Reduplication of STVX– bases: /stu-/ → [i<u>stu</u>-stu-]

/RED, stu-/		*#ST	ANCHOR-L-BR	CNTG-BR	DEPV-IO	ALIGN-ROOT-L
a.	<u>stu</u>-stu-	*!				***
b. ☞	*i*<u>stu</u>-stu-				*	****
c.	<u>su</u>-stu-			*!		**
d.	<u>tu</u>-stu-		*!			**

The simple cluster-copying candidate (25a), which is equivalent to the pattern exhibited by TRVX– bases, is impossible here, because that would result in an initial ST cluster (prohibited by undominated *#ST). Copying a non–base-initial consonant (25d) is again suboptimal, as it violates ANCHOR-L-BR. Lastly, the candidate which just copies the root-initial *s* (25c) is not permitted because it violates CONTIGUITY-BR, as the *s* and the vowel are not adjacent in the base while they are in the reduplicant. (It does also violate *PCR, but this constraint must be lower ranked; see below.) This leaves (25b), which is equivalent to the cluster-copying candidate (25a), except that it additionally has prothesis before the reduplicant. As long as DEPV-IO and ALIGN-ROOT-L (and indeed ONSET, which is also violated in the winning candidate due to its prothetic vowel) are dominated by the three highest-ranked constraints, this candidate remains optimal.

(26) **Hittite Ranking:**
        *#ST, ANCHOR-L-BR, CONTIGUITY-BR ≫ DEPV-IO, ONSET, ALIGN-ROOT-L

It is crucial that the epenthetic *i* does not belong to the reduplicant proper, as this would lead to a fatal ANCHOR-L-BR violation. This shows that the constraint which militates for the reduplicant to be at the left edge of the word, ALIGN-RED-L (defined in (27)), is violable, and indeed violated in service of prothesis. (ALIGN-RED-L must outrank ALIGN-ROOT-L in order to ensure that the reduplicant precedes rather than follows the root.)

(27) **ALIGN-RED-L**
        Assign one violation mark * for each segment that intervenes between the left of the reduplicant and the left edge of the word.

The admission of ALIGN-RED-L into the grammar does not make any perverse predictions, as we already know that CONTIGUITY-IO is active in the grammar, and this constraint will forestall any candidates with an infixal reduplicant (e.g. *TRVX– → T-RV-RVX–*), even if there were some potential phonotactic benefit to doing so. (An infixal candidate for STVX– bases without further modification, i.e. *S-TV-TVX–*, would be eliminated by *#ST.)

(28)   Reduplication of STVX– bases: /stu-/ → [i<u>stu</u>-stu-]

/RED, stu-/		*#ST	ANCHOR-L-BR	CONTIGUITY-IO	ALIGN-RED-L
a.	<u>stu</u>-stu-	*!			
b. ☞	*i*<u>stu</u>-stu-				*
c.	*i*<u>stu</u>-stu-		*!		
d.	*i*s-<u>tu</u>-tu-			*!	**

(29)   **Hittite Ranking:** *#ST, ANCHOR-L-BR ≫ ALIGN-RED-L ≫ ALIGN-ROOT-L

Before proceeding, let us consider one additional problematic candidate: [s*i*-stu-], with copying of just the base-initial consonant and epenthesis between the base and the reduplicant. Contrast this with [su-stu-] (25c), which has the same CV shape, but with copying of the base vowel, and thus a fatal CONTIGUITY-BR violation. By epenthesizing its vowel rather than obtaining it via copying, [s*i*-stu-] avoids that CONTIGUITY-BR violation. It also avoids creating the cluster which needs to be repaired through prothesis, while simultaneously improving satisfaction of ALIGN-RED-L, ALIGN-ROOT-L, and ONSET. It does clearly incur extra violations of MAX-BR — as well as the constraint *PCR (see below) — but these constraints are lower ranked than the ones it improves on.

The way to rule out this problematic candidate is using ANCHOR-L-BR. If we assume that the base of reduplication (i.e. the "B" in the BR-correspondence relation over which ANCHOR-L-BR is defined) is *all material to the right of the reduplicant*, then the epenthetic vowel would be parsed as part of the base in this candidate. The reduplicant [s] would still be in correspondence with the root/base [s], but that root/base [s] is no longer base-initial: [s_2-$i_1s_2t_3u_3$-]. This would definitively incur a violation of this formulation of ANCHOR-L-BR. Note that this interpretation of ANCHOR-L-BR winds up being virtually equivalent to Nelson's (2003) LOCALITY constraint. I continue to employ ANCHOR-L-BR for consistency with the rest of the dissertation. Nevertheless, further consideration of the relationship between ANCHOR-L-BR and LOCALITY is a worthy topic for future investigation.[18]

3.4.4 VCX– Bases in Hittite

Vowel-initial roots/bases in Hittite show VC copying: for example, *ark-* 'mount' → *ar-ark-isk–*. This pattern follows completely from the ranking necessary to generate the *iSTV-STVX–* pattern above. This is demonstrated in the tableau in (30) below. In addition to revealing several new rankings (ONSET ≫ ALIGN-ROOT-L ≫ *PCR), this pattern for the first time gives us occasion to consider the status of *PCR in Hittite.

(30)   VCX– bases: *ark-* 'mount' → *ar-ark-*

/RED, ark-/	ANCH-L-BR	CONTIG-BR	ONSET	ALIGN-RT-L	*PCR
a. ark-ark-			*	***!	
b. ☞ ar-ark-			*	**	*
c. a-ark-			**!	*	
d. ak-ark-		*!	*	**	
e. r-ark-	*!			*	*
f. k-ark-	*!			*	

Copying from non–root-initial position (30e,f) provides ideal syllable structure (i.e. no ONSET violations), but incurs a fatal ANCHOR-L-BR violation. Copying the vowel and the second root consonant (30d) violates CONTIGUITY-BR. Copying just the root-initial vowel (30c) creates hiatus, and thus an additional ONSET violation. (This justifies the ranking ONSET ≫ ALIGN-ROOT-L.) Copying the full post-nuclear cluster (30a) leads to having three segments in the reduplicant, and thus three violations of ALIGN-ROOT-L. Since copying just the root-initial vowel and the first

---

[18] Ryan Sandell (personal communication) raises another potentially problematic candidate: [is-*i*stu-], where the epenthetic vowel intervenes between reduplicant and base, and is copied into the reduplicant. If the reduplicant [i] counts as epenthetic relative to the input, then the ranking DEPV-IO ≫ ALIGN-RED-L would properly rule it out. If not, then we might need to invoke DEPV-IR, a constraint penalizing non-input segments appearing in the reduplicant.

post-nuclear consonant (winning candidate (30b)) only incurs two ALIGN-ROOT-L violations and doesn't violate any of the higher-ranked constraints, ALIGN-ROOT-L selects it as optimal.

This *VC*-*VCX*– pattern is very interesting from an Indo-European perspective. As mentioned at the beginning of this chapter (see also Chapter 1), the standard Indo-European distribution of C_1-copying to TRVX– roots/bases (*TV*-*TRVX*–) vs. an alternative pattern to other cluster types (always including STVX– roots/bases) can be explained by the constraint *PCR in (31) (repeated from (2) above). (See Chapter 6 for fuller discussion of this constraint.)

(31)   **NO POORLY-CUED REPETITIONS (*PCR) [≈ *$C_\alpha VC_\alpha$ / _C[-sonorant]]**
For each sequence of repeated identical consonants separated by a vowel ($C_\alpha VC_\alpha$), assign a violation * if that sequence immediately precedes an obstruent.

As indicated in winning candidate (30b), the *VC*-*VCX*– pattern, when applied specifically to VRTX– bases, violates *PCR, because it places a consonant repetition (*rar*) before an obstruent (*k*). In order to select this pattern, *PCR must be outranked by ANCHOR-L-BR, CONTIGUITY-BR, and ONSET. This places it at the very bottom of the ranking of the constraints relevant for reduplication, rendering it essentially *inactive* in the grammar. The same statement holds of its status in Luwian, which shows the same behavior (e.g. CLuw. *ilḫa*– 'wash' → *il-ilḫa*–). This is particularly interesting because *PCR is active in the grammars of almost all the other Indo-European languages with reduplication, *including Proto-Anatolian*, as will be shown in Section 3.7 below. I will argue in Section 3.8 that the unexpectedly low ranking of *PCR in both Hittite and Luwian, which allows for the independent emergence of the *VC*-*VCX*– pattern in the two languages, is the result of independent phonological changes that eliminated *PCR's explanatory power, which led to its demotion in the rankings.

3.4.5   Hittite Summary

The Hasse diagram in (32) summarizes the rankings needed to generate the copying patterns of Hittite. We have thus now seen that the Hittite reduplication patterns sketched in Section 3.3 can be analyzed with a consistent constraint ranking, largely making use of the same constraints employed for reduplication in the other Indo-European languages.

(32)   **Hittite total ranking**
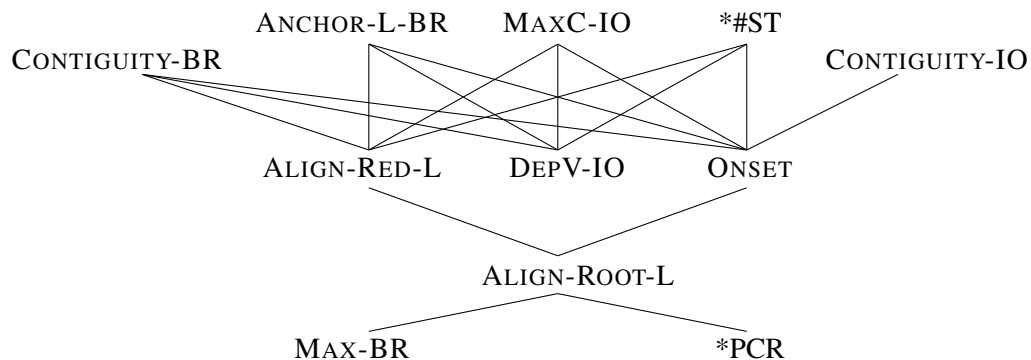
3.5 Synchronic Analysis of Luwian Copying Patterns

Hittite and Luwian display the same surface reduplicative patterns for CVX– and VCX– bases (*CV-CVX–* and *VC-VCX–*, respectively). They diverge only in their treatment of cluster-initial bases. Unlike Hittite, Luwian reflects the Proto-Indo-European C_1-copying pattern for TRVX– bases, i.e. *TV-TRVX–*. Though Luwian exhibits forms of the shape *TV-STVX–* (with apparent C_2-copying), these forms ought not be analyzed as synchronically generated via reduplication; as will be argued in Section 3.6.1 below apropos of the reconstruction of Proto-Anatolian, Luwian synchronically lacks /STVX–/ roots in its lexicon, and thus lacks any attested synchronic pattern associated with STVX– bases. This brief section provides the analysis of the Luwian *TV-TRVX–*, and confirms that this poses no problems for the analysis of the *VC-VCX–* pattern.

3.5.1 TRVX– Bases in Luwian

Whereas Hittite shows cluster-copying for TRVX– bases (*TRV-TRVX–*), Luwian shows the more typical Indo-European C_1-copying pattern: *TV-TRVX–*. We can select the Luwian pattern by taking the ranking proposed for Hittite (cf. (18) above) and reversing the ranking of CONTIGUITY-BR relative to ALIGN-ROOT-L, as illustrated in (34). This is the pattern that will be reconstructed for Proto-Anatolian in the following section, and we will use the same ranking to generate it there.

(33)   **Luwian Ranking:** ALIGN-ROOT-L ≫ CONTIGUITY-BR

(34)   TRVX– bases: *para-* 'carry off' → *pa-pra–*   (cf. Hittite *prai-* → *pri-prai–*)

/RED, pra-/		ANCHOR-L-BR	ALIGN-ROOT-L	CONTIGUITY-BR
a.	pra-pra-		***!	
b. ☞	pa-pra-		**	*
c.	ra-pra-	*!	**	

3.5.2 VCX– Bases in Luwian, and the Ranking of *PCR

Just like in Hittite, vowel-initial bases in Luwian show VC copying; e.g., *ilḫa–* 'wash' → *il-ilḫa–*. As shown in (35), this reduplicative pattern can be analyzed just like the identical Hittite pattern (cf. Section 3.4.4). The difference in relative ranking of ALIGN-ROOT-L and CONTIGUITY-BR between Hittite and Luwian does not affect the outcome of the derivation. Each ranking shown in (35) is crucial, and represents the total ranking for Luwian reduplication (with the addition of the ranking ALIGN-ROOT-L ≫ MAX-BR).

(35)   VCX– bases: *ilḫa–* 'wash' → *il-ilḫa–*

/RED, ilχa-/		ANCH-L-BR	ONSET	ALIGN-RT-L	CONTIG-BR	*PCR
a.	il̲χ̲-ilχa-		*	***!		
b. ☞	il̲-ilχa-		*	**		*
c.	i̲-ilχa-		**!	*		
d.	i̲χ̲-ilχa-		*	**	*!	
e.	l̲-ilχa-	*!		*		*
f.	χ̲-ilχa-	*!		*		

3.6 Reconstructing Proto-Anatolian

The reduplication patterns of Hittite and Luwian — which have been analyzed above in Sections 3.4 and 3.5, respectively — are schematized in (36) below. Luwian's pattern for STVX– bases is in parentheses because, as I will argue below, it cannot be interpreted as a synchronically productive pattern, but rather only as a fossilized relic shaped by sound change not synchronically perceived as reduplicated. We can reconstruct the reduplicative behavior of Proto-Anatolian, the last common ancestor of Hittite and Luwian, on the basis of these patterns.

(36)   Reduplication patterns of Hittite and Luwian

Base type	CVX–	TRVX–	STVX–	VCX–
Hittite	C̲V̲-CVX–	T̲R̲V̲-TRVX–	*i*S̲T̲V̲-STVX–	V̲C̲-VCX–
Luwian	C̲V̲-CVX–	T̲V̲-TRVX–	(T̲V̲-STVX–)	V̲C̲-VCX–

There is no question that Proto-Anatolian displayed *CV* copying for CVX– bases, as this pattern is found in all of the attested Anatolian languages, and is easily reconstructible for Proto-Anatolian's parent language, Proto-Indo-European. With respect to the other three types of bases discussed in this chapter (TRVX–, STVX–, and VCX–), however, the task of reconstructing their reduplicative patterns in Proto-Anatolian is non-trivial. Using a combination of sound change evidence and parsimony, I will argue in this section (following Yates & Zukoff 2016a) for the reconstruction in (37).

(37)   Proto-Anatolian reconstruction

Base type	CVX–	TRVX–	STVX–	(VCX–)
Reduplication pattern	C̲V̲-CVX– (C_1-copying)	T̲V̲-TRVX– (C_1-copying)	S̲T̲V̲-STVX– (cluster-copying)	*does not exist*

This reconstruction claims that Proto-Anatolian possessed the proto-typical Indo-European C_1-copying pattern (*T̲V̲-TRVX–*) for TRVX– bases (meaning that Hittite's cluster-copying pattern is an innovation), but had cluster-copying (*S̲T̲V̲-STVX–*) for STVX– bases (with Luwian's apparent C_2-copying pattern deriving from this via sound change, not change in the reduplicative grammar). This distribution is entirely parallel to Gothic (see Chapter 4), as illustrated in (38). As will be

demonstrated for the Gothic pattern, such a distribution is motivated by *PCR, which diverts the derivation away from the target C_1-copying pattern just when the base begins in an ST cluster. Therefore, despite no evidence for its activity in either Hittite or Luwian, an active *PCR constraint must be reconstructed for Proto-Anatolian.

(38) Proto-Anatolian and Gothic

		CV-CVX–	TV-TRVX–	STV-STVX–
a.	Proto-Anatolian	*ǵi-ǵis–	*bV-brV–	*stu-stu–
b.	Gothic	hɛ-hɛːt	gɛ-groːt	stɛ-stald

Note also the claimed absence of the _VC-VCX–_ pattern in Proto-Anatolian (37), despite its presence in both Hittite and Luwian. I will argue below that this pattern is an independent development in the two languages, and that this has a significant impact on our understanding of the development of *PCR within Anatolian.

3.6.1 Reconstructing the Behavior of STVX– Bases

Hittite and Luwian seem to show completely incompatible behavior for STVX– bases. As provided in (39), both Hittite and Luwian attest a reduplicated stem to the Proto-Anatolian root $\sqrt{}$*stu 'become evident'. The Hittite form shows copying of the entire ST cluster (just like in Gothic), plus a prothetic _i_ (cf. Section 3.4.3). Luwian, on the other hand, appears to have just copied the second member of the cluster (i.e. the non-sibilant obstruent). On the surface, this would seem to be the C_2-copying pattern which is productive for STVX– roots in Sanskrit: for example, $\sqrt{}$*stamb^h* 'prop' → perfect _ta-stamb^h-a_. Given that both the Hittite cluster-copying pattern (minus the prothesis) and the apparent Luwian C_2-copying pattern are attested elsewhere in Indo-European, either one would _a priori_ be a possible reconstruction for Proto-Anatolian. However, once we consider the way that sound change interacts with this question, it becomes clear that the Hittite pattern must be closer to the original situation.

(39) Reduplication with STVX– bases (repeated from (9) above)

	Gloss	Root/Base	Reduplicated stem	
Hitt.	'become evident'	$\sqrt{}$*stu* ([istu–])	_išdušduške–_	[istu-stu-]
CLuw.	'become evident'	PA *stu–	_dušduma/i–_	[tu-stu-]
	'bind'	PA *sh_2(o)i–*	_ḫišḫi(ya)–_	[χi-sχi-]

In Section 3.4.3, we saw that the prothesis which marks reduplicated STVX– bases in Hittite must be a synchronic process; that is to say, if the prothetic vowel were analyzed as part of the underlying representation of the root, we would incorrectly predict these roots to show the _VC-VCX–_ reduplication pattern, rather than the cluster-copying pattern. While the process of prothesis to initial ST clusters remained synchronically active in Hittite, we can also view it from a diachronic perspective as a sound change relative to Proto-Anatolian. Not all the other Anatolian languages have this process (see immediately below),[19] and thus it must be an innovation within the development of

---

[19] Lydian and Palaic likely also attest prothesis in this environment (Melchert 1994:206, 371), but, given the alternative treatment in Luwian and elsewhere, this is to be analyzed as a post-Proto-Anatolian development, either independently from Hittite or as a common innovation with Hittite in a post-Proto-Anatolian subgroup (Kavitskaya 2001:294–295).

Hittite. This suggests that Hittite might have inherited a pattern from Proto-Anatolian which was equivalent to the *iSTV-STVX*– pattern but without the prothetic vowel, i.e. *STV-STVX*–.

If Hittite diverges from the inherited form only by the application of regular sound change, might the Luwian form be analyzable in the same way? Yes. When we consider the way in which Luwian treated initial ST clusters, i.e. the environment for prothesis in Hittite, we see that it shows deletion of the initial *s*. This is illustrated in (40a) where we can see non-reduplicated Luwian forms and their Hittite cognates. Where the Hittite form shows an initial #*i*STV... sequence, Luwian instead just shows a #TV... sequence. The same holds in (40b) of reduplicated forms: where Hittite would have an initial #*i*STV... sequence, Luwian has a #TV... sequence.[20] The table in (41) shows that these treatments are restricted to word-initial position; intervocalic -ST- sequences are retained faithfully in both languages.

(40)   Treatment of PA #ST (cf. Melchert 1994:30–32, 2016:187–188, Yates 2014, 2016)

PA		CLuw.			Hitt.	
a.	*⟦sp⟧*or*–	>	⟦p⟧*arritti*	'spreads'	cf.	⟦*išp*⟧*āri*	(Kloekhorst 2008:406–408)
	⟦st⟧(e)h₃men–	>	⟦t⟧*ummān*	'ear'		⟦*išt*⟧*āmanan*	(Kloekhorst 2008:411–413)
b.	*⟦sh₂⟧i(-sh₂i)*–	>	⟦ḫ⟧*išḫiyanti*	'bind'		⟦*išḫ*⟧(a)i–	
	⟦st⟧u(-stu)–	>	⟦d⟧*ušdu(miš)*	'manifest'		⟦*išt*⟧*u*–	

(41)   Treatment of PA medial ST-clusters

PA		CLuw.			Hitt.	
h₁é⟦sh₂⟧-r̥	>	*ā⟦šḫ⟧ar(-sa)*	'blood'	cf.	*ē⟦šḫ⟧ar*	
–o⟦s-t⟧i–	>	*lump-a⟦št⟧i*–	'regret'		*dalug-a⟦št⟧i*– 'length'	
h₁é⟦s-t⟧i	>	*ā⟦št⟧i*	'is'		*ē⟦šz⟧i*	

We can summarize these processes as in (42), where (42a) gives the diachronic correspondences (i.e. sound changes), and (42b) gives the synchronic phonological processes that were operative to achieve these sounds correspondences. The prothesis rule for Hittite was synchronically active both in Hittite and all periods subsequent to the adoption of the sound change within Pre-Hittite. We can only demonstrate that the Luwian deletion rule was operative in some period of Pre-Luwian, as this process seems to have altered underlying representations by the time of attested Luwian.

(42)   a.   **Sound Changes:**
        i.   Proto-Anatolian *#ST > Hittite #*i*ST
        ii.   Proto-Anatolian *#ST > Luwian #T
      b.   **Synchronic Processes:**
        i.   (Pre-)Hittite   Ø → [i] / #__ST
        ii.   Pre-Luwian   /s/ → Ø / #__T

---

[20] Rieken (2010) argues that *#*sk* is retained as such in Luwian. Based on the analysis developed here, we would predict a synchronic /skVX–/ root in Luwian to reduplicate as *sV-skVX*– (see footnote 22 below). Since we have no evidence (in either language) for the treatment of inherited *#*sk* in reduplication, we cannot probe this question here.

As the innovative Luwian deletion rule became categorical, the /s/ of Proto-Anatolian *ST-initial roots would no longer surface in simplex verbal forms. This /s/ might have been recoverable if supported by alternations, but Luwian has no productive prefixing morphology other than reduplication. In principle, reduplicating verbal stems like (40b) could have sustained an underlying root-initial /s/, but the historical simplex verbs corresponding to the attested reduplicated forms of Proto-Anatolian *ST-roots appear to have been lost; for example, while Luwian attests reduplicated *ḫišḫi(ya)–*, it does not attest a simplex *ḫāi–* equivalent to Hittite *išḫ(a)i–*. The lack of direct evidence for [s] led to *restructuring* of historically *ST-initial roots, with /s/ uniformly lost from URs — i.e. PA */STVX–/ > Luw. /TVX–/. *STVX* and *TVX* roots would then merge as /TVX/ synchronically.[21] It is thus appropriate to assume that *ḫišḫi(ya)–* came to be interpreted as a unitary stem/root in Luwian, rather than one derived via reduplication from an independently occurring root.

Therefore, Luwian does not actually provide direct evidence for a synchronic treatment of STVX– bases, as its lexicon lacked them entirely.[22] Nonetheless, the fact that both languages would display the diachronically regular outcome of a Proto-Anatolian *STV-STVX–* pattern — especially given the necessary non-productivity of that output within Luwian — is strong evidence in favor of reconstructing this pattern for Proto-Anatolian.

3.6.2 Reconstructing the Behavior of TRVX– Bases

Hittite and Luwian disagree also on the treatment of TRVX– bases, as illustrated in (43) below. Hittite shows full copying of the initial TR cluster: *TRV-TRVX–*, for example *pri-prai-* 'blow' (**pi-prai-*). Luwian, on the other hand, copies only the initial obstruent: *TV-TRVX–*, for example *pa-pra-* 'carry off' (**pra-pra-*). The primary argument for reconstructing the Luwian pattern rather than the Hittite pattern for Proto-Anatolian comes down to parsimony.

(43)   Reduplication with TRVX– bases (repeated from (8) above)

	Gloss	Base	Reduplicated stem	
Hitt.	'blow'	*par(a)i–*	*parip(p)ar(a)i–*	[priːpːr(a)i-]
	'kneel'	*ḫal(a)i–*	*ḫaliḫal(a)i–*	[χli-χl(a)i-]
CLuw.	'carry off'	*par(a)–*	*papra–*	[pa-pra-]

Luwian reflects the pattern that is easily reconstructible for Proto-Indo-European based on the comparative evidence outside of Anatolian. This C_1-copying pattern is clearly found in Ancient Greek (e.g. *ke-kri-* not ***kre-kri-*), Sanskrit (e.g. *pa-prac^h-* not ***pra-prac^h-*), Gothic (e.g. *ge-grōt* not ***gre-grōt*), and elsewhere. If we were to reconstruct the Hittite cluster-copying pattern to Proto-Anatolian, this would require us to posit a change between Proto-Indo-European and Proto-Anatolian (*C_1-copying > cluster-copying*), and then posit another change — in fact, the exact opposite change — between Proto-Anatolian and Luwian (*cluster-copying > C_1-copying*). On the other hand, if we reconstructed the Luwian C_1-copying pattern to Proto-Anatolian, this would require

---

[21] A prediction of this analysis is that synchronically generated reduplicative forms to roots of either historical shape would at this stage show the same *CV-CVX–* pattern. That is to say, PA */sTVX/ > Luwian /TVX/ → *TV-TVX–*. Clear evidence for or against this claim is so far lacking; a possible positive example is CLuw. *dūp(a)i–* 'strike' (< PIE **(s)teup–*; cf. Rix et al. 2001:602–603) → *dūdupa–* (hapax; see Melchert 1993:238), but the issue is confounded by cognate Lycian *tub(e)i–* 'id.' and the problem of **s*-mobile.

[22] The analysis developed in Section 3.5 predicts that, if STVX– roots/bases were input to the grammar (via richness of the base), the evaluation would select C_1-copying (*SV-STVX–*). This does not appear to be a testable prediction.

positing only a single change between Proto-Anatolian and Hittite (C_1-*copying* > *cluster-copying*), allowing us to assume no change in this domain between PIE and Proto-Anatolian.

In Chapter 7, I will argue that PIE should be reconstructed as having displayed C_1-copying to STVX– roots (i.e. _SV-STVX–_). If we take this to be true, then Proto-Anatolian's treatment of STVX– bases (i.e. the cluster-copying _STV-STVX–_ pattern) is innovative relative to PIE, representing the change C_1-*copying* > *cluster-copying*. Adopting the C_1-copying pattern for the reconstruction of Proto-Anatolian TRVX– bases would then provide a unitary direction of change within the development of Anatolian, as shown in (44), from C_1-copying to cluster-copying: the STVX– bases change first (explainable with *PCR), then the TRVX– bases follow in the same direction (explainable with CONTIGUITY-BR). While the adoption of cluster-copying to satisfy *PCR at that particular moment in time (i.e. at the point when the innovative cluster-copying pattern arises) may be mysterious, it is not unexpected given the pan–Indo-European context — i.e., the same mystery would seem to hold for all the systems that show basic *PCR effects. However, the subsequent change to cluster-copying for TRVX– bases into Hittite could be motivated by a mis-analysis of the new evidence provided by the cluster-copying pattern for STVX– bases. This hypothesis will be explored further in Section 3.8.

(44) Change in cluster-initial bases from PIE to Hittite (assuming PA TRVX– C_1-copying)

	PIE	>	PA	>	Hittite
TRVX–	C_1-copying	=	C_1-copying	>	**Cluster-copying**
STVX–	C_1-copying	>	**Cluster-copying**	=	Cluster-copying

If instead we were to reconstruct cluster-copying for TRVX– bases in Proto-Anatolian, as outlined in (45) below, we would be left with a diachronic "Duke of York" scenario for Luwian: PIE ***A** > Proto-Anatolian ***B** > Luwian **A**. We would have to assume a wholesale change from across-the-board C_1-copying in PIE to across-the-board cluster-copying in Proto-Anatolian. This can be effected with a single change in ranking: PIE ALIGN-ROOT-L ≫ CONTIGUITY-BR *changes to* PA CONTIGUITY-BR ≫ ALIGN-ROOT-L.[23] However, there is no obvious motivation for such a change. This cannot be a terribly substantive argument, because one must posit this change in ranking somewhere along the chain regardless; I attribute it to the development from PA to Hittite. Nevertheless, the lack of motivation for this change is matched by the lack of motivation for its undoing, which must then be posited between Proto-Anatolian and Luwian: PA CONTIGUITY-BR ≫ ALIGN-ROOT-L *changes to* Luwian ALIGN-ROOT-L ≫ CONTIGUITY-BR.[24] Given the doubly unmotivated change in both directions that this would require, positing Proto-Anatolian cluster-copying is thus an unparsimonious and unattractive solution.

(45) Change in cluster-initial bases from PIE to Luwian (assuming PA TRVX– cluster-copying)

	PIE	>	PA	>	Luwian
TRVX–	C_1-copying	>	**Cluster-copying**	>	**C_1-copying**
STVX–	C_1-copying	>	**Cluster-copying**	(>)	**(Ø)**

---

[23] Incidentally, this would deny any role for *PCR in Proto-Anatolian, which would, in some ways, make it easier to explain the rise of the _VC-VCX–_ pattern.

[24] We cannot know how this change would interact with the treatment of STVX– bases, as they were eliminated from the lexicon within this same period of development; see above.

Beyond an argument from parsimony, we might have evidence from archaisms. Hittite has several forms which may preserve traces of a _TV_-_TRVX_– C_1-copying pattern. One is _tatrant_– 'sharp-edged; prone to goring' (cf. Melchert 1984:33, fn. 68, Kloekhorst 2008:857), which serves as the basis for a verbal stem _tatraḫḫ_– 'incite'. Another is _paprant_– 'impure', which serves as the basis for a verbal stem _papraḫḫ_–/_papre(šš)_– 'make/be(come) impure' (cf. footnote 10 above). If these adjectives were to go back to earlier reduplicated verbal formations (perhaps participial formations, though of exactly what sort, I don't know), they could be said to be reflecting the expected treatment of TR-clusters in verbal reduplication at that stage. While it is certainly conceivable that the stage they would be reflecting is Proto-Indo-European, if it could be shown that they were properly of Proto-Anatolian vintage (i.e. being productively generated up through or for the first time in Proto-Anatolian), this would be evidence for reconstructing the C_1-copying pattern to TRVX– bases in Proto-Anatolian. Since these forms' status is anything but certain, I leave these as possible pieces of evidence for future verification.

3.6.3   Reconstructing the Absence of VCX– Bases

Since the _VC_-_VCX_– pattern is found in both Hittite and Luwian, parsimony would favor reconstructing it for Proto-Anatolian. However, I must nonetheless claim that this pattern was absent from Proto-Anatolian. The abiding motivation for this claim is that the pattern is _incompatible_ with the analysis of Proto-Anatolian to be presented below in Section 3.7. Namely, generation of the _STV_-_STVX_– alongside the _TV_-_TRVX_– pattern requires *PCR to be ranked high, while the _VC_-_VCX_– pattern would require *PCR to be ranked low. Therefore, if this pattern were present in the language, we would have an insurmountable ranking paradox.

While the motivating factor for making this unparsimonious claim is purely theory internal, there is independent reason to believe that it could be correct. Proto-Indo-European lacked vowel-initial roots (see Rix et al. 2001, among others). Within Anatolian, vowel-initial roots first could have arisen only after the loss of *h_1 in word-initial pre-vocalic position (#_V). All the Anatolian languages do lack *h_1,[25] and do have vowel-initial roots; this would again by parsimony suggest that *h_1 was lost already in Proto-Anatolian. However, there are two pieces of evidence that may suggest that *h_1 was retained (at least in certain positions) beyond the break-up of Proto-Anatolian proper: nasal assimilation in Hittite (46), and assibilation in Proto-Anatolian (47). If it is correct that *h_1 was retained beyond Proto-Anatolian, then it must be the case that vowel-initial roots were still lacking in Proto-Anatolian, and therefore the _VC_-_VCX_– pattern must have been a later development in the individual languages.[26,27]

First, inherited nasal-sibilant clusters regularly undergo assimilation to –_šš_– [-sː-] in Hittite (46a), but inherited *–_nh$_1$s_– sequences appear to instead yield Hittite –_nz_– [-nts͡-] (46b).[28] That is to say, nasal assimilation fails to occur if, at the stage during which the process is active, h_1 intervenes between the nasal and the sibilant. Melchert (1994:63) argues that *–_ns_– assimilation is a post-PA development. If this is correct, then (46b) would argue for the retention of *h_1 beyond Proto-Anatolian and thus into (Pre-)Hittite — and, by the same logic, into (Pre-)Luwian.

---

[25] Here I follow the traditional view. This has recently been challenged by Kloekhorst (2004, 2006, 2008) and Simon (2013), who claim that, in both Hittite and Luwian, the use of <V(C)-> signs in word-initial position actually indicates [ʔ(V)C-]. This position has been refuted by Weeden (2011:61–68).

[26] I am indebted to Tony Yates for identifying and assembling the evidence on these points. All mistakes are of my own doing.

[27] Melchert (2015) now rejects the evidence for "limited Čop's Law," which would entail the loss *h_1 / #_V in Proto-Anatolian (cf. Melchert 1994:65).

[28] On (46a), see Melchert (1994:163), and on (46b), see Kloekhorst (2008:468–469) (for the morphology, cf. Hitt. _tepšu_– 'small' < PIE *$d^héb^h$-_su_–). See Byrd (2015:100–102) on the non-deletion of PIE *h_1 / n_C.

(46) Nasal-sibilant (non)assimilation in Hittite
    a.   PA *dé/óns-u–   > Hitt. *daššu*–   'strong'
          PA *$h_2$óns-o–   > Hitt. *ḫāšša*–   'offspring'
    b.   PA *ǵénh$_1$-su   > Hitt. *kēnzu*   'lap'   (not **kēššu*)

Second, PIE/PA coronal stops (abbreviated here as T, not to be confused with the use of T elsewhere to mean obstruents in general) display sibilant reflexes across Anatolian when they preceded a palatal glide: *T / _ yV > Hitt. –z–, Luw. –z–, Lyc. –z– (= [t͡s]) (47a). Given the pan-Anatolian agreement on this point, this assibilation must be a Proto-Anatolian feature (cf. Melchert 1994:62). On the other hand, when h_1 intervened between the coronal stop and the palatal glide, assibilation was blocked: *–Th$_1$y– > Hitt. –Ty– (47b) (Kimball 1999:404; cf. Melchert 1983:14).[29] This shows that h_1-loss must post-date the period during which this assibilation process was active. While this sequence of events could be contained within the internal development of "Proto-Anatolian" (most precisely, assibilation was operative in a particular period of Pre-Proto-Anatolian and h_1-loss occurred between this stage and Proto-Anatolian proper), it could be suggestive of a post-PA date for h_1-loss — i.e. separate, parallel loss in the individual languages.

(47) Coronal (non)assibilation in Proto-Anatolian
    a.   PIE/PA *–tyo–   > Hitt. *šarāzziya*–, Lyc. *hrzze/i*– 'upper'
          PIE/PA *$h_2$et-ye/o–   > Hitt. *ḫazziya*– 'strike (an instr.)', HLuw. *ḫazi*– 'incise; write'
    b.   PIE/PA *dh$_1$-ye/o–   > Hitt. *tiya*[30] 'bind' (not **ziya*)
          PIE/PA *d$^{(h)}$h$_1$y-énti > Hitt. *tianzi* 'place' (not **zianzi*)

In view of this evidence, it is at least plausible, if not strictly necessary,[31] to assume that the loss of *h_1 in *h_1-initial roots occurred subsequently to the other changes affecting reduplication in Hittite and Luwian (specifically the demotion of *PCR). After the loss of *h_1, these roots were subject to reduplication in accordance with the new synchronic grammar, which yielded the *VC*-*VCX*– pattern.

3.7 Synchronic Analysis of Proto-Anatolian

Now that we have established the reconstruction for Proto-Anatolian, we can proceed to its analysis, which can be handled quite quickly. The logic of the system — which is, as mentioned above, functionally identical to the reduplicative system of Gothic — can be summarized as follows: reduplicate with C_1V- (ALIGN-ROOT-L ≫ CONTIGUITY-BR, MAX-BR; as demonstrated by (49), with MAX-BR omitted for space), unless doing so would yield a *PCR violation, in which case, reduplicate the full cluster *STV*- (ANCHOR-L-BR, *PCR ≫ ALIGN-ROOT-L; as demonstrated by (48)).

The most notable difference between this system and those of Hittite and Luwian is the high-ranked *PCR constraint. This blocks the default C_1-copying candidate (48b) for STVX– bases. With the ranking ANCHOR-L-BR ≫ ALIGN-ROOT-L, mis-anchoring (48c) is worse than copying the entire cluster, so (48a) is selected as optimal. When C_1-copying would not lead to a *PCR

---

[29] Assibilation of *T / _ i (e.g. Hitt. 3SG.PRES –*zi* < PIE/PA *–*ti*; cf. CLuw. –*tti*) is a separate Hittite development.

[30] Cf. Sanskrit *dyáti*, Old Avestan *(ni.)diiātąm*, Greek *déō*.

[31] While it would be quite likely for *h_1 to be retained in initial pre-vocalic position as long or longer than in inter-consonantal position, since initial pre-vocalic position would presumably have been a more robustly cued environment, this is not guaranteed. Therefore, these arguments are at best suggestive.

violation, as in (49) for TRVX– bases, the ranking ALIGN-ROOT-L ≫ CONTIGUITY-BR allows for the C$_1$-copying pattern to win out, selecting candidate (49b).

(48)   PA *STVX– → *<u>STV</u>-STVX–   (PA *stu– → *<u>stu</u>-stu– > Hitt. *i<u>šdu</u>šdu–*, CLuw. *dušdu–*)

/RED, STVX–/	ANCH-L-BR	*PCR	ALIGN-RT-L	CNTG-BR
a. ☞ <u>STV</u>-STVX– (*<u>stu</u>-stu–*)			***	
b. <u>SV</u>-STVX– (*<u>su</u>-stu–*)		*!	**	*
c. <u>TV</u>-STVX– (*<u>tu</u>-stu–*)	*!		**	

(49)   PA *TRVX– → *<u>TV</u>-TRVX–   (e.g. PA *brV– → *<u>bV</u>-brV– > Luw. *pa-pra–*)

/RED, TRVX–/	ANCH-L-BR	*PCR	ALIGN-RT-L	CNTG-BR
a. <u>TRV</u>-TRVX– (*pra-pra–*)			***!	
b. ☞ <u>TV</u>-TRVX– (*pa-pra–*)			**	*
c. <u>RV</u>-TRVX– (*ra-pra–*)	*!		**	

(50)   **Proto-Anatolian Ranking:**
        ANCHOR-L-BR, *PCR ≫ ALIGN-ROOT-L ≫ CONTIGUITY-BR, MAX-BR

3.8   Constraint Re-ranking and the Demise of *PCR in Anatolian

The Anatolian languages show different reduplicative patterns — and thus independent constraint re-ranking — with respect to Proto-Anatolian. Generating the set of changes that characterize the attested Anatolian languages (as summarized above in (44)) requires the separate re-ranking of just three constraints — *PCR, ALIGN-ROOT-L, and CONTIG-BR — in Hittite and Luwian, as outlined in (51).

(51)   Constraint rankings in Anatolian

PA	*PCR	≫	ALIGN-ROOT-L	≫	CONTIG-BR
Luwian	ALIGN-ROOT-L	≫	CONTIG-BR	≫	*PCR
Hittite	CONTIG-BR	≫	ALIGN-ROOT-L	≫	*PCR

The set of diachronic developments can thus largely be characterized by two changes in rankings. First, Hittite shows a reversal of CONTIG-BR and ALIGN-ROOT-L, generating cluster-copying as the default pattern for *all* cluster-initial roots, i.e. the <u>TRV</u>-TRVX– pattern. Second, *PCR is rendered *inactive* in both Hittite and Luwian, requiring its demotion to the bottom of the grammar; this allows for the emergence of the <u>VC</u>-VCX– pattern. This raises the following question: *why* does *PCR cease to be operative between Proto-Anatolian and the Anatolian daughter languages?

The development of the <u>VC</u>-VCX– pattern for (synchronically) vowel-initial roots (< *h$_1$eC–) demonstrates the demotion of *PCR in the Hittite and Luwian grammars, but this is an effect rather than a cause of these changes in ranking. I propose that the demotion of *PCR can be attributed to the nature of the learning input and learning process following the Hittite- and Luwian-internal phonological changes affecting *TRVX–* and *STVX–* bases in reduplication which were discussed

already in Section 3.6.1 and 3.6.2 in relation to the reconstruction of Proto-Anatolian. After the post-PA loss of (pre-vocalic word-initial) *h_1*, as motivated in Section 3.6.3, the innovative grammar arrived at through the new learning conditions would have productively generated the *VC-VCX*–reduplication pattern for newly vowel-initial roots.

The remainder of this section details the proposed stages of development and changes that take Proto-Anatolian into Hittite and Luwian, respectively. To properly derive the various changes in ranking — specifically those relating to *PCR — I will propose a slight revision to the Recursive Constraint Demotion learning algorithm, termed here *Maximally Informative Recursive Constraint Demotion* (MIRCD), which provides a principled means of demoting *PCR even in the absence of evidence for its violation.

3.8.1 The Relative Chronology of Constraint Re-Rankings into Hittite

The crucial innovation for the development of Hittite was the change for TRVX– bases from the inherited C$_1$-copying pattern (*<u>*TV*-TRVX</u>*–) to full cluster-copying reduplication (<u>*TRV*-TRVX</u>–). In terms of the ranking of the constraints, the adoption of this pattern amounts to the promotion of CONTIGUITY-BR over ALIGN-ROOT-L. This change is represented in the transition between the stage in (52) below and the stage in (53) below. (The ✗ symbol indicates a diachronically prior stage's winner which now loses under the new constraint ranking.)

(52)   Proto-Anatolian

/RED, prai–/	ALIGN-ROOT-L	CONTIG-BR
a. pri-prai–	***!	
b. ☞ pi-prai–	**	*

⇓ *Re-rank* ALIGN-ROOT-L & CONTIG-BR ⇓

(53)   Pre-Hittite I

/RED, prai–/	CONTIG-BR	ALIGN-ROOT-L
a. ☞ pri-prai–		***
b. ✗ pi-prai–	*!	**

Can we find any rationale for this change? Despite its prevalence within Indo-European, the C$_1$-copying pattern is cross-linguistically rare, with direct parallel (perhaps only) in Klamath (Barker 1964; see Steriade 1988, and Chapter 6).[32] This potential typological asymmetry might reflect a bias towards contiguous copy, i.e., a bias towards the high ranking of CONTIGUITY-BR. From a learning perspective, one might view this through the lens of the learning bias towards Output-Output faithfulness constraints (McCarthy 1998, Hayes 2004:188), as Base-Reduplicant correspondence could well be subsumed under the broader category of Output-Output correspondence.

One could also wonder whether the persistence of the *STV*-*STVX*– pattern might have itself contributed to the change. If, for some reason, learners failed to take in enough evidence to determine the intended treatment of TRVX– bases, but did take in enough evidence to determine the intended treatment of STVX– bases, they would have had to generalize from STVX– to TRVX–.

---

[32] Though it must be noted that having the combination of conditions under which this pattern could be observed is itself cross-linguistically rare.

In the absence of the C_1-copying pattern for TRVX– bases, there are two possible ways a learner might analyze cluster-copying for STVX– bases: *PCR or CONTIGUITY-BR. If learners opted for the CONTIGUITY-based solution, generalizing this to TRVX– bases yields cluster-copying, as well. If the extent of attestation reflects the actual distribution of relevant forms in the language, then there certainly would have been a dearth of evidence for both types, and this sort of generalization from one type to the other would not be unimaginable. In any event, the ultimate attribution of the cluster-copying patterns to CONTIGUITY-BR must play a role in the eventual demotion of *PCR, as outlined below.

Alternatively, it is also hypothetically possible that this change could have been directly motivated by a change in speakers' analysis of the reduplicative vowel. If, in Proto-Anatolian, the reduplicative vowel had been analyzed as morphologically fixed (i.e., not arising through copying; see Alderete et al. 1999 and Chapter 1), then the ranking of CONTIGUITY-BR would have been irrelevant (cf. Chapter 2 on Greek). That is to say, the correspondence relationships in a C_1-copying form would be $\underline{T_1}\text{-}V_i\text{-}T_1R_2V_3X\text{–}$, where the reduplicative vowel (V_i) stands in correspondence only with the input, not with the base of reduplication. Since there is only one segment in the reduplicant that has a correspondent in the base, CONTIGUITY-BR is inherently vacuously satisfied. If the reduplicative vowel later gets re-analyzed as a copy vowel (whether a perfect copy or an imperfect copy), then CONTIGUITY-BR suddenly becomes relevant, as it would now be violated by a C_1-copying candidate ($*\underline{T_1}V_3\text{-}T_1R_2V_3X\text{–}$). If speakers had covertly ranked CONTIGUITY-BR high enough in the ranking (i.e. above ALIGN-ROOT-L), then such a change in the analysis of the reduplicative vowel would trigger this change in copying pattern. However, as discussed briefly in Section 3.3.1, both Hittite and Luwian seem to point to identical vowel copy as the primary pattern for the reduplicative vowel. Parsimony of reconstruction would then suggest that Proto-Anatolian itself also had identical vowel copy as its primary pattern; but the presence of exceptions within the languages certainly leaves open the possibility of some different behavior at the prior stage. Since I am not prepared to make a strong claim about the nature of the reduplicative vowel at any of these stages, I leave this as nothing more than a suggestion.

Whatever its motivation, the change from $\underline{TV}\text{-}TRVX\text{–}$ to $\underline{TRV}\text{-}TRVX\text{–}$ — i.e., the change from the Proto-Anatolian grammar in (52)/(54) to the Pre-Hittite grammar in (53)/(55) — had significant implications for speakers' *analysis* of the pattern for STVX– bases, though not for the surface properties of the pattern itself. At the stage in (55), the losing C_1-copying candidate (b) violates both CONTIGUITY-BR and *PCR. But unlike *PCR, CONTIGUITY-BR is independently necessary to account for $\underline{TRV}\text{-}TRVX\text{–}$ reduplication. In the absence of unambiguous evidence for the activity of *PCR, Hittite learners converged on a simpler analysis: the $\underline{STV}\text{-}STVX\text{–}$ pattern was reanalyzed as being driven by CONTIGUITY-BR, and *PCR, with no forms requiring its activity, ends up being demoted to the bottom of the grammar, resulting in the stage in (56). It is crucial that this process results in the total demotion of *PCR, all the way below ALIGN-ROOT-L, in order to derive the behavior of VCX– bases at the following stage. (On the nature of this ranking change, see the discussion in Section 3.8.2 below.)

(54) Proto-Anatolian

/RED, stu-/	*PCR	ALIGN-ROOT-L	CONTIG-BR
a. ☞ stu-stu-		***	
b. su-stu-	*!	**	*

⇓ *Re-rank* ALIGN-ROOT-L & CONTIG-BR (cf. (52) > (53)) ⇓

(55) Pre-Hittite I

/RED, stu-/	*PCR	CONTIG-BR	ALIGN-ROOT-L
a. ☞ stu-stu-			***
b. su-stu-	*!(?)	*!(?)	**

⇓ *Demote* *PCR ⇓

(56) Pre-Hittite II

/RED, stu-/	CONTIG-BR	ALIGN-ROOT-L	*PCR
a. ☞ *(i)*stu-stu-		***	
b. su-stu-	*!	**	*

Provided that the innovations in (55) and (56) chronologically *precede* the loss of $*h_1$ (a plausible assumption based on the discussion in Section 3.6.3), the new grammar generates *VC*-*VCX*–reduplication straightforwardly, as shown in (58). Note again that *PCR must be demoted not just below CONTIG-BR but also below ALIGN-ROOT-L in order to generate the *VC*-*VCX*– pattern.

(57) PIE/Proto-Anatolian

/RED, $h_1Vrǵ^h$–/	*PCR	ALIGN-ROOT-L	CONTIG-BR
a. ☞ h_1V-$h_1Vrǵ^h$-		**	
b. h_1Vr-$h_1Vrǵ^h$-		***!	

⇓ ... ⇓

(58) Pre-Hittite III (= Hittite)

/RED, ark–/	CONTIG-BR	ALIGN-ROOT-L	*PCR
a. ☞ ar-ark-		**	*
b. ark-ark-		***!	
c. ak-ark-	*!	**	

Prior to the changes between Proto-Anatolian and "Pre-Hittite III", $*h_1VC$– roots would have reduplicated like ordinary *CVX*– roots, as shown in (57). But, after (i) the constraint re-ranking motivated by the changes for cluster-initial bases, and (ii) the loss of $*h_1$, these newly vowel-initial roots in Hittite are correctly predicted to show *VC*-*VCX*– reduplication, i.e. (58). A summary of the proposed changes from Proto-Anatolian to Hittite and their relative chronology is given in (59).

(59)   Hittite relative chronology

Stage	Ranking
(I) Proto-Anatolian	*PCR ≫ Align-Root-L ≫ Contig-BR

- *TRVX– roots: C_1-copying pattern changes to cluster-copying pattern*
- *Indeterminacy about ranking of *PCR vis-à-vis STVX– roots*

| (II) Pre-Hittite I | *PCR ?? Contig-BR ≫ Align-Root-L |

- **PCR is unnecessary to account for STVX– roots, so it is demoted*

| (III) Pre-Hittite II | Contig-BR ≫ Align-Root-L ≫ *PCR |

- *$*h_1$ deletes / #_V*
- *Newly vowel-initial roots fed into grammar, generate VC-VCX– pattern*

| (IV) Pre-Hittite III / Hittite | Contig-BR ≫ Align-Root-L ≫ *PCR |

3.8.2 Maximally Informative Recursive Constraint Demotion (MIRCD)

To generate the full set of changes posited between Proto-Anatolian and Hittite, we need *PCR to go from the top-ranked constraint (among the three under discussion) in Proto-Anatolian to the bottom-ranked constraint in Hittite, despite there being a period during which *PCR remained a surface true constraint (at least within the realm of reduplication). Notably, this scenario would seem to constitute a counterexample to the "subset principle".

The subset principle states that, when learners are choosing between multiple possible grammars consistent with the positive evidence, they ought to select the grammar which is most restrictive (allows in the fewest possible unseen forms), because doing otherwise has the potential to overgenerate relative to the target language and make it impossible to later arrive at the more restrictive target language (cf. Prince & Tesar 2004 and references therein). In phonology, this reduces mainly to a preference for the higher ranking of markedness constraints than faithfulness constraints, as implemented in, for example, Biased Constraint Demotion (BCD; Prince & Tesar 2004) and Low Faithfulness Constraint Demotion (LFCD; Hayes 2004). It is standardly assumed that it is a desideratum of a phonological learning procedure for it to capture the subset principle. Capturing the subset principle is thus taken as one of the key arguments in favor of BCD and LFCD over simple Recursive Constraint Demotion (RCD; Tesar 1995, Tesar & Smolensky 1998, 2000).

However, if the current question regarding the diachronic development of *PCR within Anatolian is being framed in the correct manner, this is a case where learners have failed to learn the subset language, as a superset language emerges in a diachronically subsequent stage. That is to say, despite having no positive evidence that *PCR was violable in their language, (Pre-)Hittite learners learned a grammar that tolerated the emergence of a pattern (the VC-VCX– pattern) that violated *PCR, rather than the more restrictive grammar (appropriate to a prior stage) that would have outlawed forms of this type. Therefore, resolute adherence to the subset principle would in fact *not* be consistent with the empirically available data, and we would be required to develop/adopt a learning procedure that, at least under particular circumstances, learns a grammar other than the most restrictive one. I propose below a modification to the RCD algorithm (which can be made consistent with BCD as well) that will be capable of accounting for the *PCR problem in Anatolian without completely undermining the restrictiveness of the approach.

The logic of the argument presented here is that *PCR *lost its explanatory power* in the development of Hittite with the advent of the <u>*TRV*-*TRVX*–</u> pattern, and this lack of explanatory power is ultimately responsible for its complete demotion. As alluded to above, this type of logic is inconsistent with most established procedures for phonological learning. With no evidence of its violation, RCD — which is concerned only with whether or not a constraint prefers a loser — would install the *PCR constraint in the top stratum, not the bottom one. This holds all the more for BCD and LFCD, where the fact that *PCR is a markedness constraint would further bias it towards high ranking.[33] Nor will this work in standard error-driven weighted constraint learning models like the Gradual Learning Algorithm (GLA; Boersma 1997, Magri 2012), because weight accrued by erroneously picking a *<u>*SV*-*STVX*–</u> output will be assigned to both *PCR and Contiguity-BR.[34] While I will entertain an alternative solution in Section 3.8.6 below based on inherent bias for Base-Reduplicant faithfulness constraints (in the mode of BCD), I believe that the most successful solution is the following.

I propose here that a slight adjustment to the Recursive Constraint Demotion algorithm can capture the logic of the problem and properly derive the ranking issues. What is needed is a procedure that prefers to install constraints with greater explanatory power. Namely, rather than RCD installing *all* constraints that prefer no losers (or even only winners) among the current *Winner∼Loser* pairs which have not yet received a **W** from an installed constraint (the "support"), RCD installs only the constraint or constraints which prefer *the most* winners. (Consult again Prince 2002 on the comparative tableau format and the use of **W**, **L**, and **e**.) Such a system can be described as aiming to explain observed *Winner∼Loser* pairs using the fewest constraints possible.[35]

The standard formulation of the RCD algorithm collects all constraints which prefer no losers among the current support — i.e. those constraints which have only **W**'s and/or **e**'s in their column — and installs them (recursively) in the highest stratum.[36] The way that Becker (2009:164) formalizes the algorithm, which is reproduced in (60) below, is very slightly different, but in an interesting way (though this divergence is made without comment, as far as I can tell). Rather than selecting all the constraints that prefer no losers, this formulation selects only those constraints which prefer at least one winner and no losers (60a). For the *PCR ranking problem in Pre-Hittite that we are currently trying to solve, either version would result in *PCR being placed in the top stratum, because it prefers no losers and one winner (see (62) below). (However, only Becker's version would be able to account for the Pre-Luwian situation without adjustment; see Section 3.8.5.) Since we are trying to generate a ranking where *PCR is at the bottom, this approach is not going to be sufficient.

The relative difference between the Becker approach and the standard approach hints at a solution to the problem. My proposal — which I term Maximally Informative Recursive Constraint Demotion (MIRCD), formalized in (61) below — is that only the *maximally informative* winner-preferring constraints are installed in the highest stratum. I formalize this by introducing a new category into the RCD algorithm, the *maximally-informative-winners* (61b). This selects from among

_____

[33] In the ultimate cue-based version of *PCR that I will develop and advocate in Chapter 6, it is less clear that *PCR is to be thought of as a normal markedness constraint.

[34] There is some indication that, using the GLA, assigning *PCR a sufficiently lower initial weight than all the other constraints involved could generate the result, but I can see no reason for making such an assumption.

[35] Hayes's (2004) "Favour Autonomy" preference in LFCD — which prefers the installation of constraints which account for the most data that no other remaining constraints could also account for — is a principle of much the same sort, though applying with a somewhat different priority and scope than what is necessary for the present case. Hayes's "Favour Activeness" preference — which favors the installation of constraints whose column contains at least one **W** but no **L**'s (given the current support) to constraints whose column has only **e**'s — also bears some similarity.

[36] See Tesar (1995) for the original formulation of and argument for RCD. See also Hayes (2004:169, ex. 2) for a description of the algorithm in plainer English.

(60) RCD Algorithm (Becker 2009:164)

Given a Support *S*,
Given a set of constraints in *S*, *not-yet-ranked constraints*,
H := a new constraint hierarchy.

While *S* is not empty, repeat:

 a. *current-stratum* := all the constraints in *not-yet-ranked constraints* that have at least one W and no L's in their column in *S*

 b. If *current-stratum* ≠ Ø,

  i. remove winner-loser pairs that are assigned a W by any constraint in *current-stratum*.

  ii. put *current-stratum* as the next stratum in *H*, and

  iii. remove *current-stratum* from *not-yet-ranked constraints*

Put *not-yet-ranked constraints* as the next stratum in *H*.
Return *H*.

---

the winner-preferring constraints (or indeed the non–loser-preferring constraints) the one or more constraints which prefer the most winners. These constraints are maximally informative in the sense that their installation will explain the greatest number of data points, i.e. *Winner~Loser* pairs. It is then the *maximally-informative-winners* which are the set of constraints that get installed (61c), rather than the entire "*current-stratum*" as in the standard algorithm.

(61) Maximally Informative Recursive Constraint Demotion (MIRCD)

Given a Support *S*,
Given a set of constraints in *S*, *not-yet-ranked constraints*,
H := a new constraint hierarchy.

While *S* is not empty, repeat:

 a. *current-stratum* := all the constraints in *not-yet-ranked constraints* that have (at least one W and) no L's in their column in *S*

 b. **maximally-informative-winners := all the constraints in current-stratum for which no other constraint in current-stratum has more W's in their column in S**

 c. If *maximally-informative-winners* ≠ Ø,

  i. remove winner-loser pairs that are assigned a W by any constraint in *maximally-informative-winners*.

  ii. put *maximally-informative-winners* as the next stratum in *H*, and

  iii. remove *maximally-informative-winners* from *current stratum*

  iv. return *current stratum* to *not-yet-ranked constraints*

Put *not-yet-ranked constraints* as the next stratum in *H*.
Return *H*.

  This proposed change will only be compatible with Biased Constraint Demotion if the "biased" part of BCD — namely the preferential installation of Markedness constraints before Faithfulness constraints — takes the *maximally-informative-winners* as its input, not if the *maximally-informative-winners* are chosen from among the markedness-biased set of constraints in *current-*

stratum. This is because *PCR is a markedness constraint while CONTIGUITY-BR is a faithfulness constraint (though see the discussion in Section 3.8.6 below regarding these assumptions), yet we need some aspect of the system to prefer CONTIGUITY-BR over *PCR in the present case. Therefore, the preference for markedness over faithfulness must be located after the step in (61b), not the step in (61a). Nonetheless, this shows that the adoption of MIRCD is not completely incompatible with the mechanisms that advocate for the subset grammar, only that it overrides this mechanism in one particular case, namely, when multiple winner-preferring constraints differ in their explanatory power. Further consideration of exactly how this impacts our view of the subset principle in phonological learning and what other empirical facts bear on this question is an important direction for future research.

3.8.3 From Proto-Anatolian to Hittite

With the MIRCD algorithm in place, I will now demonstrate how this proposal actually derives the proper result for *PCR in Pre-Hittite. The violation profile in (62) shows two candidate comparisons for the two cluster-initial base types at the stage following the change from C_1-copying to cluster-copying in TRVX– bases, i.e. "Pre-Hittite I" from Section 3.8.1 above. The comparisons are (i) between the winning cluster-copying candidate and the losing C_1-copying candidate, and (ii) between the winning cluster-copying candidate and the losing "over-copying" candidate.

It will be helpful to consider the relationship between ALIGN-ROOT-L (abbreviated ALIGN) and MAX-BR (abbreviated MAX), so these derivations will assume that the base has additional copyable material after the first base vowel — i.e. specifically CCVCV– rather than just CCVX–. The "over-copying" candidate is the one which has copied the second syllable, and thus incurs extra violations of ALIGN-ROOT-L relative to the winning cluster-copying candidate. In addition to ALIGN-ROOT-L and MAX-BR, the violation profile in (62) and the tableaux that follow include the violation profile of these *Winner∼Loser* pairs with respect to *PCR and CONTIGUITY-BR (abbreviated CONTIG).

(62)   MIRCD for Pre-Hittite I: Initial Support

		*PCR	CONTIG	ALIGN	MAX
i.	TRVCV- → TRV-TRVCV- ≻ TV-TRVCV-	e	W	L	W
ii.	TRVCV- → TRV-TRVCV- ≻ TRVCV-TRVCV-	e	e	W	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	W	L	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	e	W	L

With traditional RCD, both *PCR and CONTIG would be installed in the first stratum, as both of them prefer no losers, and indeed both prefer at least one winner (i.e., their columns have only **W**'s and **e**'s). This is not what we want, since this ranking would not be consistent with the *VR-VRT*– pattern. However, if we employ the MIRCD update, which installs only the *maximally-informative-winners*, we achieve a different result, namely, the one we are looking for. While *PCR has one **W** in its column, CONTIG has two. This means that only CONTIG belongs to the *maximally-informative-winners* set, and it alone gets installed. This is shown in (63).

(63) MIRCD for Pre-Hittite I: Install *maximally-informative-winners*, i.e. CONTIG

		CONTIG	*PCR	ALIGN	MAX
i.	TRVCV- → TRV-TRVCV- ≻ TV-TRVCV-	W	e	L	W
ii.	TRVCV- → TRV-TRVCV- ≻ TRVCV-TRVCV-	e	e	W	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	W	L	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	e	W	L

The gray rows in (63) indicate the *Winner~Loser* pairs which are removed from the support by the installation of CONTIG. Crucially, the single *Winner~Loser* pair which provides *PCR with its **W** (STV-STVCV- ≻ SV-STVCV-) is among those removed from the support. This means that, in subsequent iterations of the algorithm, *PCR will *never* get installed in a stratum above another constraint, since it will never again be a constraint that actively prefers a winner — even though, of course, it prefers no losers. Put another way, installing CONTIG first has entirely robbed *PCR of its *informativity*: all remaining *Winner~Loser* pairs in the support fare equally well on *PCR.

Among the remaining *Winner~Loser* pairs, ALIGN is now a winner-preferring constraint. Since it is in fact the only winner-preferring constraint, it is identified as a member of the *maximally-informative-winners* set, and gets installed, as shown in (64) below. Installing ALIGN takes care of the remaining support, and thus MIRCD completes without having installed *PCR (or MAX) in an upper stratum. *PCR is ranked at the very bottom of the grammar. This is exactly the result we were trying to derive. The ranking in (64) is precisely the one we require in order to generate the *VR-VRT–* pattern at the subsequent stages.

(64) MIRCD for Pre-Hittite I: Install *maximally-informative-winners*, i.e. ALIGN

		CONTIG	ALIGN	*PCR	MAX
i.	TRVCV- → TRV-TRVCV- ≻ TV-TRVCV-	W	L	e	W
ii.	TRVCV- → TRV-TRVCV- ≻ TRVCV-TRVCV-	e	W	e	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	L	W	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	W	e	L

To summarize, the MIRCD update of RCD has a property we might call "informativity-based constraint bounding". If a constraint $\mathbb{C}_1$ is informative (i.e. winner-preferring) about some non-zero set of *Winner~Loser* pairs, and another constraint $\mathbb{C}_2$ is informative about a proper *superset* of those *Winner~Loser* pairs (i.e. all the *Winner~Loser* pairs about which $\mathbb{C}_1$ is informative, and one or more additional *Winner~Loser* pairs), $\mathbb{C}_1$ will always be installed in the very lowest stratum of the ranking. This is because the preference for installing *maximally-informative-winners* will always install $\mathbb{C}_2$ first, which will remove all **W**'s from $\mathbb{C}_1$'s support.[37]

3.8.4 (MI)RCD in Proto-Anatolian

When the situation of informativity-based constraint bounding is not present in the support, MIRCD may be functionally equivalent to standard RCD. We can briefly illustrate one case of this with Proto-Anatolian. In this stage, where TRVX– bases show C_1-copying but STVX– bases show

---

[37] It might be the case that $\mathbb{C}_1$ could potentially avoid this fate if there is another constraint $\mathbb{C}_3$ which (i) is a uniquely winner-preferring constraint, (ii) has **W**'s in all of *Winner~Loser* pairs for which $\mathbb{C}_2$ has a **W** but $\mathbb{C}_1$ does not, and (iii) has a greater overall number of **W**'s than $\mathbb{C}_2$. Such a $\mathbb{C}_3$ would be installed before $\mathbb{C}_2$ and remove the support which differentiates $\mathbb{C}_2$ from $\mathbb{C}_1$. Whether this, or other formal properties of the proposal, have any significant ramifications, I leave as a question for future inquiry.

cluster-copying, *PCR and CONTIGUITY-BR do not have the same scope, as can be seen in (65). CONTIGUITY-BR is now not a uniquely winner-preferring constraint, because it is violated by the winning C_1-copying candidate for TRVX– bases. This lets *PCR, with its one **W** and no **L**'s, be installed on the first round, since there is no other winner-preferrer with more **W**'s (in fact, there are no other uniquely winner-preferring constraints in the initial support). This is reflected in (66). Lastly, ALIGN takes care of the remaining support, and thus MIRCD completes, as shown in (67). This properly yields the ranking needed for Proto-Anatolian, as developed in Section 3.7.

(65)   MIRCD for Proto-Anatolian: Initial Support

		*PCR	CONTIG	ALIGN	MAX
i.	TRVCV- → TV-TRVCV- ≻ TRV-TRVCV-	e	L	W	L
ii.	TRVCV- → TV-TRVCV- ≻ TRVCV-TRVCV-	e	L	W	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	W	L	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	e	W	L

(66)   MIRCD for Proto-Anatolian: Install *maximally-informative-winners*, i.e. *PCR

		*PCR	CONTIG	ALIGN	MAX
i.	TRVCV- → TV-TRVCV- ≻ TRV-TRVCV-	e	L	W	L
ii.	TRVCV- → TV-TRVCV- ≻ TRVCV-TRVCV-	e	L	W	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	W	L	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	e	W	L

(67)   MIRCD for Proto-Anatolian: Install *maximally-informative-winners*, i.e. ALIGN

		*PCR	ALIGN	CONTIG	MAX
i.	TRVCV- → TV-TRVCV- ≻ TRV-TRVCV-	e	W	L	L
ii.	TRVCV- → TV-TRVCV- ≻ TRVCV-TRVCV-	e	W	L	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	L	W	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	W	e	L

Notice that the invocation of *maximally-informative-winners* was never required for this data, as at each step of the process, there was only ever one constraint which was uniquely winner-preferring. Therefore, in at least cases of this sort, MIRCD and RCD are completely equivalent.

3.8.5   From Proto-Anatolian to Luwian

In the development into Hittite, the crucial means of demoting *PCR to the bottom of the grammar was getting the independently necessary high ranking of CONTIGUITY-BR to depress the ranking of *PCR via MIRCD. In other words, *PCR's status as a winner-preferring constraint was eliminated when the support which was explained by CONTIGUITY-BR was removed after CONTIGUITY-BR's installation. The total demotion of *PCR into Luwian can also be explained in terms of the removal of support for *PCR. However, rather than being removed in the process of MIRCD, it is (at least in part) removed from the language via sound change.

As argued for in Section 3.6.1 above, (Pre-)Luwian lost the synchronic contrast between roots of the shape /STVX/ and /TVX/, due to the operation of a categorical rule of /s/-deletion. Insofar as any remnants of the earlier treatment of */STVX/ roots/bases remain in the language, they are treated as frozen archaisms, perhaps not even analyzed as having ever been reduplicated.

This means that, when learners were constructing their reduplicative grammar with respect to cluster-initial bases, the only data which they could have taken into account was the data for TRVX– bases, which still show C_1-copying.[38]

The initial support for (MI)RCD would, then, be equivalent to that of Proto-Anatolian (65), *but without any Winner~Loser pairs for STVX– bases*. This is reflected in (68) by the grayed out rows. Here it is now important whether basic RCD selects all non–loser-preferring constraints (the standard formulation) or just all winner-preferring constraints (Becker 2009's formulation). If we assume that it selects only those constraints which actively prefer winners, then nothing further needs to be said. This will uniquely select ALIGN for installation in the top stratum. This explains the entirety of the support, so *PCR and the other constraints get installed in the bottom stratum, as shown in (69).

(68)   MIRCD for Pre-Luwian: Initial Support

		*PCR	CONTIG	ALIGN	MAX
i.	TRVCV- → TV-TRVCV- ≻ TRV-TRVCV-	e	L	W	L
ii.	TRVCV- → TV-TRVCV- ≻ TRVCV-TRVCV-	e	L	W	L
i.	STVCV- → STV-STVCV- ≻ SV-STVCV-	W	W	L	W
ii.	STVCV- → STV-STVCV- ≻ STVCV-STVCV-	e	e	W	L

(69)   MIRCD for Pre-Luwian: Install ALIGN

		ALIGN	*PCR	CONTIG	MAX
i.	TRVCV- → TV-TRVCV- ≻ TRV-TRVCV-	W	e	L	L
ii.	TRVCV- → TV-TRVCV- ≻ TRVCV-TRVCV-	W	e	L	L

If, however, we used the standard formulation of RCD, which installs all constraints which have no **L**'s among the current support, this would treat ALIGN and *PCR equally on the first run. This would result in installing *PCR in the top stratum, crucially ranked above CONTIG. This is the opposite of the ranking required for Luwian (see Section 3.5), and cannot generate the *VR-VRT*– pattern. Therefore, the apparent facts of (Pre-)Luwian are inconsistent with the standard formulation of RCD. They require some update to the algorithm that prefers the installation of winner-preferrers over constraints which have only **e**'s. The Becker approach is one such solution; MIRCD is another. MIRCD will here be equivalent to Becker's version of RCD, because they both have a mechanism for selecting a constraint with **W**'s (or **W**'s and **e**'s) over a constraint with just **e**'s. Insofar as MIRCD can be viewed as a ramped up version of Becker's RCD, Luwian therefore also provides evidence in favor of adjusting RCD in the direction of MIRCD.[39]

MIRCD is thus consistent with the development of Luwian argued for above. A summary of the proposed changes from Proto-Anatolian to Luwian and their relative chronology is given in (70).

---

[38] They would of course also be considering the behavior of CVX– bases, but this will give no additional information beyond TRVX– bases: the only relevant *Winner~Loser* pair is the one which demonstrates the preference for not copying post-nuclear segments, a preference equally observable from TRVX– bases.

[39] Note that the relative ranking of CONTIG and *PCR is under-determined based purely on applying either Becker's RCD or MIRCD to the data from TRVX– bases. If we want to claim that anything more significant than chance was responsible for fixing the ranking as CONTIG ≫ *PCR, something further will need to be said.

(70)　Luwian relative chronology

Stage		Ranking
(I)	Proto-Anatolian	$\boxed{\text{*PCR} \gg \text{ALIGN-ROOT-L} \gg \text{CONTIG-BR}}$

- *Emergence of rule deleting *s in *#ST*
- *Deletion rule becomes categorical, eliminating evidence for /s/*
- **/STVX/ roots are restructured as /TVX/*

(II)	Pre-Luwian I	$\boxed{\text{*PCR ?? ALIGN-ROOT-L} \gg \text{CONTIG-BR}}$

- *No synchronic contrast in cluster-initial roots, thus no evidence for *PCR*
- **PCR is demoted to the bottom of the grammar*

(III)	Pre-Luwian II	$\boxed{\text{ALIGN-ROOT-L} \gg \text{CONTIG-BR} \gg \text{*PCR}}$

- **h₁ deletes / #_V*
- *Newly vowel-initial roots fed into grammar, generate <u>VC</u>-VCX– pattern*

(IV)	Pre-Luwian III / Luwian	$\boxed{\text{ALIGN-ROOT-L} \gg \text{CONTIG-BR} \gg \text{*PCR}}$

3.8.6　MIRCD or a Bias for BR-Faithfulness?

There is one way in which the demotion of *PCR might be generable without a change to the RCD algorithm *per se*, namely, by appealing to Biased Constraint Demotion, or some other learning procedure that can incorporate ranking bias by constraint type. As mentioned earlier, it is frequently assumed (McCarthy 1998, Hayes 2004) that learners have a bias towards the high ranking of Output-Output faithfulness constraints, over not only Input-Output faithfulness constraints, but also over markedness constraints. The two constraints whose relative ranking is at stake in the development from Proto-Anatolian to Hittite and Luwian, respectively, are *PCR and CONTIGUITY-BR. If CONTIGUITY-BR — and Base-Reduplicant faithfulness constraints generally — are to be classified as Output-Output faithfulness constraints for the purposes of learning biases, and *PCR is to be classified as a markedness constraint, then a learning procedure which implements the bias for Output-Output faithfulness constraints will install CONTIGUITY-BR before *PCR when they explain the same set of data. This is the ranking we are trying to derive in both cases. Therefore, if these assumptions hold, it would seem that we might not need to revise our learning procedures, and that this might not really represent a counterexample to the subset principle. However, there is (to my knowledge) little to no work on the question of whether Base-Reduplicant faithfulness constraints should truly be afforded the Output-Output faithfulness bias.

One reason to doubt this approach may be the behavior of MAX-BR. MAX-BR plays no role in any of the Indo-European reduplicative systems, and indeed is wantonly violated in all languages with partial reduplication patterns. If there were an underlying bias towards the high ranking of BR-faithfulness constraints, and MAX-BR is properly characterized as a member of that set, then we should perhaps expect partial reduplication (or at least the sort of minimal partial reduplication evident in Indo-European; see also Hendricks 1999 on other sorts of minimal reduplication patterns) to be much less frequent than it actually is.

Note also that, even if we switch to BCD, the Luwian facts still require Becker's version of the constraint selection step, i.e. preferentially selecting active winner-preferrers. If not, then *PCR would incorrectly be selected in the first run. Given that the difference between the Becker update

and MIRCD is fairly small, this suggests that MIRCD could still be a reasonable proposal even if a ranking bias for BR-faithfulness constraints is justified.

3.8.7 Additional Support for the Late Demotion of *PCR (in Hittite)

The diachronic account developed in this chapter argues that *PCR was demoted to the bottom of the rankings independently in Hittite and Luwian. Additional support for this notion comes from the existence of several consonant-initial *PCR-violating reduplicative forms that seem to have arisen within the internal (pre-)history of Hittite. These are given in (71) below.

(71)   Consonant-initial *PCR-violating forms
      a. *titḫa*   [tit̬χa]   'thunders' (3SG.NPST.MID)
      b. *lilḫuwai* [lil̬χʷai] 'pours'   (3SG.NPST.MID)   (simplex *laḫ(ḫ)u–* 'pour')

In the case of (71a), the base-initial cluster [tχ] that leads to a *PCR violation is demonstrably not reconstructible to PIE/Proto-Anatolian, since $*h_2$ should have been lost in that sequence: PIE $*Th_2V$ > PA $*TV$, as evidenced by Hittite *paltāna–* 'shoulder' < $*pl̥th_2\text{-}eno–$ (cf. Melchert 1994:69). As argued by Oettinger (1979:514) and Dempsey (2015:304–306), (71a) Hittite *titḫa* is cognate with Vedic √*tan^i–* 'thunder', and both derive from PIE √$*(s)tenh_2–$. The Hittite form can thus be derived via the developments in (72).

(72)   PIE $*ti\text{-}tn̥h_2\text{-}o$ > Pre-Hittite $*tit\text{:}aχ\text{:}a$ > Hittite *titḫa*

If we assume that the syncope of the vocalized syllabic nasal ($*a < *n̥$)[40] in (72) is a Hittite-internal development, this explains the synchronically unusual stem-final cluster *–tḫV–*. We might think that, if *PCR were active at the stage when this syncope sought to apply, the syncope might have been blocked, as it creates a new *PCR-violating sequence [titχ]. At an early prehistoric stage (PA/Pre-Hittite I; see (59) above), syncope would likely have been blocked by *PCR. But when *PCR was subsequently demoted within the internal development of Hittite, the blocking effect ceased to apply, and syncope occurred. While this is again suggestive of late *PCR demotion, the conditions for syncope in (Pre-)Hittite remain too poorly understood for any certainty.[41]

A similar set of developments may have lead to the emergence of (71b) Hittite *lilḫuwai–* (< PIE √$*leh_3\text{-}w–$). Reduplicative forms of (at least) the shape *LV-LT–* (L = liquid) are systematically unattested in the Indo-European languages (cf. Sandell 2014b).[42] For instance, there is no **lelǵ–* (to the root of Latin *lēgī*, Tocharian B *lyāka*, etc.) although there are several morphological contexts (e.g. PIE perfect weak stems) where this configuration theoretically could have arisen.[43] If this gap is non-accidental and driven by *PCR at the PIE level, then [lilχʷ-] could only have emerged within Hittite after the demotion of *PCR (either via historical syncope or generated synchronically). Potentially relevant here is (hapax) spelling <li-la-ḫu-i>, the oldest attested form of the verb (cf. Güterbock & Hoffner 1989:57), which *could* directly reflect a pre-syncope stage.[44]

It must be noted that, at least under the analysis advanced in this chapter, it is unclear whether these forms can be taken as evidence of the productive reduplication patterns of Hittite. As argued

---

[40] See Melchert (2012:181–182) for evidence in favor of this treatment of syllabic nasals in Hittite.

[41] An alternative possibility (suggested by Ryan Sandell, personal communication) is that the application of syncope itself in Pre-Hittite I opacified *PCR at this stage and contributed to its ultimate demotion.

[42] Perhaps the only exceptions are Old English dialectal forms like *leort-* < *le̲-lt-* (Jasanoff 2007); see Chapter 4.

[43] See Jasanoff (2012) for discussion, with a different conclusion.

[44] As shown by Melchert (2011:130), the etymologically related [i]-reduplicated stem CLuw. *li-lūwa–* is a Luwian-internal creation to the (laryngeal-metathesized) base *lūwa–* 'pour' (< PIE $*luh_3–$ ← $*lh_3u\text{-}C–$), and thus provides no evidence for an inherited reduplicated [i]-present with zero-grade of the root (as argued by Dempsey 2015:294).

for throughout, the normal treatment of cluster-initial bases in Hittite is cluster-copying. In the case of STVX– bases, where cluster-copying would result in an illicit word-initial string, prothesis is additionally employed to repair it. Assuming that the base-initial clusters for the forms in (71) are not phonotactically licit in word-initial position (see again Kavitskaya 2001, Yates 2014, 2016 on Hittite cluster phonotactics), we should thus probably expect here **[itχi-tχa] and **[ilχi-lχʷai] with cluster-copying plus prothesis. If, though, it could be claimed that, in these cases alone, the reduplicative vowel is not a copy of the base vowel but rather either specified underlyingly or arising via epenthesis, then we could actually generate the C_1-copying forms synchronically, since CONTIGUITY-BR would not motivate copying the second member of the cluster. Absent this, we must assert that these two forms are not synchronically generated reduplicated forms, but rather unanalyzed archaisms. Regardless of their synchronic status, though, they could not have entered the language if their creation pre-dated the demise of *PCR. These forms thus lend support to the overall chronology developed in this chapter.

3.8.8 Hittite *šip(p)and-* and the Ranking of *PCR before Proto-Anatolian

In Chapter 7, I will argue, based on the comparative evidence from across the Indo-European languages, that Proto-Indo-European exhibited C_1-copying for STVX– bases — i.e., *SV-STVX–*. This would require that, for PIE proper, *PCR was not high-enough ranked to divert STVX– bases away from the default C_1-copying pattern. Nevertheless, Proto-Anatolian comes to have a *PCR constraint that is indeed high-enough ranked to force cluster-copying (*STV-STVX–*) in these circumstances.

Evidence for this reconstruction of the PIE stage, and for the transition between these two stages, may come from the Hittite root √*šip(p)and-* 'libate'. Melchert (2016) argues that this root should be analyzed as deriving from an earlier reduplicated formation *se-spónd-*, which then underwent deletion of root-C_1 (possibly with compensatory lengthening) to *se(:)pónd-*. This process would perhaps be equivalent to what we observe in the CēC patterns in Sanskrit and Germanic (see Chapter 5). Such a process can be analyzed as being driven by *PCR, as the sequence repaired with deletion is exactly that which would induce a *PCR violation.

If this analysis is correct, this requires us to posit that Anatolian inherited the *SV-STVX–* C_1-copying pattern from PIE, and then later underwent the *PCR-driven deletion process prior to changing to the *STV-STVX–* cluster-copying pattern. While the evidence is perhaps too limited to assert much more about the transitional stage between PIE and Proto-Anatolian, this would suggest that the process which shaped *šip(p)and-* might have been the first evidence Anatolian learners received for the activity of *PCR, and that this was later generalized in a different way to create the cluster-copying pattern.

3.9 Conclusion

The reduplicative patterns of the Anatolian languages differ from those reconstructible for Proto-Anatolian. This is indicative of *grammar change*. Like the other Indo-European languages examined in this dissertation, Proto-Anatolian reduplication shows effects of *PCR. Yet, by the period of attested Hittite and Luwian, *PCR has been demoted to the bottom of the grammar, as shown especially by the innovative *VC-VCX–* reduplicative pattern. I have argued that these developments can be accounted for with the following diachronic scenario.

First, independent phonological changes in both Hittite and Luwian affected the reduplication of cluster-initial roots: the development of across-the-board cluster copying in Hittite; the deletion of word-initial *s in Luwian. This produced ambiguities in the learning data with respect to *PCR. By adopting a version of Recursive Constraint Demotion that favors installation of maximally informative winner-preferring constraints over less informative winner-preferring constraints — which I have termed Maximally Informative Recursive Constraint Demotion (MIRCD) — this ambiguity leads directly to the demotion of *PCR to the bottom of the rankings, in both languages. Second, the *VC*-*VCX*– reduplicative pattern emerged independently in each language after the post-Proto-Anatolian loss of pre-vocalic word-initial *h_1, when newly vowel-initial roots were input into the innovative synchronic grammar for the first time. This innovation was made possible (and indeed perhaps necessary) by MIRCD's total demotion of *PCR.

Chapter 4

Germanic

4.1 Introduction

The reduplicated verbal forms attested in Gothic are but one piece of the complex system of verbal morphology and morphophonology found across the Germanic languages. The verbal systems of the Germanic languages are traditionally divided into two types of verbs: "weak" and "strong" (cf. Prokosch 1939:159–203). The classification of a given verb is principally determined by how it forms its preterite stem(s). Weak verbs form their preterites through affixation of the "dental preterite" suffix. On the other hand, as previewed in (1), strong verbs form their preterite stems through a variety of phonological changes applied to the verbal root.

(1)   Example Gothic Strong Verb paradigms

Root Shape	1SG.PRES		3SG.PRET		3PL.PRET	
/CiRC/	*binda*	[bind-a]	*band*	[band]	*bundun*	[bund-un]
/CiC/	*giba*	[giv-a]	*gaf*	[gaf]	*gebun*	[geːβ-un]
/CaC/	*daba*	[dav-a]	*dof*	[doːf]	*dobun*	[doːβ-un]
/CaRC/	*haita*	[hɛːt-a]	*haihait*	[h̲ɛhɛːt]	*haihaitun*	[h̲ɛhɛːt-un]
/CoːC/	*floka*	[floːk-a]	*faiflok*	[f̲ɛfloːk]	*faiflokun*	[f̲ɛfloːk-un]

Traditional descriptions identify seven classes of strong verbs, each with a somewhat different pattern of morphophonological marking. However, these classes can also be defined in terms of the phonological properties of the roots involved. In this chapter, I will show that the particular phonological change which marks a given strong preterite stem can be directly *predicted* by the phonological properties of the verbal root. For this reason, I will propose that the strong preterites select for a null underlying representation of the morpheme PRETERITE (and also of the morpheme SINGULAR in the context of PRETERITE and INDICATIVE), and that differentiation of stems is induced by a family of constraints that require overt exponence of morphosyntactic features, REALIZE MORPHEME (RM; Kurisu 2001), and thus phonological contrast between stems which are morphologically related in a particular way. The nature of the changes undergone to satisfy RM, of which reduplication is only one of many (and indeed a sort of "last resort" synchronically),

---

falls out from the interaction between the phonological properties of individual roots and the ranking of markedness and faithfulness constraints.

The chapter is structured as follows. Section 4.2 introduces the Gothic verbal system, focusing on the distribution of stem formation patterns in the strong preterites. Section 4.3 defines and motivates the REALIZE MORPHEME constraints that will be central to the proposal, and previews how they will deployed in the analysis. Section 4.4 then discusses how various sound changes within the (relatively late) pre-history of Gothic and Proto-Germanic have opacified some of the relationships that are important for the analysis. In considering which changes should or should not be included in the representations over which the analysis will be constructed, which will be identified as those of "Pre-Proto-Germanic", I will argue that only properties which are phonologically contrastive can play a role in the determination of stem contrast. The section concludes with a statement of the Pre-Proto-Germanic representations of the strong verb patterns, and the details of the vowel system assumed for that stage.

The analysis, set in Pre-Proto-Germanic, is laid out in detail in Section 4.5, accounting for the full range of strong preterite stem formation patterns (with the exception of the subclasses VIIb and VIId, which are problematic for various reasons, and postponed until Section 4.7). Section 4.6 demonstrates that the patterns of reduplicant shape, at least as attested in Gothic, are completely consistent with the stem-formation analysis developed in this chapter, and furthermore exhibit the effects of the NO POORLY-CUED REPETITIONS (*PCR) constraint that was shown to be crucial for the analysis of Ancient Greek reduplication in Chapter 2 and (Proto-)Anatolian in Chapter 3. This section also reviews the evidence of reduplication from Northwest Germanic, but concludes that there is not sufficient evidence to draw any significant conclusions.

Section 4.7 considers the synchronic analysis and the diachronic source of the problematic Class VIId pattern, but ultimately is unable to find an analysis which is consistent with the larger proposal. Section 4.8 considers an alternative analysis of the stem formation patterns based on context-sensitive allomorphy in the underlying representations of the morphemes involved in the preterite derivations. It is shown that such an analysis has a number of shortcomings, and should not be preferred to the REALIZE MORPHEME-based analysis, at least for the relevant stage of Pre-Proto-Germanic.

4.2 The Germanic Verbal System

In Gothic, and Germanic generally, the verbal system is divided into two types: *weak* verbs, which are morphologically derived formations; and *strong* verbs, which are root formations. This section lays out the morphological and phonological details of both types, both in Gothic terms and in (Pre-)Proto-Germanic terms. The goal of this will be to show that the morphophonological patterns observable within the strong verb system represent a coherent synchronic morphophonological system that is capable of being analyzed in a consistent and exhaustive fashion.

4.2.1 Verbal Categories and Inflection in Gothic

Before proceeding to the interaction between morphological categories and the phonology of the language, it will be helpful to enumerate the morphological distinctions made within the Gothic verbal system. The description of Gothic is based primarily on Lambdin (2006:esp. 15–17).[1] These

---

[1] Many forms in this section are drawn from Jared Klein's lecture notes (Klein 2012).

generalizations largely apply also to the reconstructed state of Proto-Germanic (see Ringe 2006), which will ultimately be the primary locus of the analysis.

The most central distinction within the Gothic verbal system, and the one which will be most significant for this chapter, is between two tense stems: the PRESENT (which is sometimes, and perhaps more aptly, referred to as the "non-past") and the PRETERITE. Both present and preterite stems form indicatives and subjunctives, distinguished primarily by their subject agreement suffixes. Additionally, the present stem forms the basis of the passive, the infinitive, and the imperative.[2]

As has long been noted (see, e.g., Meid 1971), in all cases, the present stem is distinct from the preterite stem. For the weak verbs, this is effected via the addition of the "dental preterite" suffix. For the strong verbs, these distinctions are reflected instead via direct phonological alternations affecting the verbal root (and/or reduplication). Across both strong and weak verbs, the present stem is invariant across person/number agreement forms, as well as across indicative and subjunctive (see more immediately below). The preterite indicative, however, generally shows a distinction between the stem of the singular and the stem of the plural/dual. In the weak verbs, this is effected through allomorphy of the "dental preterite" suffix ([-d-] in the singular, [-deːd-] elsewhere). On the other hand, in the strong verbs, this is instead effected through vocalic alternations (with the principled exception of Classes VI & VII, which show no alternation between preterite stems because of the phonological properties of their roots).

The following fact is important for understanding the preterite: for both strong and weak verbs, the stem that surfaces in the plural/dual indicative is the one which is used for the subjunctive, in all person/number categories (including the singular). This strongly implies that the plural/dual stem is the *default* preterite stem, as advocated by Ringe (2006). Therefore, the indicative singular preterite stem is the "marked" preterite stem, and thus the one which must be, in a particular sense, more morphologically derived than the other. The REALIZE MORPHEME-based analysis I develop in this chapter requires that there be a subset~superset relationship with respect to morphosyntactic features comprising the two preterite stems. *A priori*, this requirement could be compatible with either of the two stems being identified as the subset and the other as the superset.[3] This stem distribution, however, answers the question: the singular indicative stem must be the superset, since the other stem is found in a heterogeneous set of morphological contexts.

4.2.2 The Gothic Weak Verbs

In Gothic, as throughout Germanic, the weak and strong verbs are differentiated both morphologically and phonologically (consult generally Prokosch 1939, Lambdin 2006, Ringe 2006, *a.o.*). The weak verbs are morphologically derived (e.g., denominal, deadjectival, causative, fientive, etc.), and display a stem-forming suffix between the root and the inflectional suffixes. This stem-forming suffix appears both in the present, as shown in (2), and in the preterite, as shown in (3).[4] (Full inflectional paradigms for Weak Classes I–IV, respectively, in both present and preterite, indicative and subjunctive, can be found in Appendix II.)

---

[2] There are also participles in the present (active) and preterite (passive). I do not address these categories in this chapter.

[3] In Zukoff & Sandell (2015), failing to take the subjunctive facts into account, we indeed proposed the opposite relationship. This still permitted a consistent analysis.

[4] A small group of strong verbs display a stem-forming element in the present (e.g. /j/ as in [hafj-an] 'seize') while lacking it in the preterite. In general, however, these stem-forming elements are restricted to the weak verbs.

(2) Weak Verbs in the present

Class	Morphological Marker	1PL PRESENT INDICATIVE
Weak Class I	/-j-/	[nas-j-am] 'we save'
Weak Class II	/-oː-/	[salb-oː-m] 'we anoint'
Weak Class III	/-a(i)-/	[hab-a-m] 'we have'
Weak Class IV	/-n-/ (in pres.) ∼ /-noː-/ (in pret.)	[full-n-am] 'we fill'

(3) Weak Verbs in the preterite

Class	1SG PRET INDICATIVE		1PL PRET INDICATIVE	
Weak Class I	*nasida*	[nas-i-ð-a]	*nasidedum*	[nas-i-ðeː-ð-um]
Weak Class II	*salboda*	[salb-oː-ð-a]	*salbodedum*	[salb-oː-ðeː-ð-um]
Weak Class III	*habaida*	[hav-ai-ð-a]	*habaidedum*	[hav-ai-ðeː-ð-um]
Weak Class IV	*fullnoda*	[full-noː-ð-a]	*fullnodedum*	[full-noː-ðeː-ð-um]

Weak verbs build their preterites with the "dental preterite" suffix, which appears as [-d-] in the preterite singular indicative and [-deːd-] elsewhere, represented below with the plural indicative. This attaches outside of the stem-forming derivational suffix, as shown in (3). The strong preterites, on the other hand, lack the dental suffix, and instead form their preterite stems with phonological changes applied to the root. I turn now to the strong preterites.

4.2.3 The Gothic Strong Verbs

The system of strong verbs is traditionally divided into seven classes (with approximately four sub-types within Class VII, which are traditionally referred to as the "reduplicating preterites" when discussing Gothic). These divisions are based on the phonological properties of the verbal root — specifically, the quality and quantity of the root vowel, and the number and types of consonants that follow the root vowel. These properties in turn determine how the preterite is formed.

The table in (4) below provides example partial surface paradigms for each of the seven+ classes as they are observable in Gothic. The first column of forms shows the first person singular present forms, which represents the basic stem, and, by hypothesis, the verbal root itself (see Section 4.8.2 for discussion); the same stem is also found in the infinitive. The second and third columns represent the stems of the preterite singular and preterite plural, respectively. In some of the cases (Classes I–V), the two stems are distinct from each other, while, in others (Classes VI & VII), they are identical. Orthographic forms are given in italics, while brackets enclose the phonetic transcription (as best as I can reconstruct it, per *communis opinio*; consult, e.g., Marchand 1973).[5]

---

[5] Note that my division into lettered subclasses in Class VII is a matter of convenience and does not (knowingly) correspond to any traditional notation. This lettering happens to partially overlap with that used by Jasanoff (2007:247), but this is accidental. When comparing this work to other previous works, the reader should expect to find Class VII taken as a single (if internally diverse) unit.

(4)   Gothic Strong Verbs: surface forms[6]

Root Shape	Class	1SG.PRES		3SG.PRET		3PL.PRET	
/CiRC/	I	*beita*	[biːt-a]	*bait*	[bɛːt]	*bitun*	[bit-un]
	II	*kiusa*	[ki̯u̯s-a]	*kaus*	[kɔːs]	*kusun*	[kus-un]
	III	*binda*	[bind-a]	*band*	[band]	*bundun*	[bund-un]
/CiC/	IV	*nima*	[nim-a]	*nam*	[nam]	*nemun*	[neːm-un]
	V	*giba*	[giv-a]	*gaf*	[gaf]	*gebun*	[geːβ-un]
/CaC/	VI	*daba*	[dav-a]	*dof*	[doːf]	*dobun*	[doːβ-un]
/CaRC/	VIIa	*haita*	[hɛːt-a]	*haihait*	[h̪ɛhɛːt]	*haihaitun*	[h̪ɛhɛːt-un]
/CVːC/	VIIb	*slepa*	[sleːp-a]	*saislep*	[s̪ɛsleːp]	*saislepun*	[s̪ɛsleːp-un]
	VIIc	*floka*	[floːk-a]	*faiflok*	[f̪ɛfloːk]	*faiflokun*	[f̪ɛfloːk-un]
	VIId	*leta*	[leːt-a]	*lailot*	[l̪ɛloːt]	*lailotun*	[l̪ɛloːt-un]

The diversity of patterns among the seven classes can be reduced substantially when we group the classes by the broad phonological shape of the root. Classes I–III represent the three different patterns that a from roots of the shape /CiRC/, where R represents the set of sonorants: /j/ (Class I); /w/ (Class II); and /r,l,m,n(,ŋ,ŋʷ)/ (Class III). In the preterite plurals of these classes, the underlying vowel of the root is absent (i.e. deleted). This places the root's medial sonorant between two consonants, where it consequently vocalizes. For underlying glides (Classes I & II), sonorant vocalization yields the corresponding short high vowel: [bitum] ← //bjt-um//, [kusum] ← //kws-um//. The situation in Class III is somewhat less transparent, but precisely equivalent. Syllabic sonorants which are [+consonantal] are realized with a preceding epenthetic [u]: [bundum] ← //bnd-um//.

Similarly, Classes IV and V can be collapsed as roots of the shape /CiC/: Class IV roots are those in which the root-final consonant is a sonorant (/CiR/), while Class V roots are those in which that consonant is an obstruent (/CiT/). Unlike in Classes I–III, the root vowel is not deleted in the preterite plural; on the contrary, it instead lengthens to [eː]. I will demonstrate below that this lengthening is due to the absence of a vocalizable sonorant.

Classes I–V all share the property of having an underlying root vowel /i/ in Gothic terms (*/e/ in Proto-Germanic; see immediately below). This underlying vocalism largely correlates with a preterite singular stem vowel in [a]. The surface exceptions to this — Class I *bait* [bɛːt] and Class II *kaus* [kɔːs] — are illusory; just as suggested by the orthography, these long vowels are the result of coalescence and monophthongization of [a] with the root's glide.[7] Roots with a different underlying root vowel, i.e. Classes VI and VII, do not form their preterite singular by changing the root vowel to [a]. Instead, these roots form their preterite singular stems (and indeed their preterite plural stems, as well) by other means: lengthening of the underlying root vowel /a/ (which surfaces as [oː]) in Class VI, and reduplication in Class VII.

---

[6] The root shapes associated with each class are the canonical root shapes for those classes. Most freely permit having anywhere from zero to two consonants before the root vowel. Some additional forms which do not match the canonical root shapes are found for most classes. See Appendix II for a complete list of strong verb roots in Gothic.

[7] Jay Jasanoff (personal communication) informs me that some Germanicists do not believe that the Gothic spellings *<ai>* and *<au>* should be interpreted as monophthongs [ɛː] and [ɔː], but rather as the (diachronically original) diphthongs [aj] and [aw]. This would be a welcome simplification of the problem.

In this chapter, I will demonstrate how each of these distinct patterns — subject to some slight revision to the data under consideration (see Section 4.4) — can be generated by the drive to establish non-identity between stems for roots that select for a null preterite morpheme. I now proceed to outline the core components of the analysis I will propose.

4.3 Null Morphemes Plus REALIZE MORPHEME

As I will demonstrate explicitly in Section 4.8 when I consider an alternative analysis, any analysis of this system based on underlying allomorphy (i.e. competing context-sensitive Vocabulary Insertion rules) of the PRETERITE and SINGULAR morphemes has major shortcomings. Instead, I propose that this system is to be analyzed with the following two components. First, the underlying representations of the preterite morpheme and the singular morpheme (in the appropriate morphosyntactic context) are phonologically null (i.e. /$\emptyset_{\text{PRET}}$/ and /$\emptyset_{\text{SG}}$/). Second, phonological exponence in the preterite is the result of unfaithful mappings driven by the operation of constraints on contrast between stems (REALIZE MORPHEME constraints), and their interaction with markedness and faithfulness constraints.

The rationale for this approach comes largely in the form of two broad generalizations that hold across the entire verbal system, both strong and weak. First, it has long been noted that the preterite stem is always phonologically distinct from the present stem (see, e.g., Meid 1971). In the weak verbs, this contrast is effected by the addition of the dental suffix without any concomitant phonological changes to the root (or even the derived stem, excepting Weak Class IV). In the strong verbs, this differentiation is effected by the phonological changes to the root described in Sections 4.2.3–4.4.2. Second, there is also a strong tendency for the stem of the (indicative) preterite singular to be distinct from the stem of the preterite plural (which is the default preterite stem; see Section 4.2.1 for the arguments). This universally holds of the weak verbs, reflected in the number-conditioned allomorphy of the dental suffix (plural /-deːd-/ vs. singular /-d-/). It holds also of the strong verbs of Classes I–V, which each have [a] in the preterite singular, but some other phonological differentiation from the present stem in the preterite plural. Strong Classes VI and VII do not follow this generalization, but for principled reasons that will be made explicit below.

These generalizations suggest that *contrast between stems* is a crucial part of the verbal system of Gothic and Proto-Germanic. This contrast can be effected even in the absence of segmental material belonging to some underlying affix, provided that the need for contrast is encoded in the constraint grammar. I propose that this need for contrast is driven by constraints of the type REALIZE MORPHEME, defined as follows:

(5) **REALIZE MORPHEME** (RM; Kurisu 2001:39)
  Let α be a morphological form, β be a morphosyntactic category, and F(α) be the phonological form from which F(α+β) is derived to express a morphosyntactic category β. Then RM is satisfied with respect to β iff F(α+β) $\neq$ F(α) phonologically.

This constraint states that any morphological form containing a strict superset of features of another morphological form must be phonologically distinct from said form. A similar effect could be achieved by employing Base-Derivative "Anti-Faithfulness" constraints (Alderete 2001; cf. Benua 1997);[8] however, the RM approach directly builds in the rationale for the constraint's existence, namely, the realization of morphosyntactic features.

---

[8] See also Crosswhite's (1999) "ANTI-IDENT" proposal, among others.

RM derives the two generalizations discussed above if we make the following assumption: the morphosyntactic features PRETERITE and SINGULAR are visible in the output of the morphological component, but PRESENT and PLURAL (at least in the preterite) are not. This "invisibility" could be due either to privativity of PRETERITE and SINGULAR (i.e. there are no morphosyntactic features PRESENT and PLURAL at any point in the post-syntactic derivation), or deletion of the feature values PRESENT and PLURAL in the morphological component (whether they are themselves privative in the morphosyntax, or rather one value of a binary feature with PRETERITE and SINGULAR, respectively).[9] (See Arregi & Nevins 2012 — especially their Chapter 4 — for details on morphosyntactic deletion and impoverishment operations.) Therefore, RM will be violated if, for a given root, there is no phonological distinction between present stem and preterite stem, or no distinction between a tense stem in the singular and that tense stem in the plural.[10]

However, as mentioned above, the extent to which these contrasts are actually realized across the system varies. The desire for a preterite stem distinct from the present is always actualized, but the desire for a preterite singular stem distinct from the preterite plural fails to be actualized in Strong Classes VI and VII. Therefore, REALIZE MORPHEME must be broken up into constraints on *individual morphosyntactic features*, such that they can be variably ranked in the grammar and thereby generate different distributions. (The constraint family might, therefore, be more aptly named REALIZE MORPHOSYNTACTIC FEATURE.) These are defined in (6).

(6) a. **REALIZE MORPHEME: PRETERITE (RM:PRET)**
         Assign a violation mark * for any preterite stem which is not phonologically distinct from the present stem formed from the same root.

     b. **REALIZE MORPHEME: SINGULAR (RM:SG)**
         Assign a violation mark * for any preterite indicative singular stem which is not phonologically distinct from the default preterite stem (i.e. the preterite plural stem) formed from the same root.

When these constraints are active in the phonological evaluation, they will disfavor the faithful mapping from the underlying form just in case some fixed phonological content from an affix denoting the relevant morphosyntactic property is not available. In the weak verbs, such affixal material is available in the form of the dental preterite suffix. Nonetheless, RM:SG seems to be relevant, in that it could drive the singular~plural allomorphy of the dental suffix:[11]

(7) Allomorphy of the dental preterite suffix
     a. PRETERITE ↔ /-d-/   / __SINGULAR
     b. PRETERITE ↔ /-deːd-/   elsewhere

This pattern could be analyzed in one of two ways. First we could simply employ the Vocabulary Insertion rules shown in (7). The distribution of allomorphs which is output by these rules happens to satisfy RM:SG, so no unfaithful mapping is required. Alternatively, we could assume a view of allomorphy where the morphology provides both allomorphs to the phonological input,

---

[9] Assembling the precise morphological analysis that is required in order to provide the right input to the various phonological derivations is a non-trivial task, which I must unfortunately leave to future work.

[10] This distinction is not observed in the present or in the preterite subjunctive. Therefore, the enforcement of this number contrast must in some way be restricted to the preterite indicative.

[11] Some additional condition on Vocabulary Insertion must determine when one of the allomorphs of the dental suffix is inserted as opposed to the null allomorph. This is the distinction between weak verbs and strong verbs. I believe that this difference can be reduced to the presence of a *v*-domain functional head between the root and T in weak verbs, versus its absence in strong verbs. In other words, the null allomorph is inserted just in case T takes Root (or a complex head consisting of Root + *v* but nothing else) as its complement.

and the phonological grammar selects the candidate output which has the allomorph that allows for the most harmonic mapping. This can be implemented with USE:X constraints (cf. MacBride 2004's "FIAT" constraints). We would require two USE:X constraints (cf. MacBride 2004:26–34), one for each allomorph, with the USE:X constraint for the preferred allomorph (USE:/-deːd-/) outranking the USE:X constraint for the dispreferred allomorph (USE:/-d-/) .

This ranking is sufficient to select the preferred allomorph /-deːd-/ for the default case (exemplified by the preterite plural), as shown in (8) below. (On the motivation for asserting that RM:SG plays no role in the derivation of the default stem, see Section 4.5.4 below.) However, as long as RM:SG outranks USE:/-d-/, the derivation of the preterite singular will be diverted away from the default allomorph — since using it would make the two stems identical — and toward the restricted allomorph /-d-/. This is demonstrated in (9) below. While either approach to the allomorphy is sufficient to generate the data, the USE:X analysis directly encodes the generalization about stem contrast into the derivation (via the application of RM:SG), whereas using contextual Vocabulary Insertion rules captures the generalization only by happenstance.

(8)   Preterite Plural of Weak Verbs (Weak Class I √*nas-j-* 'save')

INPUT: /nas, j, {deːd, d}$_{PRET}$/	RM:SG	USE:/-deːd-/	USE:/-d-/
a. ☞ nas-i-deːd-	*n/a*		*
b. nas-i-d-		*!	

(9)   Preterite Singular of Weak Verbs (Weak Class I √*nas-j-* 'save')

INPUT: /nas, j, {deːd, d}$_{PRET}$, Ø$_{SG}$/ BASE: PRET [nas-i-deːd-]	RM:SG	USE:/-deːd-/	USE:/-d-/
a. nas-i-deːd-	*!		*
b. ☞ nas-i-d-		*	

On the other hand, when no affixal phonological material at all is available — as I claim to be the case for the strong verbs — the manner by which the phonological contrast is effected will be determined purely by the ranking of relevant markedness and faithfulness constraints with respect to the RM constraints. The variety of surface patterns in the strong verbs derives from the way in which the relevant constraint set affects roots of different shapes differently.

4.4   Reconstructing Back to Pre-Proto-Germanic

The generalizations observable from the surface patterns of Gothic itself (i.e. the paradigms given in (4) above) are very nearly sufficient for constructing an analysis based on the components just outlined in Section 4.3. However, as alluded to earlier, there are some (relatively shallow) diachronic phonological developments that have made the system more complex than it might have been when it originated. Given that the majority of the phonological changes that mark the strong preterites are vocalic alternations ("ablaut"), sound changes and new phonological processes involving the vowel system have the potential to significantly affect the transparency (or lack thereof) of the relationship between stem forms, which is crucial to the analysis to be presented.

While I will return in Section 4.5.8.2 to the question of how the analytical system that will be developed can be tweaked to properly account for Gothic itself, it will be significantly more

straightforward to work out the analysis for the system that precedes the relevant sounds changes — or, for those which are synchronic processes, a system which fails to reflect the effects of those changes (see immediately below). Furthermore, since most aspects of the system are, to one degree or another, preserved not just in Gothic but across Germanic, the system must have originated prior to the breakup of the Germanic language family. Therefore, it is desirable to establish an analysis of the system within the development of Proto-Germanic. Since some of the complicating sound changes must be reconstructed to Proto-Germanic, the precise stage of the language at which I will be framing the analysis should be referred to as Pre-Proto-Germanic.[12]

To this end, I first lay out the sound changes which crucially impact the system of strong verb stem formation in Section 4.4.1. Section 4.4.2 will then illustrate the phonological properties of the strong verbs prior to the advent of these sound changes. This will be the data set for which the analysis will be constructed. Lastly, Section 4.4.3 will summarize the vowel system of Pre-Proto-Germanic, which will lay the groundwork for understanding the features over which the constraints will be defined.

But first, an important inference about the nature of the system comes from the interaction (or lack thereof) between consonantal alternations and the requirement for contrast between stems, as required by REALIZE MORPHEME. Consider the consonantal alternations detailed in (10), exemplified for Gothic (though at least the former is reconstructible for Proto-Germanic).

(10)   Consonantal processes in(to) to Gothic

      a.   *Post-vocalic spirantization*
          Voiced obstruents spirantize in post-vocalic position ([-son,+voi] → [+cont] / _V)
          Example: Gothic infinitive /gib-an/ 'to give' → [giv-an]
      b.   *Final devoicing*
          Voiced obstruents devoice word-finally ([-son,+voi] → [-voice] / _#).[13]
          Example: Gothic preterite singular /gab/ 'I gave' → Gothic [gaf]

The analysis to be developed in this chapter is based around constraints that require *phonological* contrast between *morphologically* related stems. Without proper contextualization, we would presume that the operation of the consonantal alternations involving [±continuant] and/or [±voice], when applying to only one member of a related pair, should be enough to satisfy the need for phonological distinctness. But this is never the case. Take, for example, the final devoicing example above: 1SG.PRESENT [giv-a] (< (Pre-)Proto-Germanic *[gev-a]) ∼ 1SG.PRETERITE [gaf]. The analysis will claim that the vocalic alternation is triggered by REALIZE MORPHEME's desire to differentiate the two stems. In order for this approach to succeed, the voicing alternation [v] ∼ [f] controlled by final devoicing must *not* be sufficient to render the two stems distinct in the sense relevant for REALIZE MORPHEME. If a voicing alternation of this sort were sufficient for this purpose, we would expect to see a preterite singular exhibiting just final devoicing, i.e. **[gif] (< (Pre-)Proto-Germanic *[gef]); that is to say, the application of vowel backing over and above

---

[12] "Pre-languages" are entities which are (usually) not definable in absolute terms, but rather only in relative terms (consult generally Fox 1995 on linguistic reconstruction). That is to say, "Pre-Gothic" could in theory refer to any stage in the prior history of Gothic that pre-dates one or more changes that are reconstructible through internal reconstruction. Proto-Germanic could thus also technically be referred to as a Pre-Gothic stage, since it is the stage at which all of the changes affecting Gothic subsequent to Proto-Germanic have not yet taken effect. With this in mind, I use "Pre-Proto-Germanic" to refer to the specific stage that pre-dates and post-dates the relevant sound changes laid out in Section 4.4.1, all of which (other than those uniquely affecting Gothic) are known via internal reconstruction of Proto-Germanic.

[13] This counter-bleeds spirantization, which might suggest we should take the voiced obstruents to be underlyingly specified as [+cont], and reverse the spirantization process to a hardening process in non-post-vocalic position.

devoicing would be a superfluous violation of faithfulness, as the change in voicing is required independently by whatever markedness constraint triggers final devoicing.

Within this system, we never find consonantal alternations of this sort preempting the application of the vocalic changes that typically mark the preterite. This may have significant implications for our understanding of phonological contrast/distinctness with respect to REALIZE MORPHEME. Namely, if we were to generalize from this case, we might posit that automatic alternations are not sufficient to render stems phonologically distinct; that is, some (morpho)phonological process must apply over and above those processes which are triggered by the regular phonology. This makes sense from the point of view of *phonological* contrast. The fact that the final devoicing markedness constraint is in full effect means that the feature [±voice] is not contrastive in final position. It is then reasonable to infer that an alternation involving this feature in this position should not be sufficient to render morphological elements contrastive either, insofar as they rely on phonological contrast for their distinctiveness.

This is bolstered by a property of the analysis that will be proposed below: the types of unfaithful mappings which are sufficient to render stems distinct (i.e. satisfy REALIZE MORPHEME) are all definable in terms of changes to contrastive features/properties. The faithfulness constraints that govern the selection of different mappings (given different phonological properties of inputs) exclusively refer to the following contrastive elements: (i) the contrastive vowel features [±high] and [±back] — all other vowel features present in the inventory (namely [±round] and [±low]) are non-contrastive (see Section 4.4.3 below); (ii) moras, which govern a three-way length distinction in Pre-Proto-Germanic; and (iii) the presence vs. absence of segments (i.e. MAXV, etc.). Therefore, this seems to be additional indication that stem contrast relies on phonological contrast in a meaningful way.

If this notion holds true, then many of the sound changes that I will "reconstruct past" in Section 4.4.1 immediately below might independently be ignored by the stem-contrast system. Other than the change from Proto-Germanic to Gothic that yields a total merger between */e/ and */i/, these changes will largely be processes that neutralize contrasts in particular environments. If there is active contrast neutralization in a position, and stem contrast relies on phonological contrast, then it follows that alternations involving these neutralization processes will not be sufficient to license stem contrast. Therefore, while I will ultimately refer to the representations I analyze as belonging to "Pre-Proto-Germanic", i.e. the stage before all these changes enter the language, the considerations of contrast may allow us to ignore those processes with regard to REALIZE MORPHEME and the stem formation system even after they come into effect, meaning that we might instead be able to refer to the stage simply as Proto-Germanic.

4.4.1 Relevant Sound Changes and Phonological Processes

The problematicity of particular diachronic changes will come into focus only within the context of the full analysis. Discussion of the precise nature of the problems will thus follow the presentation of the analysis, in Section 4.5.8. Nevertheless, it will be important to be explicit about the ways in which the data I analyze differs from that directly observable for Gothic. I will discuss only the changes that are particularly relevant to the cases at hand. For a thorough accounting of the sound changes and phonological processes affecting Proto-Germanic, and their relative chronology, I direct the reader to Ringe (2006:§3.2, esp. Fig. 3.1).

4.4.1.1 The Phonemic Merger of */e/ and */i/ in Gothic

Among the changes I describe in this section, only one affects the underlying phonemic inventory. This is the merger of the Proto-Germanic short front vowels */e/ and */i/, resulting in Gothic /i/, as laid out in (11).

(11)   *Merger of short front vowels*
       Proto-Germanic short */e/ raises and merges with short /i/ in Gothic.
       Example: Proto-Gmc infinitive */nem-an/ → *[neman] > Gothic /nim-an/ → [niman]

This merger significantly impacted the transparency of the system, in that it opacified certain alternations between stem forms, causing them to become saltatory in nature (cf. Hayes & White 2015). That is to say, after the change of the root vowel from [-high] */e/ to [+high] /i/, alternations like Class V *sit-and* (< **set-and*) 'they sit' ∼ *seːt-un* 'they sat' involve the change of two features ([+high,-long] ∼ [-high,+long]) rather than just one ([-long] ∼ [+long]),[14] even though the phonological inventory contains a licit segment which can be arrived at via a change of just one feature: [+high,-long] → [+long] = [iː]. Similarly, the alternation in, for example, Class III *hilp-a* (< **help-a*) 'I help' ∼ *halp* 'I helped' involves the change of two features ([+high,-back] ∼ [-high,+back]) when one would have been sufficient: [+high,-back] → [+back] = [u].[15] The reason why these saltatory alternations are problematic is that, under the system I will propose, these alternations are to be understood as minimal violations of faithfulness compelled by the constraints on contrast between stems. Saltatory changes are non-minimal in their faithfulness violation, and thus cannot be generated without special mechanisms. Working at a stage in which the stem vowel of the present is still [e] is thus crucial to developing a transparent system.[16] Since this merger is specific to Gothic, this requires that the transparent system be located in Proto-Germanic.

4.4.1.2 Raising Processes in (Pre-)Proto-Germanic

While total phonemic merger occurred only in Gothic, there were several conditioned phonological processes that neutralized the /e/∼/i/ contrast (in favor of [i]) already in Proto-Germanic: unstressed raising (12a), *i*-umlaut (12b), and pre-nasal raising (12c). (The language abbreviations used below are: PIE = Proto-Indo-European, PGmc = Proto-Germanic, OE = Old English, Eng. = Modern English, OHG = Old High German.)

(12)   Raising within Proto-Germanic
       a.   *Unstressed raising* (Ringe 2006:122–126, §3.2.5.iii)
            /e/ raises to [i] when unstressed (i.e. in non-initial position).
            Example: 'mice' PIE *múːs-es* > PGmc *ˈmuːs-**iz** (> Pre-OE *ˈmyːs-iz > Eng. *mice*)
       b.   *Umlaut* (Ringe 2006:126–128, §3.2.5.iv)
            /e/ raises to [i] when followed by [i] or [j].
            Example: 'he bears' PIE *bʰér-eti* >> PGmc *ˈbir-ið (> OHG *birid*)

---

[14] I will ultimately adopt a moraic rather than featural approach vowel length for Pre-Proto-Germanic, but the featural parlance will be sufficient to discuss the relevant points here.

[15] This change does also involve a change in [±round], but this feature is non-contrastive (see below).

[16] One could argue that, based on the nature of the alternations, Gothic speakers might have been able to learn underlying representations with /e/ despite this vowel undergoing absolute neutralization on the surface. I will return to questions of this sort in Section 4.5.8.

c. *Pre-nasal raising* (Ringe 2006:149–150, §3.2.7.ii)
   Proto-Germanic /e/ → [i] / __N{C,#}.
   Example: PIE 'five' *$pénk^we$ > PGmc *$fimf$ (> Gothic *fimf*, OE *fīf*)

While any of these processes in theory had the potential to alter underlying representations (especially unstressed raising, since stress was fixed on the initial syllable, and thus there could be no alternations), none would strictly prohibit learning /e/ as the underlying vowel when alternations were involved. That is to say, even if, for example, umlaut applied to the vowel of the verbal root in cases like PGmc 3SG.PRES 'bir-ið, other paradigmatically related forms where umlaut would not apply (e.g. PGmc 1SG.PRES 'ber-a) would support positing underlying /e/ rather than /i/, since [e] could only come from /e/ in the non-neutralizing context.

But more to the point, it must be the case that the umlauted present forms effectively never serve as bases in the calculation of preterite stems. None of the preterite endings happen to have [i], and so no preterite stem vowel could ever undergo umlaut. Therefore, if an umlauted present stem form were entered as the base to be assessed by REALIZE MORPHEME: PRETERITE, the faithful mapping with stem vowel [-e-] should be permitted to surface in the preterite, as it would be distinct from its base. Given the earlier discussion, we have a clear solution to this problem: properties which are non-contrastive in their context are not sufficient to satisfy requirements of stem contrast. Therefore, since the [-i-] of the present base is not contrastively distinct from the potential [-e-] of the preterite, REALIZE MORPHEME will not be satisfied.

It is a thoroughly non-trivial question how exactly to enforce this notion. That is to say, how exactly does the preterite derivation know that the [-i-] arose through neutralization? If it has access to the entire surface form when making this determination, then it could see that there was an umlaut trigger following the stem vowel. However, there is at least one case (to be discussed in Section 4.5.3.2 below) where it is necessary to say that the preterite inflectional ending is not taken into account when calculating the preterite stem itself. This would seem to make it unlikely that that same derivation should take into account the inflectional ending of its base.

This might lead us to take a different approach to this realm of questions, namely, to assume that we are dealing in these cases with a strictly "stem-level" derivation (see Sections 4.5.3.2 and 4.5.8.1 for further discussion and references), where inflectional endings, and the phonological processes they might trigger, are not yet present. This would be sufficient to rule out problems in the umlaut case, since this would excise the umlaut trigger from the derivation,[17] but it would not on its own be sufficient to avoid the problems in the pre-nasal raising case, since the triggering environment is completely contained within the stem. In Section 4.5.8, I will suggest that this particular case (and perhaps others like it), and its attendant problems (which actually go beyond its impact on stem contrast), can be explained with reference to the P-map (Steriade 2009). But overall, some appeal to phonological contrast, and its transference between base and derivative, would seem to a general solution to problems of this sort. I leave further consideration of the relevant mechanisms as a question for future research.

4.4.1.3 Other Vowel Changes

I reconstruct past several additional changes. These are of less import, so I discuss them only briefly. First, there are two relevant, though questionable, types of monophthongization. One of these, which is internal to Gothic (if it is actually correct to describe Gothic as having undergone this change at all; see footnote 7) is the monophthongization of diphthongs involving [a]. This is shown in (13).

---

[17] The one exception to this might be the Strong Class V verb *bidjan*, where the *-j-* is a stem-forming suffix (of the type generally restricted to the weak verbs) that appears only in the present.

Prior to this change, the preterite singulars of Classes I–V transparently displayed a stem vowel [a]. It is important to the analysis to be developed that these forms can be viewed as synchronically representing a mapping from /e/ ([-back]) → [a] ([+back]). Therefore, it is ideal to work at a stage when this is transparently the case.

(13) *Monophthongization of diphthongs with first member [a]*
      The diphthongs *[aj,aw] become [ɛː,ɔː].
      Example: Proto-Germanic 1SG.PRET. *[bajt] 'I bit' > Gothic [bɛːt]
                  Gothic 'young woman' NOM.SG *mawi* [ma.wi] ∼ NOM.PL *maujos* [mɔː.j-os]

The other has to do with the tautosyllabic treatment of the sequence *-ej-/-ij-*. It is standardly assumed that Pre-Proto-Germanic tautosyllabic *-ej-/*-ij-* had become a long monophthong *[iː] by the period of Proto-Germanic proper, since there is no clear evidence that any of the Germanic languages attest outcomes other than [iː]. However, inscriptional evidence (from Helmet B of Negau; see Reichardt 1953:307)[18] suggests that some Germanic variety spoken in the last centuries BC, the approximate time period typically identified for Proto-Germanic, retained pronunciation of *-ej-* as [ej]. If this is the correct interpretation, then monophthongization to [iː], if reconstructable to Proto-Germanic at all, must be a very late development, certainly after the period of Pre-Proto-Germanic we are considering.

This question is not actually problematic for the analysis either way. The relevant context is the Class I presents: for example, Gothic *beitan* 'to bite' [biːt-an]. As long as the long vowel (once it exists) is phonologically decomposable and separable as [ej]/[ij], it will be able to serve its proper function as the base for the preterite.

A relatively late process within Proto-Germanic is the treatment of *VNh* sequences. In such a sequence, the nasal consonant is lost, with compensatory lengthening and nasalization of the preceding vowel, as shown in (14). (This must diachronically follow / synchronically counter-bleed pre-nasal [e]-raising (12c).)

(14) *Nasal Deletion before [h]*
      Short /VN/ → [Ṽː] / __h .
      Example: 'to hang/suspend' PGmc/Gothic /hanh-an/ → [hãːhan]

This is the sole source of long [aː] in Gothic, which is actually [ãː]. This outcome remains distinct from [oː], which is the normal lengthened correspondent of short [a]. In at least some instances, the nasal is easily recoverable synchronically based on alternations between [h] and [g]: for example, Gothic present infinitive *briggan* [briŋg-an] ∼ preterite singular *braht-*. It is thus reasonable to assume that the nasal is retained in the underlying representation, and the long nasalized vowels are derived by rule. Even if, in some or all cases, the long nasalized vowels get attributed to the underlying representation, the fact that they are long will be sufficient to generate the proper behavior with respect to the formation of their preterite stems.

One other vowel-related process is "breaking", as shown for Gothic in (15). This is a process that lowers high vowels before [r] and [h]. It is unclear to me whether this process is to be reconstructed for Proto-Germanic,

(15) *"Breaking"*
      Short /i,u/ lower to [ɛ,ɔ] before [r,h,hʷ].
      Example: Pre-Goth. infinitive */ber-an/ 'to bear' (> */bir-an/) → (Pre-)Goth. [bɛran]

---

[18] Thank you to Jay Jasanoff for bringing this to my attention.

I believe that including the result of breaking in the phonological representations employed here has no deleterious effects on the analysis, but this has not been exhaustively confirmed. One relevant question which I will not seek to answer regards the reduplicative vowel. Somewhat unexpectedly, it frequently surfaces as the breaking outcome in the daughter languages (Gothic *ai* [ɛ], Old English *eo*). The traditional explanation for this fact is one based on "analogy": to reduplicating preterites with root-initial [r,h,hʷ] (e.g. Gothic √*hait* 'call' → PRET.SG *haihait-* [h̥ɛh̥ɛːt-]), the reduplicative vowel would be in the breaking environment; this vowel quality "spreads" analogically from these sources.

4.4.2 The Strong Verbs in Pre-Proto-Germanic

Once we take into account the changes/processes detailed in Section 4.4.1, we arrive at the Pre-Proto-Germanic forms provided in (16) below. These are the representations I will employ for the analysis. (I list only the form of the stem, as inflectional endings play no role in the stem-formation process.) The only significant difference (beyond the sound changes we have reconstructed past) between the reconstruction for this stage and what we observe for Gothic is the preterite plural of Class VIId. Jasanoff (2007) argues that the evidence from Northwest Germanic requires positing for Proto-Germanic a zero-grade stem for the preterite plural of the ablauting Strong Class VII verbs (i.e. Class VIId). Forms like this reconstructed **lelt-* (< PIE **le-lh₁d-*) are (arguably) directly attested in Old English (Anglian dialectal forms): e.g. *leort-* (< **lelt-*), *reord-* (Jasanoff 2007:244–245). Class VIId poses significant problems, both synchronic and diachronic, for the account developed in this chapter; see Section 4.7.2 for full discussion.[19]

(16)   Pre-Proto-Germanic Strong verbs

Root Shape	Class	PRES	PRET.SG	PRET.PL	
/CeRC/	I	bejt-	bajt-	b̥it-	(= //bjt-//)
	II	kews-	kaws-	k̥us-	(= //kw̥s-//)
	III	bend-	band-	b̥und-	(= //bn̥d-//)
/CeC/	IV	nem-	nam-	neːm-	
	V	geb-	gab-	geːb-	
/CaC/	VI	dab-	doːb-	doːb-	
/CaRC/	VIIa	hajt-	h̲e̲hajt-	h̲e̲hajt-	
/CVːC/	VIIb	sleːp-	s̲e̲sleːp-	s̲e̲sleːp-	
	VIIc	floːk-	f̲e̲floːk-	f̲e̲floːk-	
	VIId	leːt-	l̲e̲loːt-	**lelt-**	(≠ Gothic l̲e̲loːt-)

Note that I do not here reflect Verner's Law — a process of fricative voicing conditioned by the inherited PIE accent (consult Ringe 2006:102–105, §3.2.4.ii, among others) — which must have applied to base-initial fricatives in reduplicated forms in Proto-Germanic (see Jasanoff 2007). These voicing alternations have been leveled out in Gothic (with the exception of the byforms of

---

[19] Jasanoff (2007) also provides evidence from Northwest Germanic that makes less certain the reconstruction to Proto-Germanic of cluster-copying as the reduplicative behavior of ST-initial roots. See Section 4.6 for discussion.

sesleːp- in *sezleːp-*).[20] Since the voicing value of the base-initial segment plays no role in the analysis of reduplication (as developed in Section 4.6), I omit this distinction for the sake of clarity.

4.4.3   The Vowel System of Pre-Proto-Germanic

Given that the stem alternations described above primarily involve vocalic changes, it will be important to understand the vowel system at the relevant stage, i.e. Pre-Proto-Germanic. At this stage of the language, the vowel system consists of ten vowel phonemes, two of which are marginal (again, consult generally Ringe 2006:§3). The phonemes are given in the table in (17).

(17)   Distinctive features in the Pre-Proto-Germanic vowels

	1μ		2μ		3μ	
	[-back]	**[+back]**	**[-back]**	**[+back]**	**[-back]**	**[+back]**
[+high]	i	u	iː	uː	–	–
[-high]	e	a	eː	oː	eːː	oːː
		([+low,-rnd])		([-low,+rnd])		([-low,+rnd])

The system is based around a two-way feature contrast for the features [±high] and [±back]. The features [±low] and [±round] are distributed differently for different phonemes (most notably in the non-high back vowels, as notated in (17)), but are non-contrastive. The featural contrast intersects with a three-way length distinction, which can be characterized in terms of moras. Trimoraicity is restricted to the non-high vowels,[21] and is itself a marginal contrast which is ultimately neutralized in all the daughter languages.

While three-way vowel length contrasts are exceedingly rare cross-linguistically,[22] it is virtually certain that Proto-Germanic had two types of long vowels, as there is evidence of distinct correspondence sets in the Germanic daughter languages (see Ringe 2006:73–75). Long vowels which were inherited from Proto-Indo-European as such are retained as normal (i.e. bimoraic) long vowels, whereas the result of vowel contraction across a lost PIE laryngeal resulted in what are typically referred to as "overlong" or "trimor(a)ic" vowels. It is clear that the latter type was longer than the former based on, among other things, the outcomes of the back versions of these types in final position in Gothic. Inherited (post-)PIE [oː,aː], which merged as Proto-Germanic [aː], yield Gothic short [a] in final position (and before final [z]).[23] In these same environments, the type resulting from contraction yields Gothic long [oː]. This strongly implies that the contrast was indeed based on length, and can be characterized as a two vs. three mora distinction, as assumed in (17).

---

[20] This root attests preterites with both voiceless [s] and voiced [z] in root-initial position (see, e.g., Ringe 2006:191): with [s], *saíslep* 'he was asleep' and *ana-saíslep-un* '(who) have fallen asleep'; with [z] *ga-saízlep* 'has fallen asleep'. The voiced variant seems to be the historically expected Verner's Law outcome (see, e.g., Ringe 2006:102–105,§3.2.4.ii) from PIE **se-sléh₁p-*. In all other cases, Gothic has completely leveled in favor of the voiceless, non-Verner's variant (Ringe 2006:191–192).

[21] That is to say, we cannot reconstruct any contrast between bimoraic and trimoraic high vowels. This could either be because trimoraic high vowels never existed, or because they completely merged with the bimoraic high vowels without leaving behind distinct outcomes for any sound changes.

[22] To my knowledge, underlying three-way contrasts of this sort are attested only in Estonian (Prince 1980) and North Saami (Baal, Odden, & Rice 2012). Stanton & Zukoff (to appear) argue that Scottish Gaelic and Ho-Chunk require three-way vowel length distinctions on the surface to explain facts involving copy epenthesis in these languages, but that these languages contain only a two-way contrast underlyingly.

[23] Unexpectedly, word-final PIE ablauting [oː], as opposed to [oː] < **oh₂#,eh₂#*, appears to have merged instead with the trimoraic type (Jasanoff 2002:35–38).

Using these representations of the vowels, the phonological alternations involved in the strong verb preterites in Pre-Proto-Germanic can be derived by constraints referring to just the two distinctive features ([±high] and [±back]) and the number of moras.

4.5 Synchronic Derivation of Pre-Proto-Germanic Strong Preterites

In this section, I provide a complete analysis of the Pre-Proto-Germanic avatars of the numerous surface patterns seen in the Gothic strong preterites (with the exception of Classes VIIb and VIId, whose discussion I postpone until Section 4.7). The analysis is built around the hypotheses developed above, namely, phonologically null underlying exponents of the morphosyntactic features PRETERITE and SINGULAR, and active REALIZE MORPHEME constraints that induce the surface contrasts. For concreteness, I will refer to the default preterite stems as preterite plurals, and the marked preterite stems as preterite singulars.

4.5.1 Strong Class I–III Preterite Plurals

The optimal strategy for marking the preterite in the strong verbs is *vowel deletion*. This is observed in the basic case of the preterite plurals of Classes I–III; for example, Class II √*kews* → PRET.PL *kus-* (//kw̥s-//).[24] I will first focus on what is driving this change, and then turn to the question of how to select particular changes over others in the appropriate contexts.

As the change observed in the Class I–III preterite plural involves vowel deletion (calculated relative to the root, though also relative to the present stem), the constraint violated by the optimal form will be MAXV-IO (18). This violation is compelled by the need to phonologically differentiate the preterite stem from the present stem, i.e., the operation of RM:PRET (as defined in (6a) above). Therefore, we know that we have the ranking in (19): RM:PRET ≫ MAXV-IO. As demonstrated in (20), the faithful candidate (20a) fails to be differentiated from the PRESENT base, and thus fatally violates RM:PRET. This compels an unfaithful mapping, allowing for the selection of the deletion candidate (20b).

(18) **MAXV-IO**
Assign one violation mark * for each vowel in the input which lacks a correspondent in the output.

(19) **New Rankings:** RM:PRET ≫ MAXV-IO

(20) Preterite Plural of Class II (also Class I & III)

INPUT: /kews, Ø_{PRET}/ ; BASE: PRES [kews-]	RM:PRET	MAXV-IO
a.　　　kews-	*!	
b. ☞　kus-　(← //kws-//)		*

This ranking shows that RM:PRET can indeed induce unfaithful mappings for the purpose of creating contrast between preterite stem and present stem. Now let us consider how the faithfulness constraints are interacting with one another to determine the nature of the unfaithful mapping.

---

[24] Note that deletion in Class III, where the root contains a [+consonantal] post-vocalic sonorant, results in [u] epenthesis on the surface: for example, Class III √*help* → PRET.PL *hulp-* (//hl̥p-//). I treat these as if the surface form truly contained the vocalized sonorant (i.e. the intermediate morphophonemic form).

We know that elsewhere in the system one of the attested mappings is *vowel backing* (the preterite singular of Class I–V; see Section 4.5.3). This means that the constraint which is violated by changing underlying /e/ to surface [a] is violable within the system, and thus we need to make sure that such a mapping would not be optimal in the current case.

To regulate the feature changes of this sort, I will employ MAXFEATURE and DEPFEATURE constraints, following the work of Casali (1996, *et seq.*). (The rationale for employing faithfulness constraints of this kind is made clear below.) These are constraints that reference specified feature values, and take the feature as the locus of correspondence. The segment remains a unit of correspondence as well, just not the one which is relevant for these constraints. The constraint penalizing vowel backing will thus be DEP[+back]-IO, defined in (21). The active constraint here must be the DEP[F] constraint rather than the inverse MAX[F] constraint (MAX[-back]-IO) because of the way it would interact with MAXV (see below). DEP[+back]-IO is violated if the feature value [+back] surfaces in the output despite not being contained in the input. Since backing is not the preferred repair in this basic case, we know that DEP[+back]-IO must dominate MAXV-IO, as shown in (22). This ranking properly selects deletion as the preferred repair for RM:PRET, as demonstrated in (23).

(21)   **DEP[+back]-IO**
        Assign one violation mark * for each [+back] feature in the output which was not present in the input.

(22)   **New Rankings:** RM:PRET ≫ MAXV-IO

(23)   Ruling out alternative mappings for Class II preterite plurals

INPUT: /kews, Ø_PRET/ BASE: PRES [kews-]	RM:PRET	DEP[+back]-IO	MAXV-IO
a. kews-	*!		
b. ☞ kus- (← //kws-//)			*
c. kaws-		*!	

As mentioned above, if we attempted to use MAX[-back]-IO, we would fail to derive the correct result, because this constraint is equally violated by the suboptimal backing candidate (24c) as by the desired deletion candidate (24b). Since backing does not incur violation of MAXV-IO, the deletion candidate would be harmonically bounded, and thus obviously not selected as the winner.

(24)   Ruling out alternative mappings for Class II preterite plurals

INPUT: /kews, Ø_PRET/ BASE: PRES [kews-]	RM:PRET	MAX[-back]-IO	MAXV-IO
a. kews-	*!		
b. ☺ kus- (← //kws-//)		*	*!
c. 💣 kaws-		*	

This illustrates why, for a number of the interactions that are present in the system, MAX[F] constraints will not be effective, and we will instead need the inverse DEP[F] constraint to do the job. There is one major exception to this generalization, where we actually do need a MAX[F] constraint to simultaneously prevent feature change and total deletion, just in case that feature value is present in the input. Namely, in Class VIIa (see Section 4.5.7), MAX[+back]-IO will be responsible for

preventing vowel fronting for roots with underlying /a/ (i.e. /-a-/ ↛ **[-e-]), while simultane-ously pre-empting the otherwise preferred vowel deletion mapping (i.e. /-a-/ ↛ **[-Ø-]). This also shows why we must be using MAX/DEP[F] constraints rather than asymmetric IDENT constraints (i.e. IDENT constraints with an underlying value specified). IDENT[+back]-IO would be vacuously satisfied by deletion of /a/, and thus would not be able to have the desired effect.[25]

4.5.2 Strong Class IV & V Preterite Plurals

In Classes IV & V, the preterite plural is formed not by vowel deletion, but rather by *vowel lengthening*; for example, √*geb* → PRET.PL *geːb-*. Why do these forms not show deletion like Class I–III? This behavior can be attributed to an emergent markedness pressure against creating new consonant clusters.

First, let us pause to note the behavior of two-consonant clusters in the language generally: they are tolerated (subject to phonotactic restrictions on particular types of clusters). This can be seen from representative examples from Gothic: /skip/ → [skip] 'ship', /broθar/ → [broθar] 'brother'. This means that the constraint which would penalize them — i.e. *CLUSTER (25a) — must be subor-dinated to the faithfulness constraints which militate against possible ways to avoid clusters, namely MAXC-IO (25b) and DEPV-IO (25c). This ranking is provided in (26) and illustrated in (27).

(25) a. ***CLUSTER (*CC)**
          Assign one violation mark * for each sequence of two adjacent non-syllabic conso-nants in the output.
      b. **MAXC-IO**
          Assign one violation mark * for each consonant in the input which lacks a correspon-dent in the output.
      c. **DEPV-IO**
          Assign one violation mark * for each vowel in the output which lacks a correspondent in the input.

(26) **New Rankings:** MAXC-IO, DEPV-IO ≫ *CC

(27) Faithful realization of underlying clusters: /skip/ → [skip] 'ship'

/skip/		MAXC-IO	DEPV-IO	*CC
a. ☞	skip			*
b.	sip / kip	*!		
c.	sikip		*!	

---

[25] The MAX/DEP[F] approach allows for feature mobility (features underlyingly associated with one segment migrating to a different segment in the output), which is penalized by the constraint LINEARITY[F]-IO, defined below. Feature mobility is not observed in this system, so this constraint must be highly ranked.

(i) **LINEARITY[F]-IO**
     Given two features F_i and G_i in the input with correspondents F_o and G_o in the output, assign a violation mark if the precedence relation between F and G is not the same as the precedence relation between F_o and G_o:
     a. If F_i precedes G_i, then assign a violation if F_o follows or is simultaneous with G_o
     b. If F_i is simultaneous with G_i, then assign a violation if F_o precedes or follows G_o

Despite being ranked below the general faithfulness constraints MaxC-IO and DepV-IO, I argue that *CC is still having an effect in the language, specifically on the patterns of preterite stem formation. (The same exact emergent preference against clusters must apply also in determining the shape of the reduplicant in Gothic/Proto-Germanic and many of the other languages discussed in this dissertation; see Chapter 1 and Section 4.6 below.) We can see this when we consider what the result of vowel deletion would be in Class IV–V roots compared to Class I–III roots. In Class I–III, the root vowel is followed by a sonorant consonant, followed by another consonant: /CeRC/. When deletion occurs (//CRC//), there is a sonorant consonant in a position where it can be vocalized (i.e. inter-consonantal position; see footnote 26 immediately below). This allows for a syllabifiable string without any new consonant clusters. The situation for Class IV–V roots, which are of the shape /CeC/ (/CeT/ in Class V, /CeR/ in Class IV), is different. For these roots, if vowel deletion were to occur (//CC//), it would create a consonant cluster which was not present underlyingly, and thus a new violation of *CC.[26] Since we do not observe vowel deletion in these cases, we can identify *CC as the motivation.[27]

So, to avoid the cluster that would have arisen via the default vowel deletion mapping, the Class IV–V preterite plurals are instead formed by vowel lengthening. I will identify the faithfulness constraint militating against this repair as Dep-μ-IO, as defined in (28).[28] As long as Dep-μ-IO is dominated by both RM:PRET and *CC, as shown in (29), lengthening will be the optimal way of effecting stem contrast for /CeC/ roots (and indeed also /CaC/ roots; see Section 4.5.6 on Class VI). Furthermore, we know that Dep[+back]-IO ≫ Dep-μ-IO, as lengthening is selected even though backing could have effected the contrast without creating a cluster. This is demonstrated in (30).[29]

(28) **Dep-μ-IO**
      Assign one violation mark * for each mora in the output which lacks a corresponding mora in the input.

(29) **New Rankings:** RM:PRET, *CC, Dep[+back]-IO ≫ Dep-μ-IO

---

[26] This analysis will work straightforwardly only if we assume that sonorants are not vocalizable at stem edges, as otherwise √CeR → //CR// and √ReC → //RC// deletion mappings should be able to vocalize a sonorant and escape a *CC violation. It does seem that Proto-Germanic words where we would expect word-initial vocalization frequently show some other outcome, especially when the stem is not paradigmatically isolated (consult Ringe 2006:§3). This question deserves additional attention, but I will not address it further here.

[27] Note that what is at issue here cannot be whether the resulting consonant cluster would or would not phonotactically licit, as the blocking of deletion applies in either case: √ber ↛ **br- (cf. *broþar* 'brother'), just as √geb ↛ **gb- (**#gb in general).

[28] For (Pre-)Proto-Germanic, the three-way length contrast makes it impossible to represent length with a single feature [±long], and thus inappropriate to use Max/Dep[+/-long] as the constraints regulating changes in length. Nevertheless, Max/Dep-μ constraints are of very much the same nature as Max/Dep[F] constraints, as they govern correspondence of elements other than the segment itself.

[29] We cannot establish a ranking between Dep-μ-IO and MaxV-IO, because lengthening in /CeRC/ will be ruled out independently by the constraint *V:RC (see (34) below), which must be active in shaping the singulars of Class I–III (Section 4.5.3.2) and both plural and singular in Class VII (Section 4.5.7).

(30)   Preterite Plural of Class V (also Class IV)[30]

INPUT: /geb, Ø_PRET/ BASE: PRES [geb-]	RM:PRET	*CC	DEP[+back]	DEP-μ	MAXV
a. geb-	*!				
b. gb-		*!			*
c. gab-			*!		
d. ☞ ge:b-				*	

4.5.3   Strong Class I–V Preterite Singulars

We have now computed the preterite plurals (i.e. the default preterite stems) for both Class I–III and Class IV–V, so let us turn to the corresponding preterite singulars — first to Class IV–V, then to Class I–III. We can observe that, in both types, the preterite singular stem is formed by *vowel backing*, i.e. changing the root vowel /e/ to [a]. I argue that this is driven by the simultaneous activity of both RM constraints, i.e. RM:PRET *and* RM:SG. That is to say, this alternative mapping comes about because the preterite singular stem will aim to be distinct not only from the present stem, but also from the preterite plural stem. (On the notion of base priority in this analysis, see Section 4.5.4 below.)

4.5.3.1   Strong Class IV–V Preterite Singulars

For the /CeC/ roots of Class IV–V, the preterite plural stem has already claimed vowel lengthening as the optimal mapping, so the preterite singular must settle for the next best option, which turns out to be *vowel backing*: for example, $\sqrt{geb} \rightarrow$ PRET.SG *gab-*. In this case, not only is the faithful mapping (33a) ruled out by RM:PRET as usual, but also the normally optimal unfaithful mapping — vowel lengthening (33c) — is blocked by RM:SG, since that mapping is already in use for the plural. Deletion is again blocked by *CC, just as it was for the plural. The evaluation then chooses the candidate which violates the next lowest ranked faithfulness constraint, which, as evidenced by this interaction, must be DEP[+back]-IO. This constraint must be dominated by both RM constraints (and also *CC), as shown in (32), or else it would be preferable to fail to create the *singular~plural* contrast and instead just select the lengthening mapping again.

The constraint DEP[+high]-IO (defined in (31)), and a candidate that violates it (the vowel raising candidate (33e)), are now included. DEP[+high]-IO must outrank DEP[+back]-IO in order to avoid selecting the raising candidate. DEP[+high]-IO thus also outranks DEP-μ-IO and MAXV-IO by transitivity. The necessary high ranking of DEP[+high]-IO in this analysis poses a problem, as Proto-Germanic had several processes that involve raising of front mid vowels (which would violate this constraint) even though at least some of the markedness problems that raising is solving could in theory probably have been repaired by backing instead. I will return to the ramifications of this problem below in Section 4.5.8.

(31)   **DEP[+high]-IO**
       Assign one violation mark * for each [+high] feature in the output which was not present in the input.

(32)   **New Rankings:** RM:PRET, RM:SG, DEP[+high]-IO, *CC ≫ DEP[+back]-IO

---

[30] I omit the IO designation from the faithfulness constraints in all remaining tableaux for reasons of typographical space.

(33) Preterite Singular of Class V (also Class IV)

INPUT: /geb, Ø_PRET, Ø_SG/ BASES: PRES [geb-], PRET.PL [geːb-]	RM:PRET	DEP[+high]	*CC	RM:SG	DEP[+back]	DEP-μ	MAXV
a. geb-	*!						
b. gb-			*!				*
c. geːb-				*!	*		
d. ☞ gab-					*		
e. gib-		*!					

4.5.3.2 Strong Class I–III Preterite Singulars

Virtually the same interaction that produces the Class IV–V preterite singulars also determines the outcome for the /CeRC/ roots of Class I–III, which also show a preterite singular stem with backing to /a/: for example, √*kews* → PRET.SG *kaws-*. The difference from the Class IV–V derivation just examined is which mapping has already been claimed by the plural and what markedness constraint is in play. Whereas the Class IV–V plurals had claimed lengthening to satisfy RM:PRET, the Class I–III plurals claimed deletion. Therefore, RM:SG now rules out deletion, rather than ruling out lengthening as it did above. Why, then, does the Class I–III singular derivation not opt for lengthening, which seemed to be the next best option for Class IV–V when deletion had been blocked (in that case, by *CC)?

I posit that there is an additional markedness constraint operative in the system: *V:RC, defined in (34) (akin to *SUPERHEAVY, Sherer 1994's *σ_{μμμ}, Zec 1995's "Bimoraicity Constraint"). There is independent evidence for the activity of this constraint in the form of "Osthoff's Law", a general term for various instances of vowel shortening before *sonorant + consonant* sequences in a number of the Indo-European languages (especially in Greek; cf. Collinge 1985:127–131). The evidence for the operation of Osthoff's Law within (Pre-)Proto-Germanic is not particularly robust (see Ringe 2006:75–78), but it does appear to hold at least in the form of a static phonotactic generalization. Forms which violate this constraint are innovated in the Germanic daughter languages, largely through syncope processes — for example, Gothic *doːms* 'judgment' (Lambdin 2006:325) < *doːmis — so the ranking of this constraint must be lower at later stages (although not necessarily totally inactive).

Unlike with /CVC/ roots, when the root is of the shape /CVRC/, vowel lengthening will lead to the sequence banned by the *V:RC constraint because of the post-vocalic *sonorant + consonant* sequence. Therefore, for Class I–III, even though the faithfulness constraint against lengthening (DEP-μ-IO) is ranked relatively low, that mapping will be blocked by markedness. So, the evaluation again selects vowel backing (36d) as the optimal repair for RM:SG in the preterite singular.

(34) ***V:RC**
      Assign one violation mark * for each output V:RC sequence (≈ *SUPERHEAVY).

(35) **New Rankings:** *V:RC ≫ DEP[+back]-IO (≫ DEP-μ-IO)

(36)   Preterite Singular of Class II (also Class I–III)

INPUT: /kews, Ø_PRET, Ø_SG/ BASES: PRES [kews-], PRET.PL [kus-]	RM:PRET	*V:RC	RM:SG	DEP[+back]	DEP-μ	MAXV
a. kews-	*!					
b. kus- (← //kws-//)			*!			*
c. keːws-		*!			*	
d. ☞ kaws-				*		

One further note on the Osthoff's Law constraint is in order before moving on. The specific formulation of the constraint as *V:RC, rather than a more general *SUPERHEAVY constraint, is designed to avoid questions regarding syllable structure and consonant moraicity (and how these questions interact with morphological constituency and the potential need for derivational levels).[31] Specifically, we need to be concerned with the lengthening pattern in the Class IV preterite plurals (/CeR/ → [CeːR-]) and the preterite plurals and singulars of Class VI roots with post-nuclear sonorant consonants (/CaR/ → [CaːR-]) (see Section 4.5.6). If we were to assume that the relevant markedness constraint is simply *SUPERHEAVY, then, if the root-final sonorants in these cases were syllabified and moraic, then *SUPERHEAVY would penalize these mappings in exactly the same way as it does for the Class I–III preterite singulars, which is not what we want. Rather than digging deeper into the questions that would need to be answered in order to determine whether or not the root-final sonorants actually are syllabified and moraic in the relevant way, I adopt the *V:RC formulation, which, for the most part, circumvents these problems.

The one place where this move cannot escape these questions entirely is in the 2nd person singular preterite inflected form of /CaR/ roots. The 2nd person singular strong preterite ending is /-t/. We thus have surface forms in this category like [foːr-t]. This surface string would very clearly violate *V:RC. Rather than throwing out the markedness constraint, I believe that the answer to this has to do, in one way or another, with the interaction of the phonology and the morphology. For one, the phonological mappings we are analyzing in the preterite are specifically those of the verbal stem. Except in exactly this case (as far as I can tell), the properties of the inflectional endings never even have the potential to factor into the determination of the preterite stem. This implies that, in a certain sense, the system we are constructing is that of the "stem-level" phonology — as in Lexical Phonology (Pesetsky 1979, Kiparsky 1982, Mohanan 1982, *et seq.*) or Stratal OT (consult Kiparsky 2015, among others) — and that the inflectional endings are not present at the point in the derivation when the stem shape is determined. Alternatively, we could view this as a sort of paradigm uniformity effect. In the two other inflected forms of the preterite singular, the personal ending is null. Therefore, the word form in its entirety does not violate *V:RC. If one of these forms serves as the "base" of the paradigm, then the effects of *V:RC could underapply in the 2nd person. Modulo this problem, *V:RC will be sufficient to rule out lengthening in /CVRC/ roots.

4.5.4 *Excursus*: REALIZE MORPHEME and Base Priority

A point that must be addressed here is why the plural is derivationally *prior* to the singular, insofar as RM:SG violations have been assessed in the tableaux for the plural in (20) and (30), but not in the tableaux for the singular in (33) and (36). I intend this to follow from the assumption that SINGULAR in this system is a privative feature (at least with respect to the stem, if not necessarily with respect to the agreement suffixes), and thus the derivation for the "plural" is really the derivation for the

---

[31] See Sandell & Byrd (2014, 2015) for recent discussion of some of these issues in Proto-Indo-European.

unmarked/default stem of the preterite (see Sections 4.2.1 and 4.3). The singular stem is morphosyn-tactically marked for number at the end of the post-syntactic derivation, and thus stands in the sort of superset relation relative to another morphological output (i.e. the default preterite stem) that is necessary to activate REALIZE MORPHEME (here RM:SG).

The same applies to the relation between present and preterite. If the two stems must be distinct, why is it the preterite rather than the present that receives the unfaithful mapping? I intend this to reduce to the same type of interaction found with the singular and plural in the preterite: the feature PRETERITE remains at the output of the post-syntactic derivation, but a feature PRESENT (*vel sim.*) does not. As alluded to earlier, this could be due to a morphological feature-deletion operation. Alternatively or in addition, it could be that the "present" stem is really just an unmarked verbal stem. This would be consistent with the fact that it can have either present or future semantics (probably governed by the use of the perfective prefix *ga-*; Lambdin 2006:16–17), and is the stem used for the infinitive. If the derivation of the "present" stem lacks a specification for any tense feature, then the preterite (which clearly does have a tense feature, at least given the analysis pursued here) stands in the sort of superset relation with the "present" stem that is necessary to active REALIZE MORPHEME (here RM:PRET).

This notion of base priority raises another question. In order to know whether a preterite stem is phonologically distinct from its corresponding present stem, that present stem must already have been derived. The same holds all the more for preterite singular stems relative to their default preterite plurals, since those are derived via grammatically controlled unfaithful mappings, rather than being equivalent to the root as in the present. This question is not specific to this analysis, but rather a general property of Base-Derivative correspondence constraints (cf. Benua 1997), in which properties of the derived word may never be transferred to the base. REALIZE MORPHEME is essen-tially a Base-Derivative correspondence constraint, since it specifies its relation as that between a morphologically less derived word and a morphologically more derived word based on it, so we would expect the same condition to apply, and it does.

4.5.5 Interim Summary

The ranking proposed thus far properly derives both the singular and the plural of Classes I–V. Next to consider are the forms of Class VI, which display vowel lengthening in both singular and plural, and Class VII, which display reduplication in both singular and plural. These distri-butions differ from Classes I–V in that they don't differentiate their singular and plural stems. Crucially, Classes VI & VII differ from Classes I–V in their root vocalism: Classes I–V have root vowel /e/; Classes VI & VII include all other permissible root vowels (/a, eː, oː/). This will form the basis for that distinct behavior.

4.5.6 Strong Class VI Preterite Singulars and Plurals

4.5.6.1 Strong Class VI Preterite Plurals

We have already seen in Class IV–V that lengthening is the preferred repair for RM:PRET in /CVC/ roots, so lengthening to effect *present∼preterite* contrast for the plural of Class VI, which has the root shape /CaC/, should be fully expected. The fact that these roots have a different vowel, /a/ rather than /e/, allows us to determine the ranking of a new faithfulness constraint. In the Class I–V preterite singulars, altering the value for [±back] meant a change from [-back] to [+back], and this was, under the right circumstances, tolerated. For Class VI, where now the underlying value is [+back], altering the value for this feature means changing from [+back] to [-back], and thus

would violate the opposite MAX/DEP[F] constraints. (DEP[+back]-IO will here always be vacu-ously satisfied, since [+back] actually is present in the input.) The relevant constraint here could be either DEP[-back]-IO or MAX[+back]-IO. Both constraints are capable of ruling out feature change; however, only one of them — MAX[+back]-IO, as defined in (37) — is capable of also ruling out total vowel deletion. This capability is not needed for Class VI, because the vowel deletion candidate (39d) is independently blocked by high-ranked *CC. However, once we get to Class VIIa in Section 4.5.7.1, we will need a faithfulness constraint to prevent deletion, as the structure that would result from deletion is phonotactically acceptable. Since MAX[+back]-IO can handle both roles while DEP[-back]-IO can only handle one, I adopt MAX[+back]-IO as the constraint respon-sible for prohibiting vowel fronting (candidate (39d)) in Class VI. Once this constraint is ranked above DEP-μ-IO (as shown in (38)), we properly derive vowel lengthening in the Class VI preterite plural. This yields mappings like $\sqrt{dab} \rightarrow$ PRET.PL *do:b-*, as demonstrated in (39).

(37) **MAX[+back]-IO**
      Assign one violation mark * for each [+back] feature in the input which is not present in the output.

(38) **New Rankings:** MAX[+back]-IO ≫ DEP-μ-IO

(39) Preterite Plural of Class VI

INPUT: /dab, Ø_PRET/ BASE: PRES [dab-]	RM:PRET	*CC	MAX[+back]	DEP[+high]	DEP[+back]	DEP-μ	MAXV
a. dab-	*!						
b. db-		*!	*!				*
c. ☞ do:b- (← //da:b//)						*	
d. deb-			*!				
e. dub-				*!			

4.5.6.2 Strong Class VI Preterite Singulars

As noted earlier, Class VI (and also Class VII) does not accord with the generalization that the preterite singular is distinct from the preterite plural. This implies that RM:SG is lower ranked than any further faithfulness constraints (e.g., DEP[+high]-IO) that could be violated to generate a distinct output form. One such faithfulness constraint which must outrank RM:SG is INTEGRITY-IO (*'no multiple correspondence'*, defined in (40)), a constraint which will be particularly relevant for Class VII. This yields mappings like $\sqrt{dab} \rightarrow$ PRET.SG *do:b-*, as demonstrated in (42). Note that the optimal output in such cases bears a RM:SG violation.

(40) **INTEGRITY-IO**
      Assign one violation mark * for each segment in the input which stands in correspondence with multiple segments in the output.

(41) **New Rankings:** RM:PRET, MAX[+back]-IO, INTEGRITY-IO, DEP[+high]-IO ≫ RM:SG

(42) Preterite Singular of Class VI

INPUT: /dab, Ø_PRET, Ø_SG/ BASES: PRES [dab-], PRET.PL [doːb-]	RM:PRET	*CC	MAX[+back]	DEP[+high]	INTEGRITY	RM:SG	DEP-μ	MAXV
a. dab-	*!							
b. db-		*!	*!					*
c. ☞ doːb- (← //daːb-//)						*	*	
d. deb-			*!					
e. $d_ie_jd_ia_j$b-					*!*			
d. dub-				*!				

4.5.7 Strong Class VIIa and VIIc Preterite Singulars and Plurals

Finally, we turn to the Class VII strong verbs, which form their preterites through *reduplication* (or, more precisely, phonological copying/splitting). As discussed in Chapter 1 (see also Section 4.6 below), the reduplication is prefixal, and generally of the shape [Ce-], where C is a copy of the first consonant of the root/base. Class VII includes roots with underlying root vowel /a/ followed by two consonants, specifically *sonorant + obstruent* (Class VIIa), and also roots with underlying long vowels, specifically /oː/ (Class VIIc) and /eː/ (Class VIIb and Class VIId). What these different types have in common is that, if they were to undergo lengthening, the result would be (roughly) a superheavy syllable. In this subsection, I will treat the Class VIIa plurals and singulars, and then the Class VIIc plurals and singulars, both of which are straightforwardly analyzable with the same sorts of markedness and faithfulness interactions observed thus far. I postpone, however, examination of Class VIIb and Class VIId (roots with underlying /eː/) until Section 4.7, as both of these types (Class VIId especially) pose significant difficulties for the proposed analysis (and indeed virtually all analyses of the current problem).

4.5.7.1 Strong Class VIIa Preterite Singulars and Plurals

Class VIIa forms both its preterite plurals and its preterite singulars with *reduplication*: for example, √*hajt* → PRET.PL *hehajt-*, PRET.SG *hehajt-*. The roots of Class VIIa are of the shape /CaRC/. This shape is the same as Class I–III (/CeRC/), modulo the quality of the vowel. The two are thus equivalent with respect to the (im)possibility of vowel lengthening as an optimal unfaithful mapping, as they would both equally violate the constraint *VːRC (defined in (34)) if lengthening were to occur. The current constraint ranking properly derives reduplication, once we fix the ranking of INTEGRITY-IO as the lowest among previously undominated constraints, as shown in (43). This is demonstrated in (44) below.[32]

(43)   **New Rankings:** RM:PRET, *VːRC, MAX[+back]-IO, DEP[+high]-IO ≫ INTEGRITY-IO

---

[32] Further consideration of how the precise form of the reduplicant is derived within the system will be undertaken in Section 4.6 below.

(44)   Preterite Plural of Class VIIa

INPUT: /hajt, Ø_PRET/ BASE: PRES [hajt-]	RM:PRET	*V:RC	MAX[+back]	DEP[+high]	INTEGRITY	RM:SG	DEP-μ	MAXV
a.　　hajt-	*!							
b.　　ho:jt- (← //ha:jt-//)		*!					*	
c.　　hit- (← //hjt-//)			*!					*
d.　　hejt-			*!					
e. ☞ $h_i e_j h_i a_j$jt-					**			
f.　　hujt-				*!				

Just as in Class VI, MAX[+back]-IO rules out the fronting candidate (44d). Unlike in Class VI, as alluded to earlier, MAX[+back]-IO must also be responsible for eliminating the deletion candidate (44c). The fact that deletion is not observed here (where no other high-ranked markedness constraints — i.e. *CC — are in danger of being violated as a result of deletion) indicates that there is a preference to *maintain* the underlying [+back] feature (via MAX[+back]-IO), not simply to not insert the [-back] feature (via DEP[-back]-IO).[33] Also unlike in Class VI, the lengthening candidate (44b) would incur a *V:RC violation, so it is not permitted. As long as MAX[+back]-IO and *V:RC dominate INTEGRITY-IO, the reduplicative candidate (44e) will correctly be selected as the winner.

Note that "reduplication" is emerging as a phonological repair for a RM:PRET violation, without the presence of a /RED/ morpheme in the input. This means that, at the Pre-Proto-Germanic stage (and presumably in Gothic as well), this pattern may more aptly be understood as *phonological copying* rather than reduplication proper, insofar as there is a difference (see Yu 2005 on phonological copying of a similar sort, which he refers to as "compensatory reduplication").

Class VIIa, like Class VI, does not show a stem contrast between preterite plural and preterite singular. This follows from the currently established ranking (see (41) and (43)). RM:SG is ranked below all the faithfulness constraints whose violation could be employed to effect stem contrast. Therefore, even though a higher-ranked faithfulness constraint, INTEGRITY-IO, could be and is violated in order to effect the contrast between present and preterite in service of RM:PRET (in the derivation of the singular as well), it would be suboptimal to do any such thing just in service of RM:SG. Thus, the derivation for the Class VIIa preterite singular in (45) yields the identical result as the derivation for the Class VIIa preterite plural in (44) above. The only difference between the two tableaux is the new RM:SG violation in the winning candidate (45e), but this violation is low-enough ranked to have no impact on the evaluation.

---

[33] That we need this constraint to be MAX[+back]-IO rather than DEP[-back]-IO is a positive result, as high-ranked DEP[-back]-IO might pose problems in generating the emergent [e] vowel in the reduplicant.

(45)   Preterite Singular of Class VIIa

INPUT: /hajt, Ø$_{PRET}$, Ø$_{SG}$/ BASE: PRES [hajt-], PRET.PL [hehajt-]	RM:PRET	*V:RC	MAXI[+back]	DEP[+high]	INTEGRITY	RM:SG	DEP-μ	MAXV
a. hajt-	*!							
b. hoːjt- (← //haːjt-//)		*!					*	
c. hit- (← //hjt-//)			*!					*
d. hejt-				*!				
e. ☞ $h_i e_j h_i a_j$jt-					**	*		
f. hujt-				*!				

Class VIIa provides a good opportunity to consider the way that two additional constraints interact with the stem formation system. In Section 4.5.2, I showed that the segment-level faithfulness constraints MAXC-IO and DEPV-IO had to dominate *CC in order to permit clusters in the language generally. The only constraint that *CC has been shown to rank above is DEP[+back]-IO, so there is thus far little evidence about the ranking of MAXC-IO and DEPV-IO from transitivity. While we could imagine that deletion or epenthesis at either root edge might cause problems over and above their basic faithfulness violations, deletion or epenthesis root-internally would not encounter any such problems. Therefore, the /CVRC/ shape of Class VIIa lets us reasonably consider stem-formation candidates that violate MAXC-IO and DEPV-IO.

Candidate (47c) represents deletion of the root's post-vocalic glide (taken to be a consonant, as would be unambiguous in Class VIIa roots with a post-vocalic liquid, e.g. √*falþ* 'fold'). Candidate (47d) represents insertion of a vowel into the root's post-vocalic cluster. Since neither of these candidates is observed, and the reduplication mapping (47b) is indeed selected as optimal, this means that MAXC-IO and DEPV-IO must dominate INTEGRITY-IO, as shown in (46).[34]

(46)   **New Rankings:** MAXC-IO, DEPV-IO ≫ INTEGRITY-IO

(47)   Preterite Plural of Class VIIa

INPUT: /hajt, Ø$_{PRET}$/ BASE: PRES [hajt-]	RM:PRET	MAXC	DEPV	INTEGRITY	RM:SG
a. hajt-	*!				
b. ☞ $h_i e_j h_i a_j$jt-				**	
c. hat-		*!			
d. hajat- / hajet-			*!		

4.5.7.2   Strong Class VIIc Preterite Singulars and Plurals

Class VIIc also has reduplicated mappings for both preterite plural and preterite singular, such as √*floːk* → PRET.PL *fefloːk-*, PRET.SG *fefloːk-*. Like Class VIIa, Class VIIc also has a (non-high) back root vowel, but this time it is long ([oː]) rather than short ([a]) (recall that the featural differences

---

[34] Note that both of these candidates improve with respect to *CC over the faithful candidate (47a) and the optimal reduplication candidate (47b). However, since we already know that MAXC-IO and DEPV-IO dominate *CC, it is impossible to tell based on these interactions whether *CC dominates INTEGRITY-IO. *CC thus has been omitted from the tableau.

are non-contrastive). Thus, just as also for Class VIIa, MAX[+back]-IO simultaneously prevents both vowel deletion (candidate (51d)) and vowel fronting (candidate (51e)) as possible repairs for RM:PRET for Class VIIc. In Class VIIa, lengthening was ruled out by *V:RC. Given the formulation of this constraint adopted above, this will not apply to roots of Class VIIc, since they uniformly have (at most) one post-nuclear consonant. (If we had adopted a general *SUPERHEAVY, this would have equally well applied to both types. See the end of Section 4.5.3.2 for discussion of why this route was avoided.) We must then appeal to a different markedness constraint to prevent lengthening (as in candidate (51b)) in this case; I will employ $*V_{\mu\mu\mu}$, which is defined in (48). This is well-motivated on typological grounds, as very few languages show a three-way vowel length contrast (see footnote 22). As long as $*V_{\mu\mu\mu}$ dominates INTEGRITY-IO, as shown in the ranking in (50a), the desired reduplication mapping (candidate (51f)) will be preferred to the lengthening mapping (candidate (51b)).

The presence for the first time of an underlying long vowel presents a new type of possible candidate: vowel shortening (candidate (51c)). This type of repair is penalized by MAX-μ-IO, as defined in (49). We can use this constraint to rule out shortening in this case, but only if we make a particular assumption about the distribution of moras in the input and mora migration in the output. The potential problem lies in the vowel deletion mapping which is optimal for the Class I–III preterite plurals: /CeRC/ → [CR̩C-]. Vowel deletion would presumably involve mora deletion as well, and thus violation of MAX-μ-IO; if so, the high ranking of MAX-μ-IO necessary to block shortening in the current case should block deletion in the basic case as well. To avoid this, we must posit that the vowel's underlying mora persists in the output, and surfaces associated to the newly syllabic sonorant: i.e. /Ce⟦μ⟧RC/ → [CR̩⟦μ⟧C-].[35]

This probably requires us also to say that, in this language, consonants are not underlyingly associated with moras, and can only acquire them in the output either by mora migration or via Weight-by-Position (cf. Hayes 1989).[36] Otherwise, we might expect the post-nuclear sonorant to be associated with a mora underlyingly, which would mean that either that mora or the one from the vowel would be deleted in the Class I–III preterite plural forms. If we adopt these assumptions, then we can assert that the Class I–III preterite plurals do not actually involve mora deletion, and therefore that MAX-μ-IO can indeed be used to rule out shortening in Class VIIc.[37] This is demonstrated in (51) below.

---

[35] If these forms involve mora migration, this requires that LINEARITY-μ-IO (cf. LINEARITY[F]-IO in footnote 25) be ranked quite low — specifically, below DEP[+back]-IO (see Section 4.5.1). I will simply assume that proper locality conditions can be imposed to ensure the proper outcomes in such cases of mora migration, though this is something worth confirming explicitly.

[36] It appears as though most if not all of the geminates found in the early Germanic languages can be traced to post-Proto-Germanic sound changes, and thus that (Pre-)Proto-Germanic did not have underlying geminates, i.e. consonants which would be required to bear a mora underlyingly.

[37] If these assumptions cannot be upheld, and simple MAX-μ-IO can thus not be ranked high, we can use a special version of this constraint, that is specific to vowels that underlyingly possess more than one mora, to rule out shortening in the present case:

(i) **MAX-μ/μ-IO**
For each mora in the input that was associated to a vowel with more than one mora in the input, assign one violation mark * if that mora is not present in the output.

Alternatively, if we adopted a featural view rather than a moraic view of the three-way length contrast — for example, using [±long] and [±overlong] (*vel sim.*) to describe short ([-long,-overlong]), long ([+long,-overlong]), and overlong ([+long,+overlong]) vowels — then a constraint MAX[+long]-IO would be sufficient. For Gothic, where the three-way distinction has been fully reduced to a two-way distinction, the use of one feature [±long] could fully describe the length oppositions.

Given, though, that trimoraic vowels are indeed tolerated (at least in certain positions) in the language (see Section 4.4.3), we know that MAX-μ-IO must also outrank the markedness constraint against trimoraic vowels, $*V_{\mu\mu\mu}$; otherwise, input trimoraic vowels would always shorten to basic long vowels. This ranking is shown in (50b), and reflected in the tableau in (51).[38]

(48) **$*V_{\mu\mu\mu}$**
Assign one violation mark * for each trimoraic vowel in the output.

(49) **MAX-μ-IO**
Assign one violation mark * for each mora in the input which lacks a corresponding mora in the output.

(50) **New Rankings:**

a. $*V_{\mu\mu\mu} \gg$ INTEGRITY-IO
b. MAX-μ-IO $\gg *V_{\mu\mu\mu}$

(51) Preterite Plural of Class VIIc

INPUT: /floːk, Ø_PRET/ BASE: PRES [floːk-]	RM:PRET	MAX[+back]	MAX-μ	$*V_{\mu\mu\mu}$	INTEGRITY	RM:SG	DEP-μ	MAXV
a. floːk-	*!							
b. floːːk-				*!			*	
c. flak- (← //flŏk-//)			*!					
d. fulk- (← //fl̩k-//)		*!	*!					*
e. fleːk-		*!						
f. ☞ $f_i e_k f_i l_j o\!ː_k k$-					**			

As in Class VIIa, the relatively low ranking of RM:SG prevents the *plural~singular* contrast from being actualized, and the same output is selected for the preterite singular stem as well.

4.5.8 Summary of Analysis

This concludes the main analysis of the phonology of strong preterite stem formation in Pre-Proto-Germanic. The Hasse diagram in (52) below summarizes the constraint rankings used in the above analysis. (Consult the "new rankings" examples throughout the preceding section for ranking arguments. Rankings that also follow from transitivity are omitted.) I have yet to discuss the analysis of Class VIIb and Class VIId. Both of these types are problematic, especially Class VIId. As such, I will refrain from drawing significant conclusions from their behavior, and thus I do not include constraints relevant to only their analysis in this ranking summary. Several lingering questions remain to be addressed. I take up several of these below.

---

[38] Given that there are likely contextual restrictions on where trimoraic vowels were licensed, it is probable that context-specific versions of $*V_{\mu\mu\mu}$ outrank MAX-μ-IO. It is perhaps noteworthy that trimoraic vowels seem not to be allowed within roots.

(52) **Hasse diagram of rankings**

MAXC-IO DEPV-IO MAX-μ-IO *VːRC RM:PRET MAX[+back]-IO DEP[+high]-IO

$*V_{μμμ}$

*CC   INTEGRITY-IO

RM:SG

DEP[+back]-IO

DEP-μ-IO   MAXV-IO
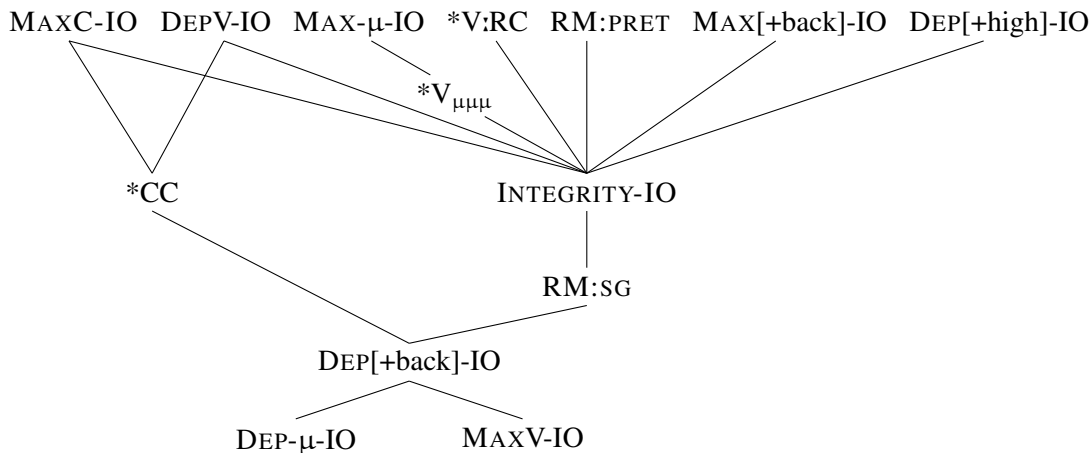
4.5.8.1 Stem Formation and the Regular Phonology

One significant question that follows from this analysis relates to the interaction of this system with the regular phonology of the language. In some respects, it would be ideal if the ranking required to generate the stem formation patterns followed entirely or even partially from independently observable phonological processes and conditions. There is one respect in which this does hold for the current analysis: (Pre-)Proto-Germanic does not allow VːRC sequences, and these are prevented from occurring in the stem formation system as well. However, to the extent that we have evidence of synchronic phonological processes in (Pre-)Proto-Germanic, they seem to at least partially contradict the ranking that is required for stem formation. Specifically, Proto-Germanic had several processes that involved the raising of /e/ to [i], as listed in (53), which is repeated from (12) above (see Section 4.4.1).

(53)   Raising within Proto-Germanic

   a.   *Unstressed raising* (Ringe 2006:122–126, §3.2.5.iii)
        /e/ raises to [i] when unstressed (i.e. in non-initial position).
        Example: 'mice' PIE *múːs-es* > PGmc *ˈmuːs-**i**z* (> Pre-OE *ˈmyːs-iz* > Eng. *mice*)
   b.   *Umlaut* (Ringe 2006:126–128, §3.2.5.iv)
        /e/ raises to [i] when followed by [i] or [j].
        Example: 'he bears' PIE *bʰér-eti* >> PGmc *ˈb**i**r-ið* (> OHG *birid*)
   c.   *Pre-Nasal raising* (Ringe 2006:149–150, §3.2.7.ii)
        Proto-Germanic /e/ → [i] / __N{C,#}.
        Example: PIE 'five' *pénkʷe* > PGmc *f**i**mf* (> Goth *fimf*, OE *fiːf*)

Let us first consider pre-nasal raising. The markedness problem in this context is specific to [e], as underlying /aN{C,#}/ sequences do not raise to [u] or [i]. Given the ranking DEP[+high]-IO ≫ DEP[+back]-IO, which was required in order to derive vowel backing in Classes I–V, we would incorrectly predict that /e/ should map to **[a] in this context rather than to [i]. For this particular example, we might appeal to the P-map (Steriade 2009), and say that mapping /e/ → [i] in the pre-nasal context is a less perceptually salient change than mapping /e/ → [a]. (The pre-nasal vowel was almost certainly nasalized.) Then we could posit two different context-sensitive DEP[+high]-IO constraints — a low-ranked one contextually restricted to the pre-nasal environment

(Dep[+high]-IO/__NC), and a high-ranked one contextually restricted to all other environments (Dep[+high]-IO/¬__NC) — which could be differently ranked with respect to Dep[+back]-IO: Dep[+high]-IO/¬__NC ≫ Dep[+back]-IO ≫ Dep[+high]-IO/__NC.

Unstressed raising admits to a similar analysis: since this is definitely a reduction process, changing [e] to a lower and longer vowel [a] would not be a sensible mapping given the P-map. We might be able also to shoe-horn the umlaut case into an analysis of this sort as well: it might be reasonable to assume that there is less of a perceptual change in mapping /e/ to [i] than to [a] when followed by a high front vocoid. If we take all these together, then the Dep[+high]-IO constraints in the contexts where change is perceptually minimal can be ranked above (context-free) Dep[+back]-IO, while the Dep[+high]-IO constraints covering the complement set of environments are ranked below it.

This shows that it may be possible to make the stem-formation grammar consistent with the general phonology of the language, though it does not really accomplish the presumed desideratum of having the ranking follow from independent principles. However, upon further reflection, this might not be such a desideratum after all. Consider the ramifications of the umlaut case (and the impact of the consonantal alternations) discussed in Section 4.4: the way this process interacts with stem formation suggests that non-contrastive properties (which is what [±high] would be in that context) are not taken into account in the evaluation of REALIZE MORPHEME.

This suggests that application of the regular phonology is actively ignored by the stem formation system, presumably because the application of regular phonology undermines the effectiveness of contrast between stems. That is to say, if two stems differ only in whether a regular phonological process applied in its regular phonological context, that is not a good cue that the two stems are systematically distinct. For this reason, the stem formation system must actually manipulate features/properties that are not subject to regular phonological processes in that context if it is to properly effect morphological contrast. As such, it follows that the rankings employed in stem formation should not necessarily follow from independent facts about the phonology.

This would be consistent with a phonological component that is separated into distinct levels, as alluded to in Section 4.5.3.2 above apropos of the problem posed by the inflectional ending /-t/ for the formulation of the *V:RC constraint. If we adopted a version of Stratal OT (consult Kiparsky 2015, among others) where there was a "stem-level" phonological grammar that was independent from the "word-level" phonological grammar, then we would expect a division of exactly the sort we seem to be observing here. Furthermore, we should expect "irregular" morphophonology of this sort to reside at the stem level (cf. Bermúdez-Otero 2013), as it is a restructuring of the formerly regular accent-based ablaut system of Proto-Indo-European (see Sandell & Zukoff 2017, and the brief discussion in Chapter 5). I do not believe, however, that a wholesale adoption of standard Stratal OT will be sufficient for the current problem, as the Gothic/Pre-Proto-Germanic stem formation system requires simultaneous reference to bases (by the REALIZE MORPHEME constraints) and to the input (by the faithfulness constraints). (See Section 4.8 for some additional discussion on this point.) I leave it as a question for future research the extent to which these theoretical frameworks can be satisfactorily integrated.

4.5.8.2 Applying the Analysis to Gothic

In this section, I used Pre-Proto-Germanic forms as the basis for the analysis, rather than Gothic. This was because several of the sound changes that applied between the two stages diminished the transparency between related stem forms, and thus would have required a more complicated analysis (see Section 4.4 for discussion). I will now focus on the result of one of these changes: the merger of Proto-Germanic */e/ and */i/ as Gothic /i/ (see Section 4.4.1).

In Pre-Proto-Germanic, the root vowel for roots of Class I–V was /e/. Its (contrastive) features were [-back], [-high], and having one mora (which I will here simplify as [-long]). In, for example, Class V, the mapping that marked the preterite stems consisted of changing [-long] to [+long] in the plural, and [-back] to [+back] in the singular. This is shown in (54a). These mappings each violated one faithfulness constraint: DEP[+long] (= DEP-μ) in the plural; DEP[+back] in the singular. The DEP[+long] violation was compelled by RM:PRET, and the DEP[+back] violation was compelled by the combined force of RM:PRET and RM:SG. Both mappings represent the optimal minimal faithfulness violations in their particular contexts.

(54)   Root-to-Stem mappings in Pre-Proto-Germanic vs. Gothic (Class V)

	Root		Stem	Feature(s) Changed	Constraint(s) Violated
a	PPGmc √geb-	↗	PRET.PL *geːb-*	[±long]	DEP[+long]
		↘	PRET.SG *gab-*	[±back]	DEP[+back]
b.	Gothic √gib-	↗	PRET.PL *geːb-*	[±long], [±high]	DEP[+long], DEP[-high]
		↘	PRET.SG *gab-*	[±back], [±high]	DEP[+back], DEP[-high]

This transparency is destroyed after the root vowel changes to /i/ in Gothic, as shown in (54b). Among the surface stem forms, the sound change only directly affects the present, which now has the mapping √gib- → [gib-] (versus PPGmc √geb- → [geb-]). The faithful mapping in the present serves as the base for the preterite derivations. The problem here is that, in both preterite stem mappings (which did not change via regular sound change), there is now a two feature difference from the root (as well as the present base). The REALIZE MORPHEME constraints only require the change of a single feature. Therefore, if a licit structure can be derived by changing a single feature, that should be selected as the optimal preterite mapping.[39] In both cases, it would seem as though at least one such mapping exists.

The preterite plural both maps [-long] → [+long] and [+high] → [-high]. Applying just the latter change, i.e. √gib- → ✗ **[geb-], does not yield a licit string, as [e] is not permitted outside of the breaking context (where it is more properly [ɛ]; though this never contrasts with [e]; cf. (15)). However, applying just the former change does: √gib- → ✓ **[giːb-]. I see no way to deploy basic markedness and/or faithful constraints in such a way that a mapping to surface [-iː-] could be ruled out here without also incorrectly ruling out the the present of Class I: e.g., Gothic √bijt- → PRES [biːt-] (= //bijt-//) (formerly PPGmc √bejt- → [bejt-]).

We might attempt to get around this by employing markedness constraints that are lexically restricted to applying in the preterite. If there were a constraint *[iː]PRET, this could block a mapping to **[giːb-] in the preterite plural while allowing [biːt-] to surface in the present. As long as both DEP[+long] and DEP[-high] were both ranked below *[iː]PRET, and also were both ranked below any other faithfulness constraints whose violation could satisfy REALIZE MORPHEME for the relevant root shape, then a mapping to [geːb-] would be optimal. However, this constraint is probably too powerful. While no Gothic strong verbs have preterites with [iː], weak verbs freely have [iː] in their roots, and this surfaces faithfully in the preterite: e.g., infinitive *(ga)leikon* 'to liken, compare' (Lambdin 2006:331), whose stem is [liːk-oː-], would have a preterite first person singular *(ga)leikoda*, whose stem is [liːk-oː-ð-]. Therefore, unless we were willing to index the marked-

---

[39] We encounter an exactly equivalent problem in Class VIId already in Pre-Proto-Germanic; see Section 4.7.2.

ness constraint not to the PRETERITE generally, but somehow exclusively to the *strong* preterites, this avenue of explanation will not be successful.[40]

This would significantly reduce the appeal of the analysis, since this reifies the strong preterites, which is exactly the opposite of the goal of the analysis: the goal was to derive the distinct behavior of the strong preterites entirely from independent (if not independently motivated) machinery. Nevertheless, there must be some point in the history of the strong verbs within Germanic at which speakers no longer employ this sort of analytical system; after all, I am not claiming that Modern English speakers have any such grammar in their heads. If these problems are insurmountable, than perhaps already in Gothic such a system did not exist in its purest form.

One last ditch effort to circumvent this problem would be to claim that, while */e/ and */i/ totally merged on the surface, this did not actually result in restructuring of the underlying representations of the roots of Class I–V; that is to say, Gothic present stem [gib-], for example, is actually still derived from underlying root /geb-/. The claim would be that, despite absolute neutralization on the surface, the nature of the stem alternations in the strong verb system was sufficient to allow speakers to posit underlying /-e-/. The present is then derived by an unfaithful mapping triggered by a general markedness constraint *[e] (i.e. *[-back,-high,-long]).

This *[e] constraint has to be highly ranked in the language's regular phonology anyway, since surface [e] is banned in virtually all contexts. The exception again is the breaking context (/__{r,h,h^w}), where [i] (and likewise short high back [u]) is not permitted to surface, and [e] appears instead ([ɔ] surfaces when the vowel in question is back). This means that, regardless of our analysis of the underlying vowel of the strong verbs, we have a totally allophonic ranking: *[+high,-long] ≫ *[-back,-high,-long] ≫ FAITH[±high]. This will deterministically map underlying /i/ to [i] or [e] depending on the context, and do the same for any potential underlying /e/'s — which are in principle available given richness of the base.

The unfaithful mapping in the present would not pose any problems for the base with respect to REALIZE MORPHEME. It would take as its base [gib-]. This would mean that, in the preterite, the faithful mapping **[geb-] would not be ruled out by REALIZE MORPHEME; rather, it is ruled out by the same markedness constraint that prevents it from surfacing in the present. The normally optimal mapping of /-e-/ to [-i-], then, is what is ruled out by REALIZE MORPHEME.

It thus seems as though it may be possible to tailor the REALIZE MORPHEME-based analysis of preterite stem formation to Gothic itself, if we assume that certain sound changes did not result in restructuring of underlying representations. Many details of this analysis would still need to be worked out, but I leave this as a topic for future investigation. At present, then, my claim is that the analytical system developed in this chapter operated as such in Pre-Proto-Germanic, and that it is possible though not certain that it persisted as such into Gothic.

4.6   Reduplicant Shape and the Stem Formation System

The ranking developed in the previous section for the analysis of preterite stem formation in Pre-Proto-Germanic is consistent with the constraint ranking developed in Chapter 1 to account for the pattern of reduplicant shape in Class VII verbs in Gothic. I follow the standard view that Proto-Germanic should be reconstructed with the same copying patterns as displayed by Gothic. However, Jasanoff (2007) shows that Northwest Germanic seems to exhibit significantly different copying patterns for cluster-initial roots, as will be detailed in Section 4.6.2 below. This makes it less certain

---

[40] We would presumably need multiple markedness constraints of this type, as we will need also to rule out mappings like √*gib-* → PRET.SG **[gub-] rather than [gab-], and possibly others.

whether to attribute the Gothic pattern to Proto-Germanic. However, since the apparent changes are attested across Northwest Germanic (and, according to Jasanoff, the change in copying pattern contributed to additional changes in Northwest Germanic), it is certainly possible that this is a Northwest Germanic-internal innovation (i.e. post-Proto-Germanic), and that the traditional reconstruction is valid.

4.6.1 Synchronic Analysis of Reduplicant Shape (Based on Gothic Patterns)

The table in (55) provides an exhaustive list of the cluster-initial reduplicating Class VII verbs in Gothic.

(55)   Cluster-initial Class VII preterites in Gothic (forms from Lambdin 2006:115)

a.   **Obstruent-sonorant–initial roots → C$_1$-copying preterites**

	Root	Present (1SG)		Preterite (1SG)		
(Class VIIa)	'tempt'	*fraisa*	[frɛːs-a]	*faifrais*	[f<u>ɛ</u>frɛːs]	not **[fr<u>ɛ</u>frɛːs]
(Class VIIb)	'sleep'	*slepa*	[sleːp-a]	*saislep*[41]	[s<u>ɛ</u>sleːp]	not **[sl<u>ɛ</u>sleːp]
(Class VIIc)	'bewail'	*floka*	[floːk-a]	*faiflok*	[f<u>ɛ</u>floːk]	not **[fl<u>ɛ</u>floːk]
(Class VIId)	'weep'	*greta*	[greːt-a]	*gaigrot*	[g<u>ɛ</u>groːt]	not **[gr<u>ɛ</u>groːt]

b.   **Obstruent-obstruent–initial roots → cluster-copying preterites**

	Root	Present (1SG)		Preterite (1SG)		
(Class VIIa)	'possess'	*stalda*	[stald-a]	*staistald*	[st<u>ɛ</u>stald]	not **[s<u>ɛ</u>stald]
(Class VIIa)	'divide'	*skaida*	[skɛːð-a]	*skaiskaiþ*	[sk<u>ɛ</u>skɛːθ]	not **[s<u>ɛ</u>skɛːθ]

These roots show two distinct copying patterns, depending on the composition of the initial cluster. If the root begins with an obstruent-sonorant cluster (i.e. TR or SR) — as also for roots with an initial singleton consonant (not shown; see Section 4.10.7 in Appendix I for a complete list) — copying takes the shape *C$_1$e*- (55a). On the other hand, if the root begins in an obstruent-obstruent cluster (i.e. ST), the whole cluster gets copied as *C$_1$C$_2$e*- (55b). Given that the "reduplicant" arises without an underlying RED morpheme, it must be the case that the fixed "reduplicative" vowel [e] arises through *copy + reduction* (i.e. the fixed [e] stands in correspondence with the root vowel), rather than through morphological fixed segmentism (Alderete et al. 1999; see Chapter 1).[42]

The explanation for the distinction in copying behavior by cluster type, motivated empirically in Chapter 1 and to be investigated further in Chapter 5, is the NO POORLY-CUED REPETITIONS constraint (*PCR). Its definition is repeated here in (56).

(56)   **NO POORLY-CUED REPETITIONS (*PCR) [≈ *C$_\alpha$VC$_\alpha$ / _C$_{[-sonorant]}$]**
        For each sequence of repeated identical consonants separated by a vowel (C$_\alpha$VC$_\alpha$), assign a violation * if that sequence immediately precedes an obstruent.

---

[41] This root attests preterites with both voiceless [s] and voiced [z] in root-initial position. See footnote 20 above.

[42] In (55), I represent the reduplicant vowel as [ɛ], as it clearly is in Gothic. Primarily for typographical ease and for comparison with the other Indo-European languages, I use [e] in the remainder of the discussion, as it is reasonable to assume that, at the stage of Pre-Proto-Germanic we are investigating, there was no contrast between the two.

Looking at the facts of the copying pattern in isolation (that is, prior to considering how it fits into the larger preterite stem-formation system developed in this chapter), the ranking required to generate the differential behavior by cluster type is the following:

(57)   **Reduplication ranking** (Chapter 1): *PCR, ANCHOR-L-BR ≫ *CC ≫ CONTIGUITY-BR

This ranking already has one point of contact with the ranking developed above to account for the preterite stem formation patterns, namely *CC. Additionally, I argued above that INTEGRITY-IO was the constraint violated by reduplication in Class VII. Assuming that an INTEGRITY-IO violation is assessed for *each* input segment which stands in correspondence with multiple output segments, INTEGRITY-IO is also relevant in the calculation of reduplicant shape. It works towards essentially the same end as *CC, penalizing extra copying; they are both "size restrictor" constraints (see Spaelti 1997, Hendricks 1999, Riggle 2006, among others).[43] The tableaux in (58) and (59) confirm that including INTEGRITY-IO in the ranking, at the same spot as *CC, generates the right reduplicant shapes. (INTEGRITY-IO violations can be reckoned by counting a candidate's indices, which notate segments standing in BR-correspondence.) This provides the two rankings with another point of contact.

(58)   TR C_1-copying: $\sqrt{}$*floːk* → *fefloːk* 'he wept'

/floːk, Ø$_{PRET}$/	ANCHOR-L-BR	INTEGRITY-IO	*CC	CONTIGUITY-BR
a. ☞ $f_i e_k f_i loː_k$k-		**	*	*
b. $f_i l_j e_k f_i l_j oː_k$k-		***!	**!	
c. $l_j e_k f l_j oː_k$k-	*!	**	*	
d. $e_k f loː_k$k-	*!	*	*	

(59)   ST cluster-copying: $\sqrt{}$*stald* → *stestald* 'he possessed'

/stald, Ø$_{PRET}$/	ANCHOR-L-BR	*PCR	INTEGRITY-IO	*CC
a. $s_i e_k s_i ta_k$ld-		*!	**	**
b. ☞ $s_i t_j e_k s_i t_j a_k$ld-			***	***
c. $t_j e_k s t_j a_k$ld-	*!		**	**
d. $e_k s ta_k$ld-	*!		*	**

One additional reduplication pattern that is left to consider is that of Class VII vowel-initial roots. Gothic attests just two (both of Class VIIa, and, presumably coincidentally, both of the shape *a + glide + k*): $\sqrt{}$*auk* 'increase' → preterite *aiauk-* [ɛ.ɔːk-] (< *e.awk*), $\sqrt{}$*aik* 'deny' → preterite *aiaik-* [ɛ.ɛːk-] (< *e.ajk*). These forms show copying of just the root-initial vowel, creating hiatus. The constraint ranking already proposed derives this straightforwardly, as long as ONSET (defined in (60)) ranks below INTEGRITY-IO. This ranking is provided in (61), and demonstrated in (62).[44]

---

[43] To this list of size restrictors, we could also add ALIGN-ROOT-L ('*Assign a violation mark * for each segment intervening between the left edge of the root and the left edge of the word*'), which I have used for that purpose (with respect to a different morpheme) for Ancient Greek in Chapter 2 and for Anatolian in Chapter 3. In Gothic, ALIGN-ROOT-L's effects will be completely coextensive with INTEGRITY-IO, and thus must be ranked below all the constraints which also dominate INTEGRITY-IO. Since it is redundant, I omit it from further discussion.

[44] Notice that candidate (62b) is equivalent to what is observed for vowel-initial roots in Anatolian (see Chapter 3), where *PCR has ceased to be active.

(60) **ONSET**
Assign one violation mark * for each onsetless syllable.

(61) **New Rankings:** INTEGRITY-IO ≫ ONSET

(62) Reduplication in vowel-initial roots: √*auk* → *aiauk* 'he increased'

/awk, Ø$_{PRET}$/	ANCHOR-L-BR	*PCR	*CC	INTEG-IO	ONSET
a. ☞ e$_i$.a$_i$wk-			*	*	**
b. e$_i$w$_j$a$_i$w$_j$k-		*!	*	**	*
c. w$_j$aw$_j$k-	*!	*!	*	*	
d. e$_i$k$_k$a$_i$wk$_k$-			*	**!	*
e. k$_k$awk$_k$-	*!		*	*	
f. e$_i$w$_j$k$_k$a$_i$w$_j$k$_k$-			**!	**!*	*

With the constraints relevant for reduplicant shape now fully ranked with respect to *CC and INTEGRITY-IO, both of which are focal points in the stem formation rankings, we can check to make sure that the analyses of the two processes are consistent. That is to say, we might worry that constraints and rankings necessary for one aspect of the grammar might cause problems for the other. Happily, no such problems exist.

The calculation of all three types of copying behavior — the C$_1$-copying pattern for TR roots, the cluster-copying pattern for ST roots, and the vowel-copying pattern of vowel-initial roots — can be properly executed when stacked up against other stem-contrast–driven repairs. This is illustrated for ST roots in (63) below. Furthermore, reconciling the two aspects of the grammar reveals additional critical rankings among several of the constraints proposed for the stem formation patterns relative to *CC. Specifically, any constraint which is required to dominate INTEGRITY-IO, such that reduplication is chosen as the optimal mapping for Class VII as opposed to some other possible mapping, must also dominate *CC, since reduplication is effected in such cases even if it creates an additional cluster. Two types of suboptimal unfaithful mappings, and the constraints which they fatally violate, are exemplified in (63), though many more could be included. The fully reconciled ranking for reduplication and preterite stem formation is shown in (65).

(63) Preterite Singular of Class VIIa

INPUT: /stald, Ø$_{PRET}$/ BASE: PRES [stald-]	RM:PRET	*V:RC	MAX[+back]	INTEGRITY	*CC	MAXV
a. stald-	*!				**	
b. sto:ld- (← //sta:ld-//)		*!			**	
c. stuld- (← //stld-//)			*!		*(*)	*
d. steld-			*!		**	
e. ☞ s$_i$t$_j$e$_k$s$_i$t$_j$a$_k$ld-				***	***	

(64) **New Rankings:** RM:PRET, *V:RC, MAX[+back]-IO, *etc.* ≫ *CC

(65)   **Total ranking for stem contrast and reduplicant shape**

$$\text{MAX-}\mu\text{-IO} \qquad \left\{ \begin{array}{c} \text{MAXC-IO} \\ \text{DEPV-IO} \\ \text{*V:RC} \end{array} \right\} \qquad \left\{ \begin{array}{c} \text{ANCHOR-L-BR} \\ \text{RM:PRET} \\ \text{*PCR} \end{array} \right\} \qquad \left\{ \begin{array}{c} \text{MAX[+back]-IO} \\ \text{DEP[+high]-IO} \\ \textit{etc.} \end{array} \right\}$$

$\text{*V}_{\mu\mu\mu}$

*CC   INTEGRITY-IO

CONTIGUITY-BR   RM:SG   ONSET

DEP[+back]-IO

DEP-μ-IO   MAXV-IO
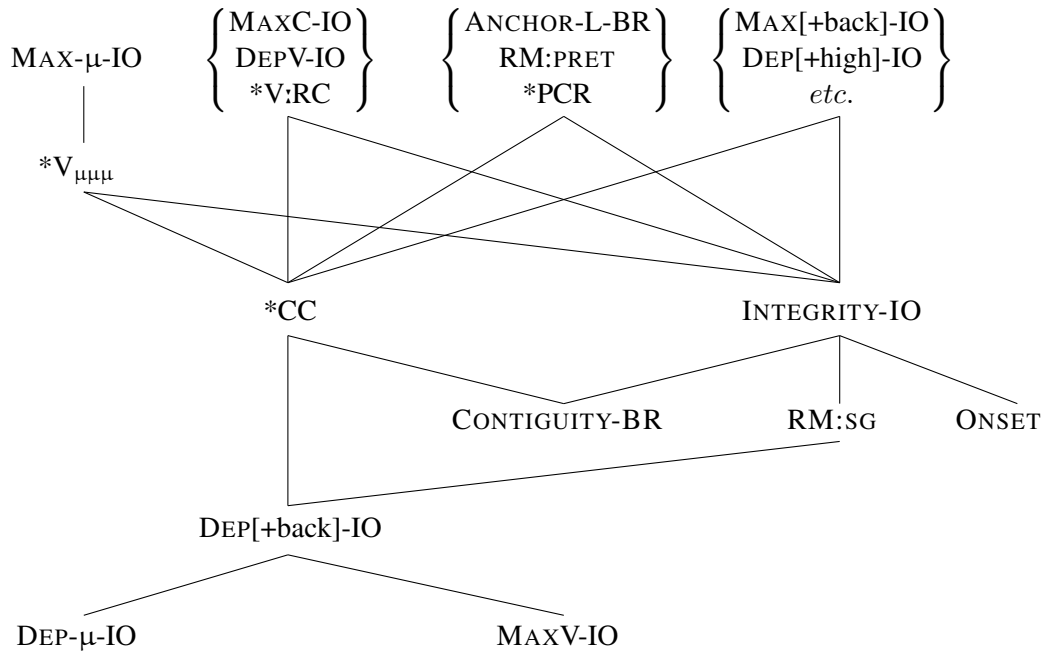
4.6.2   Evidence of Cluster-Initial Reduplication Patterns in Northwest Germanic

As mentioned above, this analysis is based on the assumption that Pre-Proto-Germanic displayed the same copying patterns as are attested in Gothic. This requires that any divergent patterns in Northwest Germanic be identified as post-Proto-Germanic developments. This subsection lays out the Northwest Germanic data, and argues that it is consistent with these assumptions.

The evidence of reduplication to cluster-initial roots in Northwest Germanic is scant and ambiguous. Nevertheless, it almost certainly does not reflect the copying system attested in Gothic. The table in (66) below lists the forms adduced by Jasanoff (2007) in Old Norse (ON), Old High German (OHG), and Old English (OE) (see also Fullerton 1991). (Note that Old High German orthographic <z> represents phonetic [ts]. The *z in reconstructed prior forms does represent phonetic [z].)

(66)   Reduplication in Northwest Germanic cluster-initial roots (Jasanoff 2007:246, 255, 262)

	Infinitive		"Reduplicated" Preterite	
ON	*gróa*	→	*grera*	
	snúa	→	*snera*	(< *sneza < *snesa ?)
	gnúa/bnúa	→	*gnera/bnera*	("analogical"; Jasanoff 2007:245f., 255)
	slá	→	*slera*	("clearly secondary"; Jasanoff 2007:245f., 255f.)
OHG	*pluozan*	→	*pleruzzun*	(< *blerōt- < *blelōt-)
	(ana-)stōzan	→	*(ana-)steroz*	(< *stezaut)
	(ki-)scrōtan	→	*(ki-)screrot*	(< *skreraud- or < *skrezaud- < *skresaud-)
OE	*(on-)drǣdan*	→	*(on-)dreord*	
	spātan	→	*speoft*	(= //speopt//)

The Old English forms are part of a regularity that persists (at least dialectally) in Old English — perhaps extended from a core of forms reflecting the original behavior of Class VIId preterite plurals in Proto-Germanic — whereby all *bases* which are cluster-initial, not just roots which are cluster-initial, display reduplication of the type shown in (66). This will be detailed further in Section 4.7.2 below. Since this is quasi-productive in Old English, it is less secure that the evidence definitively reflects the situation of Proto-Northwest Germanic (such as it is). Old Norse and Old High German do not seem to reflect the quasi-productivity of Old English (or at least not the same kind), and so the characteristics of their forms are more likely to reflect Proto-Northwest Germanic. Limited as they may be, all the Old Norse and Old High German preterite forms in (66) can be characterized by the generalizations in (67).

(67) Generalizations about Northwest Germanic copying patterns

    a. The root-initial cluster is maintained as such
    b. The root-initial cluster contains either an [s] or a liquid [l,r] (or both)
                (except ON *gnera*, which Jasanoff 2007:255 identifies as "analogical")
    c. The root-initial cluster is followed by [er]

The predominance of [r] in these forms is clearly important (cf. Jasanoff 2007:246–247). In Northwest Germanic, in intervocalic position, [r] has two possible sources: (i) original *r*, or (ii) *z*, which arose as the voiced allophone of *s*. Given this, and the generalization in (67b) that all the attested clusters contain a liquid or an [s], at least one of the two sources of [r] is present in each root cluster. This suggests that we want to seek a solution that is based on the copying of these segments.

But first let us show that a possibly simpler solution does not work: if we are to ascribe maximal unity to this class, the choice of reduplicated consonant cannot be based on position. Consider a potential comparandum: reduplication in Latin ST-initial roots (see Chapter 6 for more detailed analysis). As shown in (68), ST-initial roots display an infixal reduplication pattern, where the reduplicant is displaced from the left edge, and the second member of the cluster (i.e. the stop) is what gets copied.

(68) Latin infixing perfect reduplication to ST roots (forms from Weiss 2009:410)

Root		Perfect		
√*st*	'stand/stop'	s-*te*-t-ī	not **se-st-ī	(but present *si*-st-ō)
√*spond*	'promise'	s-*po*-pond-ī	not **so-spond-ī	
√*scid*	'cut'	s-*ci*-cid-ī	not **si-scid-ī	(<c> = [k])

If we were to try to claim that Northwest Germanic was displaying a similar pattern, we might describe it as follows: place a reduplicative *Ce-* string before the last member of the initial consonant cluster, and have the *C* be a copy of that member of the cluster.[45] This would transparently match Old Norse *grera* and Old High German *screrot*, as well as Old High German *pleruzzun* (and Old

---

[45] The Latin infixation pattern is motivated by *PCR, as prefixation would induce a *PCR violation. *PCR, at least in its standard version, cannot be employed here, since the repetitions being avoided are *obstruent-sonorant*. However, if we posited a more general *PCR constraint for this stage that dispreferred consonant repetitions before any consonant, rather than just obstruents, this could induce infixation here. Alternatively, positing an ANCHOR-L-IO constraint could do so as well, since prefixal copying would mean that the "reduplicant" consonant is at the left edge, rather than the root consonant. (If we were dealing with an actual reduplicative morpheme — which I have explicitly claimed we are not, at least in Proto-Germanic — then we could use ALIGNMENT to trigger infixation: ALIGN-ROOT-L ≫ ALIGN-RED-L.)

Norse *slera*, secondary as it may be) once we take into account *l*-dissimilation. This would not, however, explain Old Norse *snera* and *gnera/bnera*, which it would predict to be **snena and **gnena/**bnena, or Old High German *steroz*, which it would predict to be **stetoz. Therefore, it is clear that a rule of exactly this sort is not capable of directly deriving the full range of Old Norse and Old High German forms.

It would seem, then, to be about segment type. If we posit that the (Proto-)Northwest Germanic reduplication pattern is specifically seeking out liquids and [s]'s for copying, then we can explain the choice of copied segment given the available data, except of course for Old Norse *gnera/bnera*, the only form to have neither a liquid nor [s] in the root. To this, however, we can add a singleton-initial apparently reduplicated form that also shows [r] unexpectedly: Old High German *būan* → *biruun* (Jasanoff 2007:246–247). This has a similar sort of structure to the other forms under discussion (though note that the first vowel is *-i-* as opposed to *-e-*), except that where the others have an initial cluster including a liquid or [s], this form has just [b]. When viewing this form alongside *gnera/bnera*, it becomes clear that we cannot identify all the relevant forms as coming from a phonologically regular reduplication pattern based on target segment type; there must be some component of the system which is favoring the realization of [r] even if there is no source from which it could be copied (quasi-)faithfully.[46]

I will therefore have to leave the Northwest Germanic forms without a conclusive analysis. Nonetheless, the inconclusive nature of the data itself allows for conclusions about the status of reduplication in Northwest Germanic relative to Proto-Germanic and Gothic. Based on the available data, it seems clear that at least some TR roots showed an infix-like reduplication pattern in Northwest Germanic, rather than the C_1-copying pattern observed in Gothic. This would have to be treated as an innovation against Proto-Germanic, based on parsimony of diachronic change: since PIE very clearly has C_1-copying and Gothic very clearly has C_1-copying, it would not be parsimonious to assume a change between PIE and Proto-Germanic (that feeds Northwest Germanic) and a subsequent reversive change into Gothic.

The treatment of ST-initial roots in Northwest Germanic is completely unclear. Given that there is clearly extension of the medial [r] pattern, it is not at all clear that Old High German *(ana-)stōzan* → *(ana-)steroz*, or even *(ki-)scrōtan* → *(ki-)screrot*, reflects the original Northwest Germanic state of affairs; that is, these might be later innovations. As such, it is no appropriate to say that attested Northwest Germanic provides evidence that Proto-Northwest Germanic did not retain a cluster-copying pattern for ST akin to Gothic; it simply provides no evidence in any direction. Since the only concrete evidence for the reduplicative treatment of ST-roots in Germanic comes from Gothic, it is reasonable to assume that Proto-Germanic had the Gothic treatment.

The only other reasonable assumption is that it retained the PIE treatment, which I will argue in Chapter 7 to have been C_1-copying. This would not materially change the analysis of preterite stem formation in Pre-Proto-Germanic, but it would negate the additional ranking arguments that come from the permission of cluster-copying (namely the ranking of *CC below the other high-ranked markedness and faithfulness constraints).

Getting *screrot* rather than **scecrot is somewhat trickier under any formulation. Appealing to some sort of consonant-contiguity could help, though this might lead us toward preferring an interpretation of the string as *screrot* not *screrot*, which may ultimately be more appealing anyway.

[46] There are also two Old Norse singleton-initial reduplicated forms have medial [r] as expected: *róa* → *rera*, *sá* → *sera* (Jasanoff 2007:246–247). These are frequently invoked as sources that contributed to the spread of medial [r] in reduplicated forms.

4.7 The Problems with Strong Class VIIb and VIId

I now return to the last set of strong verbs: Class VIIb and Class VIId. Both types have underlying root vowel /eː/, and no properties that would obviously distinguish one from the other. As such, we should expect them to behave the same. The fact that they don't is problematic. Even more problematic is the behavior of Class VIId itself. It appears to show "too many repairs" simultaneously in its response to RM:PRET. This cannot be derived within the current analysis.

In this section, I first treat Class VIIb, and provide an analysis of its behavior that is consistent with the larger system, if different in a crucial way. I then treat the more problematic Class VIId, considering both its diachrony and its synchrony, but it will ultimately be left still wanting of a solution.

4.7.1 Strong Class VIIb Preterite Singulars and Plurals: A Minor Problem

Gothic has one example of a root with /eː/ where the vowel appears as such consistently across the three verbal stems, i.e. present, preterite plural, and preterite singular: √*sleːp* → present *sleːp-*, preterite plural *sesleːp-*, preterite singular *sesleːp-*. Both preterite stems display one change relative to the root/present: the addition of reduplication. However, upon close inspection of the rankings developed in Section 4.5 (see (52) and (65) for complete rankings), this is actually not exactly what we would predict. This is shown in (69), where the ranking INTEGRITY-IO ≫ DEP[+back]-IO leads to the incorrect selection of the vowel backing candidate *sloːp-* (69e) over the desired reduplication candidate *sesleːp-* (69f). Furthermore, since INTEGRITY-IO also outranks RM:SG, the preterite singular will (incorrectly) receive the same mapping as well.

(69)   Preterite Plural of Class VIIb

INPUT: /sleːp, Ø_PRET/ BASE: PRES [sleːp-]	RM:PRET	MAX[+back]	MAX-μ	*V_μμμ	INTEGRITY	DEP[+back]	DEP-μ	MAXV
a. sleːp-	*!							
b. sleːːp-				*!			*	
c. slep-			*!					
d. sulp- (← //slp-//)			*!					*
e. ☞ sloːp-						*		
f. ☺ sᵢeₖsᵢlⱼeːₖp-					*!*			

Descriptively, the device we need in order to block vowel backing is something that says the following: *don't add a [+back] feature if the vowel has more than one mora*. We can refer to this as DEP[+back]/μμ-IO. If we assume that vowel contrasts (such as [±back]) are more robustly perceptible on long vowels than short vowels, such a constraint, and its ranking above the general constraint DEP[+back]-IO, might be said to follow from the P-map (Steriade 2009). If this is a reasonable constraint, and it ranks above INTEGRITY-IO, then we do indeed derive reduplication in such a case, as shown in (70).

(70) Preterite Plural of Class VIIb

INPUT: /sleːp, $\emptyset_{\text{PRET}}$/ BASE: PRES [sleːp-]	RM:PRET	DEP[+back]/μμ	INTEGRITY	DEP[+back]
a. sleːp-	*!			
b. sloːp-		*!		*
c. ☞ $s_i e_k s_i l_j eː_k$p-			**	

However, it must be noted that Class VIIb is comprised of just this one case. Roots with underlying /-eː-/ much more frequently display a different sort of behavior, namely that of Class VIId (to which I turn immediately below). Therefore, it is unclear whether we want to adopt this approach as a general property of the system, or if it should instead be encoded somehow as an exceptional pattern (whether or not it involves a constraint along these lines).

4.7.2 Strong Class VIId Preterite Singulars and Plurals: A Bigger Problem

The relatively much more frequent type of behavior for roots with underlying /-eː-/ is that of Class VIId. At least in Gothic terms, Class VIId shows both reduplication and backing in both of its preterite stems: for example, $\sqrt{leːt\text{-}} \rightarrow$ present *leːt-*, preterite plural <u>*leloːt-*</u>, preterite singular <u>*leloːt-*</u>. Understanding Class VIId is one of the most intractable problems in Germanic historical linguistics (see Jasanoff 2007 and citations therein). It remains largely intractable within the current analysis as well. In this subsection, I will first lay out the diachronic facts, and put forward some suggestions which may help advance our understanding of the Proto-Germanic situation generally; but then I will show that, even with these updated assumptions, it is not clear that the the system being proposed is at all capable of handling the data in a transparent way.

4.7.2.1 Diachrony of Class VIId: From PIE to Pre-Proto-Germanic

There is evidence to suggest that the Gothic pattern for Class VIId may not fully reflect the behavior of this type in Pre-Proto-Germanic (see below). It will thus be instructive to first determine what the expected outcomes of the verbs of Class VIId would be, given the application of regular sound change to the forms of these roots that we can reconstruct for Proto-Indo-European. The verbs of Class VIId in Gothic can all be reconstructed as having PIE *-eh_1- in the root (cf. Rix et al. 2001); this is the source of the root vowel [eː] in Proto-Germanic terms. The table in (71) below provides the Class VIId roots which are attested in Gothic, coupled with their presumed etymologies (based on Rix et al. 2001).

The Germanic preterite is derived from the Proto-Indo-European perfect. In the PIE perfect, reduplication — which always took the shape *Ce-* (see Chapter 7 for discussion) — was obligatory. Different kinds of accentually-conditioned vowel alternations (i.e. ablaut) affected the singular and the plural (see Chapter 5 for details): the plural underwent vowel deletion ("zero-grade ablaut"), while the singular underwent a featural change of /e/ $\rightarrow$ [o] ("*o*-grade ablaut"). For roots of the shape *$C_i eh_1 C_j$, which I will exemplify using *$\sqrt{leh_1 d\text{-}}$, these conditions would yield the paradigm in (72a). Once we account for the diachronic developments of the laryngeals — deletion with compensatory lengthening in V_C position (affecting the present and the perfect/preterite singular),

(71) Strong Class VIId roots in Gothic (forms from Lambdin 2006:93)

Root (Gothic << PIE)			Preterite	
$\sqrt{gr\bar{e}t}$-	<	*g^hreh_1d-	*gaigrōt*	'weep'
$\sqrt{l\bar{e}t}$-	<	*leh_1d-	*lailōt*	'let, allow, leave'
$\sqrt{r\bar{e}d}$-	<	*$(H)reh_1d^h$-	*rairōþ*	'reflect on'
$\sqrt{t\bar{e}k}$-[47]			*taitōk*	'touch'
$\sqrt{s\bar{e}}$- (inf. [sɛː.an] ← /seː-an/)	<	*seh_1-	*saisō*	'sow'
$\sqrt{w\bar{e}}$- (inf. [wɛː.an] ← /weː-an/)	<	*weh_1-	*waiwō*	'blow'

and total loss (perhaps via a stage of vocalization) in C_C position (affecting the perfect/preterite plural) — we predict the paradigm in (72b).

(72) The perfect paradigm in PIE and its outcome in (Pre-)Proto-Germanic

		Root	Present	Perf./Pret. Sg.	Perf./Pret. Pl.
a.	PIE	*$\sqrt{leh_1d}$-	*leh_1d-	*$le\text{-}loh_1d$-	*$le\text{-}lh_1d$-[48]
b.	Pre-PGmc	*$\sqrt{leːt}$-	*$leːt$-	*$leloːt$-	*$le\text{-}lt$-

The preterite singular form is exactly what we find in the Gothic preterite singular *leloːt-*; but notably, in Gothic, this same stem shape also surfaces in the plural, which is not what we would predict from (72). The type of preterite plural that we predict in (72), *_lelt_-*, is arguably attested in the Old English (Anglian) dialectal form *leort-*, presumably from *_lelt_-* with liquid dissimilation (Jasanoff 2007:244–245). A number of additional preterites of this shape are also found among Old English dialectal forms, attested beside standard preterite stems with a long vowel. These are listed in (73). Note, however, that, among these, only *reord-* ← $\sqrt{r\bar{æ}d}$- << PIE *$\sqrt{(H)reh_1d^h}$- (and perhaps *dreord-* ← $\sqrt{dr\bar{æ}d}$-, though with an innovative treatment of clusters in reduplication, as seen also in *spātan → speoft-*; see Jasanoff 2007:esp. 262f. and the discussion in Section 4.6.2 above) can be directly traced back to the Class VIId root shape in *$C_ieh_1C_j$; the others have been secondarily attracted into this pattern sometime after Proto-Germanic (Jasanoff 2007:266).

The variation found within Old English might be suggestive of a partial solution to the problem in (Pre-)Proto-Germanic. In Old English, the $C_iVC_iC_j$- reduplicated forms stand beside $C_iV{:}C_j$-non-reduplicated forms. In Chapter 5, I will argue that the process of *deletion + compensatory lengthening under consonant repetition* ($VC_i \rightarrow V{:} \,/\, C_i\_C_j$), which is to be viewed as an effect of the anti-repetition constraint *PCR (see Section 4.6 above for its impact on reduplicant shape in Germanic), is operative in shaping the perfect/preterite plural stems of CeT roots in Germanic (i.e. the origin of the preterite plural of Class V) and in Sanskrit. The Old English alternants can be linked to each other by the operation of exactly this process; i.e., the dialectal reduplicated forms match the structural description of this rule, while the standard non-reduplicated forms reflect the structural change.

Setting aside for the moment what ramifications this would have for the rest of the system, let's consider what would happen if we assumed that this variation — i.e. the optional/variable applica-

---

[47] Rix et al. (2001:616–617) derive $\sqrt{t\bar{e}k}$- from PIE *teh_2g-, but this ought to yield initial *þ*- not *t*- via Grimm's Law. Potential external comparanda aside, a root with *$-eh_1$- would be sufficient for deriving the root vowel in Germanic.

[48] It is possible that the *$-lh_1d$- sequence (and *-CHC- more generally) would not have been phonotactically licit in PIE, and thus undergone some synchronic repair (see Byrd 2015 for recent discussion).

(73)   Old English Class VII forms (Fulk 1987:esp. 164, Jasanoff 2007:245)

Root		Standard Preterite	Dialectal Preterite
rǣdan	'to advise'	*rēd-*	_reord-_
lǣtan	'to let'	*lēt-*	_leort-_ (< *lelt-* via dissimilation)
hātan	'to call'	*hēt-*	_heht-_
lācan	'to play'	*lēc-*	_leolc-_
bēatan	'to beat'	*bēot-*	_beoft-_ (= //be̲o̲bt-//)
(on-)drǣdan	'to dread'	*drēd-*	*(on-)dreord-*
spātan	'to spit'	*spēt-*	*speoft-* (= //s̲p̲eopt-//)

tion of this deletion process — were reconstructible to Proto-Germanic. This would mean that, beside the form like *le̲l̲t-* which is predicted by regular sound change, we should also predict a form like *leːt-*, via the operation of the deletion process. This preterite plural *leːt-* would be identical to the present stem arrived at diachronically through laryngeal lengthening. At whatever stage where RM:PRET is active, the mini-paradigm of present *leːt-* ~ preterite plural *leːt-* would not be permitted to stand.

Faced with this problem, speakers would have to develop a new strategy for forming preterite plurals to these roots.[49] Evidently, the strategy they landed upon was importing the stem form of the preterite singular, i.e. *leloːt-*. While the mechanism by which such a form is allowed within the RM-based system in the first place is still mysterious (see below), whatever licenses it in the singular will certainly license it in the plural.

(74)   Variation in the (Pre-)Proto-Germanic preterite paradigm

	Root	Present	Pret. Sg.	Pret. Pl.
Pre-PGmc	*√/leːt-	*leːt-	*le̲l̲oːt-	*l̲e̲-lt- ~ *l̲e̲loːt-

Since the analysis being proposed in this chapter is essentially one that describes the situation of Gothic but within the phonological system of Pre-Proto-Germanic, it is likely that the evidence I am asserting is unduly influenced by Gothic. That is to say, certain assumptions about the structure of the data may be biased towards the grammar that generates preterite plural variant *le̲l̲oːt-* and biased against the grammar that generates preterite plural variant *le̲-lt-*. Put another way, if this dimension of variation correlated with other variable properties of the grammar, these other properties may well have been completely filtered out of the system by the time we arrive at Gothic. If this is the case, then it may in fact not be appropriate to try to shoe-horn the *le̲-lt-* variant into the precise synchronic grammar we are trying to construct here, since it is implicitly trying to model the set of forms which ultimately feed Gothic. As such, I leave further consideration of the status of the *le̲-lt-* variants as a topic for much-needed future investigation.

---

[49] I'm assuming here that the variation is not intra-speaker variation but rather inter-speaker variation, such that the speakers who are trying to produce *leːt-* have a grammar which will not allow them to produce *le̲l̲t-*, presumably because of a highly ranked *PCR constraint.

4.7.2.2 Class VIId and the Synchronic System of Pre-Proto-Germanic

We now have a somewhat clearer, though still exceedingly murky picture of the status of Class VIId going into Pre-Proto-Germanic. Given the discussion about variability and the nature of the data, it seems reasonable to focus on deriving the paradigm of Class VIId that is ultimately reflected in Gothic, where it shows both reduplication and backing in both of its preterite stems: √*leːt-* → present *leːt-*, preterite plural *leloːt-*, preterite singular *leloːt-*. However, when we approach this problem from a strictly synchronic perspective — namely, considering how the grammar ought to treat roots with underlying /-eː-/ — we will see that progress on this front is elusive. Nonetheless, the endeavor will reveal that there are certain respects in which we are close to an answer, even if it is not clear what that answer will ultimately be.

There is a fundamental reason why the RM-based synchronic analysis is simply not capable of deriving this pattern: the Class VIId preterite mappings show greater unfaithfulness than is strictly necessary. In apparent service of just RM:PRET (since no contrast is effected between singular and plural), the root undergoes *two* changes — incurring violations of both INTEGRITY-IO and DEP[+back]-IO (and indeed the more specific, higher-ranked version DEP[+back]/μμ-IO employed for Class VIIb in Section 4.7.1 above, if we were to adopt that approach) — rather than just the seemingly requisite one. This means that, whether we do adopt the DEP[+back]/μμ-IO constraint (as in tableau (75)) or we do not (as in (76)), the desired candidate [leloːt-] (75d)/(76d) is *harmonically bounded* by two other candidates: the one that shows just reduplication **[leleːt-] (75c)/(76c), which is selected in tableau (75); and the one that shows just vowel backing **[loːt-] (75b)/(76b), which is selected in tableau (76). Both of these types of surface strings are allowed elsewhere in the system: reduplication with stem vowel [eː] is reflected in Class VIIb (such as it is); stem vowel [oː] without reduplication is reflected in Class VI (albeit arising through a different mapping). Whichever tableau we adopt, the ranking of RM:SG is low enough that no distinct mapping would be selected in the singular.

(75) Preterite Plural of Class VIId

INPUT: /leːt, Ø_PRET/ BASE: PRES [leːt-]	RM:PRET	DEP[+bk]/μμ	INTEGRITY	RM:SG	DEP[+back]	DEP-μ	MAXV
a. leːt-	*!						
b. loːt-		*!			*		
c. 💣 $l_i e_j l_i e{:}_j t$-			**				
d. ☺ $l_i e_j l_i o{:}_j t$-		*!	**		*		

(76) Preterite Plural of Class VIId

INPUT: /leːt, Ø_PRET/ BASE: PRES [leːt-]	RM:PRET	INTEGRITY	RM:SG	DEP[+back]	DEP-μ	MAXV
a. leːt-	*!					
b. 💣 loːt-				*		
c. $l_i e_j l_i e{:}_j t$-		*!*				
d. ☺ $l_i e_j l_i o{:}_j t$-		*!*		*		

Notice that these two rankings (i.e. with and without DEP[+back]/μμ-IO) respectively select the two unfaithful mappings that together characterize the actual form. This could lead one to entertain a scenario in which the two variants were present among the population, and somehow learners eventually converged on the combination of the two. This could have been bolstered by the evidence from the identical form inherited in the preterite singular, though it is unclear why that form should not also have been subject to immediate remodeling. The one Class VIIb root — $\sqrt{sle:p}$- → preterite $\underline{se}sle:p$- — would thus be an outlier, reflecting purely the grammar of the speakers who had high-ranked DEP[+back]/μμ-IO, rather than the mixed type.

Alternatively, one could pursue an approach based on the following observation: the preterite stems of Class VIId, taken by themselves, look just like those of Class VIIc. That is to say, they display reduplication and a stem vowel [o:]. The way this was generated in Class VIIc was by virtue of that [o:] vowel being proper to the underlying representation of the root and the realization of the present stem. Therefore, one way to get the derivation of the Class VIId preterite to generate the correct stem forms would be to treat them in exactly this way. This is illustrated in (77).

(77)   Preterite Plural of Class VIId

INPUT: /**lo:t**, Ø$_{\text{PRET}}$/ BASE: PRES [**lo:t-**]	RM:PRET	MAXI[+back]	INTEGRITY	RM:SG	DEP[+back]	DEP-μ	MAXV
a. le:t-		*!					
b. lo:t-	*!				*		
c. l$_i$e$_j$l$_i$e:$_j$t-		*!	**				
d. ☞ l$_i$e$_j$l$_i$o:$_j$t-			**		*		

If were to try to develop such an approach, we would want to say that these roots for some reason have two distinct underlying allomorphs: one in /-e:-/ and one in /-o:-/. Vocabulary Insertion would have to govern the distribution: the /-o:-/ allomorph is inserted in the context of PRETERITE, the /-e:-/ allomorph is inserted elsewhere. But even with this move, we would still need to assume that the derivation for the preterite containing /-o:-/ takes as its base a virtual present also formed from the /-o:-/ allomorph. This could have been a way to accommodate the inherited preterite singular, but again, it is not clear why speakers would have felt the need to do so rather than simply remodel, as was done throughout the rest of the system (see Chapter 5). This is therefore not an explanation in and of itself.

Thus we see that the analytical system developed for the rest of the strong verb paradigms is not directly consistent with what we observe for Class VIId, certainly in Gothic terms and likely in Pre-Proto-Germanic terms as well. Some other mechanism will have to be brought to bear to bring them in line, though I cannot at this time see what such a mechanism would be beyond the suggestions raised above.

4.8   In Defense of Null Morphemes in the Preterite

The core claim of this chapter is that, despite the fact that the Gothic/Pre-Proto-Germanic preterite stems clearly exhibit phonological properties not found in the root, these preterite stems are actually built from underlying representations with phonologically null Vocabulary Items for the preterite morpheme and the singular morpheme, and changes are triggered by constraints requiring phono-

logical contrast between stems (REALIZE MORPHEME). In order to justify adopting this *prima facie* unexpected and complex approach, it ought to be demonstrated that it is not possible to identify reasonable phonologically non-null underlying representations for these morphemes that could result in the output distributions we observe. I take this up in this section.

While an earnest attempt at identifying substantive underlying representations of such a sort can yield an allomorphy-based system that seems fairly plausible on its face (and might not be inappropriate for the less phonologically transparent system of Gothic), such a system cannot fully achieve its goal while sufficiently capturing the relevant generalizations (in Pre-Proto-Germanic terms, at least). Notably, if the system is to economically encode that there is consistently no difference between the preterite plural stem and the preterite singular stem in Classes VI and VII, then it necessarily must posit a null underlying allomorph of the singular morpheme. This is precisely the result that this endeavor was trying to avoid. Furthermore, in order to capture the fact that the preterite plural of Class I–III involves deletion relative to the root/present stem, this approach would have to posit an underlying allomorph that is simply an instruction to delete a vowel. If the goal of taking such a tack is to find underlying allomorphs that contain the phonological substance observable in the output stems, this certainly seems like a step in the wrong direction.

Another problem with this approach is in the way it has to treat the relationship between the two preterite stems. Without access to constraints that can assess similarity between stems, the preterite singular stem must be literally derived from the output preterite plural stem in order to capture the generalization about identity between preterite stems in Classes VI and VII. However, this type of output derivation introduces a clear redundancy: the backing effect in the preterite singular of Class IV–V must directly undo the lengthening which has been undertaken to create the preterite plural. If we have constraints that evaluate the similarity between stems, then we can derive both preterite stems independently from the root, and avoid this type of redundancy.

Given these results, which will be demonstrated below, it is clear that an allomorphy-based system is not superior to one based on null morphemes and constraints on stem contrast. And since the latter inarguably does a better job at directly encoding the relevant phonological generalizations, it should be viewed as the preferred approach to the problem.

4.8.1 Pre-Requisites for Finding Substantive Underlying Representations

I take it as given that the wide variety of preterite stem formation patterns found in Gothic/Pre-Proto-Germanic (as laid out in full in (4)/(16) above), particularly among the plural stems, makes it clear that simple, consistent affixation will not be capable of deriving the system in full.[50] That is to say, it seems clear that, regardless of which stem form (if any) we take to be basic and what content we ascribe to the root, there is no single non-null morpheme (by which I mean one containing specified segmental/featural content or a reduplicative morpheme RED) or consistent combination of non-null morphemes that can simultaneously capture all the relations in the system: namely, *bejt- ↔ bit-*, *geb- ↔ geːb-*, *dab- ↔ doːb-*, *hajt- ↔ hehajt-*, *etc*. In other words, the differences between related stems, while perhaps sharing certain similarities in particular cases, is on the whole too non-uniform to reduce to the addition of consistent specified phonological content to one or the other of the stem forms. As such, any reasonable attempt to identify non-null underlying representations for these cases will have to settle for a system with multiple distinct "morphs" for particular morphemes, selected in specified ways by the different root types present in the lexicon. If it could be shown that a small set of morphs could straightforwardly derive the range of stem patterns, this could be consid-

---

[50] The sort of affixation would have to be infixal, and in some cases replacive, as differences between stems largely applies to the vowels in stem-medial position. This on its own would require non-trivial technology to deploy properly.

ered a more attractive approach than the complicated one based on null underlying representations and constraints on stem contrast developed in this chapter.

By "morphs" I mean the following (based on the usage in Wolf 2008). For a particular morphosyntactic feature or feature bundle — which is what I intend by the term *morpheme* — that is input to the process of Vocabulary Insertion, there may be multiple Vocabulary Items associated with that morphosyntactic feature or feature bundle. The distinct Vocabulary Items, which are selected in particular instances of the Vocabulary Insertion process as determined by their morphosyntactic or morphophonological context, are "morphs". (These are not necessarily allomorphs in the traditional sense, as they may be subject to phonological adjustment in the course of the phonological derivation, such that they surface in a slightly different form in the actual phonological output.)

To provide a somewhat abstract example with partial relevance to the question at hand, one might posit that the morphosyntactic feature PRETERITE has a variety of Vocabulary Items associated with it, and each Vocabulary Item is associated with a particular context. This is shown in (78). An effort to apply this approach (as will be spelled out below) might take the context of Vocabulary Insertion rule (78a) PRETERITE ↔ /+long/ to be *"in the context of a root whose phonological shape is CVC"*. Since the context would have to contain phonological information, this would require us to adopt an approach that allows Vocabulary Insertion rules to have access to phonological information (at least belonging to roots). This would be consistent with a cyclic spellout approach that takes inwardly sensitive allomorphy to be exclusively sensitive to phonological information and not syntactic information.

(78)   a.   PRETERITE ↔ /+long/   /   __Context$_i$
       b.   PRETERITE ↔ /RED/   /   __Context$_j$
       c.   PRETERITE ↔ /a/   /   __Context$_k$
       d.   *etc.*

An immediate worry with this type of approach is that the link between root shape and the preterite morph it selects is, as far as the synchronic system goes, arbitrary. As was clear from the analysis I developed above in Section 4.5, the phonological shape of a root and the phonological property that characterizes a particular stem formation pattern are directly correlated; that is to say, by knowing the phonological properties of the root and the properties of the phonological grammar (i.e. CON), we can *predict* the nature of the phonological "process" that relates a root to its preterite stem(s).

Even if the morphological system has the capability of conditioning its competing Vocabulary Insertion rules on the class of roots exhibiting particular constellations of phonological properties (rather than, say, just separate lists of roots, which happen to be characterized by specific constellations of relevant phonological properties), by ascribing the heavy lifting to the morphological component, there is nothing to guarantee that a morph with some particular phonological substance is selected by the phonologically appropriate class of roots, and not some other phonologically inappropriate class. Since there is in fact a strong set of phonological generalizations relating root shape and stem formation pattern, it seems much more desirable to locate the grammatical determination of the stem formation patterns *in the phonology*, rather than in the morphology, where these generalizations can only emerge by stipulation.

4.8.2   Arguments for Identifying the Present Stem as the Root

Given that all the stem alternations (other than reduplication) involve changes to the vowel, a reasonable way to begin searching for morphs could be to assume that the root itself consists only of

its consonantal segments, and that the tense-stem–forming material resides solely in the vowels.[51] However, even the most cursory examination of the distribution shows that this is not a reasonable tack to pursue. The distribution of phonological shapes among the three different stems under discussion — i.e. those evident in (i) the present, (ii) the preterite plural, and (iii) the preterite singular (though each of these but the preterite singular has a wider variety of categories in which it appears; consult Section 4.2.1 above) — for all the different Classes is provided in (79).[52]

(79) Stem distribution by Class

Class	PRESENT	PRET. PLURAL	PRET. SINGULAR
I–III	[CeRC]	[CR̩C]	[CaRC]
IV–V	[CeC]	[Ce:C]	[CaC]
VI	[CaC]	[Co:C]	[Co:C]
VIIa	[CaRC]	[CeCaRC]	[CeCaRC]
VIIb	[Ce:C]	[CeCe:C]	[CeCe:C]
VIIc	[Co:C]	[CeCo:C]	[CeCo:C]

Looking purely at the set of present stems, we can recognize that they attest the fully crossed set of possible stem vowels ([e,a,e:,o:])[53] and possible root shapes (/C...C/ and /C...RC/),[54] subject to the ban on V:RC strings (see Section 4.5.3.2), which rules out the shapes C{e:,o:}RC, and the dispreference for trimoraic vowels (see Sections 4.4.3 and 4.5.7.2), which rules out the shapes C{e::,o::}(C). This can only be interpreted as meaning that the stem vowel of the present stem is *unpredictable* for any given root shape, and thus must be part of the underlying representation of the root, rather than being contributed by some other morpheme. Similar though less spectacular arguments to this effect can be made from the other stem categories as well.

4.8.3   The Allomorphy Analysis

Granting that the stem vowel of the present must be interpreted as belonging to the root, and thus that the present stem reflects the faithful realization of the root itself, we now know that any and all morphemic contributions to the stem formation process would have to be ascribed to the preterite.

To simplify matters, I will now invoke the generalization about the distribution of preterite stems made in Section 4.2.1: the stem reflected in the preterite plural is the default preterite stem, as it appears in all preterite categories other than the preterite singular indicative, where we find a distinct stem. This means that we must identify two morphemes that are relevant to preterite stem formation: (i) PRETERITE, and (ii) SINGULAR (in the context of PRETERITE and INDICATIVE, or, perhaps more appropriately, the absence of SUBJUNCTIVE). The preterite plural stem (i.e. default preterite stem) is thus derived simply by adding the PRETERITE morpheme to the root.

---

[51] Note that this situation would precisely mirror what is commonly assumed for Semitic nonconcatenative verbal morphology; see McCarthy (1979), among many others.

[52] I do not include as a separate type Class VIId, which shows [Ce:C] in the present (à la Class VIIb) but [CeCo:C] in the preterite (à la Class VIIc). See Section 4.7.2 for discussion.

[53] High vowels are not generally permitted as root/stem vowels in (Pre-)Proto-Germanic, except perhaps in rare and difficult cases; consult Ringe (2006).

[54] Initial position of the root in most if not all Classes can actually range from zero consonants (i.e. /...(R)C/) to two consonants (i.e. /CC...(R)C/. This variation is not relevant for this discussion, as only the rhyme segments factor into a stem's behavior. See Appendix I for the list of strong roots in Gothic.

The exact nature of the special preterite singular stem is more complicated to characterize: *Is it the addition of a portmanteau morpheme* PRETERITE.SINGULAR*? Is it the addition of the* SINGULAR *morpheme in the proper morphosyntactic context? Is it the addition of* SINGULAR *alongside* PRETERITE*?* These possibilities effectively reduce to two types of derivations: one in which both preterite stems are derived directly from the root (80);[55] or one in which the default preterite stem is derived from the root, while the special preterite singular stem is formed via output derivation from the default preterite stem (81).

The tables in (80) and (81) factor out the morphemic contributions of the two preterite stems in the two respective types of derivation. Note that I employ the feature [±long] to capture vowel length alternations, rather than the three-way mora-based characterization employed above. This is a matter of simplicity, as it will be easier to reason about the various aspects of this proposal in terms of features than moras. Employing the mora-based approach would only negatively impact the allomorphy analysis, since it would require an morph that specifies deletion of a mora (akin to "DELETE V" below) rather than just the change of a feature value. So the feature-based representation effectively improves the evaluation of the approach I am arguing against.

It is immediately evident that the purely root-based derivation of both stems shown in (80) has a duplication problem. Both preterite stem types for Class VI and Class VIIa–c would require the same underlying representation for the morphemic contributions of the two stem types. This misses the generalization that these classes do not possess a special preterite singular stem distinct from the default preterite stem.

(80)   Independent root-based derivation of plural and singular stems

Class	ROOT	PRET(PL)		SG(PRET)	
I–III	/CeRC/	DELETE V	[CR̩C]	/+back/	[CaRC]
IV–V	/CeC/	/+long/	[Ce:C]	/+back/	[CaC]
VI	/CaC/	/+long/	[Co:C]	/+long/	[Co:C]
VIIa	/CaRC/	/RED/	[CeCaRC]	/RED/	[CeCaRC]
VIIb	/Ce:C/	/RED/	[CeCe:C]	/RED/	[CeCe:C]
VIIc	/Co:C/	/RED/	[CeCo:C]	/RED/	[CeCo:C]

This problem is avoided if we employ an output derivation for the preterite singular stem. Once we assume that the preterite singular is derived from the output default preterite stem rather than directly from the root, we can identify the morphemic contribution of the preterite singular stem in Class VI and Class VIIa–c as null (i.e. /Ø/). This avoids the duplication problem, as is evident from (81) below.

The table in (81) demonstrates that deriving the special preterite singular stem from the default preterite stem has the potential to reduce the number of morphs for the two categories down to three apiece. The underlying phonological contribution of the default preterite stem (i.e. the morph for PRETERITE) is /+long/ for roots of the shape /CeC/ (Class IV–V) and /CaC/ (Class VI), united as /CV̆C/. It is /RED/ for roots of the shape /CaRC/ (Class VIIa), as well as /Ce:C/ (Class VIIb) and /Co:C/ (Class VIIc), the latter two united as /CV:C/. Note though that the two broader shapes /CaRC/ and /CV:C/ cannot be further united under the banner of roots with three timing slots (two moras?) in the rhyme, because Class I–III /CeRC/ would also fit the bill, but shows

---

[55] Root-based derivation will here be completely equivalent to output-based derivation from the surface present stem, as these are phonologically identical. Which conception of the derivation we employ ultimately has no bearing on the outcome of this discussion.

(81)   Default preterite stem derived from root; preterite singular stem derived from default

Class	ROOT	PRET(PL) →		SG(PRET)	
I–III	/CeRC/	DELETE V	[CR̥C]	/a/	[CaRC]
IV–V	/CeC/	/+long/	[Ceː C]	/+back,-long/	[CaC]
VI	/CaC/	/+long/	[Coː C]	/Ø/	[Coː C]
VIIa	/CaRC/	/RED/	[C̲e̲CaRC]	/Ø/	[C̲e̲CaRC]
VIIb	/Ceː C/	/RED/	[C̲e̲Ceː C]	/Ø/	[C̲e̲Ceː C]
VIIc	/Coː C/	/RED/	[C̲e̲Coː C]	/Ø/	[C̲e̲Coː C]

a different morph. Lastly, the Vocabulary Item for PRETERITE for /CeRC/ roots (Class I–III) would have to somehow contain an instruction that leads to deletion of the root vowel (represented in (81) as "DELETE V"). If our original goal was to find morphs that contained phonological substance, this is certainly not a positive result in that direction.

We also observe three morphs for the preterite singular stem. The most common is /Ø/, which holds of roots of the shape /CaC/ (Class VI), /CaRC/ (Class VIIa), /Ceː C/ (Class VIIb), and /Coː C/ (Class VIIc). These do all share the property of having a root vowel that is not short /e/. This could suggest that /Ø/ is an elsewhere morph, while the other two are contextually restricted. These other two morphs are /a/ for /CeRC/ roots (Class I–III) and /+back,-long/ for /CeC/ roots (Class IV–V).

Note that the /a/ morph for Class I–III must be supplemented with specific instructions regarding where it is to be linearized in the string it attaches to: since it is a full segment, and not just floating vowel features, it has multiple options for linearization site, namely either before or after the medial sonorant (assuming that reasonable independent conditions prevent it from surfacing exterior to the root). This is different than the other morphs encountered thus far (except, in certain respects, for the /RED/ morph), which were all floating vowel features (or transformational instructions necessarily applying to vowels), and thus could reasonably be assumed to have no other docking possibilities than the root vowel.

It ought also be noticed that the preterite singular morphs for Class I–III and Class IV–V are substantially similar: segmental /a/ in the first case, featural /+back,-long/ in the second. Segmental /a/ is just /+back,-long(,-high)/ attached to an independent root node. We might then be led to consider whether the two morphs can be reduced to one.

One option in this direction would be to extend the /+back,-long/ morph to Class I–III. This might be appealing, since it would allow us to restrict the inventory of morphs in this domain to floating features (or floating feature combinations), but it ends up being wholly insufficient. In Class II, the vocalized sonorant of the default preterite stem — which, remember, is the base for the preterite singular stem, under our current working assumptions — surfaces as [u]. Without a specification for an additional vowel slot, we might think this should undergo no change in the preterite singular (i.e. **[CuC]), since the vowel already bears the features [+back] and [-long]. To get around this, one might try appealing to a REALIZE MORPHEME (Kurisu 2001) effect — i.e., that the presence of an additional morpheme requires some overt phonological distinction relative to the base — to trigger the addition of a new V slot to host it, yielding desired [CawC]. Note that this will make bad predictions for all of the preterite singulars with the /Ø/ morph, since these show no overt phonological distinction relative to their base. In any event, this was the sort of analytical move that the allomorphy approach was trying to avoid; if appeal to REALIZE MORPHEME and null morphs is required, then why not go all the way in that direction, as in the analysis adopted in this chapter. A more substantive objection to the REALIZE MORPHEME lifeline in the current case

is that it should not apply in Class I even though we would need it to. In Class I, the vocalized sono-
rant is [i]. The [+back] feature should be able to dock here, and have the preterite singular surface
as **[CuC], rather than the desired [CajC]. Reduction of the singular morphs in this direction is
thus not feasible.

The other option for collapsing together the preterite singular morphs would be to extend /a/
to Class IV–V. This would require us to assume that an underlying /a/ could simply replace a
base [e:]. This raises anew the question of why the /a/ would not also simply replace the [i] and
[u] of Class I & II. If the /a/ were generally replacive in nature, we might expect it to turn both
Class I and Class II into **[CaC] rather than [CajC] and [CawC], respectively. The divergent special
conditions on the linearization of /a/ in Class I–III and Class IV–V thus argue against collapsing
the morphs into one.

Therefore, there are minimally three distinct morphs for both categories. This presents nine
possible combinations for plural~singular allomorphy patterns in the preterite. These are repre-
sented by the nine cells in (82). There are minimally five (maximally six) root types: among those
listed in (81), only Class VIIb and Class VIIc can be reconciled both with respect to the allomorphy
they take and their phonological structure. This leads us to the distribution in (82) below, where
four out of the nine cells are instantiated. Class VIIa and ClassVIIb–c (which, recall, could not
be collapsed in a way that did not also include Class I–III) select the same morphs (/RED/ for
the default preterite, /Ø/ for the preterite singular). Class VI finds itself partially overlapping with
Class IV–V (they have the same morph for the default preterite: /+long/) and partially overlap-
ping with Class VIIa and ClassVIIb–c (they have the same morph for the preterite singular: /Ø/).
Class I–III shares no morphs with any of the other types.

(82)   Distribution of morphs by Class

| | | SG(PRET) | | |
		/a/	/+back,-long/	/Ø/
PRET(PL)	DELETE V	/CeRC/ (I–III)		
	/+long/		/CeC/ (IV–V)	/CaC/ (VI)
	/RED/			/CaRC/ (VIIa), /CV:C/ (VIIb–c)

This distribution can be formalized within the allomorphy-based approach with the Vocabulary
Insertion rules provided in (83) below. It requires three rules for the PRETERITE morpheme (or four,
if we do not collapse the /CeC/ and /CaC/ types), and three rules for the SINGULAR morpheme.
(In both cases, the first two rules need not be ordered.)

(83)   a.   i.   PRETERITE ↔ DELETE V   / __CeRC
            ii.   PRETERITE ↔ /+long/   / __CV̌C (or, separately, CeC and CaC)
            iii.   PRETERITE ↔ /RED/   / __elsewhere

       b.   i.   SINGULAR ↔ /a/   / __CeRC
            ii.   SINGULAR ↔ /+back,-long/   / __CeC
            iii.   SINGULAR ↔ /Ø/   / __elsewhere

173

These Vocabulary Insertion rules do, to some extent, capture the phonological generalizations discussed in Section 4.2.3. For one, the fact that both types of roots of the shape CV̌C show lengthening in the preterite plural can be encoded in a single rule. However, as mentioned earlier, the fact that /+long/ is associated to all and only the root shapes in which lengthening would result in licit syllable structure (i.e. not create a superheavy syllable, *vel sim.*) is not directly encoded. This system does bring to the fore the generalization that /RED/ and /Ø/ are the respective elsewhere morphs. In the stem contrast analysis, this generalization is encoded in a slightly more indirect though roughly equivalent way, to be understood as the result of constraint ranking.

4.8.4 Evaluating the Allomorphy Approach

It thus seems that an approach based on allomorphy is not completely unworkable. Nevertheless, there are at least several strikes against it in relation to the null morpheme plus stem contrast analysis proposed in this chapter.

First and foremost, the allomorphy analysis cannot get around the need to posit the existence of null morphs. It has to do so in order to capture the lack of differentiation between preterite plural and singular in Class VI and ClassVIIa–c. If the objection to the analysis proposed in Section 4.5 was that it relied on null morphemes, that same critique has to be leveled at the allomorphy analysis. But beyond that, the allomorphy analysis actually requires "DELETE V", an morph that not only has no substance, but must specify a process. If the goal of setting up an allomorphy analysis was to bring greater substance into the underlying representations, it seems as though it has succeeded in doing the opposite. More to the point, it seems like a desideratum for phonological processes to be located in the phonology. This is achieved in the stem contrast analysis, but not in the allomorphy analysis.

Second, it must distribute a generalization that seems like it ought to be unitary. The allomorphy analysis requires two separate Vocabulary Insertion rules for the SINGULAR morpheme whose result is the same surface [a] vowel. Furthermore, these surface [a] vowels are derived from the same underlying root structure, namely a root vowel /e/. These [a] rules apply only to root shapes with /e/, and not to root shapes with any other vowel. The reason that this generalization has to be distributed across distinct rules is because the allomorphy analysis requires a strictly output-based derivation of the singular, because it has no access to constraints that evaluate the similarity between related stems.

Here we can notice another shortcoming of the allomorphy analysis. In Class IV–V, when the default preterite stem (i.e. the plural) is derived from the root (or the output present stem), it adds the feature [+long]; yet when the preterite singular is built from this, it must remove the [+long] feature (by adding [-long]) and simultaneously add the [+back] feature. This is illustrated in (84a). On the other hand, the stem contrast analysis permits both preterite stems to be derived directly from the root, with the effects of output derivation recreated as stem contrast effects. If both stems can be derived directly from the root, we do not need to be undoing the effects of the first derivation; the plural is formed by adding [+long] to the root, the singular is formed by adding [+back] to the root. This is illustrated in (84b).

(84)   a.   Allomorphy analysis with output-based distribution:
        CeC → PRET.PL /+long/ [Ce:C] → PRET.SG /+back,**-long**/ [CaC]

    b.   Stem contrast analysis with root-based derivation
        i.   CeC → PRET.PL /+long/ [Ce:C]
        ii.   CeC → PRET.SG /+back/ [CaC]

There thus seem to be indications that certain aspects of the system reflect a root-based derivation while others suggest an output-based derivation. The stem contrast analysis allows for both of these approaches to be integrated in a consistent and principled way, namely, all phonological evaluations take the root as their input, but certain morphologically circumscribed evaluations can also consider information from particular morphological bases.

In sum, the allomorphy analysis has a number of shortcomings, both in its own right and relative to the stem contrast analysis. For this reason, the stem contrast analysis should be preferred.

4.9 Conclusion

The analysis developed in this chapter has shown that the pattern of morphophonological markings found among non-derived ("strong") verbs in Gothic (or, rather, its historical precursor Pre-Proto-Germanic) can be generated through the interaction of markedness and faithfulness constraints with REALIZE MORPHEME constraints that require phonological contrast between morphologically related stems. This system of verbal morphology thus provides compelling evidence that specific morphosyntactic features are visible to the phonology, and that the extent to which such morphosyntactically distinct forms remain formally distinct can be modeled in terms of language-specific constraint grammars. Furthermore, the morphophonological system that controls the patterns of stem formation is fully compatible with the grammar that controls the shape of reduplication, in the one corner of the language where it is retained.

4.10 Appendix I: Strong Verbs in Gothic

In this appendix, I compile additional details about the collection of strong verbs attested in Gothic, largely drawn from Lambdin (2006). First, I provide the *present ∼ preterite singular ∼ preterite plural* paradigm for each of the seven strong verb classes. In the tables where I present these paradigms, each member of the paradigm will include three forms: (i) the Gothic orthography of the surface form (in italics), (ii) the Gothic phonetic transcription of the stem (in square brackets [...]), and (iii) the Pre-Proto-Germanic (morphophonemic) phonological representation of the stem that I have assumed in the analysis in this chapter (in double slashes //...//). Following Lambdin, I have given separate entries for verbs that show "breaking" (see (15) above), the pan-Germanic lowering process, which, in Gothic terms, is as follows: /i,u/ → [ɛ,ɔ] / _[r,h,hʷ(,V)]. These entries are equivalent to the non-breaking types except for the application of that process.

Second, I reproduce the list of Gothic roots, represented by their infinitives, which Lambdin (2006) assigns to each of those classes (occasionally supplemented and/or modified based on Ringe 2006:§4.3.3.i). Many of the verbs have limited attestation in the preterite, and so some of the category assignments may be partially inferential. Additional strong verb roots are attested in Northwest Germanic, but this will serve to illustrate the range of phonological properties associated with each type. Where possible, I make note of exceptional behaviors and other noteworthy properties exhibited by particular members of the various classes, but this is not wholly systematic.

4.10.1 Strong Class I

(85)   Basic paradigm of Strong Class I (Lambdin 2006:65)

	Infinitive		Preterite Sing. (3rd pers.)		Preterite Pl. (1st pers.)	
a.	*beidan*	[biːð-] //bejd-//	*baiþ*	[bɛːθ-] //bajd-//	*bidum*	[bið-] //bjd-//
b.	*teihan*	[tiːh-] //tejh-//	*taih*	[tɛːh-] //tajh-//	*taihum*	[tɛh-] //tjh-//

(86)   Strong Class I roots (infinitives) (Lambdin 2006:65)

CejC					
beidan	'await'	*beitan*	'bite'	*(bi)leiban*	'remain'
(ga)leiþan	'go'	*leihwan*	'lend'	*(ur)reisan*	'arise'
(ga)teihan	'tell'	*þeihan*	'thrive'	*weihan*	'fight'
weipan	'crown'	*(in)weitan*	'greet, worship'		

CCejC					
dreiban	'drive'	*greipan*	'seize'	*hneiwan*	'bow'
skeinan	'shine'	*(dis)skreitan*	'tear'	*(ga)smeitan*	'smear'
sneiþan	'cut, harvest'	*speiwan*	'spit'	*steigan*	'ascend'
sweiban	'cease'	*þreihan*	'press, crowd'		

There is also a present *dig-* 'knead' (attested only in the dative singular of the present participle *digandin*; Ringe 2006:239–240), which shows the expected Class I preterite stems (singular *daig-*, plural *dig-*) but has [-i-] in the present rather than [-iː-] < *-ej-. This is problematic, as the preterite plural stem *dig-* is identical to the present. I would predict a present like *dig-* to pattern with Class V (yielding preterite plural **deːg-*), not Class I. Other than admitting this as an exception, the only

way to explain this away would be to say that, as the attestation might suggest, the root really did form a basic present, and speakers calculated the preterite relative to a supposed normal present base *diːg-/*dejg-*, presuming the participial form to be anomalous or altogether dissociated.

4.10.2 Strong Class II

(87)   Basic paradigm of Strong Class II (Lambdin 2006:72)

	Infinitive	Preterite Sing. (3rd pers.)	Preterite Pl. (1st pers.)
a.	*driusan* [driu̯s-] //drews-//	*draus* [drɔːs-] //draws-//	*drusum* [drus-] //dru̯s-//
b.	*tiuhan* [tiu̯h-] //tewh-//	*tauh* [tɔːh-] //tawh-//	*tauhum* [tɔh-] //tu̯h-//

(88)   Strong Class II roots (infinitives) (Lambdin 2006:72)

CewC					
(ana)biudan	'bid, offer'	*biugan*	'bend'	*giutan*	'pour'
hiufan	'mourn'	*(ga)kiusan*	'choose'	*liudan*	'grow, spring up'
liugan	'tell a lie'	*(fra)liusan*	'lose'	*niutan*	'enjoy'
siukan	'be sick'	*tiuhan*	'lead'		

CCewC					
driugan	'be a soldier'	*driusan*	'fall'	*(dis)hniupan*	'break apart'
kriusan	'gnash'	*sliupan*	'slip'	*(af)skiuban*	'push aside'
þliuhan	'flee'	*(us)þriutan*	'trouble'		

There is also *(ga)lukan* 'close', with unexpected present [-uː-] (Ringe 2006:241; cf. Lambdin 2006:72) not **-iu-*. This presents the same problem as *dig-*, but with no possible explanation via lack of a real present.

4.10.3 Strong Class III

(89)   Basic paradigm of Strong Class III (Lambdin 2006:87)

	Infinitive	Preterite Sing. (3rd pers.)	Preterite Pl. (1st pers.)
a.	*bindan* [bind-] //bend-//	*band* [band-] //band-//	*bundum* [bund-] //bn̩d-//
b.	*siggwan* [siŋɡʷ] //seŋɡʷ//	*saggw* [saŋɡʷ] //saŋɡʷ//	*suggwum* [suŋɡʷ] //sn̩ɡʷ//
c.	*filhan* [filh-] //felh-//	*falh* [falh-] //falh-//	*fulhum* [fulh-] //fl̩h-//
d.	*wairþan* [wɛrθ-] //werθ-//	*warþ* [warθ-] //warθ-//	*waurþum* [wɔrθ-] //wr̩θ-//

(90)   Strong Class III roots (infinitives) (Lambdin 2006:87)

C(C)erC					
bairgan	'hide'	*(uf)gairdan*	'gird up'	*hwairban*	'walk'
(af)swairban	'wipe off'	*(ga)þairsan*	'wither'	*wairpan*	'throw'
wairþan	'become'				

C(C)elC					
filhan	'conceal'	*(us)gildan*	'pay back'	*hilpan*	'help'
swiltan	'be dying'	*wilwan*	'rob'		

C(C)eNC (*g* = [ŋ] / _{*k,q,g,gw*})					
bindan	'bind'	*drig(g)kan*	'drink'	*finþan*	'learn of'
(fra)hinþan	'capture'	*siggwan*	'sing'	*sigqan*	'sink'
(fra)slindan	'devour'	*stigqan*	'thrust'	*(ana)trimpan*	'tread'
(at)þinsan	'attach'	*windan*	'wind'		

C(C)enn					
brinnan	'burn'	*(du)ginnan*	'begin'	*(af)linnan*	'depart'
rinnan	'run'	*spinnan*	'spin'	*winnan*	'suffer pain'

CReTT					
bliggwan	'beat'	*þriskan*	'thresh'	*(ga)wrisqan*	'bear fruit'
([-ggʷ-] < *-ww-*)					

While preterite plurals of *þriskan* and *(ga)wrisqan* are not attested in Gothic, Lambdin (2006) ascribes them to Class III "on the basis of forms attested elsewhere in Germanic" (p. 87). I take this to mean that their preterite plurals are expected to be *þruskan* and *(ga)wrusqan*, with a [u] that cannot be derived from (normal) sonorant vocalization. They could, however, be derived from sonorant vocalization if the onset /r/ behaved differently when forced to become syllabic, i.e. [ru] not [ur] (cf. Ringe 2006:226). (This may bear some logical connection to Steriade's 1988 "syllable transfer" effects in Sanskrit.) Alternatively, this could be thought about in terms of epenthesis rather than sonorant vocalization: an epenthetic [u] is inserted to break up the consonant cluster.

4.10.4 Strong Class IV

(91)   Basic paradigm of Strong Class IV (Lambdin 2006:51)

	Infinitive	Preterite Sing. (3rd pers.)	Preterite Pl. (1st pers.)
a.	*niman* [nim-] //nem-//	*nam* [nam-] //nam-//	*nemum* [neːm-] //neːm-//
b.	*bairan* [bɛr-] //ber-//	*bar* [bar-] //bar-//	*berum* [beːr-] //beːr-//

(92)   Strong Class IV roots (infinitives) (Lambdin 2006:51)

CeR					
bairan	'bear'	*niman*	'take'	*qiman*	'come'
(ga)tairan	'tear'	*(ga)timan*	'suit'		

CCeR			CReT		
stilan	'steal'		*trudan*	'tread'	*brikan* 'break'

There is also a present *wulan*, with unexpected [-u-] (Ringe 2006:244 presumes this to be a zero-grade, i.e. *w̥l-), but it does not attest a preterite. Note that *trudan* shows the same vowel, and could have the same explanation as a zero-grade from *tr̥d- (Ringe 2006:246), with abnormal, though maybe not unexpected (see the discussion for Class III above), placement of the epenthetic vowel to the right of the sonorant. Assuming underlying root vowel /i/ (< *e) for the root of *trudan* is problematic, as this should be allowed to surface in the preterite plural if the present is, for whatever reason, [-u-] synchronically. If the root underlyingly has /u/, then arriving at preterite stems in [-a-] and [-eː-] is problematic, as we might expect faithfulness to select different mappings.

4.10.5   Strong Class V

(93)   Basic paradigm of Strong Class V (Lambdin 2006:51–52)

	Infinitive		Preterite Sing. (3rd pers.)		Preterite Pl. (1st pers.)	
a. *giban*	[giv-]	//geb-//	*gaf*	[gaf-] //gab-//	*gebum*	[geːv-] //geːb-//
b. *qiþan*	[kʷiθ-]	//kʷeθ-//	*qaþ*	[kʷaθ-] //kʷaθ-//	*qeþum*	[kʷeːθ-] //kʷeːθ-//
c. *saihwan*	[sɛhʷ-]	//sehʷ-//	*sahw*	[sahʷ-] //sahʷ-//	*sehwum*	[seːhʷ-] //seːhʷ-//

(94)   Strong Class V roots (infinitives) (Lambdin 2006:51)

CeT					
bidjan	'pray'	*diwan*	'die'	*fitan*	'be in labor'
giban	'give'	*(bi)gitan*	'find'	*ligan*	'lie'
lisan	'gather'	*(ga)nisan*	'be saved'	*niþan*	'help'
qiþan	'say'	*rikan*	'heap up'	*saihwan*	'see'
sitan	'sit'	*(ga)widan*	'unite'	*(in)widan*	'reject'
(ga)wigan	'shake'	*wisan*	'be'		

CReT					
fraihnan	'ask'	*hlifan*	'steal'	*wrikan*	'persecute'

eT			CReR		
itan	'eat'		*sniwan*	'hasten'	

The *-j-* of *bidjan* and the *-n-* of *fraihnan* are not retained outside of the present/infinitive. The calculation of their preterite stems must exclude the stem-forming element (i.e. their present bases are [bid-] and [freh-], respectively). Lambdin includes *sniwan* in his list of Class V roots, though Ringe (2006:227–228ff.) classifies it as a Class IV (as would be more appropriate for a sonorant-final root), with generally unusual properties. I follow Ringe here, though this makes no difference for the analysis.

itan has an anomalous preterite singular *et* [eːt], which Ringe (2006:244–245) reconstructs with a trimoraic vowel *êt* [eːːt]. This would have to be the result of contraction with the earlier reduplicant vowel across the lost root-initial laryngeal: Proto-Germanic *êt < (post-)PIE *h₁e-h₁ed-. This, however, requires an unexpected e-grade at the (post-)PIE stage, as the expected o-grade PIE *h₁e-h₁od- should be expected to retain its o-quality as Proto-Germanic **ôt > Gothic **ōt. In Gothic, the plural has the expected stem *et* [eːt], which is identical to the unexpected singular. Historically, this ultimately derives from PIE zero-grade *h₁e-h₁d-. The long vowel can either come simply from lengthening of the reduplicant vowel by the root-initial laryngeal; or it can derive from *PCR-driven deletion with compensatory lengthening, as in most other members of Class V.

4.10.6 Strong Class VI

(95)   Basic paradigm of Strong Class VI (Lambdin 2006:93)

	Infinitive			Preterite Sing. (3rd pers.)			Preterite Pl. (1st pers.)		
a. *faran*	[far-]	//far-//	*for*	[foːr-]	//foːr-//	*forum*	[foːr-]	//foːr-//
b. *fraþjan*	[fraθ-j-]	//fraθ-j-//	*froþ*	[froːθ-]	//froːθ-//	*froþum*	[froːθ-]	//froːθ-//
c. *standan*	[sta-n-d-]	//sta-n-d-//	*stoþ*	[stoːθ-]	//stoːθ-//	*stoþum*	[stoːθ-]	//stoːθ-//

(96)   Strong Class VI roots (infinitives) (Lambdin 2006:93)

CaC					
(ga)daban	'be fitting'	*faran*	'go, travel'	*hafjan*	'raise'
malan	'grind'	*(ga)raþjan*	'count'	*sakan*	'dispute'
wakan	'wake'				

CCaC					
(ga)draban	'hollow out'	*(ga)dragan*	'heap up'	*fraþjan*	'understand'
graban	'dig'	*hlahjan*	'laugh'	*(af)hlaþan*	'load'
skaþjan	'injure'	*slahan*	'strike'	*skaban*	'shave'
(ga)skaþjan	'create'	*sta**n**dan*	'stand'	*swaran*	'swear'
þwahan	'wash'				

aC				CaTT	
alan	'grow'	*(us)asan*	'die, expire'	*wahs**j**an*	'grow'

In all cases that display them (as in previous classes, though with greater frequency here), the stem-final -*j*- (bolded in the above table) is not retained outside of the present stem. Likewise, the stem-internal -*n*- of *standan* is also not retained outside of the present stem. These must effectively treat their present base as the string excluding the -*j*- or -*n*-.

Ringe (2006:247) reconstructs the preterite singular for *alan* (and other vowel-initial roots of Class VI attested elsewhere in Germanic), i.e. *ol* [oːl], with a trimoraic vowel in Proto-Germanic, *ôl* [oːːl], resulting from contraction with the reduplicated vowel (as in *itan* above).

4.10.7 Strong Class VII

(97)   Basic paradigm of Strong Class VII (Lambdin 2006:115)

	Infinitive	Preterite Sing. (3rd pers.)	Preterite Pl. (1st pers.)
a.	*falþan* [falθ-] //falθ-//	*faifalþ* [f̪ɛfalθ-] //f̪ɛfalθ-//	*faifalþum* [f̪ɛfalθ-] //f̪ɛfalθ-//
b.	*slepan* [sleːp-] //sleːp-//	*saislep* [s̪ɛsleːp-] //s̪ɛsleːp-//	*saislepum* [s̪ɛsleːp-] //s̪ɛsleːp-//
c.	*flokan* [floːk-] //floːk-//	*faiflok* [f̪ɛfloːk-] //f̪ɛfloːk-//	*faiflokum* [f̪ɛfloːk-] //f̪ɛfloːk-//
d.	*tekan* [teːk-] //teːk-//	*taitok* [t̪ɛtoːk-] //t̪ɛtoːk-//	*taitokum* [t̪ɛtoːk-] //t̪ɛtoːk-//

(98)   Strong Class VIIa (C$_0$aRC) roots (infinitives and preterite singulars) (Lambdin 2006:93)

falþan	*faifalþ*	'fold'
haldan	*haihald*	'tend (cattle); hold'
(ga)staldan	*(ga)staistald*	'possess, acquire'
fahan ([fã:han] < *fanhan*)	*faifah* ([fɛfã:h] < *fefanh*)	'seize'
hahan ([hã:han] < *hanhan*)	*haihah* ([hɛhã:h] < *hehanh*)	'hang'
(af)aikan ([ɛːkan] < *ajkan*)	*(af)aiaik* ([ɛ.ɛːk] < *e.ajk*)	'deny'
fraisan	*faifraisan*	'tempt'
haitan	*haihait*	'call, name'
laikan	*lailaik*	'play, leap'
maitan	*maimait*	'cut'
skaidan	*skaiskaiþ*	'divide'
aukan ([ɔːkan] < *awkan*)	*aiauk* ([ɛ.ɔːk] < *e.awk*)	'increase'

(99)   Strong Class VIIb–d roots (infinitives and preterite singulars) (Lambdin 2006:93)

VIIb – C$_0$eː(C) roots

slepan	*saislep*	'sleep'

VIIc – C$_0$oː(C) roots

*flokan**	*faiflok*	'bewail'
*lauan** ([lɔː.an] ← /loː-an/)	*lailo*	'revile'
hwopan	*hwaihwop* ([hʷɛhʷoːp])	'boast'

VIId – C$_0$eː(C) roots with -*oː*- preterite stem vowel

gretan	*gaigrot*	'weep'
letan	*lailot*	'let, allow, leave'
(ga)redan	*(ga)rairoþ*	'reflect on'
tekan	*taitok*	'touch'
saian ([sɛː.an] ← /seː-an/)	*saiso*	'sow'
waian ([wɛː.an] ← /weː-an/)	*waiwo*	'blow'

4.11 Appendix II: Gothic Verbal Inflectional Paradigms

This appendix provides the full finite paradigms (present and preterite, indicative and subjunctive), in Gothic terms, for the four classes of weak verbs, and for the strong verbs (represented by Strong Class IV *niman* 'take'). I have indicated morpheme boundaries. Many of the inflected forms of weak verbs would have vowels in hiatus across the boundaries if they surfaced in full. This is never tolerated in this domain, and usually repaired by elision of one of the vowels. Morpheme boundaries are thus approximate.

4.11.1 Weak Verbs

(100) Weak Class I: *nasjan* 'save'

Indicative

		Present	Preterite
Sg	1	*nas-j-a*	*nas-i-d-a*
	2	*nas-j-is*	*nas-i-d-es*
	3	*nas-j-iþ*	*nas-i-d-a*
Du	1	*nas-j-os*	*nas-i-ded-u*
	2	*nas-j-ats*	*nas-i-ded-uts*
Pl	1	*nas-j-am*	*nas-i-ded-um*
	2	*nas-j-iþ*	*nas-i-ded-uþ*
	3	*nas-j-and*	*nas-i-ded-un*

Subjunctive

		Present	Preterite
Sg	1	*nas-j-au*	*nas-i-ded-jau*
	2	*nas-j-ais*	*nas-i-ded-eis*
	3	*nas-j-ai*	*nas-i-ded-i*
Du	1	*nas-j-aiwa*	*nas-i-ded-eiwa*
	2	*nas-j-aits*	*nas-i-ded-eits*
Pl	1	*nas-j-aima*	*nas-i-ded-eima*
	2	*nas-j-aiþ*	*nas-i-ded-eiþ*
	3	*nas-j-aina*	*nas-i-ded-eina*

(101) Weak Class II: *salbon* 'anoint'

Indicative

		Present	Preterite
Sg	1	*salb-o*	*salb-o-d-a*
	2	*salb-o-s*	*salb-o-d-es*
	3	*salb-o-þ*	*salb-o-d-a*
Du	1	*salb-o-s*	*salb-o-ded-u*
	2	*salb-o-ts*	*salb-o-ded-uts*
Pl	1	*salb-o-m*	*salb-o-ded-um*
	2	*salb-o-þ*	*salb-o-ded-uþ*
	3	*salb-o-nd*	*salb-o-ded-un*

Subjunctive

		Present	Preterite
Sg	1	*salb-o*	*salb-o-ded-jau*
	2	*salb-o-s*	*salb-o-ded-eis*
	3	*salb-o*	*salb-o-ded-i*
Du	1	*salb-o-wa*	*salb-o-ded-eiwa*
	2	*salb-o-ts*	*salb-o-ded-eits*
Pl	1	*salb-o-ma*	*salb-o-ded-eima*
	2	*salb-o-þ*	*salb-o-ded-eiþ*
	3	*salb-o-na*	*salb-o-ded-eina*

(102) Weak Class III: *haban* 'have, hold'

Indicative

		Present	Preterite
Sg	1	*hab-a*	*hab-ai-d-a*
	2	*hab-a-is*	*hab-ai-d-es*
	3	*hab-a-iþ*	*hab-ai-d-a*
Du	1	*hab-os*	*hab-ai-ded-u*
	2	unattested	*hab-ai-ded-uts*
Pl	1	*hab-a-m*	*hab-ai-ded-um*
	2	*hab-a-iþ*	*hab-ai-ded-uþ*
	3	*hab-a-nd*	*hab-ai-ded-un*

Subjunctive

		Present	Preterite
Sg	1	*hab-a-u*	*hab-ai-ded-jau*
	2	*hab-ai-s*	*hab-ai-ded-eis*
	3	*hab-ai*	*hab-ai-ded-i*
Du	1	*hab-ai-wa*	*hab-ai-ded-eiwa*
	2	*hab-ai-ts*	*hab-ai-ded-eits*
Pl	1	*hab-ai-ma*	*hab-ai-ded-eima*
	2	*hab-ai-þ*	*hab-ai-ded-eiþ*
	3	*hab-ai-na*	*hab-ai-ded-eina*

(103) Weak Class IV: *fullnan* 'become full'

Indicative

		Present	Preterite
Sg	1	*full-n-a*	*full-no-d-a*
	2	*full-n-is*	*full-no-d-es*
	3	*full-n-iþ*	*full-no-d-a*
Du	1	*full-n-os*	*full-no-ded-u*
	2	*full-n-ats*	*full-no-ded-uts*
Pl	1	*full-n-am*	*full-no-ded-um*
	2	*full-n-iþ*	*full-no-ded-uþ*
	3	*full-n-and*	*full-no-ded-un*

Subjunctive

		Present	Preterite
Sg	1	*full-n-au*	*full-no-ded-jau*
	2	*full-n-ais*	*full-no-ded-eis*
	3	*full-n-ai*	*full-no-ded-i*
Du	1	*full-n-aiwa*	*full-no-ded-eiwa*
	2	*full-n-aits*	*full-no-ded-eits*
Pl	1	*full-n-aima*	*full-no-ded-eima*
	2	*full-n-aiþ*	*full-no-ded-eiþ*
	3	*full-n-aina*	*full-no-ded-eina*

4.11.2 Strong Verbs

(104) Strong Class IV: *niman* 'take'

Indicative

		Present	Preterite
Sg	1	*nim-a*	*nam*
	2	*nim-is*	*nam-t*
	3	*nim-iþ*	*nam*
Du	1	*nim-os*	*nem-u*
	2	*nim-ats*	*nem-uts*
Pl	1	*nim-am*	*nem-um*
	2	*nim-iþ*	*nem-uþ*
	3	*nim-and*	*nem-un*

Subjunctive

		Present	Preterite
Sg	1	*nim-au*	*nem-jau*
	2	*nim-ais*	*nem-eis*
	3	*nim-ai*	*nem-i*
Du	1	*nim-aiwa*	*nem-eiwa*
	2	*nim-aits*	*nem-eits*
Pl	1	*nim-aima*	*nem-eima*
	2	*nim-aiþ*	*nem-eiþ*
	3	*nim-aina*	*nem-eina*

Chapter 5

Sanskrit and the $C_1\bar{e}C_2$ Pattern

5.1 Introduction

The Sanskrit perfect displays reduplicative behavior very much in line with the other Indo-European languages examined in this dissertation thus far.[1] Sanskrit exhibits the proto-typical Indo-European C_1-copying pattern with stop-sonorant–initial roots (1a), but an alternative pattern — C_2-copying — just in case the root begins in an *s*-obstruent cluster (1b). This basic difference is again driven by the No Poorly-Cued Repetitions (*PCR) constraint. What is different about Sanskrit is the way it treats would-be *PCR-violating base-initial clusters resulting from the application of zero-grade ablaut in the so-called "perfect weak stem". In Sanskrit, rather than showing a reduplicative pattern here, an apparently non-reduplicative root allomorph of the shape $C_1\bar{e}C_2$ surfaces instead, as illustrated in (1d) (see Sandell 2015b, 2017 for extensive recent discussion). In this chapter, I will demonstrate that, while this is a distinct type of solution to the problem, this behavior is likewise driven by *PCR.

(1)   Reduplication patterns for Sanskrit cluster-initial bases

	Root		Perfect Tense	
a.	$\sqrt{prac^h}$-	'ask'	*pa-prāc^h-a*	
b.	$\sqrt{stamb^h}$-	'prop'	*ta-stamb^h-a*	(not **sa-stamb^h-a*)
c.	$\sqrt{par}$-	'fill'	*pa-pr-ur*	
d.	$\sqrt{sap}$-	'serve'	*sēp-ur*	(not **sa-sp-ur*)

This $C_1\bar{e}C_2$ pattern in Sanskrit, which is further exemplified in (2),[2] finds an almost exact match in the preterite plural forms of the Germanic Strong Class V verbs, as illustrated in (3) with data from Gothic. (See Chapter 4 for extensive discussion of the Germanic strong verb system.) While the correspondence in vowel quality is illusory (Germanic *ē* corresponds with Sanskrit *ā*), they match in every other respect: the patterns hold of CVT roots in a category which goes back

---

[1] For description of the reduplication patterns of Sanskrit, see Whitney (1889:esp. 279–290/§782–801), Steriade (1988), Kulikov (2005). For recent analyses, see Steriade (1988), McIntyre (1992:Ch. 4), Kennedy (2011).

[2] This table shows only CaC roots where the second root-consonant is an obstruent (CaT). These roots consistently show the $C_1\bar{e}C_2$ pattern. Several other types of CaC roots also show the $C_1\bar{e}C_2$ pattern. The full distribution and its ramifications will be discussed in detail in Chapter 6.

to a zero-grade perfect in Proto-Indo-European, and the pattern consists of a long vowel between root-C_1 and root-C_2.

(2)   Sanskrit $C_1\bar{e}C_2$ perfects of CaT roots

Root		Perfect (weak stem)
√*sap-*	'serve'	*sēp-úr*
√*dab^h-*	'burn'	*dēb^h-úr*
√*b^{h}aj-*	'divide'	*b^{h}ēj-úr*
√*pac-*	'cook'	*pēc-úr*
√*rab^h-*	'take hold'	*rēb^h-úr*
√*ras-*	'roar'	*rēs-úr*
√*nad-*	'sound'	*nēd-úr*

(3)   Gothic Class V preterite plurals (forms from Lambdin 2006:51)

	Infinitive		Preterite Plural (1PL)	
'sit'	*sitan*	[sit-an]	*setum*	[seːt-um]
'give'	*giban*	[gib-an]	*gebum*	[geːb-um]
'say'	*qiþan*	[kʷiθ-an]	*qeþum*	[kʷeːθ-um]
'heap up'	*rikan*	[rik-an]	*rekum*	[reːk-um]
'be saved'	*nisan*	[nis-an]	*nesum*	[neːs-um]

In this chapter, I will develop an analysis of the $C_1\bar{e}C_2$ pattern, set within the context of Sanskrit's basic reduplicative pattern. I will first present an analysis based on allomorph selection, where the underlying representation leading to the $C_1\bar{e}C_2$ form is selected in the phonological component just in case a licit reduplicative structure that satisfies *PCR and the Input-Reduplicant faithfulness constraint LINEARITY-IR (cf. McCarthy & Prince 1995, 1999) cannot be obtained. Subsequently, I will argue that the $C_1\bar{e}C_2$ pattern originates in a *PCR-driven phonological repair of an overtly reduplicative structure (specifically //C̲1̲e-C_1C_2-// → [$C_1\bar{e}C_2$-]). While the phonological solution is somewhat simpler to analyze in derivational/serial terms, the pattern does permit an analysis in strictly parallel OT (cf. Prince & Smolensky 1993/2004) as long as we include relatively more complex, though not unprecedented, theoretical machinery.

From here, I will explore the link between the Sanskrit $C_1\bar{e}C_2$ pattern and the virtually identical one found in the Germanic Strong Class V preterite plurals, as mentioned above in (3). I will argue that the Germanic $C_1\bar{e}C_2$ pattern arose in the history of Germanic through virtually the same diachronic pathway as it did in the history of Sanskrit. However, differences in the accentual systems of the two languages led to differences in the way these forms got situated within the broader morphophonological system; specifically, the change within Germanic from a mobile accent system to one with fixed initial stress opacified the reduplicative origin of the $C_1\bar{e}C_2$ forms. I will then lay out preliminary work on a learning model that uses this development as a pivot from the Proto-Indo-European perfect to the Germanic strong preterite system laid out in Chapter 4.

The patterns and distributions discussed in this chapter are very much dependent on the application of zero-grade ablaut. Many questions still remain about the synchronic analysis and diachronic development of this phenomenon; see Lundquist & Yates (forthcoming:§3) for a summary of the

current state of the literature on these questions. A full treatment of these problems would require a dissertation unto itself; see the on-going line of research by Kiparsky (2010a, 2016, forthcoming) on formalizing the analysis of accent and ablaut in Sanskrit and Indo-European generally. For this reason, my discussion of ablaut as it relates to the reduplicative phenomena under discussion will have to remain largely abstract and informal. Clarification of these issues will be left to future work.

5.2   Reduplication for Cluster-Initial Roots in Sanskrit

For cluster-initial *roots* — crucially distinct from other cluster-initial bases (see below) — Sanskrit shows a very typical Indo-European division into two types of reduplicative patterns based on the composition of the initial cluster. Roots beginning in stop-sonorant (TRVX–) show the standard Indo-European C_1-copying pattern, as illustrated in (4) below. On the other hand, roots beginning in sibilant-obstruent (STVX–),[3] show an alternative pattern, namely C_2-copying. This is illustrated in (5) below.[4] Note that the vowel of the reduplicant is always identical in quality to a segment in the underlying root: [i] or [u] if it contains an underlying (vocalizable) /i,y/ or /u,v/, [a] otherwise (see, e.g., Steriade 1988).

(4) C_1-copying perfects to TRVX– roots (forms from Whitney 1885 [1988])

Root		Perfect Tense
√$b^h raj$-	'shine'	*<u>ba</u>-$b^h r\bar{a}j$-a* not **<u>ra</u>-$b^h r\bar{a}j$-a*
√$prac^h$-	'ask'	*<u>pa</u>-$pr\bar{a}c^h$-a* not **<u>ra</u>-$pr\bar{a}c^h$-a*
√dru-	'run'	*<u>du</u>-druv-ē* not **<u>ru</u>-druv-ē*
√tviṣ-	'be stirred up'	*<u>ti</u>-tviṣ-ē* not **<u>vi</u>-tviṣ-ē*

(5) C_2-copying perfects to STVX– roots (forms from Whitney 1885 [1988])

Root		Perfect Tense
√sparç-	'touch'	*<u>pa</u>-spr̥ç-ē* not **<u>sa</u>-spr̥ç-ē*
√$st^h\bar{a}$-	'stand'	*<u>ta</u>-$st^h\bar{a}$-u* not **<u>sa</u>-$st^h\bar{a}$-u*
√$stamb^h$-	'prop'	*<u>ta</u>-$stamb^h$-a* not **<u>sa</u>-$stamb^h$-a*

---

[3]   Sanskrit has a larger inventory of sibilants than most of the other languages discussed in the dissertation. It has three sibilants: dental/alveolar [s], (alveo)palatal/post-alveolar [ç], and retroflex [ṣ]. In addition to a five-way place distinction for stops (labial, dental/alveolar, palatal, retroflex, velar), where the palatal might be an affricate [tʃ] (which I represent as < c >; consult footnote 4 on transcription conventions). The sibilant place distinction is neutralized (at least on the surface) before coronal obstruents, always matching the following obstruent in place.

[4]   In representing Sanskrit data in this chapter, I use a mix of traditional orthographic conventions and IPA transcription, as follows. For the palatal [-continuant] obstruents — which could either be identified as affricates [tʃ,dʒ] or stops [c,ɟ] — I employ traditional transcription as < c, j >. For the sibilant fricatives, I will use IPA transcription as in footnote 3, where [ç] = < ś > and [ṣ] = < ṣ >. For the glides I use traditional transcription as < y, v >: the value of < y > is certainly IPA [j], while the value of < v > is somewhat less clear — its phonological behavior is partially glide-like (pointing to [w] or [ʋ]), but in other ways it is more fricative-like (pointing to [v]). In Chapter 6, I argue that it is actually a "narrow approximant" [ʋ] (cf. Padgett 2002), a segment type intermediate between fricatives and glides. This question is not significant for the topics examined in this chapter. The long mid front vowel is traditionally transcribed as < e >, but I will represent it interchangeably as [ē] and [eː].

As in the other languages with this sort of split behavior, we can analyze this division as being motivated by the NO POORLY-CUED REPETITIONS (*PCR) constraint. While the most precise version of this constraint will be laid out in Chapter 6, we can continue using the working definition employed previously, which is repeated in (6) below. This constraint in effect penalizes the C_1-copying pattern for STVX– roots, because doing so would result in a consonant repetition ($S_\alpha VS_\alpha$) preceding an obstruent (the root-second T).

(6)   **NO POORLY-CUED REPETITIONS (*PCR)** [$\approx *C_\alpha VC_\alpha$ / $\_C_{[\text{-sonorant}]}$]
      For each sequence of repeated identical consonants separated by a vowel ($C_\alpha VC_\alpha$), assign a violation * if that sequence immediately precedes an obstruent.

In addition to *PCR, we will need four other constraints to fully analyze this basic pattern. These are essentially the same constraints employed in Chapter 1 and throughout the dissertation. ALIGN-ROOT-L (7a) in effect penalizes copying extra segments into the reduplicant, here ruling out copying the full root-initial cluster.[5] ONSET (7b) penalizes onsetless syllables. ANCHOR-L-BR (7c) is violated when the leftmost segment of the reduplicant does not correspond to the leftmost segment of the base (i.e. the root-initial consonant). Lastly, CONTIGUITY-BR (7d) penalizes discontiguous copying, such as the sort found in the C_1-copying pattern. This constraint is freely violated in Sanskrit. To generate the correct distribution, we require the ranking in (8) below.

(7)   Other basic constraints relevant for reduplication
      a.   **ALIGN-ROOT-L**
           Assign a violation mark * for each segment that intervenes between the left edge of the root and the left edge of the word.[6]
      b.   **ONSET**
           Assign a violation mark * for each onsetless syllable.
      c.   **ANCHOR-L-BR**
           Assign a violation mark * if the segment at the left edge of the reduplicant does not stand in correspondence with the segment at the left edge of the base.
      d.   **CONTIGUITY-BR**
           Assign one violation mark * for each pair of segments that are adjacent in the reduplicant but have non-adjacent correspondents in the base (i.e. no X_1X_3-$X_1X_2X_3$).

(8)   Ranking for Sanskrit cluster-initial reduplication
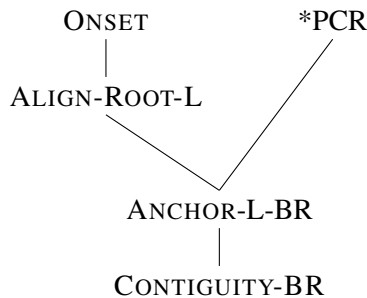
As demonstrated in (9), when *PCR is not in danger of being violated, namely, for TRVX–roots, this ranking selects the C_1-copying candidate (9a). Candidate (9d) copies just the base vowel. This succeeds in minimizing the reduplicant, and thus maximizing satisfaction of ALIGN-ROOT-L, but it leaves the reduplicant vowel without an onset; this fatally violates ONSET. Candidate (9b) unnecessarily copies the entire root-initial cluster, incurring an extra violation of ALIGN-ROOT-L, which is fatal. In candidate (9c), the segment at the left edge of the reduplicant stands in correspondence with a non–base-initial segment (the root-second consonant), and so it is eliminated by ANCHOR-L-BR. (Candidate (9d) also violates ANCHOR-L-BR, but it is independently ruled out by higher-ranked ONSET.) The evaluation thus selects the C_1-copying candidate (9a), despite its CONTIGUITY-BR violation.

(9) C_1-copying in TRVX– roots (no *PCR violation): $\sqrt{prac^h}$- $\to$ *pa-prāc^h-a* 'he has asked'

/RED, prac^h, a/	*PCR	ONSET	ALIGN-RT-L	ANCHOR-L-BR	CONTIG-BR
a. ☞ pa-prāc^h-a			**		*
b. pra-prāc^h-a			***!		
c. ra-prāc^h-a			**	*!	
d. a-prāc^h-a		*!	*	*	

On the other hand, the C_1-copying candidate for STVX– roots (10a) fatally violates *PCR, because the consonant following the repetition is an obstruent. Therefore, these roots will have to display an alternative copying strategy. Given the ranking ALIGN-ROOT-L ≫ ANCHOR-L-BR, this strategy will have to involve mis-anchored copying. High-ranked ONSET blocks the vowel-copying candidate (10d), so it is the other mis-anchored mapping, the C_2-copying candidate (10c), that is selected as optimal. (*PCR need only be crucially ranked above ANCHOR-L-BR, as it is a *PCR violation that compels mis-anchoring.)

(10) C_2-copying in STVX– roots (*PCR violation): $\sqrt{stamb^h}$- $\to$ *ta-stamb^h-a* 'he has propped'

/RED, stamb^h, a/	*PCR	ONSET	ALIGN-RT-L	ANCHOR-L-BR	CONTIG-BR
a. sa-stamb^h-a	*!		**		*
b. sta-stamb^h-a			***!		
c. ☞ ta-stamb^h-a			**	*	
d. a-stamb^h-a		*!	**	*	

5.3   The C_1ēC$_2$ Pattern in Sanskrit Zero-Grade Perfects

The fact that the cluster in the cases detailed above is a cluster *underlyingly* turns out to be crucial. The C_2-copying pattern only applies to roots whose underlying representation is /CCVX/, where CC is the relevant type of cluster — namely ST. If a (would-be) reduplicative base should become cluster-initial in the course of derivation, namely by means of root vowel-deletion, and the cluster is of the disfavored sort, then a different pattern, the C_1ēC$_2$ pattern, emerges. While it is a very different kind of formation on the surface, it is motivated in exactly the same way as C_2-copying: by the need to satisfy *PCR.

189

In the perfect active plural (and dual) and the perfect middle, "zero-grade ablaut" applies to the root. This refers to a process whereby the vowel of the root is deleted in particular morphological categories. Zero-grade ablaut is frequently believed to be correlated with the position of the accent (see Lundquist & Yates forthcoming:§3 and references therein for discussion); namely, at some stage in the development of ((Pre-)Proto-)Indo-European, if not still in Sanskrit, vowels were subject to a regular phonological deletion process when they surfaced without the word accent. While this relationship no longer transparently holds in many categories in Sanskrit, it does seem to still hold transparently in the perfect: the inflectional categories in which the suffix bears an accent (and thus causes the root to surface without an accent) are exactly those in which zero-grade applies to the root vowel (consult Kiparsky 2010a, 2016, forthcoming, Lundquist & Yates forthcoming, among many others).

When the root contains a medial sonorant, e.g. CRaC or CaRC, zero-grade ablaut generally applies (though there are exceptions; see Steriade 1988) and that sonorant consequently vocalizes. This leads to a surface reduplicative base of the shape CVC ($//\underset{\circ}{w}// \rightarrow$ [u], $//\underset{\circ}{y}// \rightarrow$ [i], $//\underset{\circ}{m},\underset{\circ}{n}// \rightarrow$ [a]) or C$_1$rC ($//\underset{\circ}{r}// \rightarrow$ [r̩]), all of which pattern with the full-grade forms of underlying CVC roots in generating C$_1$-copying. However, in roots which lack a vocalizable sonorant, namely, C$_1$aC$_2$ roots, zero-grade ablaut leads to (or would lead to) a newly cluster-initial reduplicative base:

(11)   Zero-grade for CaC roots: $\sqrt{}$C$_1$aC$_2$- $\xrightarrow[\text{ablaut}]{\text{zero-grade}}$ -C$_1$C$_2$-

This possibility of creating a base-initial cluster sets up the conditions for a potential *PCR-based distribution of reduplication patterns for cluster-initial zero-grade bases, because there are restrictions on the types of repetitions permitted before a consonant. This is borne out, but through a different process than what is observed for cluster-initial roots. When the resulting C$_1$C$_2$-cluster is a TR-cluster, C$_1$-copying is observed, as illustrated in (12).[7] If, however, this new cluster would be an ST-cluster, as for the roots in (13) below,[8] this allomorph would yield a *PCR violation if accompanied by C$_1$-copying.[9] To avoid this, C$_1$-copying is blocked, just as in cluster-initial roots. But the alternative treatment is not C$_2$-copying; instead what we see is an (apparently) non-reduplicative pattern, the "C$_1$ēC$_2$ pattern": $\sqrt{}$C$_1$aC$_2$ $\rightarrow$ perfect [C$_1$ēC$_2$-].

(12)   C$_1$-copying perfects to TR zero-grade bases in Sanskrit

Root		Perfect	
$\sqrt{}$_par-_	'fill'	_pa-pr-úr_	(not **_pēr-úr_, **_ra-pr-úr_)
$\sqrt{}$_bʰar-_	'bear'	_ba-bʰr-ḗ_	(not **_bʰēr-ḗ_, **_ra-bʰr-ḗ_)
$\sqrt{}$_dʰar-_	'hold'	_da-dʰr-ḗ_	(not **_dʰēr-ḗ_, **_ra-dʰr-ḗ_)

---

[7] There are two stop-liquid roots that take the C$_1$ēC$_2$ pattern: _tērúr_ ← $\sqrt{}$_tar-_ 'pass' and _pʰēlúr_ ← $\sqrt{}$_pʰal-_ 'burst; fruit'. These must be treated as exceptions. See Chapter 6 for further discussion.

[8] There are actually many more disfavored clusters than just ST, as seen in (2) above. Because there are relatively few restrictions on what consonants can appear in initial and final position in CaC roots (see Cooper 2009, Sandell 2015a), zero-grade ablaut has the potential to create clusters of virtually any sort, many of which are not permissible as root-initial clusters. The details of this distribution will be laid out in Chapter 6, apropos of the definition of *PCR.

[9] A form like **_sasdúr_ can be ruled out independently on phonotactic grounds: obstruent clusters in Sanskrit display regressive voicing assimilation, which would turn /sd/ into **[zd]; however, [z] is categorically disallowed. Likewise, the examples with /ç/ might also be phonotactically illicit based on the would-be medial -çT- cluster. (See Kessler 1994, Sandell 2015b:230–232, 2017:5–6 for formal discussion of Sanskrit cluster phonotactics.) Nonetheless, there are many other cluster types which undergo the C$_1$ēC$_2$ pattern despite being phonotactically licit, including the [sp] of **_saspur_ (as observed also by Kümmel 2000:19–20). I will return to the role of phonotactics for the C$_1$ēC$_2$ pattern in Chapter 6 when considering the scope of *PCR in Sanskrit.

(13) $C_1\bar{e}C_2$ perfects to ST zero-grade bases in Sanskrit

Root		Perfect	
$\sqrt{}$*sap-*	'serve'	*sēp-úr*	(not **sa-sp-úr, **pa-sp-úr)
$\sqrt{}$*sad-*	'sit'	*sēd-úr*	(not **sa-sd-úr, **da-sd-úr)
$\sqrt{}$*çap-*	'curse'	*çēp-úr*	(not **ça-çp-úr, **pa-çp-úr)
$\sqrt{}$*çak-*	'be able'	*çēk-úr*	(not **ça-çk-úr, **ca-çk-úr)

The most direct way to encode the fact that C_2-copying is not available for these roots is to employ the Input-Reduplicant (IR) faithfulness constraint LINEARITY-IR (cf. McCarthy & Prince 1995, 1999), as defined in (14).

(14)   **LINEARITY-IR**
For every pair of segments in the reduplicant x_R, y_R such that x_R precedes y_R, assign one violation mark * if they have correspondents in the underlying root x_I, y_I, and x_I does not precede y_I.

This constraint will be violated by copying into reduplicant-initial position a root consonant that underlyingly *follows* the root vowel (as in these cases), but not by copying a root consonant that underlyingly *precedes* the root vowel (as is the case for C_2-copying from root-initial clusters). I assume that the reduplicant vowel corresponds to a segment in the underlying root (see Steriade 1988 for arguments in favor of this approach), such that LINEARITY-IR violations are assigned as in (15).

(15)   LINEARITY-IR violations: cluster-initial root vs. CaC root

	Zero-grade category (underlying vowel is deleted in root)	LINEARITY-IR
Cluster-initial Roots:	/RED, $s_1t_2^h\bar{a}_3$, úr/ $\rightarrow$ [$\underline{t_2a_3}$-$s_1t_2^h$-úr]	
CaC Roots:	/RED, $s_1a_2p_3$, úr/ $\rightarrow$ [$\underline{p_3a_2}$-s_1p_3-úr]	*

LINEARITY-IR therefore can, and does, block C_2-copying for these bases, and forces the use of a secondary repair strategy for *PCR, which results in the $C_1\bar{e}C_2$ pattern. A preview of this analysis is sketched in (16), where $\mathbb{C}$ stands in for the constraint(s) violated by the $C_1\bar{e}C_2$ mapping. The identity of this constraint and the full nature of the mapping itself are discussed in the following sections. This demonstrates that, by employing *PCR and LINEARITY-IR, we can arrive at the total distribution of C_1-copying, C_2-copying, and the $C_1\bar{e}C_2$ pattern as laid out in (17) below.

(16)   Blocking reduplication with *PCR and LINEARITY-IR: $\sqrt{}$*sap-* $\rightarrow$ *sēp-ur* 'they have served'

/RED, sap, úr/		*PCR	LINEARITY-IR	$\mathbb{C}$
a.	<u>sa</u>-sp-úr	*!		
b.	<u>pa</u>-sp-úr		*!	
c. ☞	sēp-úr			*

(17)   Distribution of stem-formation patterns in the Sanskrit perfect

	ST cluster (*PCR-violating)	TR cluster (*PCR-satisfying)
Zero-grade Base	$C_1\bar{e}C_2$	C_1-copying
Cluster-initial Root	C_2-copying	C_1-copying

While these constraints circumscribe the conditions under which the $C_1\bar{e}C_2$ pattern surfaces synchronically, they still leave open the question of exactly what mechanism or mechanisms lead to the derivation of exactly this pattern. Potential solutions fall into one of two camps (as per Sandell 2017:8–11): a morphology-centered approach based on phonologically *conditioned* allomorphy, whereby competing derivations based on distinct underlying representations are selected by constraints in the phonological grammar; or a phonology-centered approach based on phonologically *driven* allomorphy, whereby phonological constraints trigger an alternative mapping directly from the otherwise expected underlying representation. (On the terminological distinction, see Carstairs 1988.)

I will first lay out the morphological analysis in Section 5.4, as it is perhaps the simpler means of deriving the result. Subsequently, in Section 5.5, I will present what must be the phonological origin of the pattern, and use this as a spring board for potential phonological analyses. I will show that serial/derivational phonological analyses can capture the pattern in a relatively straightforward way, but that a fully parallel OT analysis is also capable of deriving the pattern, if requiring some unusual machinery.

5.4   An Allomorphy Analysis of the Sanskrit $C_1\bar{e}C_2$ Pattern

In the preceding section, I showed that the phonological constraints *PCR and LINEARITY-IR dictate which type of "allomorph" surfaces in the perfect: a reduplicative allomorph or the $C_1\bar{e}C_2$ allomorph. Therefore, if we are to employ a phonologically conditioned allomorphy approach, the analysis ought to be located in the *phonological* component, where these constraints can exert their force directly (cf. Wolf 2008). It cannot be located in the morphological component, where allomorph selection can interact with phonology only insofar as it can specify particular phonological contexts in its Vocabulary Insertion rules. There is no proposal known to me that would allow reference to subsequent satisfaction or violation of phonological constraints to determine Vocabulary Insertion in the morphology; in any event, this would entail significant look-ahead.

In order to construct an analysis of this sort in the phonological component, we need a mechanism that can select between competing underlying representations — "morphs", in the terminology of Wolf (2008) — for the same morpheme (i.e. morphosyntactic feature or feature bundle) based on the respective outcomes of their phonological derivations. A mechanism of exactly this sort has been proposed by Mascaró (2007) (followed contemporaneously by Bonet, Lloret, & Mascaró 2007), in the form of his "PRIORITY" constraint. Mascaró proposes that the morphological component may output a set of morphs for a given morpheme, and that this set may be (fully or partially) ordered in terms of preference. When a morpheme that has multiple morphs is present in a phonological derivation, GEN produces candidates based on each morph, with candidates indexed to the morph employed. The constraint PRIORITY (reproduced in (18)), which is ranked in CON in the phonological component, assigns violations to any candidate that is indexed to a dispre-

ferred morph. Since this constraint is ranked with respect to other constraints, higher-ranked phonological constraints have the potential to compel the selection of a dispreferred morph.

(18) **PRIORITY** (Mascaró 2007:726)
Respect lexical priority (ordering) of allomorphs.
Given an input containing allomorphs m_1, m_2, ... , m_n, and a candidate m_i', where m_i' is in correspondence with m_i, PRIORITY assigns as many violation marks as the depth of ordering between m_i and the highest dominating morph(s).

This constraint can be adapted to the present case by positing two morphs for PERFECT. For the overtly reduplicative forms, it is clear that the PERFECT morph is /RED/. Given that the remaining forms, i.e., those showing the $C_1\bar{e}C_2$ pattern, differ from their roots only in replacing their root vowel with [-ē-], it is reasonable to posit that the other morph of PERFECT is simply that vowel /-ē-/, or the floating features of which that vowel would be comprised, namely, [+long,-back]. (The vowel is also [-high]. This feature would in principle be available already from the root vowel /a/, but could additionally be specified in the UR of the PERFECT morph.) To indicate that this morph has special linearization conditions, I represent it in italics as /-*ē*-/. Adopting this approach in full would require more explicit analysis of the proper conditions on linearization and/or feature replacement; I leave this for future work, and here simply assume that these conditions hold.

Since the $C_1\bar{e}C_2$ pattern arises only when compelled by simultaneous violation of *PCR and LINEARITY-IR (and also ONSET; see below), we know that /RED/ is the preferred morph here. This is represented in the input to the tableaux in (20) by recording the underlying exponent of PERFECT as {RED$_1$ > -*ē*-$_2$}, i.e., an ordered set of morphs. Given this preference order, PRIORITY will assign a violation to any candidate that uses /-*ē*-/ instead of /RED/. Given the ranking in (19), this will prevent the $C_1\bar{e}C_2$ pattern from surfacing whenever there is (at least) one available reduplicative candidate that simultaneously does not violate *PCR, LINEARITY-IR, and ONSET.

(19) Total ranking for Sanskrit reduplication

<pre>
ONSET   LINEARITY-IR   *PCR
    \ | /
           PRIORITY
              |
        ALIGN-ROOT-L
              |
        ANCHOR-L-BR
              |
       CONTIGUITY-BR
</pre>
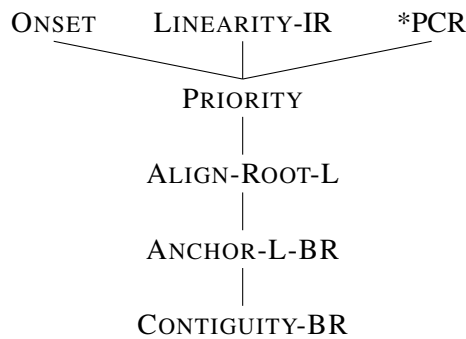
The ranking in (19), which is fully consistent with the general analysis developed in Section 5.2 for the cluster-initial roots, derives the complete distribution from (17) above, as demonstrated in the tableaux in (20) below. In these tableaux, I assume that the underlying form of the root is always the full-grade. This is perhaps a necessary assumption, as the proper operation of LINEARITY-IR requires access to the root vowel in the input; that is to say, if the input contained the zero-grade of the root, then either root-C_1 or root-C_2 should be able to surface as the sole consonant in the reduplicant without violating LINEARITY-IR. (Furthermore, it would be unclear where the vowel is coming from in the reduplicative forms like *pa-pr-úr*, since there would be no *a* in either the input or the output which it could be copying. The vowel cannot be epenthetic, since the epenthetic

vowel in Sanskrit is [i]; see, e.g., Cooper 2014:Ch. 3.) This means that ablaut of the root must be effected in the course of the phonological derivation. Note that LINEARITY-IR will be vacuously satisfied in the (c) candidates, which are those derived from underlying /-ē-/, since the reduplicant is not instantiated in the output. ONSET is omitted for reasons of space; it must dominate PRIORITY because *sēp-ur* ≻ **a-sp-ur*.

(20) Reduplication in Sanskrit: TR vs. ST

Zero-grade cluster-initial bases

	/ {RED₁ > -ē-₂}, sap, úr /		LIN-IR	*PCR	PRIOR	ALN-RT-L	ANCHOR
ST	a.	<u>sa</u>-sp-ur (1)		*!		**	
	b.	<u>pa</u>-sp-ur (1)	*!			**	*
	c. ☞	sēp-ur (2)			*		

	/ {RED₁ > -ē-₂}, par, úr /		LIN-IR	*PCR	PRIOR	ALN-RT-L	ANCHOR
TR	a. ☞	<u>pa</u>-pr-ur (1)				**	
	b.	<u>ra</u>-pr-ur (1)	*!			**	*
	c.	pēr-ur (2)			*!		

Cluster-initial roots

	/ {RED₁ > -ē-₂}, stambʰ, a /		LIN-IR	*PCR	PRIOR	ALN-RT-L	ANCHOR
ST	a.	<u>sa</u>-stambʰ-a (1)		*!		**	
	b. ☞	<u>ta</u>-stambʰ-a (1)				**	*
	c.	stēmbʰ-a (2)			*!		

	/ {RED₁ > -ē-₂}, pracʰ, a /		LIN-IR	*PCR	PRIOR	ALN-RT-L	ANCHOR
TR	a. ☞	<u>pa</u>-prācʰ-a (1)				**	
	b.	<u>ra</u>-prācʰ-a (1)				**	*!
	c.	prēcʰ-a (2)			*!		

This phonologically conditioned allomorphy analysis is to be preferred over that previously proposed in Zukoff (2015:8–9) and Sandell (2015b:230, 2017:6–7) because it clarifies the relationship between surface pattern and the morphemes (and their morphs) that generate it. These analyses employed ranked "USE:X" constraints (based loosely on MacBride 2004's "FIAT" constraints) to encode allomorph preference order; this is functionally equivalent to Mascaró's (2007) unitary PRIORITY constraint. However, the USE:X approach led to a mismatch in the mapping between surface pattern and morphs.[10]

For example, in Zukoff (2015:8–9), I employed the constraints USE:RED and USE:C₁ēC₂, with USE:RED in the position in the rankings where PRIORITY now sits. But USE:RED refers to a morph of PERFECT while USE:C₁ēC₂ would have to refer to a morph of the root (and furthermore,

---

[10] Sandell (2017:11) partially addresses this problem by tacitly comparing an input with what amounts to a portmanteau allomorph (essentially PERFECT.ROOT) to an input with separate allomorphs for PERFECT (/RED/) and ROOT. These inputs would themselves have to be derived by rule in the morphological component (Ryan Sandell, personal communication).

this morph would itself have to be derived by rule in each individual derivation[11]). Therefore, these constraints (as formulated) refer to different morphemes, and leave unexplained the morph selection for the other morpheme involved. In other words, not only would these USE:X constraints be regulating morph selection for different morphemes, these morphs would themselves have to control morph selection for the other morpheme, as well.

The USE:X-based analysis could, however, easily be updated to fix this problem. If it simply swapped out USE:$C_1\bar{e}C_2$ for USE:/-$\bar{e}$-/, and kept USE:RED in the position of PRIORITY, it would be completely equivalent to the PRIORITY-based analysis, modulo the question of whether morph preference is encoded with one constraint or multiple constraints.[12]

In the following section, I will explore possible phonological analyses of the $C_1\bar{e}C_2$ pattern that posit that the $C_1\bar{e}C_2$ forms actually do derive from an underlying representation containing /RED/. Sandell (2017:10–11) adduces an argument against such an approach (for the synchronic analysis of attested Sanskrit), and thus an argument in favor of the phonologically conditioned allomorphy approach just developed. When underlyingly aspirated consonants have a correspondent in the reduplicant, that reduplicated consonant lacks aspiration: for example, $\sqrt{d^har}$ 'hold' $\rightarrow$ __da__-d^har- not **__d^ha__-d^har-. This is normally seen as an effect of "Grassmann's Law", which prohibits aspirated consonants in adjacent syllables (see Collinge 1985:47–51 and citations therein; see also the brief discussion in Chapter 6). However, as Sandell argues, it can also be interpreted as an *emergence of the unmarked* (TETU) effect (McCarthy & Prince 1994, 1995) that effectively bans aspiration in the reduplicant.

A TETU analysis would seem to be required for the form *bapsati* (Whitney 1885 [1988]:109), a reduplicated present (with zero-grade in the plural) to the root $\sqrt{b^has}$ 'devour'. Aspiration and voicing are neutralized before [s], thus forcing the underlying root-initial /b^h/ to surface as [p]. The reduplicant consonant is faithful to the voicing of the underlying consonant, but not to its aspiration, thus surfacing as [b] not **[b^h]. (The voicing of the reduplicant consonant suggests that the relevant faithfulness dimension is Input-Reduplicant not Base-Reduplicant.) If the absence of aspiration in the reduplicant were about a prohibition on adjacent aspirated consonants on the surface (something like *[+spread glottis]...[+spread glottis]), then the failure of aspiration to surface in the base should license its surfacing in the reduplicant. Since it appears not to, this would suggest that the ban is actually specific to the reduplicant, and thus a TETU effect.[13]

However, there is contradictory evidence as well. In the reduplicated present of the root $\sqrt{d^h\bar{a}}$ 'put' (Whitney 1885 [1988]:82), aspiration *does* surface in the reduplicant just in case it is not licensed in the base: for example, 2nd singular middle __d^ha__-*t-tē* (see more on the aspiration behavior of this paradigm in Chapter 6). This form suggests that aspiration is indeed in principle allowed in the reduplicant, and that its normal absence should be attributed to a Grassmann's Law-type co-occurrence constraint.

The $C_1\bar{e}C_2$ form $b^h\bar{e}j\acute{u}r \leftarrow \sqrt{b^haj}$ 'divide' does allow the underlying aspiration of its root-initial consonant to surface. If the $C_1\bar{e}C_2$ forms were synchronically reduplicated, this would mean that the [b^h] in $b^h\bar{e}j\acute{u}r$ would be in the reduplicant. If the normal prohibition on reduplicant aspiration was indeed a TETU effect (as implied by *bapsati*), then this would prove that the [b^h] in $b^h\bar{e}j\acute{u}r$ was not in the reduplicant, and therefore that the $C_1\bar{e}C_2$ forms are not synchronically reduplicated.

---

[11] It is not sufficient to say that it surfaces only when it happens to be listed, because, as shown by Sandell (2015b:Ch. 8), the pattern very clearly expands its scope through the attested period of Sanskrit to apply to roots for which it could (presumably) not have arisen by sound change or a regular phonological process; see also Chapter 6 of this dissertation.

[12] Mascaró (2007) and Bonet, Lloret, & Mascaró (2007) argue for the unitary constraint approach on the grounds that (they claim) it is more restrictive. It is unclear to me if this is true, but such an argument would not on its own be fully decisive anyway. This question has no bearing on the analysis here.

[13] The same argument can be made from perfect *jakṣur* from $\sqrt{g^has}$ 'eat' (Whitney 1885 [1988]:42).

However, if the ban on reduplication was driven by a surface co-occurrence constraint (as implied by *dʰatte*), then we would predict the aspiration to surface regardless of whether the consonant was in a reduplicant or not. Since the independent evidence is inconsistent, this cannot be taken as a decisive argument for a synchronic analysis of the Sanskrit $C_1\bar{e}C_2$ pattern as phonologically conditioned allomorphy.

5.5 Phonological Analyses of the Sanskrit $C_1\bar{e}C_2$ Pattern

While the allomorphy approach may be sufficient to capture the synchronic distribution of the $C_1\bar{e}C_2$ pattern once it is fully developed within Sanskrit, it does not provide insight into why this pattern should be part of the system, which is otherwise entirely based around reduplication, in the first place. Given that its ultimate distribution is governed by *PCR, we should explore whether *PCR might be sufficient to motivate a *phonological* derivation of the pattern directly from a reduplicated formation (i.e. phonologically *driven* allomorphy). In this section, I demonstrate that a phonological analysis based on reduplication + consonant deletion with compensatory lengthening is capable of deriving the desired pattern.

While generating this pattern in a fully parallel phonological framework turns out to be fairly complex (though nonetheless possible), it is relatively straightforward to derive the pattern in a serial framework, as long as we have *PCR at our disposal. Therefore, I will first in Section 5.5.1 illustrate how this derivation would work in a basic ordered rule-based approach. This approach does have the conceptual benefit of directly recapitulating the diachronic developments. I will then show in Section 5.5.2 that this rule-based analysis can be directly adapted to a Stratal OT account, although the ranking of *PCR across the different strata that this analysis requires may ultimately be unappealing. Lastly and in the greatest detail, I will develop in Section 5.5.3 a parallel OT analysis of the pattern. This analysis will require the combination of several unusual (though precedented) pieces of theoretical machinery, but can nonetheless derive the desired outcome.

In the following discussion, I will make one simplifying move: for the purposes of these analyses, I take the underlying root vowel to be /e/, and likewise the surface vowel of the reduplicant to be [e]. This clearly does not hold for Sanskrit (or even Proto-Indo-Iranian), where PIE *e* has become *a* (probably phonetically [ə], *vel sim.*). However, in the virtually identical pattern which can be identified in Germanic (see Section 5.7 below), these vowels would indeed have been *e* at the relevant stage. (If such a process were attributable to Proto-Indo-European, they likewise would have been *e*.) This is a helpful simplification because it makes more transparent the notion that the surface [eː] vowel in the $C_1\bar{e}C_2$ forms should be interpreted as compensatory lengthening of the short vowel of the reduplicant. I will return to these issues of vowel quality in Section 5.6 below.

5.5.1 A Rule-Ordering Analysis of the $C_1\bar{e}C_2$ Pattern

In any of the approaches, the input will start with three elements: (i) a verbal root of the shape /$C_1\breve{V}C_2$/ (namely, /C_1eC_2/, given the above caveats); (ii) an accented inflectional ending — I will employ the Sanskrit 3PL.ACTIVE *-úr*; and (iii) a PERFECT morph that triggers reduplication — I represent this as /RED/. The rule-based analysis of the pattern involves essentially three ordered processes.[14] First, reduplication is carried out; it takes its canonical shape as a CV prefix (21.i).[15]

---

[14] This analysis mirrors almost entirely that proposed by Schumacher (2005).

[15] On the derivation of reduplicant shape in rule-based phonology, consult, for example, Marantz (1982), McCarthy & Prince (1986 [1996]), Steriade (1988).

This yields an intermediate string C_1e-C_1eC_2-$\acute{u}r$. Second, "zero-grade ablaut" applies; this is deletion of the root vowel triggered by the accent on the inflectional ending (21.ii). This yields an intermediate string $C_1eC_1C_2\acute{u}r$. The last step (which could be conceived of as two distinct steps) involves deletion of the root-initial consonant. This operation results in compensatory lengthening; namely, it deletes the melodic specifications of the consonant, but leaves its timing slot, which re-associates to the preceding vowel (21.iii). This yields the final output form [$C_1\bar{e}C_2\acute{u}r$]. The context for the deletion step in (21.iii) is the *PCR environment: it applies only when the consonant is the second member of a repetition, and only when the consonant is pre-obstruent.

(21)   Serial, rule-based derivation of the CēC pattern

		Input: $/\text{RED}, C_1eC_2, \acute{u}r/$
i.	Reduplicate: *copy CV*	C_1e-C_1eC_2-$\acute{u}r$
ii.	Zero-grade: *delete root vowel*	$C_1eC_1C_2\acute{u}r$
iii.	*PCR-driven C-deletion + CL: $VC_\alpha \to \bar{V} \ / \ C_\alpha\_C$	$C_1\bar{e}C_2\acute{u}r$
	Output:	[$C_1\bar{e}C_2\acute{u}r$]

5.5.2   A Stratal OT Analysis of the $C_1\bar{e}C_2$ Pattern

This demonstrates that the $C_1\bar{e}C_2$ pattern can, in the broadest sense, be derived phonologically. Deriving it in a fully parallel version of Optimality Theory (Prince & Smolensky 1993/2004), however, is non-trivial. While there are a number of difficulties involved, the problem boils down to the fact that this pattern appears to display greater faithfulness to a consonant which is copied into the reduplicant than to that copy's base correspondent. *A priori*, this necessitates abandoning McCarthy & Prince's (1995,1999) "Basic Model" of reduplication — which has only Input-Output (IO; or IB for Input-Base) faithfulness and Base-Reduplicant (BR) faithfulness — for some more complex model.[16] Stratal OT (consult Kiparsky 2015, among others) is one possible approach (see Kiparsky 2010b on reduplication in Stratal OT).

Stratal OT holds that there are multiple levels — called *strata* — of the phonological grammar, with potentially different constraint rankings. This will allow us to implement the serial rule-based analysis proposed above within a constraint-based phonological framework. Specifically, it has the potential to apply reduplication, accentuation, and ablaut in an earlier stratum (or across multiple earlier strata) where *PCR is not yet high-ranked enough to have any effect. Let's for now assume that this is the case (I'll return to this below), and that these earlier strata can generate an intermediate input equivalent to (21.ii) $C_1eC_1C_2\acute{u}r$. Once *PCR becomes active at the subsequent stratum (i.e. is more highly ranked in CON in that stratum), it is possible to derive deletion and compensatory lengthening in a fairly straightforward way.

When presented with an intermediate input $C_1eC_1C_2\acute{u}r$, as shown in (22) below, a high-ranked *PCR will dictate that some repair must take place; and thus the faithful candidate (22a) is ruled out. Since the segments resulting from reduplication are already present in the input at this stage, avoidance strategies which would be available in a parallel OT analysis, involving different types of copying, are inherently now unavailable. The possible repairs here reduce essentially to deletion and epenthesis (and change in the feature [±sonorant], if *PCR is here sensitive to this difference).

---

[16] Note that I have already made recourse to McCarthy & Prince's (1995,1999) "Full Model", which supplements these correspondence dimensions with Input-Reduplicant correspondence.

As long as DEPV dominates MAXC, deletion will be preferable to epenthesis; and thus a candidate like (22f) $[C_1eC_1iC_2\text{úr}]$ is ruled out.

(22)   Stratal derivation (final level)

$C_1eC_1C_2\text{úr}$		*PCR	DEPV	MAXC/_V	MAX-μ	MAXC
a.	$C_1eC_1C_2\text{úr}$	*!				
b. ☞	$C_1e\text{ː}C_2\text{úr}$					*
c.	$C_1eC_2\text{úr}$				*!	*
d.	$C_1eC_1\text{úr}$			*!		*
e.	$eC_1C_2\text{úr}$			*!		*
f.	$C_1eC_1iC_2\text{úr}$		*!			

In theory, deleting any of the three consonants (reduplicant-C_1, root-C_1, and root-C_2)[17] would alleviate the *PCR problem, since the context it penalizes necessarily includes all three. The most direct way to select root-C_1 as the target of deletion is to employ a contextual version of MAXC that prefers faithfulness to consonants that are pre-vocalic in the input: MAXC/_V (Steriade 1997, 2009; cf. Beckman 1998, McCarthy 2008, 2011). Since reduplicant-C_1 and root-C_2 are in pre-vocalic position in the input at this stage of the derivation, but root-C_1 is not, root-C_1 will be targeted for deletion when this constraint is included anywhere in the ranking. MAXC/_V thus prefers desired candidate (22b) $[C_1e\text{ː}C_2\text{úr}]$ to the candidates with other deletion sites, namely candidate (22d) $[C_1eC_1\text{úr}]$ (which deletes root-C_2) and candidate (22e) $[eC_1C_2\text{úr}]$ (which deletes reduplicant-C_1). To ensure compensatory lengthening, MAX-μ must also be present in the grammar. (This must be ranked above *Vː, which must also be dominated by *Cː to ensure the correct type of re-association; see below.)

This approach is able to derive the result with such straightforward constraints largely because it has limited the scope of possible candidates which are available to solve the *PCR problem. In a fully parallel approach, alternative copying patterns are always available. Therefore, greater precision is required in order to rule out all alternatives. For this particular case especially, that will mean employing less straightforward constraints. While the Stratal OT approach thus has the advantage of a simpler analysis, there is perhaps a reason to be skeptical of it. This relates to the nature of the relative ranking of *PCR across the different strata.

First, the Stratal OT analysis requires that zero-grade ablaut precede *PCR-driven consonant deletion. The main reason, vis-à-vis the analysis just presented, is the following. If the root vowel were still present at the stage where *PCR-driven consonant deletion occurs, all three consonants would be in pre-vocalic position in the intermediate input: $C_1eC_1eC_2\text{úr}$. This would render MAXC/_V irrelevant. While adding a constraint like ONSET or ANCHOR-L(-IO) could rule out deleting reduplicant-C_1 (i.e. candidate (22e) $[eC_1C_2\text{úr}]$), there is no obvious constraint which could distinguish between deletion of root-C_1 (desired candidate (22b) $[C_1e\text{ː}C_2\text{úr}]$) and deletion of root-$C_2$ (undesired candidate (22d) $[C_1eC_1\text{úr}]$). (Furthermore, compensatory lengthening would be less straightforward, as the consonant being deleted was in onset position in the input.) Given that ablaut must precede *PCR-driven consonant deletion, and ablaut actively creates a *PCR-violating sequence, we know that *PCR must be (relatively) low-ranked in the ablaut stratum but high-ranked in the consonant-deletion stratum.

---

[17] These labels are a matter of convenience. Stratal OT holds that there is no functional distinction between root and reduplicant at this stage of the derivation.

Now consider again the behavior of roots that are underlyingly cluster-initial (see Section 5.2). On the one hand, TRVX– roots show C_1-copying: for example, $\sqrt{prac^h} \rightarrow \underline{pa}\text{-}prac^h\text{-}$. But on the other hand, STVX– roots show C_2-copying: for example, $\sqrt{st^h\bar{a}} \rightarrow \underline{ta}\text{-}st^h\bar{a}\text{-}$. Since this distribution is governed by *PCR, *PCR must be active in the stratum where reduplicant shape is determined.[18] This implies, then, that ablaut is to be ordered in a stratum after that which contains reduplication, because ablaut freely creates *PCR violations. Therefore, we seem to have a three stratum grammar in which *PCR effectively gets "turned off" going from Stratum 1 to Stratum 2, and then "turned on" again going from Stratum 2 to Stratum 3. While a precise theory of accent and ablaut[19] is of course required in order to confirm that this is a proper characterization of the ranking differences between strata, this seems like a highly undesirable ranking pattern, and part of the general concern over Stratal OT's lack of restrictiveness regarding inter-stratum ranking patterns. Since the primary goal of this chapter is to demonstrate the feasibility of the parallel OT analysis, to which I turn next, I will leave further inquiry into the serial/stratal approach for future investigation.[20]

5.5.3 A Parallel OT Analysis of the $C_1\bar{e}C_2$ Pattern

If we eschew serialism altogether, the task becomes more complicated. Nevertheless, as I will demonstrate in the rest of this section, a fully parallel approach is indeed capable of capturing the pattern. In the fully parallel approach, the goal is to derive the output $[C_1eːC_2\text{úr}]$ directly from an input $/\text{RED}, C_1eC_2, \text{úr}/$, with no intermediate representations or levels. Since we will be dealing with a greater variety of candidates in this analysis, it will be helpful to use a non-schematic example. I will thus refer to forms based on the Sanskrit root $\sqrt{sap}$, which shows a $C_1\bar{e}C_2$ form *sēpúr*. Note though that I anachronistically treat the underlying root vowel as $/e/$ and the reduplicant vowel as $[e]$ (see the beginning of this section on operating assumptions about vowel quality). Therefore, the mapping I will aim to derive is $/\text{RED}, \text{sep}, \text{úr}/ \rightarrow [\text{seːpúr}]$.

5.5.3.1 Phonological Interpretation of the Output Form

The parallel analysis requires significant exposition regarding the phonological interpretation of the output form [seːpúr]. My interpretation is as follows:

- The [p] and the suffix segments [-úr] plainly correspond faithfully to the appropriate segments in the input via Input-Output (IO) correspondence, and need no special comment.

- The input root vowel $/e/$ has no correspondent in the output root. It has undergone total deletion, including deletion of its underlying mora/timing slot (i.e. its mora/timing slot has not been re-associated to another segment). This is induced by whatever triggers zero-grade ablaut (presumably the following accent).

---

[18] Kiparsky (2010b) is not terribly explicit about how the calculation of reduplicant shape is to be made in the Stratal OT approach to reduplication; he references markedness-defined templates, but never considers exactly how these will select between alternatives that are equivalent with respect to markedness, as in the current case. For present purposes, then, let's simply assume that (i) markedness constraints dictate a CV reduplicant, (ii) faithfulness constraints prefer a C_1-initial reduplicant, and (iii) that this preference is overridden by *PCR and a C_2-initial reduplicant is available. If these assumptions cannot be upheld, then the Stratal OT approach fails independently of the problem currently under discussion.

[19] See again Kiparsky (2010a, 2016, forthcoming) recent theoretical work in this direction.

[20] Preliminary investigation indicates that Harmonic Serialism (HS) — a mono-stratal serial OT approach (see McCarthy 2000, 2010, among others; see McCarthy, Kimper, & Mullin 2012 on reduplication in HS) — will not support a convergent analysis at all, because of this reversive behavior of *PCR. Again, however, this largely depends on the proper analysis of ablaut.

- The [s] and the [e] (absent the length) comprise the reduplicant proper. They correspond directly (and faithfully) to the /s/ and /e/ of the *input root*, along the dimension of Input-Reduplicant (IR) correspondence.

- The extra length ([ː]) which surfaces on the reduplicant [e] is the *output root* correspondent of the *timing slot* of the input root /s/. While all of the features, and indeed the root node, of that input root /s/ have been deleted (i.e. lack a correspondent in the output root), its timing slot is preserved in the output root, and it is re-associated to the [e] of the reduplicant.

- The only individual items which stand in Base-Reduplicant (BR) correspondence are the reduplicant-initial [s] and the second timing slot associated to the reduplicant [e].

These relations are represented in the diagram in (23). The X's represent timing slots associated with segments specified in the input; they are subscripted with the segments themselves. Solid downward arrows represent Input-Output mappings between timing slots. The boxes in the output represent deletions along the IO dimension: deletion of the features and the timing slot of the underlying root vowel /e/, but deletion just of the features of underlying root-C_1 /s/. The dotted lines from the input root to the output reduplicant represent the IR correspondence relation. The curved arrow from the timing slot in the output base to the [e] of the reduplicant represents the re-association that generates the surface long vowel in the output.

(23)   Diagrammatic representation of optimal mapping to [seːpúr]

$$/ \quad \text{RED} \quad - \quad \{_{\text{I-ROOT}} \quad X_s \quad X_e \quad X_p \} \quad - \quad \acute{X}_u \quad X_r \quad /$$

$$[\quad X_s \quad X_e \quad \{_{\text{O-ROOT}} \quad \boxed{X_\emptyset} \quad \boxed{\emptyset_\emptyset} \quad X_p \} \quad \acute{X}_u \quad X_r \quad] \quad = \quad \textit{seːpúr}$$
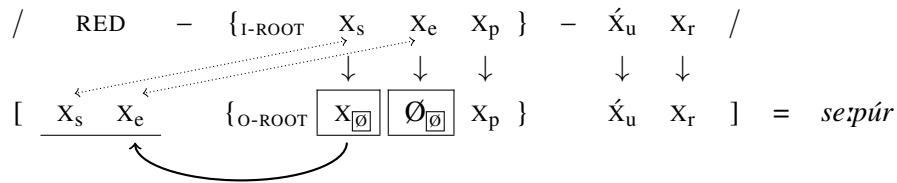
Under this interpretation, the length on the long vowel is to be understood as the output correspondent of input root-C_1. As such, I will represent this candidate as [se̲-ːp-úr], where the length mark (i.e. the underlyingly consonantal timing slot) is effectively outside the reduplicant proper. With the interpretation of the mapping itself now made explicit, we can begin to untangle the constraints that must be involved, and show why and how [se̲-ːp-úr] is superior to all other possible candidates.

5.5.3.2   Candidates and Constraints

Perhaps the most conceptually important (and unusual) constraint involved in this analysis is the faithfulness constraint MAX-X_c-IO, as defined in (24a). This constraint requires that all *consonantal* timing slots (i.e. timing slots associated with a [+consonantal], or perhaps [-syllabic], root node) in the input are retained in the output. This is the constraint that will ensure the underlying root-initial consonant (/s/) contributes its length to the output. This must be dissociated from the more familiar constraint MAX-C-IO, as defined in (24b). In this context, we must understand this constraint to be militating for the retention of input consonantal *root nodes* (and thus the features contained under such root nodes) into the output. Since a root node requires a timing slot in order to surface in the output, this retention of the root node will entail retention of the timing slot (at least under most conceivable circumstances). Due to this implicational relationship between the two constraints, they need not be critically ranked

(24) a. **MAX-X$_c$-IO**
        For each timing slot (X) associated with a consonant in the input, assign one violation
        mark * if that timing slot does not have a correspondent in the output.
    b. **MAX-C-IO**
        For each consonantal root node in the input, assign one violation mark * if that root
        node does not have a correspondent in the output.

The optimal candidate [se̲-ːp-úr] violates MAX-C-IO, since it deletes the root node of the
input /s/, but satisfies MAX-X$_c$-IO, since it has nonetheless preserved that segment's timing slot.
The other constraint that the optimal mapping violates (more than its competitors) is *Vː (25), which
penalizes long vowels.[21]

(25) ***Vː**
    Assign one violation mark * for each long vowel in the output.

Having defined MAX-X$_c$-IO, and identified the constraints violated by the optimal mapping
[se̲-ːp-úr], we can now begin to consider the ways in which this candidate is superior to other,
potentially more transparent mappings. First, the two candidate mappings that match the attested
outcomes for cluster-initial roots can be ruled out in the same way they were in the allomorphy
analysis in Sections 5.3–5.4. A candidate [se̲-sp-úr] (29b) with default C$_1$-copying is ruled out by
*PCR (whose definition is repeated below in (26)). A candidate [pe̲-sp-úr] (29c) with C$_2$-copying
is ruled out by LINEARITY-IR (whose definition is repeated below in (27)), because the redu-
plicant's consonant and vowel are in the reverse order of those segments' correspondents in the
underlying root. This shows us that *PCR and LINEARITY-IR both dominate MAX-C-IO and *Vː,
as recorded in the ranking in (28), and demonstrated in the tableau in (29).

(26) ***PCR**
    For each sequence of repeated identical consonants separated by a vowel (C$_\alpha$VC$_\alpha$), assign
    a violation * if that sequence immediately precedes an obstruent.

(27) **LINEARITY-IR**
    For every pair of segments in the reduplicant x_R, y_R such that x_R precedes y_R, assign a viola-
    tion * if they have correspondents in the underlying root x_I, y_I, and x_I does not precede y_I.

(28) **Ranking:** *PCR, LINEARITY-IR ≫ MAX-C-IO, *Vː

(29) [se̲-ːp-úr] ≻ [se̲-sp-úr], [pe̲-sp-úr]

/RED, sep, úr/		*PCR ¦ LINEARITY-IR	MAX-C-IO ¦ *Vː
a. ☞ se̲-ːp-úr		¦	* ¦ *
b. se̲-sp-úr		*! ¦	¦
c. pe̲-sp-úr		¦ *!	¦

---

[21] Alternatively, one could identify the problematicity of the long vowel as being due to the coalescence of the root-initial
timing slot with the reduplicative vowel. This might violate some sort of UNIFORMITY constraint (cf. McCarthy &
Prince 1995). However, it could only be UNIFORMITY-IO if we took the IO dimension to include reduplicant segments
(cf. Spaelti 1997, Struijke 2000). This might have ramifications for the IR constraints proposed here, so I instead stick
with *Vː as the relevant constraint.

Next to consider are the other deletion candidates, i.e. other candidates that violate MAX-C-IO. Any candidate that deletes a root consonant without maintaining its timing slot will fatally violate MAX-X$_c$-IO. This includes all permutations of one consonant in the root and one consonant in the reduplicant: [se-s-úr] (31b), [se-p-úr] (31c), [pe-p-úr] (31d), and [pe-s-úr] (31e). (The latter two additionally violate undominated LINEARITY-IR.) As long as MAX-X$_c$-IO dominates *V: (as shown in (30)), then [se-:p-úr] will be preferable to all these candidates, as demonstrated in (31).

(30)   **Rankings:** MAX-X$_c$-IO ≫ *V:

(31)   [se-:p-úr] ≻ [se-s-úr], [se-p-úr], [pe-p-úr], [pe-s-úr]

/RED, sep, úr/		LINEARITY-IR	MAX-X$_c$-IO	MAX-C-IO	*V:
a. ☞	se-:p-úr			*	*
b.	se-s-úr		*!	*	
c.	se-p-úr		*!	*	
d.	pe-p-úr	*!	*!	*	
e.	pe-s-úr	*!	*!	*	

Since the abiding factor in this derivation seems to be maintenance of underlying consonantal timing slots, we can now consider other sorts of deletion + compensatory lengthening candidates. One which can immediately be dispatched is [pe-:p-úr]: this violates LINEARITY-IR just like the non-deletion candidate [pe-sp-úr]. One group of viable competitors are those which compensatorily lengthen the surfacing root *consonant* instead of the reduplicant vowel: namely, [se-s:-úr], which is diagrammed in (32a); and [se-p:-úr], which is diagrammed in (32b). The long consonant represents a consonantal root node attached to two timing slots; [se-p:-úr] therefore need not violate something like LINEARITY-IO, since it may represent the dual association of the root node [p] to the underlying timing slots of X$_s$ and X$_p$ in that order. The most straightforward way to prefer the desired compensatory vowel-lengthening candidate [se-:p-úr] to either of these compensatory consonant-lengthening candidates is to simply appeal to the relative markedness of long consonants (i.e. geminates) and long vowels: if *C: (defined in (33)) outranks *V:, and is ranked high enough overall, then this will result in compensatorily lengthening a vowel rather than a consonant.[22] This ranking is shown in (34), and demonstrated in (35).

(32)   Compensatory consonant lengthening candidates
         a.   Diagrammatic representation of suboptimal mapping to **[ses:úr]

---
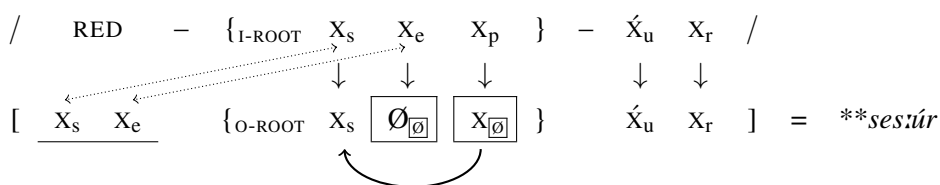
[22] This could alternatively be derived via some preference for leftward re-association over rightward re-association.

b. Diagrammatic representation of suboptimal mapping to **[sepːúr]

$$/ \quad \text{RED} \quad - \quad \{\text{I-ROOT} \quad X_s \quad X_e \quad X_p \} \quad - \quad \acute{X}_u \quad X_r \quad /$$

$$[\quad X_s \quad X_e \quad \{\text{O-ROOT} \quad \boxed{X_{\boxed{\varnothing}}} \quad \boxed{\varnothing_{\boxed{\varnothing}}} \quad X_p \} \quad \acute{X}_u \quad X_r \quad] \quad = \quad **sepːúr$$
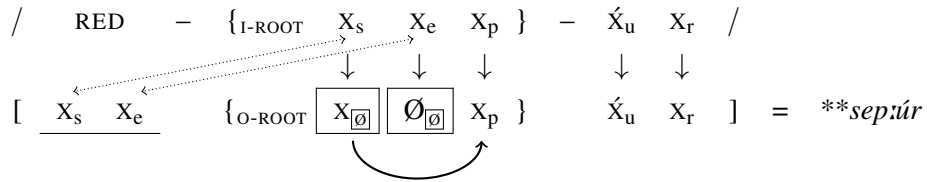
(33) ***Cː**
Assign one violation mark * for each long consonant (i.e. geminate) in the output.

(34) **Rankings:** *Cː ≫ *Vː

(35) [se̲ːp-úr] ≻ [se̲-sː-úr], [se̲-pː-úr]

/RED, sep, úr/	MAX-X$_c$-IO	MAX-C-IO	*Cː	*Vː
a. ☞ se̲ːp-úr		*		*
b. se̲-sː-úr		*	*!	
c. se̲-pː-úr		*	*!	

There is one more alternative compensatory lengthening candidate that is a bit trickier to eliminate: [se-ːs-úr]. This candidate shows compensatory lengthening of the reduplicant vowel, just like optimal candidate [se-ːp-úr]. There are two possible ways such a string might be generated. The first, as diagrammed in (36a), is by deleting the root node of underlying /p/, and re-associating its timing slot directly to the reduplicant vowel. This, however, would be non-local re-association, since the timing slot of /s/ surfaces in between. This would essentially violate the No Crossing Lines condition, and thus may reasonably be assumed to not be furnished by GEN. The only other way to generate this candidate would be by sequential re-association, as diagrammed in (36b). Here, the timing slot and root node of /s/ are again both preserved in the output, but they are dissociated from one another. The timing slot belonging to /s/ is re-associated to the reduplicant vowel. And the timing slot associated to /p/, whose root node was deleted, is itself re-associated to the root node belonging to /s/. (This would be essentially equivalent to the "double flop" compensatory lengthening found in Ancient Greek; Steriade 1982, Hayes 1989:265–266.)

(36) Diagrammatic representation of suboptimal mapping to **[seːsúr]

a. Non-local re-association

$$/ \quad \text{RED} \quad - \quad \{\text{I-ROOT} \quad X_s \quad X_e \quad X_p \} \quad - \quad \acute{X}_u \quad X_r \quad /$$

$$[\quad X_s \quad X_e \quad \{\text{O-ROOT} \quad X_s \quad \boxed{\varnothing_{\boxed{\varnothing}}} \quad \boxed{X_{\boxed{\varnothing}}} \} \quad \acute{X}_u \quad X_r \quad] \quad = \quad **seːsúr$$

b. Sequential re-association

$$/ \quad \text{RED} \quad - \quad \{_{\text{I-ROOT}} \quad X_s \quad X_e \quad X_p \quad \} \quad - \quad \acute{X}_u \quad X_r \quad /$$

$$[\quad \underline{X_s \quad X_e} \quad \{_{\text{O-ROOT}} \quad X\nleftrightarrow_s \quad \boxed{\emptyset_{\boxed{\emptyset}}} \quad \boxed{X_{\boxed{\emptyset}}} \quad \} \quad \acute{X}_u \quad X_r \quad] \quad = \quad **\textit{se:súr}$$
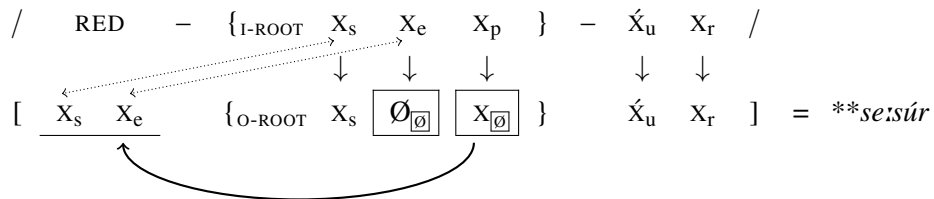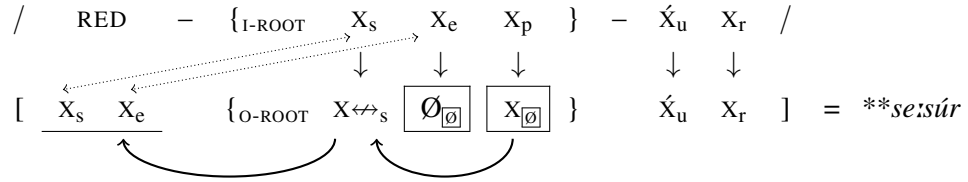
There are at least two things that are bad about this candidate relative to the optimal mapping. For one, it has more *timing slot ↔ root node* re-associations (two instead of one). Additionally, it dissociates a timing slot from its root node despite both surfacing in the output. This latter problem is unique to [se-ːs-úr] among viable candidates. Therefore, if we identify this as its fatal flaw, whatever exactly the constraint is that this mapping violates will not need to be crucially ranked, since this candidate's violation profile is otherwise identical to the winning candidate [se-ːp-úr]. I remain agnostic as to the details of this constraint, and proceed without further consideration of the candidate [se-ːs-úr].

There are three more candidates left to consider, none of which involve deletion. One is the candidate that copies no consonants at all: [e-sp-úr] (candidate (39e)). By copying just the vowel of the underlying root, the potential *PCR and LINEARITY-IR problems are avoided, as are all those relating to deletion and compensatory lengthening. However, it has one very simple problem: it violates ONSET (as defined in (7b) above).

The two remaining candidates are those which copy all the segments of the underlying root: [spe-sp-úr] (candidate (39b)) and [sep-sp-úr] (candidate (39c)). The former violates LINEARITY-IR, so it is not viable. [sep-sp-úr], therefore, is the one which bears consideration. If we are operating within Sanskrit, a medial *-TST-* sequence is not phonotactically licit (such a sequence is diachronically reduced to *-T:-* in the development from PIE to Sanskrit). If it were licit, though, there would still be a way to prefer [se-ːp-úr]: the size minimizer ALIGN-ROOT-L, as originally defined in (7a) above (repeated with slight adjustment in (37) below), will prefer the shorter reduplicant of [se-ːp-úr] (two violations) to the longer reduplicant of [sep-sp-úr] (three violations). This assumes that, for the purposes of ALIGN-ROOT-L, the left edge of the root is the timing slot corresponding to the timing slot of underlying root-initial /s/. It cannot be the case that the left edge of the root is the left edge of the segment to which that timing slot is associated, because this would drive deletion with compensatory lengthening in reduplication of all root shapes, since this is the maximal means of minimizing the reduplicant (while still providing it with surface content). Logically, then, the constraint must be evaluated over timing slots rather than segments, as in the definition in (37).

(37) **ALIGN-ROOT-L**
Assign one violation mark * for each timing that intervenes between the leftmost timing slot of the output root and the left edge of the word.

If [sep-sp-úr] were indeed phonotactically illicit, then we could instead consider a candidate that avoids the phonotactic problem through epenthesis: [sepi-sp-úr] (candidate (39d)). This would be somewhat similar to the Pre-Greek Attic Reduplication pattern discussed in Chapter 2. Whether or not DEP-V-IO is highly ranked, ALIGN-ROOT-L will be all the more effective in ruling out this candidate, as it has even more material preceding the root. Therefore, if we have the rankings shown in (38), we can properly select [se-ːp-úr] from among these candidates. This is demonstrated in (39). Note that ONSET must dominate ALIGN-ROOT-L, as otherwise the vowel-copying candidate (39e) would be preferred, since it has fewer alignment violations.

(38)   **Rankings:** ONSET ≫ ALIGN-ROOT-L ≫ MAX-C-IO, *Vː

(39)   [se-ːp-úr] ≻ [spe-sp-úr], [sep-sp-úr], [sepi-sp-úr], [e-sp-úr]

/RED, sep, úr/	LIN-IR	ONSET	ALIGN-ROOT-L	MAX-C-IO	*Vː
a. ☞ se-ːp-úr			**	*	*
b. spe-sp-úr	*!		***		
c. sep-sp-úr			***!		
d. sepi-sp-úr			***!*		
e. e-sp-úr		*!	*		

All viable alternative candidates to [se-ːp-úr] have now been dispensed with. Therefore, we can conclude that it is indeed possible to derive the $C_1\bar{e}C_2$ pattern from an underlying representation containing RED and the basic form of the root in a purely parallel evaluation. It requires Input-Reduplicant faithfulness, a dissociation of faithfulness to timing slots from faithfulness to root nodes, and parsing of the string into base and reduplicant based on timing slots rather than segments/root nodes, but otherwise employs completely standard technology. The rankings necessary for this analysis are summarized in (40), and the full analysis is illustrated in the summary tableau in (41) below.

(40)   Summary ranking for $C_1\bar{e}C_2$ pattern
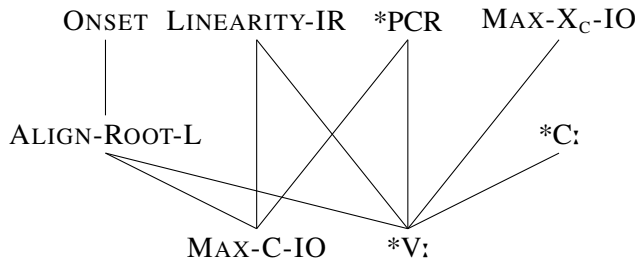
(41) Summary tableau for the $C_1\bar{e}C_2$ pattern

/RED, sep, úr/	*PCR	LINEARITY-IR	MAX-X$_c$-IO	ONSET	ALIGN-ROOT-L	*C:	*V:	MAX-C-IO
a. <u>se</u>-sp-úr	*!				**			
b. <u>pe</u>-sp-úr		*!			**			
c. <u>se</u>-p-úr			*!		**			*
d. <u>se</u>-s-úr			*!		**			*
e. <u>pe</u>-p-úr		*!	*!		**			*
f. <u>pe</u>-s-úr		*!	*!		**			*
g. <u>pe</u>-ːp-úr		*!			**		*	*
h. ☞ <u>se</u>-ːp-úr					**		*	*
i. <u>se</u>-pː-úr					**	*!		*
j. <u>se</u>-sː-úr					**	*!		*
k. <u>e</u>-sp-úr				*!	*			
l. <u>spe</u>-sp-úr		*!			***			
m. <u>sep</u>-sp-úr					***!			
n. <u>sepi</u>-sp-úr					***!*			

5.5.3.3 Cluster-Initial Roots within the Parallel Analysis

It is now necessary to confirm that the constraints and rankings which we require to generate the $C_1\bar{e}C_2$ pattern in parallel are consistent with the analysis of the basic reduplication patterns (see again Sections 5.2–5.3). Recall that TRVX– roots show C_1-copying, for example <u>pa</u>-prāc^h-a, as do TeR roots that form zero-grade -TR- stems, for example $\sqrt{par} \rightarrow$ <u>pa</u>-pr-ur; STVX– roots, on the other hand, show C_2-copying, for example <u>ta</u>-stamb^h-a. The ranking we used in order to derive the division within the cluster-initial roots is shown in (42) (repeated from (8) above).

(42) Ranking for Sanskrit cluster-initial reduplication

ONSET   *PCR

ALIGN-ROOT-L

ANCHOR-L-BR

CONTIGUITY-BR
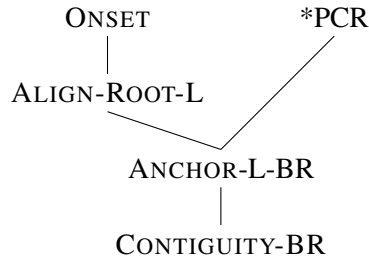
    First of all, when comparing this ranking to the ranking in (40) developed to account for the $C_1\bar{e}C_2$ pattern, we see that, for all constraints which are included in both — namely, *PCR, ONSET, and ALIGN-ROOT-L — there are no ranking contradictions. This is a good start. Among the constraints in (42), the two which were not explicitly discussed with regards to the $C_1\bar{e}C_2$ pattern are the Base-Reduplicant faithfulness constraints ANCHOR-L-BR and CONTIGUITY-BR. These were

(43)   Integrated ranking for $C_1\bar{e}C_2$ pattern and basic pattern

ONSET LINEARITY-IR *PCR MAX-X_C-IO

ALIGN-ROOT-L   *C:

MAX-C-IO   *V:

ANCHOR-L-BR

CONTIGUITY-BR
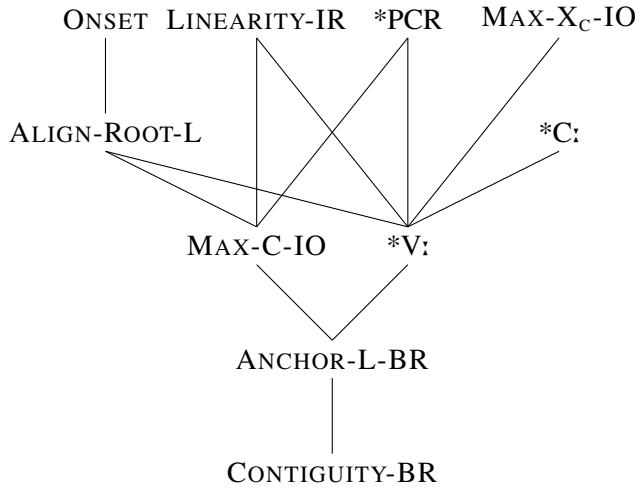
omitted because BR-faithfulness seems to play no role in the determination of the $C_1\bar{e}C_2$ pattern, as the base and reduplicant are all but completely unfaithful in that case. Nevertheless, my interpretation of optimal form [se-:p-úr] is consistent with satisfaction of both of these constraints.

CONTIGUITY-BR is vacuously satisfied, because there is only one element in the reduplicant that has a correspondent in the base: the reduplicant-initial segment [s] (or, perhaps more properly, the output timing slot associated with reduplicant-initial [s]) stands in correspondence with the base-initial timing slot (which derives from the timing slot of underlying root-initial /s/). The BR-correspondence that holds between these two elements will also be sufficient to satisfy ANCHOR-L-BR, because both of these elements are leftmost in their respective strings. Therefore, both CONTIGUITY-BR and ANCHOR-L-BR are satisfied in the derivation of the $C_1\bar{e}C_2$ pattern.

Since the optimal forms for TRVX– roots and STVX– roots respectively violate these two constraints, we now need to make sure that there exists an integrated ranking that doesn't predict the $C_1\bar{e}C_2$ to be an optimal means of avoiding such violations. The way to ensure this is to rank the constraints that the $C_1\bar{e}C_2$ mapping violates — namely, *V: and MAX-C-IO — above the BR-faithfulness constraints. This is shown in the ranking in (43) below, which is simply the ranking for the $C_1\bar{e}C_2$ pattern from (40) above with the ranking fragment ANCHOR-L-BR ≫ CONTIGUITY-BR placed below its two lowest ranked constraints. (Note that, strictly speaking, only one of *V: and MAX-C-IO needs to outrank ANCHOR-L-BR.)

To demonstrate that this ranking is sufficient, we first need to think about what candidates we ought to be comparing. First let us think about derivations involving full-grade forms, that is, those where ablaut does not advocate for deleting the root vowel. (In the tableaux in (44) and (45), this is indicated by using the Sanskrit third singular suffix /-a/, as this is a full-grade category.) To any C_1C_2eX– root, the candidate which is precisely equivalent to the $C_1\bar{e}C_2$ mapping is one which deletes root-C_1 and leaves behind its timing slot to compensatorily lengthen the reduplicant vowel: /RED, C_1C_2eX–/ → [$C_{1/2}$e-:C_2eX–]. Whether underlying root-C_1 or underlying root-C_2 is the consonant which is copied into the reduplicant, such a candidate would not violate any of the undominated constraints in (43). Regardless of which consonant is copied, it would though violate *V: and MAX-C-IO. Additionally, depending on which consonant is copied, it would violate one of the low-ranked BR-faithfulness constraints: ANCHOR-L-BR for root-C_2, CONTIGUITY-BR for root-C_1.

For the TRVX– roots in (44), these deletion + compensatory lengthening candidates are thus both harmonically bounded the candidate which copies the same consonant but does not show

207

deletion — (44a) harmonically bounds (44c), and (44b) harmonically bounds (44d) — because they too violate the low-ranked BR-faithfulness constraints but do not violate the deletion-related constraints. These candidate comparisons, then, do not provide any ranking arguments

(44) TRVX– C_1-copying

/RED, TReC, a/	*PCR	MAX-X$_c$-IO	ALIGN-ROOT-L	*Vː	MAX-C-IO	ANCHOR-L-BR	CONTIG-BR
a. ☞ Te-TReC-a			**				*
b. Re-TReC-a			**			*!	
c. Te-ːReC-a			**	*!	*!		*
d. Re-ːReC-a			**	*!	*!	*	
e. TRe-ːC-a	*!	***!	*	**			

On the other hand, this harmonic bounding relationship is disrupted for STVX– roots, because the C_1-copying candidate now has a *PCR violation which is alleviated by deletion. That is to say, in the tableau in (45) below, *PCR rules out candidate (45a), freeing the C_1-copying + deletion candidate [Se-ːTeC-a] (45c) from its harmonic bounding. The C_2-copying + deletion candidate [Te-ːTeC-a] (45d), however, remains harmonically bounded by the candidate that just exhibits C_2-copying, [Te-STeC-a] (45b). Therefore, the comparison between candidate (45b) [Te-STeC-a] and candidate (45c) [Se-ːTeC-a] provides the ranking arguments for the lower-ranked constraints: since [Te-STeC-a] (45b) ≻ [Se-ːTeC-a] (45c), (at least one of) *Vː and MAX-C-IO must dominate ANCHOR-L-BR (and CONTIGUITY-BR, by transitivity).

(45) STVX– C_2-copying

/RED, STeC, a/	*PCR	MAX-X$_c$-IO	ALIGN-ROOT-L	*Vː	MAX-C-IO	ANCHOR-L-BR	CONTIG-BR
a. Se-STeC-a	*!		**				*
b. ☞ Te-STeC-a			**		*		
c. Se-ːTeC-a			**	*!	*!		*
d. Te-ːTeC-a			**	*!	*!	*	
e. STe-ːC-a		*!	***!	*	**		

This confirms that the ranking in (43) is sufficient to generate the $C_1\bar{e}C_2$ pattern alongside the patterns for cluster-initial roots.

5.5.4 The $C_1C_2\bar{e}C_3$ Pattern?

Both of the preceding tableaux included an additional candidate which I have not yet mentioned: the (e) candidate [C_1C_2e-ːC$_3$-a]. This is a candidate which even more closely resembles the $C_1\bar{e}C_2$ pattern, in that its overall string retains the root's structure, but with long [eː] replacing short [e]. However, it actually arises through a somewhat different sort of mapping. In order to achieve this structure with a cluster-initial root, it would have to be the case that the entire root-initial cluster is copied *into the reduplicant*. This alone incurs an extra violation of ALIGN-ROOT-L which would

be fatal. But in addition, it is not one but two consonants that have to be deleted in order to clear the way for the reduplicant vowel + compensatory timing slot to end up adjacent to the underlying post-nuclear string, which is what is required in order to recapitulate the root structure. Not only does this incur an extra violation of Max-C-IO, it also incurs a violation of the more significant constraint Max-X_C-IO, because one of the two deleted consonants has not re-associated its timing slot. Furthermore, since this was meant to be a full-grade category, there shouldn't be anything triggering the deletion of the root vowel, which is also a necessary component of recapitulating the root structure.

But what if we actually were to consider a zero-grade category? If the root contained a vocalizable sonorant — for example, $\sqrt{sparç}$ 'touch' → zero-grade perfect *pa-spr̥ç-ḗ* — then there would be no phonotactic problem with C_2-copying and we should expect the C_2-copying form to surface. But if the root did not contain a vocalizable sonorant, then there is a phonotactic problem. Take, for example, $\sqrt{b^h ram}$ 'wander' (Whitney 1885 [1988]:115). While it attests a few zero-grade derivatives in b^hrm- in Vedic, its verbal system as attested starting with Epic Sanskrit has exclusively full or lengthened grade forms, indicating that neither the /r/ nor the /m/ were vocalizable by this point. If, then, zero-grade ablaut were to apply to the root, when accompanied by reduplication, we would have **ba-b^hrm-*úr*, with a phonotactically illicit -b^hrm- sequence.

One response would be to simply not apply zero-grade ablaut. Such an output is attested: *ba*-b^hram-*úr*. The alternative, however, is to apply zero-grade ablaut and let other markedness and faithfulness constraints interact to resolve the resulting marked structure. This root does appear to have such an output: b^hrēm-*úr*, closely mirroring the C_1ēC_2 pattern. It turns out that, if we were to supplement the ranking developed thus far with two new phonotactic constraints, we come fairly close to deriving this output directly.

The first relevant phonotactic markedness constraint would be *CRC, which bans medial three consonant sequences with a non-syllabic sonorant in the middle (the last C probably has to be specified as [-continuant]). When ranked above Max-X_C-IO and Align-Root-L, this rules out the non-deletion candidates (46a) and (46b). The second relevant phonotactic markedness constraint would be *V:RC, which bans long vowel + sonorant sequences in pre-consonantal position; this is the same Osthoff's Law constraint used for Germanic in Chapter 4. When ranked high, this constraint rules out any candidate that deletes just the first root consonant (46c–e).[23] This means that any optimal candidate will indeed have to delete the first two root consonants and compensatorily lengthen the reduplicant vowel. However, it seems as though this comes up just short of predicting the b^hrēm-*úr* form. The phonotactic markedness constraints are equally well satisfied by copying both members of the underlying root-initial cluster (desired candidate (46f)) as by copying just one (undesired candidate (46g)). Align-Root-L will favor the one with a short reduplicant, and thus select (46g).

---

[23] *me*-:rm-*úr* would be ruled out independently by Linearity-IR. *re*-:rm-*úr* would be ruled out independently by the final version of *PCR that will be motivated in Chapter 6. It is therefore only b^he-:rm-*úr* which is probative here.

(46) Towards a [C₁C₂e-ːC-] form

/RED, bʰrem, úr/	*CRC	*V:RC	MAX-Xc-IO	ALIGN-ROOT-L	*V:	MAX-C-IO	ANCHOR-L-BR	CONTIG-BR
a. <u>be</u>-bʰrm-úr	*!			**				*
b. <u>re</u>-bʰrm-úr	*!			**			*	
c. <u>bʰe</u>-ːrm-úr		*!		**	*	*		*
d. <u>re</u>-ːrm-úr		*!		**	*	*	*	
e. <u>me</u>-ːrm-úr		*!		**	*	*	*	
f. ☺ <u>bʰre</u>-ːm-úr			*	***!	*	**		
g. 💣 <u>bʰe</u>-ːm-úr			*	**	*	**		*

We can notice that *bʰe-ːm-úr* has no surface realization of the underlying /r/. If we could invoke a constraint that requires all underlying consonantal featural material to have a correspondent *somewhere in the output*, i.e. in at least one of the base and reduplicant, then this constraint would penalize *bʰe-ːm-úr* but not *bʰre-ːm-úr*. Such a constraint would be available if we adopted an approach to reduplication akin to Spaelti (1997) or Struijke (2000), where reduplicant and output root are subsumed under Input-Output faithfulness and the output root is subject to an additional faithfulness relation to the input root. However, I believe that such an approach would deprive the analysis of anything equivalent to LINEARITY-IR, which uniquely penalizes unfaithful ordering within the reduplicant.

This whole discussion, however, is somewhat of a wild goose chase. As will be shown immediately below, considerations relating to vowel quality make it highly unlikely that the phonologically driven allomorphy approach developed here is to be ascribed to the stage of attested Sanskrit; rather, that stage reflects the phonologically conditioned allomorphy approach developed in Section 5.4. Furthermore, the only *C₁C₂e-ːC₃*- forms that are attested in Sanskrit are very clearly late, secondary developments. Beside *bʰrēm-úr*, there is only *çrēm-úr* to √*çram* 'be weary' (Whitney 1885 [1988]:178) and *trēs-úr* to √*tras* 'be terrified' (Whitney 1885 [1988]:67).[24] In each of these cases, the forms with *-ē-* are not attested until after the Vedic period; and each root also attests earlier forms that exhibit lack of zero-grade (i.e. unexpected full-grade [a]): <u>ba</u>-bʰram-úr, <u>ça</u>-çram-úr, <u>ta</u>-tras-úr. The forms *bʰrēm-úr*, *çrēm-úr*, and *trēs-úr* thus very clearly come about through "analogical extension" of the C₁ēC₂ pattern. Sandell (2015b:Ch. 8), formalizing analogy of this sort using the Minimal Generalization Learner (MGL; Albright & Hayes 1999, Albright 2002a,b, *et seq.*), correctly predicts that these roots should develop a perfect weak stem in [-ē-],

---

[24] There are additionally several grammarian-cited *C₁C₂eːC₃*- forms, provided below. The linguistic reality of these forms is unclear.

(i) Grammarian-cited *C₁C₂eːC₃*- forms

bʰrēj-úr	←	√*bʰrāj* 'shine'	(Whitney 1885 [1988]:115)
çrētʰ-úr	←	√*çratʰ* 'slacken'	(Whitney 1885 [1988]:177)
grētʰ-úr	←	√*gra(n)tʰ* 'tie'	(Whitney 1885 [1988]:39)
svēn-úr	←	√*svan* 'sound'	(Whitney 1885 [1988]:201)
trēp-úr	←	√*trap* 'be abashed'	(Whitney 1885 [1988]:67)

based on the similarity of their structure to other roots that developed a $C_1\bar{e}C_2$ form through regular phonological means.[25]

Putting this all together, if the only $C_1C_2\bar{e}C_3$- forms are first attested late in the internal history of Sanskrit, and the phonologically driven allomorphy approach can only hold of a period prior to the first stages of attested Sanskrit, then the phonologically driven allomorphy analysis ought not to generate C_1C_2e-:C_3- forms. It does not, so this is a correct result. Furthermore, since these roots and others like it very frequently have early attested perfect weak stems that fail to show zero-grade, we can conclude that, at least by the time of attested Sanskrit, phonotactic constraints of the sort discussed in (46) dominate the constraint(s) that triggers zero-grade ablaut; that is to say, ablaut is blocked when it would result in a phonotactically illicit string that cannot be repaired except by deleting multiple consonants (or, more properly, by deleting a consonant whose timing slot cannot be re-associated).

5.5.5   Local Summary

This section has shown that the $C_1\bar{e}C_2$ pattern in Sanskrit can be generated directly in the phonology (i.e. as phonologically *driven* allomorphy), not just through allomorph selection (i.e. as phonologically *conditioned* allomorphy). While certain aspects of the phonological derivation are simpler with a serial phonological approach, a fully parallel OT approach does yield a successful analysis of the pattern, and one which is indeed consistent with the reduplication patterns of cluster-initial roots. Regardless of the precise analysis, however, the two sets of patterns can only be made consistent if the grammar has access to the *PCR constraint.

5.6   Vowel Quality and the $C_1\bar{e}C_2$ Pattern in Sanskrit

In the previous section, I made one major simplifying assumption, namely, that the vowels involved in the $C_1\bar{e}C_2$ pattern were all [e], either underlyingly or intermediately. This allowed the compensatory lengthening to yield [eː] straightforwardly. This cannot hold for any synchronic state of attested Sanskrit, as Proto-Indo-European */e/ has merged with */o/ and */a/ as (roughly) /a/ already in Proto-Indo-Iranian. That is to say, there is no short [e] at any level in Sanskrit. The relationship in Sanskrit terms between the root vowel and the vowel of the $C_1\bar{e}C_2$ pattern is thus [a] $\sim$ [eː]. Following arguments by Sandell (2017), I show briefly in this section that this [a] $\sim$ [eː] relationship cannot be one of transparent lengthening in Sanskrit, as /a/ lengthens to [aː] not [eː]. While this seems to rule out the possibility of a transparent phonological analysis for synchronic Sanskrit, it does not preclude ascribing such an analysis to earlier stages in the development into Sanskrit, or other languages which display similar patterns (namely Germanic), which still retain Indo-European */e/ as such.

In Sanskrit, there is both phonological and morphophonological evidence that the lengthened variant of /a/ is [aː] not [eː]. The best phonological evidence comes from a pattern of rhythmic lengthening adduced by Insler (1997): underlying sequences of three light syllables (open syllables with short vowels) are repaired by lengthening one of the vowels in that sequence. For example, the nominal root /vrka-/ 'wolf', with underlyingly short thematic vowel /a/, builds a denom-

---

inative verb with the suffix /-ya-/, meaning 'act like a wolf'. When the third singular present ending /-ti/ is added, the thematic vowel lengthens: /vrk**a**-ya-ti/ → [vr̩k**aː**-ya-ti] (example from Sandell 2017:10).

That this is a synchronically regular phonological lengthening can be seen especially from intra-paradigmatic vowel-length alternations where some forms meet the three-light-syllable condition while others do not. For example, the denominative to the root /rta-/ 'truth' shows lengthening in the present participle [r̩t**a**ː-ya-teː], where there would otherwise be three consecutive light syllables (**[r̩t**a**-ya-teː]), but no lengthening in the third plural [r̩t**a**-ya-nta], where the suffix yields a heavy syllable in third position and thus obviates the need for repair (examples from Insler 1997:105, cited by Sandell 2017:10).[26] Equivalent alternations are also found in compounds. These vowel lengthening facts comport with morphophonological ablaut relations, where the lengthened grade of /a/ is very clearly [aː] as well.

Therefore, within Sanskrit, it seems that a purely phonological analysis of the $C_1\bar{e}C_2$ pattern based on compensatory lengthening is probably not sustainable, as concluded by Sandell 2017. One could claim that (morpho)phonological lengthening and compensatory lengthening result in different outcomes with respect to vowel quality, but this would be a largely *ad hoc* solution. Nevertheless, there does seem to be some indication that such a distinction held at least at an earlier stage of the language. Sandell (2014a) argues that the regular, unconditioned diachronic outcome of cases which can be described as compensatory lengthening of Proto-Indo-Iranian */a/ (beside the $C_1\bar{e}C_2$ forms, this largely amounts to loss of voiced sibilants in coda position) is indeed Sanskrit [eː]. Modulo the alternative suggested immediately below, it is necessary that *at some point* (after the merger of *$\bar{e}$ and *$\bar{a}$) compensatory lengthening yielded [eː], or else there would be no explanation for the vowel quality we actually observe.

One other solution could be to claim that, for synchronic Sanskrit, the [eː] is not derived by consonant deletion + compensatory lengthening, but rather by changing the root-initial consonant to //y//, with attendant coalescence with reduplicant vowel //a// yielding [eː], in accordance with the regular treatment of tautosyllabic /-ay-/. This approach would require that both Input-Output and Base-Reduplicant faithfulness constraints on consonant features (for example, IDENT[±sonorant], IDENT[place], etc.) are ranked quite low, since the form would involve a mapping from an underlying /s/ (or indeed any number of other consonants; see Chapter 6) to surface [y] (or rather, intermediate *y*, which "subsequently" coalesces), and likewise correspondence between reduplicant [s] and base [y].

It might be possible that a ranking could be constructed such that the relatively low ranking of these faithfulness constraints does not have any deleterious effects, either in reduplication or in the language generally; but it seems likely that this would predict unfaithful mappings in unwanted corners of the grammar. Such a solution would therefore require significant additional verification. I leave this as a question for future investigation. Short of the feature change + coalescence solution, it seems then that we must adopt the phonologically conditioned allomorphy analysis developed in Section 5.4 as the means of deriving the $C_1\bar{e}C_2$ pattern within synchronic Sanskrit.

In any event, this feature change plus coalescence solution would be restricted to Sanskrit (or perhaps Indo-Iranian more generally); it would not be applicable to any other stages or related languages where the coalescence rule was not in effect. In the next section, I show that Germanic has the same pattern, and thus requires the compensatory lengthening analysis

---

[26] There are some instances of lengthening where it seems not be necessary; for example, [r̩t**a**ː-ya-n] exists beside expected [r̩t**a**-ya-n], and there are also forms like [r̩t**a**ː-ya-nt-iːs] (Insler 1997:106). Insler indicates that such forms exist mostly within paradigms where lengthening is motivated in other forms, and thus this can be viewed as paradigm uniformity effect (over-application of lengthening). It does not seem that there are any instances of under-application of lengthening.

5.7 The Development of the $C_1\bar{e}C_2$ Pattern in Germanic

5.7.1 The Evidence from Germanic

In Chapter 4, I developed an analysis of the strong preterite system of Gothic and Proto-Germanic. Within this system, we find what appears to be a phenomenologically identical $C_1\bar{e}C_2$ pattern in the Class IV–V preterite plurals: for example, Gothic $\sqrt{sit}$ 'sit' → [seːt-un] 'they sat'.[27] The long vowel in these forms is diachronically unexpected.[28]

　　These forms ought to come from Proto-Indo-European zero-grade reduplicated perfects, in the same way the preterite plurals of Class I–III unquestionably do; for example, Gothic Class II preterite plural [kus-] (← $\sqrt{kiws}$) comes from PIE *$\acute{g}e$-$\acute{g}us$- (← *$\sqrt{\acute{g}ews}$), with the only substantive change being the loss of the reduplicant. If we were to apply the same chain of developments to a Class V root like $\sqrt{sit}$, we would expect the following: PIE *$\sqrt{sed}$ → *se-zd- (> post-Grimm's Law Germanic *$\sqrt{set}$ → *se-st-) > Gothic **st- (e.g. third plural **st-un). This makes it clear that, whatever the diachronic process is that results in the loss of reduplication in other parts of the system (e.g. Class I–III preterite plurals), this cannot be the story for the Class V preterite plurals.

　　Rather, these forms should be seen as deriving from the *PCR-driven deletion + compensatory lengthening that I proposed above to account for the equivalent $C_1\bar{e}C_2$ forms in Sanskrit. That is to say, the Class V preterite plurals never underwent "de-reduplication", but rather lost their overtly reduplicated character because of the set of deletions that applied to reduplicated CeC roots in zero-grade categories. The table in (47) below repeats the Class V data from (3), augmented with the proposed derivationally intermediate / diachronically prior sources for the individual forms.

(47)　　Gothic Class V preterite plurals (forms from Lambdin 2006:51)

		Infinitive		Preterite Plural (1PL)		
'sit'	*sitan*	[sit-an]	*setum*	[seːt-um]	(as if from *se-st-um)
'give'	*giban*	[gib-an]	*gebum*	[geːb-um]	(as if from *ge-gb-um)
'say'	*qiþan*	[kʷiθ-an]	*qeþum*	[kʷeːθ-um]	(as if from *$\underline{k^w e}$-$k^w \theta$-um)
'heap up'	*rikan*	[rik-an]	*rekum*	[reːk-um]	(as if from *re-rk-um)
'be saved'	*nisan*	[nis-an]	*nesum*	[neːs-um]	(as if from *ne-ns-um)

---

[27] Given the *PCR-based explanation of the Germanic $C_1\bar{e}C_2$ forms to be proposed immediately below, it would seem that only a subset of Class IV roots can have originally participated in these changes. Namely, Class IV roots that had a stop in first position (i.e. TeR), would not have violated *PCR under any of the formulations proposed in this dissertation, since its repetition would involve a stop before a sonorant: $\underline{Te}$-TR-. Furthermore, any roots which ultimately pattern with Class IV or Class V that have two pre-vocalic consonants, e.g. Class V *fraihnan* (see Appendix I in Chapter 4 for additional examples), probably could not have participated in the original change, for the same reason.

[28] Jasanoff (2012), and several works cited therein, take a partially contrary position. Jasanoff argues that many or all of the $C_1\bar{e}C_2$ forms that comprise the Class IV–V preterite plurals ought to instead be traced back to imperfects of roots with an underlying long vowel. (This point of view is often referred to as the "Narten theory".) It is, however, quite suspicious that (almost) all of the relevant roots were of the shape *CeC in PIE, despite the fact that a number of other root shapes were in general possible (consult generally Rix et al. 2001); that is to say, this seems to be missing a phonological generalization.

　　Ultimately, my analysis is completely consistent with a claim that *some* of the Class IV–V preterite plurals go back to formations with underlying long vowels, as long as these exist beside long vowel forms derived from reduplication in the manner laid out in this chapter. In fact, having some set of pre-existing long vowel forms within the category would likely have a positive effect in the driving the changes I will advocate for below.

In comparison to Sanskrit, the analytical boon that the Germanic forms provide is that their [eː] can easily be identified as the expected result of compensatory lengthening. These forms must be part of the Germanic verbal system by the time the contrast-based stem-formation system of Pre-Proto-Germanic is in place (see again Chapter 4). In this Pre-Proto-Germanic stage (and all prior stages), the root vowel and the reduplicant vowel in the preterites/perfects of such forms remained *e*, just as in Proto-Indo-European. Therefore, the schematic/anachronistic analysis in Section 5.5.3 that used *e* across the board, which could not be properly attributed to Sanskrit, can indeed be attributed to the Germanic pattern.[29]

As alluded to in the introduction, the [eː] in Gothic/(Pre-)Proto-Germanic cannot directly match the [eː] in Sanskrit in terms of diachronic correspondence. Proto-Indo-European *ē̄* yields Germanic *ē* but Sanskrit *ā*. The $C_1\bar{e}C_2$ pattern aside, Sanskrit *ē* comes only from monophthongization of tautosyllabic PIE *ey*, *oy*, *ay*, and probably from Indo-Iranian pre-consonantal *az* sequences resulting from voicing assimilation (see Sandell 2014a for recent discussion). Therefore, there is no way in which the $C_1\bar{e}C_2$ forms in Germanic and Sanskrit can be traced back to static PIE forms via regular sound change alone.

However, it is, in theory, possible that they could reflect the same *process* operative in PIE, with changes in vowel quality resulting from independent changes in the grammars of the two branches after they diverged. More conservatively, we can at least frame this as parallel independent developments based on similar circumstances: as *PCR became more significant in the two languages, they were both forced to deal with the problem posed by the zero-grade perfects/preterites of CeC/CaC roots, and the properties that the two languages still had in common as a result of their shared inheritance led them to solve the problem in an equivalent way. I leave further questions of the relatedness and chronology of these patterns as a topic for future work.

5.7.2 The $C_1\bar{e}C_2$ Pattern and Grammar Change in Germanic

The evidence thus seems to point to the Germanic Class V preterite plurals reflecting the same (or a very similar) pattern as the Sanskrit $C_1\bar{e}C_2$ weak perfects. I've argued in this chapter that Sanskrit comes to encode this pattern with phonologically conditioned allomorphy, set within a system whose primary mode of morphophonological marking is still reduplication. In Chapter 4, I argued that Germanic developed into a significantly different sort of system. While the opacifying effect of later sound changes may likewise have driven Gothic to a phonologically conditioned allomorphy system, Pre-Proto-Germanic developed a complex system of morphophonological marking based on stem contrast, where reduplication played a peripheral role rather than a central one. Instead, the Pre-Proto-Germanic system is centered on vocalic alternations of exactly the sort represented by the $C_1\bar{e}C_2$ forms.

This raises an important question: why did Sanskrit retain reduplication as its core means of marking the perfect while Germanic changed so drastically? Put another way, why and how did the PIE system, characterized by obligatory reduplication and accentually-conditioned vowel alternations, develop into that of Pre-Proto-Germanic, where reduplication and vowel alternations complementarily effect stem contrast? I believe that it was a change in the accentual system of Germanic — which was not mirrored by any development in Sanskrit — that triggered this systemic change; specifically, it was the way that the accentual change effected the interpretation of the $C_1\bar{e}C_2$ forms that drove the sweeping changes in Germanic.

---

[29] Note that the rankings constructed to generate the $C_1\bar{e}C_2$ pattern in Section 5.5.3 were built up independently of the other rankings for Sanskrit, and therefore are not contingent on any aspect of Sanskrit phonology. I believe that this ranking is consistent with the phonology of (Pre-)Pre-Proto-Germanic, but this deserves explicit verification in the future.

As discussed briefly in Chapter 1 and at various points throughout the dissertation, the Proto-Indo-European perfect is reconstructed as having essentially the same properties as those exhibited by Sanskrit. First, it used reduplication as the morphophonological exponent of PERFECT, with the target shape for the reduplicant being CV. Second, differences in the accentual properties of suffixes triggered vocalic alternations in the root; specifically, accented endings in the perfect active plural and the perfect middle triggered vowel deletion (zero-grade ablaut). This is recorded in (48a) below.

At some point, whether within Proto-Indo-European or at some stage between PIE and early Germanic, the *PCR-driven consonant-deletion process comes into effect. The details of this rule are repeated in (48b) below. For expository purposes, we can view these two rules as standing in the feeding relationship shown in (49), where vowel-deletion feeds consonant-deletion (cf. the ordered rule-based analysis of the $C_1\bar{e}C_2$ pattern in Section 5.5.1). When these rules are applied to a form which has already undergone reduplication (i.e. reduplication feeds vowel-deletion), we derive the $C_1\bar{e}C_2$ pattern, as demonstrated again in (49). (Note that I employ the PIE third plural suffix *-$\acute{r}$, which yields directly Sanskrit -$\acute{u}r$. It is replaced in Germanic by -un.)

(48)   a.   *Zero-grade ablaut*: Deletion of pre-tonic root vowels
        b.   **PCR-driven consonant-deletion + compensatory lengthening*: $C_\alpha V C_\alpha C \rightarrow C_\alpha V{:}C$

(49)   PIE /RED, set, $\acute{r}$/ $\xrightarrow[\textit{(copy CV)}]{\text{redup.}}$ *se-set-$\acute{r}$* $\xrightarrow[(e \rightarrow \emptyset / \_C_0 \acute{V})]{\text{zero-grade ablaut}}$ *sest$\acute{r}$* $\xrightarrow[(VC_\alpha \rightarrow V{:} / C_\alpha \_ C)]{\text{C-deletion + CL}}$ [se:t$\acute{r}$]

Since the vowel-deletion process transparently feeds the consonant-deletion + compensatory lengthening process, the morphological structure of such forms, i.e., as having underlying reduplication, could be recovered even if the surface forms did not look reduplicated. However, for learners to successfully make this recovery, they must also accurately learn the vowel deletion process. In the mobile accent system of PIE, the accentual conditioning environment of the deletion rule was itself (more or less) transparent, and thus learnable. Since Sanskrit retains the PIE mobile accent system, whatever the exact mode of analysis Sanskrit speakers used to encode the pattern, they still had access to the accentual information provided by the suffixes, and thus could use that to help recover the forms' relationship (direct or indirect) to reduplication. Without these accentual conditions, however, there would be no way to accurately learn when to apply the rule, and thus, potentially, no way to recover underlying reduplication.

Between PIE and Pre-Proto-Germanic, there was a complete remodeling of the accentual system, as illustrated below in (50): mobile, lexically-determined accentuation is replaced with phonologically-fixed initial stress (cf. Halle 1997). This change in accentuation eliminated the conditioning environment for the vowel-deletion rule. That is to say, the step from intermediate *ˈse-set-un* (previously *se-set-ún*) to intermediate *ˈse-st-un* (previously *se-st-ún*) is rendered opaque, as schematized in (51) below.[30]

---

[30] Some scholars believe that fixed stress landed on the root rather than the reduplicative prefix in (Pre-)Proto-Germanic. This question is actually tangential to the point being made here. All that is required for the current account is that stress/accent is no longer mobile, and that there is now a consistent accentuation pattern across the singular and the plural; the position of that fixed accent is not significant. Cf. Fullerton (1991) for an account that does posit accentual differences between singular and plural.

(50) PIE Mobile Pitch Accent ⇒ Proto-Germanic Fixed Stress Accent
  a. PIE *[pəh₂tér]   >   PGmc. *[ˈfaðer]   'father (NOMINATIVE)'
  b. PIE *[pəh₂tr-ós] >   PGmc. *[ˈfaðraz]   'father (GENITIVE)'
  c. PIE *[póːt-s]   >>   PGmc. *[ˈfoːtiz]   'foot (NOMINATIVE)'
  d. PIE *[ped-ós]   >>   PGmc. *[ˈfoːtijaz] 'foot (GENITIVE)'
  e. PIE *[h₁és-ti]   >   PGmc. *[ˈesti]   'is'
  f. PIE *[h₁s-énti] >   PGmc. *[ˈsinti]   'are'

(51) PGmc. /RED, set, un/ $\xrightarrow[\text{stress}]{\text{redup. +}}$ $\boxed{\textit{ˈse-set-un} \xrightarrow[\text{???}]{\text{zero-grade ablaut}} \textit{ˈsestun}}$ $\xrightarrow[\text{+ CL}]{\text{C-deletion}}$ [ˈseːtun]

At the point of the Germanic stress shift, learners are thus now faced with an inexplicable alternation between the singular and the plural: CeC roots show reduplication in the singular — 3SG.PRET *[ˈse-sat] (< *[se-sót-e]), but vowel lengthening in the plural — 3PL.PRET [ˈseːt-un] (<< *[seːt-úr]). With no accentual distinctions (not of the right sort, at least) to lead them to posit a pre-tonic vowel deletion process, and no vowel deletion process to support a consonant deletion process, learners are (by hypothesis) unable to see that the forms could have a reduplicated origin, even though they are paired with reduplicated forms in the singular.

The singular form which we observe in Gothic is simply [ˈsat] without reduplication, not a diachronically expected reduplicated form **[ˈse-sat]. This unexpected lack of reduplication is representative of virtually the entire strong preterite system, as seen in Chapter 4 (the exception being just the phonologically-circumscribed case of Class VII). I suggest that the phonologically well-motivated disappearance of reduplication in the Class IV–V plurals is sufficient to trigger a general failure to learn that the preterite was a reduplicated category (i.e. that it had an underlying exponent /RED/).

In ongoing work with Ryan Sandell (an early version of which has been presented as Sandell & Zukoff 2017), we are developing a computational learning model which makes this proposal concrete. Preliminary results suggest that an agent-based multi-generational maximum entropy learning approach can generate the desired pathway of diachronic developments. I leave fuller exploration of this line of explanation for future work.

5.8 Conclusion

Sanskrit is perhaps unique among the attested Indo-European languages in the way it deals with potential *PCR violations in the perfect. Rather than showing an alternative copying pattern, like the C_2-copying pattern it employs for STVX– roots, roots that would make *PCR-violating clusters if they underwent zero-grade vowel-deletion instead show the apparently non-reduplicative $C_1\bar{e}C_2$ pattern. The $C_1\bar{e}C_2$ pattern can be derived in one of several ways. First, it could be analyzed as a case of phonologically *conditioned* allomorphy (as shown in Section 5.4), where the $C_1\bar{e}C_2$ forms are synchronically derived from a non-reduplicative morph of PERFECT. This is perhaps the most appropriate analysis for synchronic Sanskrit.

Alternatively, the $C_1\bar{e}C_2$ pattern can be treated as phonologically *driven* allomorphy approach (as shown in Section 5.5), where the $C_1\bar{e}C_2$ forms are actually the optimal mapping from an underlying representation with /RED/. While the phonological approach is simplest when cast in a derivational/serial phonological framework, because it allows for a more straightforward approach to compensatory lengthening, a parallel OT analysis is possible if we allow Input-Reduplicant faith-

fulness and faithfulness to timing slots. It undoubtedly must be the case that some version of the phonological approach was responsible for the origin of this pattern, as the allomorphy approach needs a diachronic source for the /-$\bar{e}$-/ morph that is uses to derive the $C_1\bar{e}C_2$ forms.

The $C_1\bar{e}C_2$ pattern in Sanskrit is matched almost exactly by the $C_1\bar{e}C_2$ pattern in the Germanic Strong Class V preterite plurals. The origin of this pattern can be explained in the very same way as that of Sanskrit. And furthermore, the [$\bar{e}$] vowel of the Germanic pattern can derive transparently from compensatory lengthening, as Germanic retain the Proto-Indo-European *e as such at the relevant stage. I proposed that it is exactly the $C_1\bar{e}C_2$ forms that drive the development of the contrast-based stem-formation system of Pre-Proto-Germanic proposed in Chapter 4. Specifically, the processes of zero-grade vowel-deletion and *PCR-driven consonant-deletion, which stood in a feeding order at the prior stage, are rendered opaque by the Germanic accent shift. This undermined learners' ability to construct a consistent grammar based on a reduplicated underlying representation of the preterite, leading the language ultimately down the path towards the dramatically different contrast-based system it eventually developed. The ongoing work represented by Sandell & Zukoff (2017) aims to show that this change is effectively pre-destined after the accent shift, using a computational model based on multi-generational learning.

There are potentially a few additional comparanda to the Sanskrit and Germanic $C_1\bar{e}C_2$ patterns elsewhere in Indo-European (see generally Schumacher 2005; see also Niepokuj 1997:148–164). The most robust one comes from Old Irish, which shows a number of "long vowel preterites", primarily built to CVC roots. Some examples are shown in (52) below.

(52) Old Irish $\bar{\imath}$-/$\bar{a}$-preterites (Thurneysen 1946 [1980]:429–430; cf. Niepokuj 1997:151–152)

	Present Stem		Preterite Stem		
'judge'	*mid-*	[mʲidʲ-]	*míd-*	[mʲiːdʲ-]	(as if from *_mi_-md- or *_me_-md-)
'fight'	*fich-*	[fʲixʲ-]	*fích-*	[fʲiːxʲ-]	(as if from *_fi_-fx- or *_fe_-fx-)
'fight'	*fid-*	[fʲidʲ-]	*fíd-*	[fʲiːdʲ-]	(as if from *_fi_-fd- or *_fe_-fd-)
'flee'	*tech-*	[tʲex-]	*tách-*	[taːx-]	(as if from *_ta_-tx-)
'serve'	*reth-*	[rʲeθ-]	*ráth-*	[raːθ-]	(as if from *_ra_-rθ-)
'weave'	*fig-*	[fʲigʲ-]	*fáig-*	[faːgʲ-]	(as if from *_fa_-fg-)

It is not clear (to me, at least) exactly what the origin of these forms is, and whether they have a unitary origin in the first place. Nevertheless, if Old Irish or its predecessors did inherit a grammar with similar properties as those of Sanskrit and Germanic, some or all of these are likely to be explainable as an effect of *PCR. Further work is needed to resolve this question.[31]

Lastly, Melchert (2016) has proposed that Hittite *šippand-* 'libate' is to be analyzed as resulting from *PCR-driven consonant-deletion. As mentioned in Chapter 3, Melchert argues that this root should be analyzed as deriving from an earlier reduplicated formation *_se_-spónd-, which then underwent deletion of root-C_1 — possibly with compensatory lengthening — driven by *PCR, yielding *se(ː)pónd-. (In pre-tonic position, both short *e and long *$\bar{e}$ would yield Hittite [i].) The details of whatever stage of the language generates this form are too unclear to formulate an analysis with any certainty; but it is clear that, if Melchert's (2016) interpretation is correct, the operation that produces this form is virtually identical to the one which creates the $C_1\bar{e}C_2$ pattern in Sanskrit and Germanic.

---

[31] Niepokuj (1997:155–162) also adduces similarities with reduplicated participles in Tocharian. I leave this a question for subsequent work.

As of yet, it is difficult to decide whether the process that results in $C_1\bar{e}C_2$ forms should be reconstructed to Proto-Indo-European. With just two strong comparanda, it is possible that the patterns arose independently as parallel innovations. If and when these additional comparanda are better understood and integrated into the larger picture, the answer to this question might start to come into better focus.

Chapter 6

The No Poorly-Cued Repetitions Constraint

6.1 Introduction

Throughout the preceding chapters, I have shown that the Indo-European languages consistently show distinct behavior for reduplicative bases (and would-be reduplicative bases) that begin in different sorts of consonant clusters. I have accounted for these facts with a markedness constraint that prohibits sequences of repeated consonants ($C_\alpha VC_\alpha$) in immediate pre-obstruent position ($/\_C_{[-sonorant]}$). While this constraint can essentially be stated as a contextual markedness constraint $*C_\alpha VC_\alpha / \_C_{[-son]}$, in anticipation of the arguments in this chapter, I have referred to this constraint as the NO POORLY-CUED REPETITIONS constraint, abbreviated *PCR.

The reason for generalizing beyond $*C_\alpha VC_\alpha / \_C_{[-son]}$ is that this context is not actually sufficient to capture the full range of facts when we consider the extended cluster-wise distributions of the default vs. alternative reduplication patterns found in the Indo-European languages. It is deficient in both directions: some types of consonant repetitions that are permitted are pre-obstruent, and some that are disallowed are pre-sonorant. For this reason, a more articulated theory of repetition avoidance is required. In this chapter, I argue that we can construct a constraint (or constraint family/schema) that has the precision required to capture the full range of distributions if the constraint can make reference to acoustic/auditory cues to particular consonantal contrasts (specifically the contrast between a consonant and its absence, i.e. C∼Ø), rather than traditional phonological features. The reason why a constraint referencing [-sonorant] as the context following the repetition can get us so far in the analysis is precisely because this is a locus for the lack of robust acoustic/auditory cues to those contrasts.

In this chapter, I will begin in Section 6.2 by reviewing the empirical motivation for positing a constraint against consonant repetitions in context, namely, the cluster-dependent reduplication patterns in the Indo-European languages discussed thus far, including the $C_1\bar{e}C_2$ patterns in Sanskrit and Germanic examined in Chapter 5. With the motivation for positing a constraint against consonant repetitions secured, I will then in Section 6.3 examine in much greater detail the full cluster-wise distributions of default vs. alternative reduplication strategies found in the various languages. This will delineate the dimensions of variation that the fully articulated *PCR constraint(s) will need to be sensitive to. In addition to the Indo-European languages discussed already, I will also adduce the facts of the non-Indo-European language Klamath (Barker 1964), which exhibits a reduplication pattern largely equivalent to that of Gothic and Proto-Anatolian, but with a much more robust

inventory of root-initial clusters. Despite the much richer cluster inventory, the set of consonant repetitions it permits is identical to the set permitted in Ancient Greek.

In Section 6.4, I argue that the properties which are crucial to explaining the full cluster-wise distributions of repetition avoidance effects are not phonological features or an abstract notion of sonority, but rather *acoustic/auditory cues* to particular consonantal contrasts. Motivated by the empirical evidence, I propose the POORLY-CUED REPETITIONS HYPOTHESIS — previewed in (1) — which states that repetition imposes distinct burdens on the perceptual system with regards to the licensing of contrasts (specifically the C∼Ø contrasts). Based upon this hypothesis, I formulate the NO POORLY-CUED REPETITIONS constraint (*PCR) — previewed in (2) — which penalizes repetitions that would leave the repeated consonant without sufficient cues to its contrast with Ø. Specifically, I will show that *intensity rise*, alongside *transitions* and (perhaps) *stop release burst*, are central to licensing these contrasts under repetition. The different languages, and the different stages of the different languages, in effect select from among these cues which ones will be sufficient, either on their own or in combination, to license a consonant repetition. The inclusion of this constraint in the grammar properly allows for the patterns of repetition avoidance instantiated by the languages examined in this dissertation.

(1) **THE POORLY-CUED REPETITIONS HYPOTHESIS**
     There is some property of the perceptual system which degrades listeners' ability to appre-
     hend the presence of a consonant (i.e. the contrast between that consonant and its absence)
     when that consonant is adjacent to an identical consonant.

     i.   This property diminishes the effectiveness of some or all acoustic/auditory cues to C∼Ø
          contrasts, such that some cues which are normally sufficient to license those C∼Ø
          contrasts (in otherwise equivalent positions) are no longer sufficient to license those
          contrasts under repetition.
     ii. This property diminishes the effectiveness of different cues to different extents:
          the effectiveness of cues to acoustic events which are more difficult to anchor at a
          particular point in the speech stream and/or tend to extend across multiple segments
          are diminished to a greater degree than cues to acoustic events which are more reliably
          located at their correct position in the speech stream.

(2) **The NO POORLY-CUED REPETITIONS constraint (*PCR)**
     Languages may set stricter conditions (in terms of cues) for the licensing of C∼Ø contrasts
     (i.e. the presence of C) when that C would be the second member of a transvocalic consonant
     repetition (i.e. C_α^2 in a $C_\alpha^1 V C_\alpha^2$ sequence) than in other contexts. Assign a violation mark *
     for each C_α^2 (i.e. each C∼Ø contrast where C is a C_α^2) which is not cued to the level required
     by the language-specific repetition licensing conditions.

In Sections 6.6 and 6.7, I adduce additional empirical and analytical evidence in favor of the use of the *PCR constraint. First I provide analyses of several additional reduplicative effects that require the use of *PCR. One is the infixal reduplication pattern found in Latin for ST-initial roots (Section 6.6.1), which shows that infixation is another potential solution to the *PCR problem. An infixal reduplication pattern is also found in the Sanskrit desiderative with vowel-initial roots (Section 6.6.2). Employing *PCR allows for a simple explanation of a distinction in the position of the reduplicative infix for roots with different types of post-vocalic clusters. An important subsidiary point made by this pattern is that a templatic syllable-alignment account, which otherwise could eschew the use of *PCR, appears to be inconsistent with the syllabification facts of the language; this argues for the *PCR approach, which requires additional reference to acoustic/auditory cues

in the form of cue-based faithfulness constraints. Lastly, I formalize the analysis of the alternation between C_1-copying and cluster-copying in Klamath (Section 6.6.3), in which *PCR is again crucial.

Then, I further consider several potential *PCR effects from outside of reduplication, including suffix allomorphy in Latin (Section 6.7.1) and an exception to the normal distribution of aspiration mobility in Sanskrit (Section 6.7.2). Lastly, I suggest that the facts of the English *sCVC constraint (cf. Fudge 1969) might ultimately find a *PCR-based analysis (Section 6.7.3). Since these effects can be explained in the same terms as the reduplicative effects examined throughout the dissertation, this serves as external evidence in favor of the *PCR approach.

6.2   Empirical Motivation for *PCR

In this section, I collect the empirical facts and analyses from the reduplicative systems of the Indo-European languages detailed earlier in this dissertation, which motivate positing a constraint against consonant repetitions in pre-obstruent position: $*C_\alpha VC_\alpha/\_C_{[-son]}$ (referred to in the individual chapters, and later in this chapter, as *PCR). These come from the consistent distinction between the treatment of, on the one hand, stop-sonorant–initial roots/bases (TRVX–), which show the default C_1-copying pattern, and, on the other hand, s-stop–initial roots/bases (STVX–), which show some distinct behavior, varying by language (and, in at least one case, by morphophonological category).

6.2.1   Ancient Greek: TRVX– C_1-copying, STVX– Non-copying

Ancient Greek shows a consistent distinction between the reduplicative behavior of TRVX– roots and STVX– roots in the perfect (see Chapter 2). Representative data is provided in (3) below. Roots beginning in TR clusters copy root-C_1 (3a). On the other hand, roots beginning in ST clusters copy nothing at all (3b).

(3)   Cluster-initial perfects in Ancient Greek (consult Steriade 1982, van de Laar 2000)

a.   **C_1-copying reduplication** ⇔ TRVX– roots

Root		Perfect Tense	
√*kri-*	'decide'	k̲-e-kri-	(not *_-e-kri-)
√*pneu-*	'breathe'	p̲-e-pnū-	(not *_-e-pnū-)
√*tla-*	'suffer, dare'	t̲-e-tlɛ̄-	(not *_-e-tlɛ̄-)
√*grapʰ-*	'write'	g̲-e-grapʰ-	(not *_-e-grapʰ-)

b.   **Non-copying "reduplication"** ⇔ STVX– roots

Root		Perfect Tense	
√*stel-*	'prepare'	_-e-stal-	(not *s̲-e-stal-)
√*sper-*	'sow'	_-e-spar-	(not *s̲-e-spar-)
√*skep-*	'view'	_-e-skep-	(not *s̲-e-skep-)

221

The basic analysis of this alternation is sketched in (4) and (5) below. When the consonant repetition created by C_1-copying would not be in pre-obstruent position, i.e. for TRVX– roots, the ONSET violation of the non-copying candidate (4b) is sufficient to eliminate it in favor of the C_1-copying candidate (4a). This pattern fails to emerge for STVX– roots. The reason is that, for such roots, the consonant repetition that would be created by C_1-copying *would* be in pre-obstruent position, as seen in candidate (5a). This configuration violates the proposed constraint $*C_\alpha VC_\alpha / \_C_{[-son]}$. If this constraint is ranked above ONSET, then the grammar will select non-copying in just this case. That is to say, the specter of a pre-obstruent consonant repetition is enough to divert the system away from the default reduplication pattern.[1]

(4)   Ancient Greek TRVX– C_1-copying: $\sqrt{kri\text{-}} \rightarrow$ κέκριμαι [k̲-é-kri-mai] 'I have (been) judged'

/RED, e, kri-/	$*C_\alpha VC_\alpha / \_C_{[-son]}$	ANCHOR-L-BR	*CC	ONSET
a. ☞ k̲-e-kri-			*	
b. _-e-kri-			*	*!
c. k̲r-e-kri-			**!	
d. r̲-e-kri-		*!	*	

(5)   Ancient Greek STVX– non-copying: $\sqrt{stal\text{-}} \rightarrow$ ἔσταλκα [é-stal-k-a] 'I have prepared'

/RED, e, stal-/	$*C_\alpha VC_\alpha / \_C_{[-son]}$	ANCHOR-L-BR	*CC	ONSET
a. s̲-e-stal-	*!		*	
b. ☞ _-e-stal-			*	*
c. s̲t-e-stal-			**!	
d. t̲-e-stal-		*!	*	

Ancient Greek shows the remnants of another pattern, Attic Reduplication, which is to be analyzed as the result of a similar repetition avoidance effect in Pre-Greek, though targeting different sorts of repeated consonants. This will be discussed in Section 6.3.2.6 below. (See Chapter 2 for the full synchronic and diachronic analysis.)

6.2.2   Gothic & Proto-Anatolian: TRVX– C_1-copying, STVX– Cluster-copying

Gothic shows the exact same sort of distinction as just seen in Ancient Greek, but with a different alternative pattern (see Chapter 4). The relevant data is presented in (6) below. In Gothic, among roots that display reduplication in the preterite, TRVX– roots (6a) permit C_1-copying — and thus the creation of a consonant repetition — because doing so does *not* incur a $*C_\alpha VC_\alpha / \_C_{[-son]}$ violation. This is demonstrated in (7). However, reduplicated STVX– roots (6b), which *would* be subject to a $*C_\alpha VC_\alpha / \_C_{[-son]}$ violation if they followed the default pattern, get diverted to a cluster-copying pattern in order to avoid that repetition. That is to say, a *CC violation is tolerated if C_1-copying would violate $*C_\alpha VC_\alpha / \_C_{[-son]}$. This is shown in (8). The same exact pattern is reconstructed for Proto-Anatolian in Chapter 3.

---

[1]   Note that the vowel in Greek is morphologically fixed, not arising through copying. This is relevant here for the violation profile of ANCHOR-L-BR; namely, the non-copying candidates do not violate ANCHOR-L-BR. See Chapter 2 for discussion.

(6) Reduplicated cluster-initial roots in Gothic (forms from Lambdin 2006:115)

    a. **C$_1$-copying reduplication** ⇔ TRVX– roots

	Infinitive	Preterite	
'weep'	grēt-an	ge-grōt	(not **[gre-grōt])

    b. **Cluster-copying reduplication** ⇔ STVX– roots

	Infinitive	Preterite	
'possess'	stald-an	ste-stald	(not **[se-stald])
'divide'	skaið-an	ske-skaiθ	(not **[se-skaiθ])

(7) Gothic TRVX– C$_1$-copying: √*grēt* → *ge-grōt* 'he wept'

/RED, grōt/	*C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$	ANCHOR-L-BR	ONSET	*CC
a. ☞ ge-grōt				*
b. e-grōt		*!	*!	*
c. gre-grōt				**!
d. re-grōt		*!		*

(8) Gothic STVX– cluster-copying: √*stald* → *ste-stald* 'he possessed'

/RED, stald/	*C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$	ANCHOR-L-BR	ONSET	*CC
a. se-stald	*!			*
b. e-stald		*!	*!	*
c. ☞ ste-stald				**
d. te-stald		*!		*

6.2.3 Sanskrit Cluster-Initial Roots: TRVX– C$_1$-copying, STVX– C$_2$-copying

Sanskrit cluster-initial roots show a distinction along these lines as well (see Chapter 5). Representative examples are shown in (9). TRVX– roots (9a) permit C$_1$-copying because they do not incur a *C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$ violation, as shown in (10). STVX– roots (9b) disallow C$_1$-copying, and instead exhibit C$_2$-copying, because C$_1$-copying would violate *C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$. This is shown in (11).[2]

---

[2] See Chapter 5 on the use of ALIGN-ROOT-L rather than *CC for Sanskrit.

(9) Perfects to cluster-initial roots in Sanskrit (forms from Whitney 1885 [1988])

a. **C₁-copying reduplication** ⇔ TRVX– roots

Root		Perfect Tense	
√b^hraj-	'shine'	<u>ba</u>-b^{h}rāj-a	(not *<u>ra</u>-b^{h}rāj-a)
√$prac^h$-	'ask'	<u>pa</u>-prāc^h-a	(not *<u>ra</u>-prāc^h-a)
√dru-	'run'	<u>du</u>-druv-ē	(not *<u>ru</u>-druv-ē)
√$tvis$-	'be stirred up'	<u>ti</u>-tviṣ-ē	(not *<u>vi</u>-tviṣ-ē)

b. **C₂-copying reduplication** ⇔ STVX– roots

Root		Perfect Tense	
√$sparç$-	'touch'	<u>pa</u>-spṛç-ē	(not *<u>sa</u>-spṛç-ē)
√$st^hā$-	'stand'	<u>ta</u>-st^{h}ā-u	(not *<u>sa</u>-st^{h}ā-u)
√$stamb^h$-	'prop'	<u>ta</u>-stamb^h-a	(not *<u>sa</u>-stamb^h-a)

(10) Sanskrit TRVX– C₁-copying: √$prac^h$- → pa-$prāc^h$-a 'he has asked'

/RED, prac^h, a/		*C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$	ONSET	ALIGN-ROOT-L	ANCHOR-L-BR
a. ☞	<u>pa</u>-prāc^h-a			**	
b.	<u>a</u>-prāc^h-a		*!	*	*
c.	<u>pra</u>-prāc^h-a			***!	
d.	<u>ra</u>-prāc^h-a			**	*!

(11) Sanskrit STVX– C₂-copying: √$stamb^h$- → ta-$stamb^h$-a 'he has propped'

/RED, stamb^h, a/		*C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$	ONSET	ALIGN-ROOT-L	ANCHOR-L-BR
a.	<u>sa</u>-stamb^h-a	*!		**	
b.	<u>a</u>-stamb^h-a		*!	*	*
c.	<u>sta</u>-stamb^h-a			***!	
d. ☞	<u>ta</u>-stamb^h-a			**	*

6.2.4 Sanskrit Zero-Grade Bases: TaR– C₁-copying vs. SaT– C₁ēC₂

Also in Chapter 5, I showed that the distribution of stem formation patterns for CaC roots in the zero-grade categories of the perfect (i.e., the inflectional categories where the root vowel is normally subject to vowel deletion) is likewise driven by the activity of *C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$. For TaR roots, whose zero-grade stem would be -TR-, the perfect is formed through C₁-copying reduplication, just as with the TRVX– cluster-initial roots: for example, √par 'fill' → zero-grade perfect pa-pr-$úr$. The consonant repetition is licensed here because it does not violate *C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$. On the other hand, for SaT roots (among many others; see Section 6.3.4.2), reduplication is eschewed altogether in these categories. This is because the cluster that would result from zero-grade vowel deletion is -ST-, and such a cluster would incur a *C$_\alpha$VC$_\alpha$ / _C$_{[-son]}$ violation if accompanied by

C_1-copying reduplication. To avoid this violation, a different, non-reduplicative pattern surfaces, the $C_1\bar{e}C_2$ pattern: for example, $\sqrt{sap}$ 'serve' → zero-grade perfect *sēp-úr*. (The same sort of pattern is the historical source of the Germanic Strong Class V preterite plurals; cf. Chapter 4.) While the details of the analysis are complex (see Chapter 5), the basic outline of the distribution is fully parallel to the other reduplicative distributions just discussed: the default pattern applies when it would *not* lead to a pre-obstruent repetition, but an alternative pattern surfaces when it *would*.

6.2.5 Local Summary

This review has illustrated that the reduplicative patterns of the Indo-European languages repeatedly show a distinction that can be characterized as avoidance of consonant repetitions in pre-obstruent position. This dispreference results in multiple different kinds of overtly reduplicative repairs, but also in at least one — the $C_1\bar{e}C_2$ pattern — which seems non-reduplicative in nature (though see Chapter 5 for arguments for a reduplicative origin of the pattern.) Later, in Sections 6.6 and 6.7, I will adduce several additional patterns — some involving reduplication, some in the non-reduplicative phonology; some in the ancient Indo-European languages, some from elsewhere — that can be characterized in exactly the same way.

The broad empirical coverage of an output markedness constraint against pre-obstruent consonant repetitions provides significant support that this is the proper way to conceptualize the problem. In the following sections, I present additional to contextualize the specifics of what this markedness constraint must be, and ultimately advocate that it finds its most explanatory formulation when defined in terms of acoustic/auditory cues (following Steriade 1994, 1997, 1999, *et seq.*, Flemming 1995/2002, *et seq.*).

6.3 The Cluster-Wise Distributions of Repetition Avoidance Effects

The empirical/typological data just presented motivates the inclusion of a constraint that can distinguish between TRVX– roots/bases and STVX– roots/bases in reduplication with respect to the permission of consonant repetitions.[3] The prior discussions of default vs. alternative reduplication patterns focused on the TRVX– vs. STVX– distinction because these are the only cluster types which are attested with reduplication in all of the languages under discussion (save Latin, which has TRVX– roots, but none are attested with reduplication; see Section 6.6.1). In all of the languages that treat them differently,[4] TRVX– shows the default pattern (C_1-copying), while STVX– demonstrates the alternative pattern. Therefore, it was convenient to frame the discussion in these terms, as it made for the simplest and most direct comparison between languages. However, the TRVX– vs. STVX– division represents only a subset of the total reduplicative distributions, which have thus far been partially suppressed for expository purposes.

In this section, I present the full cluster-wise distributions for each of the patterns discussed in Section 6.2, and a few others in addition. This will demonstrate that it is not actually sufficient to draw a simple dividing line between pre-sonorant (as the context which licenses repetitions) and pre-obstruent (as the context which bans repetitions). Rather, it requires a significantly more fine-grained approach. In this section, I will characterize the distributions in terms of natural class-based and/or sonority-based generalizations. In the sections that follow, I will argue that the distributions are

---

[3] It has here been formulated as a markedness constraint on sequences involving repetitions.

[4] Not discussed in this chapter are Hittite (see Chapter 3) and Old Irish (see briefly Chapter 1), which treat TRVX– and STVX– bases the same (consistent cluster-copying for Hittite; consistent C_1-copying for Old Irish).

better explained in terms of acoustic/auditory cues to particular consonantal contrasts (specifically, the C∼Ø contrast).

6.3.1 Gothic

Gothic presents the simplest (and thus least informative) case. The complete set of reduplicating cluster-initial roots in Gothic is provided in (12) below. It attests only one cluster type with reduplication other than TR or ST: fricative-liquid, where the fricative may be either *s* or *f*.

The distribution represented by this data is schematized in (13), broken down by possible C_1C_2 combinations. Clusters that allow C_1-copying are notated with ✓; these are the repetition types which are permitted, and thus the ones that *don't* violate *PCR. On the other hand, clusters that show cluster-copying are notated with ✗; these are the repetition types which are disallowed, and thus the ones that *do* violate *PCR.

(12)　　Reduplicated cluster-initial roots in Gothic (forms from Lambdin 2006:115)

　　　　a.　**C_1-copying reduplication** ⇔ *obstruent-liquid clusters*

	Infinitive	Preterite	
'weep'	grēt-an	<u>ge</u>-grōt	(not **[<u>gre</u>-grōt])
'sleep'	slēp-an	<u>se</u>-slēp	(not **[<u>sle</u>-slēp])
'bewail'	flōk-an	<u>fe</u>-flōk	(not **[<u>fle</u>-flōk])
'tempt'	frais-an	<u>fe</u>-frais	(not **[<u>fre</u>-frais])

　　　　b.　**Cluster-copying reduplication** ⇔ *sibilant-stop clusters*

	Infinitive	Preterite	
'possess'	stald-an	<u>ste</u>-stald	(not **[<u>se</u>-stald])
'divide'	skaið-an	<u>ske</u>-skaiθ	(not **[<u>se</u>-skaiθ])

(13)　　Distribution of *PCR effects in Gothic by root-initial cluster

C₁ ＼ C₂	Stop (T)	Liquid (L)
Stop (T)		✓ $(T_\alpha VT_\alpha L)$
Fricative (F/S)	✗ $(*S_\alpha VS_\alpha T)$	✓ $(S_\alpha VS_\alpha L, F_\alpha VF_\alpha L)$

The data here is obviously limited. This means that there are a number of ways in which the generalization about which clusters do or do not license a repetition could be stated. The one which will be most informative in light of the distributions found in the other languages (to be presented below) is the following:

(14)　　**Gothic repetition generalization:**
　　　　Obstruent-sonorant clusters license repetitions, other clusters do not.

While the available data contains only obstruent-*liquid* clusters as C_1-copying forms, it is reasonable to assume (especially in light of the other systems we will observe below) that the lack of other obstruent-sonorant forms (i.e. where C_2 is a glide or nasal) represents an accidental gap.

226

Gothic allows stop-glide, fricative-glide, and fricative-nasal clusters in word-initial position (it does not allow stop-nasal); however, there happen to be no such cluster-initial roots in the reduplicating category.[5] If we adopt the form of the generalization in (14), then we predict that other stop-sonorant and fricative-sonorant roots should also show C_1-copying.

Assuming that (14) is an appropriate generalization for Gothic, then the broadly-defined anti-repetition constraint $*C_\alpha VC_\alpha \, / \, \_C_{[-son]}$ would indeed be sufficient to capture the facts of Gothic. I now turn to the remaining systems, where the distribution will not always permit adherence to this version of the constraint.

6.3.2   Greek

The facts of Greek require some exposition, and probably point to some diachronic differences in the relevant domains. I first present the evidence which reflects the productive grammar of Classical and Pre-Classical Greek (which will be the stage focused on in this chapter), and then present the evidence which seems to represent other stages of the language.

6.3.2.1   The Core Facts of Greek

In Classical and Pre-Classical Ancient Greek, three types of root-initial clusters are consistently attested with non-copying reduplication: fricative-stop clusters (15a), stop-stop clusters (15b), and stop-fricative clusters (15c). (There is a set of principled exceptions to this generalization, which will be discussed below.) These contrast with stop-sonorant roots, which consistently show C_1-copying (the behavior of voiced stops requires further explication; see below). Representative examples are given in (16) (repeated from (3a) above).

---

[5] Whether or not a root exhibits reduplication in Gothic is dependent on the phonological shape of its rhyme: a root reduplicates in the preterite if it ends in a long vowel or the sequence *aRC*. See Chapter 4 for discussion and analysis of these facts.

(15) Non-copying cluster-initial perfects in Ancient Greek

Root			Perfect Tense	
a. Fricative-stop clusters (STVX–)				
√*stel-*	'prepare'	_-e-stal-	(not **s-e-stal-)	
√*sper-*	'sow'	_-e-spar-	(not **s-e-spar-)	
√*skep-*	'view'	_-e-skep-	(not **s-e-skep-)	
√*sbes-*[6]	'extinguish'	_-e-sbes-	(not **s-e-sbes-)	
√*zdeug-*	'yoke'	_-e-zdeug-	(not **z(d)-e-zdeug-)	
b. Stop-stop clusters (TTVX–)				
√*kten-*	'kill'	_-e-kton-	(not **k-e-kton-)	
√*ktis-*	'found'	_-e-ktis-	(not **k-e-ktis-)	
√*ptis-*	'pound'	_-e-ptis-	(not **p-e-ptis-)	
√*pʰtʰer-*	'destroy'	_-e-pʰtʰor-	(not **p$^{(h)}$-e-pʰtʰor-)	
c. Stop-fricative clusters (TSVX–)				
√*pseud-*	'lie'	_-e-pseus-	(not **p-e-pseus-)	
√*kses-*	'shave'	_-e-kses-	(not **k-e-kses-)	

(16) C_1-copying reduplication for TRVX– roots in Ancient Greek (consult van de Laar 2000)

Root		Perfect Tense	
√*kri-*	'decide'	k-e-kri-	(not *_-e-kri-)
√*pneu-*	'breathe'	p-e-pnū-	(not *_-e-pnū-)
√*tla-*	'suffer, dare'	t-e-tlε̄-	(not *_-e-tlε̄-)
√*grapʰ-*	'write'	g-e-grapʰ-	(not *_-e-grapʰ-)

While some additional, more complicated data will be discussed below, this is the core data in need of explanation. This distribution is indeed consistent with the $*C_\alpha VC_\alpha$ / _$C_{[-son]}$ constraint, because the clusters which exhibit non-copying are all and only those with an obstruent as C_2. This distribution is schematized in (17).

(17) Initial clusters and reduplicative behavior in Ancient Greek

C_1 \ C_2	Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Stop	✗ ($*T_\alpha VT_\alpha T$)	✗ ($*T_\alpha VT_\alpha S$)	✓ ($T_\alpha VT_\alpha N$)	✓ ($T_\alpha VT_\alpha L$)
Fricative	✗ ($*S_\alpha VS_\alpha T$)			

---

[6] This root is written σβεσ- (present σβέννῡμι; van de Laar 2000:264–265). The initial <σ> would usually represent [s] (as transcribed in the chart above). However, it is quite possible that it here really represents [z], given that Greek has regressive voicing assimilation in obstruent clusters. This question is of little consequence here.

Given this data, the nature of the repetition generalization for Ancient Greek is, like Gothic, somewhat under-determined. Focusing on the positive evidence, the generalization would be that stop-sonorant clusters license repetitions. This is provided in (18).

(18) **Ancient Greek repetition generalization:**   [to be slightly revised]
      Stop-sonorant clusters license repetitions, other clusters do not.

However, since there are no fricative-sonorant roots in reduplication (at this stage of the language; see below), we would also be at liberty to extend this to obstruent-sonorant clusters, which would then match the generalization for Gothic (14).

6.3.2.2 Voiced Stop + Sonorant Clusters

The generalization in (18) holds without exception in stop-sonorant clusters where the stop is a voiceless stop or a voiceless aspirated stop. There is, however, some amount of variation when the stop is voiced. Roots/bases beginning in *br-* and *dr-* appear to consistently display the same behavior as voiceless (aspirated) stops, i.e. C_1-copying. The evidence for *voiced stop + nasal* is limited, but fairly clear. The roots $\sqrt{dam}$ 'overpower' (present δάμνημι [dám-nḗ-mi]; van de Laar 2000:106–107) and $\sqrt{dem}$ 'build' (present δέμω [dém-ɔ̄]; van de Laar 2000:111) both attest C_1-copying perfect stems in δεδμη- [d̲-e-dmē-]. There are no roots/stems with initial *b* + nasal.

There are two perfect stems which have initial (orthographic) *gn-* <γν->, both of which show non-copying: $\sqrt{gnɔ̄}$ 'come to know' (present γιγνώσκω [g-i-gnɔ̄́-sk-ɔ̄]; van de Laar 2000:102) → perfect ἔγνωκα [_-e-gnɔ̄́-k-a] and $\sqrt{gnɔ̄}$ 'make known' (present γνωρίζω [gnɔ̄rí-zd-ɔ̄]) → perfect ἔγνώρικα [_-e-gnɔ̄́ri-k-a]. However, orthographic *g* <γ> before a nasal probably represents a velar nasal [ŋ] (Sihler 1995:207/§220), in which case these are not voiced stops at all, and these stems should instead be transcribed [_-e-ŋnɔ̄(...)-]. If this is correct, then these forms do not provide evidence for the behavior of voiced stop + nasal clusters; rather, they would indicate that nasal-nasal clusters disallow repetitions — i.e. *PCR penalizes repeated nasals before nasals in Greek (which is what we would expect given the generalization in (18)).[7] The C_1-copying *dm-* forms thus represent the totality of the evidence for voiced stop + nasal in Ancient Greek, and suggest that voicing does not alter their behavior vis-à-vis *PCR.

The evidence for the behavior of voiced stops in reduplication is less straightforward. As discussed by Steriade (1982:206–208), roots beginning in *voiced stop* (*b*, *g*) + *l* show inconsistent reduplicative behavior. (Initial *dl-* is not permitted in Greek.) The data is provided in (19) below. Most *bl-* and *gl-*initial roots that attest perfect forms exhibit C_1-copying (all but (19h)). However, some additionally attest non-copying perfects (19a–c,g), or exclusively attest a non-copying perfect (19h). There also seems to be variation in the perfects of $\sqrt{grap^h}$ 'write' (van de Laar 2000:104), though the C_1-copying form is much more common.

There are two potential distributional facts about these forms that could be relevant (though neither of them is certain at this point). First, it appears to be the case that the C_1-copying perfects tend to be attested earlier and more frequently; this could indicate a diachronic distinction, with the language moving towards non-copying with such clusters. Second, and perhaps more importantly if correct, it seems that the non-copying perfects are attested mainly with (consonant-final) preverbs: for example, $\sqrt{glup^h}$- → non-copying perfect ἐξεγλυμμένῳ [eks-_-e-glum-mén-ɔ̄(i)]

---

[7] This could alternatively be a BR-identity effect, and not be providing evidence regarding *PCR. Velar nasals are not permitted word-initially before a vowel (i.e. the markedness constraint *#ŋV is undominated). Faithful copying would yield such a sequence: **[ŋ̲-e-ŋnɔ̄(...)-]. If *#ŋV and the BR faithfulness constraints IDENT[nas]-BR (which would penalize **[g̲-e-ŋnɔ̄(...)-]) and IDENT[place]-BR (which would penalize **[n̲-e-ŋnɔ̄(...)-]) outrank ONSET, then it will be preferable to employ non-copying than to do either faithful or unfaithful copying.

229

(19) Variation in *bl-*, *gl-* roots (Steriade 1982:206–208; cf. van de Laar 2000)

	Root		C$_1$-copying perfect	Non-copying perfect
a.	√*blaiso-*	'be crooked'	b̠-e-blaisō̄-	_-e-blaisō̄-
b.	√*blap-*	'hinder'	b̠-e-blap^h-	_-e-blap^h-
c.	√*blasta-*	'sprout'	b̠-e-blastē̄-	_-e-blastē̄-
d.	√*blep-*	'look'	b̠-e-blop^h-	*not attested*
e.	√*blasp^{h}ē̄me-*	'speak irreverently'	b̠-e-blasp^{h}ē̄mē̄-	*not attested*
f.	√*blō̄-*	'go, come'	b̠-e-blō̄-	*not attested*
g.	√*glup^h-*	'carve'	g̠-e-glup-	_-e-glup-
h.	√*(kata-)glō̄ttis-*	'kiss lasciviously'	*not attested*	(kat-)_-e-glō̄ttis-

(Plato's Republic 616d). The presence of a preceding consonant could introduce phonotactic issues which are partially distinct from the repetition question. I have not yet been able to philologically examine either of these possible distributional facts to a sufficient degree to be sure of their correctness, but both could affect our understanding of these forms relative to *PCR.

If, however, explanations of this sort were not to hold up, then it would be the case that there was a *bona fide*, if marginal, difference in the behavior of stop-liquid sequences based on voicing. Given that the other voiced stop + sonorant clusters seem to pattern with the voiceless stops, such a distinction would be hard to integrate into the generalization in (18); it would essentially have to stand out as a stipulated exception. On the other hand, if we were to state our generalization in terms of acoustic/auditory cues (as I will argue for below), we might have an explanation for the marginality of repetition for these clusters in particular: Flemming (2007) shows that voiced stops are more confusable with one another before *l* than before *r*, because of the way the lateral articulation of *l* affects the stop burst properties and formant transitions out of a preceding voiced stop. This explanation will be spelled out in Section 6.5.5 below.

Regardless, this allows us to update our schematic repetition distribution from (17) above. The chart in (20) adds in the new C$_1$ types discussed in this subsection, namely, voiced stop + nasal, voiced stop + liquid, and nasal + nasal.[8] The only way in which this does not completely conform to the generalization in (18), which stated that all and only stop-sonorant clusters license repetitions, is the variation observed with voiced stop + liquid clusters. The generalization is updated in (21) to take note of this variation.

(20) Initial clusters and reduplicative behavior in Ancient Greek (expanded)

C$_1$ \ C$_2$	Vcls Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Vcls Stop (T)	✗ (*T$_\alpha$VT$_\alpha$T)	✗ (*T$_\alpha$VT$_\alpha$S)	✓ (T$_\alpha$VT$_\alpha$N)	✓ (T$_\alpha$VT$_\alpha$L)
Vcd Stop (D)			✓ (D$_\alpha$VD$_\alpha$N)	✓∼✗ ($^{(*)}$D$_\alpha$VD$_\alpha$L)
Fricative	✗ (*S$_\alpha$VS$_\alpha$T)			
Nasal			✗ (*N$_\alpha$VN$_\alpha$N)	

---

[8] Note that there is an apparent exception to the claim that nasal-nasal clusters don't tolerate C$_1$-copying. See Section 6.3.2.5 below for discussion.

(21) **Ancient Greek repetition generalization:**   [final version]
      Stop-sonorant clusters license repetitions, other clusters do not. (Variation in voiced stop +
      liquid clusters.)

This represents the entirety of the evidence for the productive reduplicative distributions of cluster-initial roots in Classical and Pre-Classical Greek. The remainder of this subsection examines the available evidence for the distributions of prior stages (and, in one case, a subsequent stage).

6.3.2.3 Fricative + Sonorant Clusters

I claimed in (17) above that Greek did not attest any *s*-sonorant roots in the perfect (indicated by the relevant grayed out cells). This is true of the Classical and Pre-Classical periods, but not true of either the prior stage (available to us through reconstruction) or the Post-Classical stage. The evidence from these two stages (such as it is) points towards opposite behaviors.

(Pre-)Classical Greek

The absence of *s*-sonorant roots in attested Classical and Pre-Classical Greek is the result of an earlier sound change. In the internal history of Greek, *s* was lost in pre-sonorant position (Sihler 1995:170–171, 216): $*s\ (>h)>\emptyset\ /\ \_C_{[+son]}$ (sometimes with compensatory lengthening of the preceding vowel, depending mainly on dialect). Prior to this sound change, *s*-sonorant roots did indeed exist, and at least a few of them seem to have made perfects which we can reconstruct. While some of the evidence is ambiguous, overall it suggests that these roots exhibited C_1-copying, and thus that repetitions of *s* before a sonorant did not violate *PCR at that stage.

The best evidence of this comes from the Ancient Greek root $\sqrt{mer}$ 'divide, share', whose present is μείρω [mēr-ō]. This root derives from the Proto-Indo-European root $*\sqrt{smer}$ (Chantraine 1980:678–679), with the initial *s* lost by the aforementioned regular sound change. While this root attests several different kinds of perfect stems (van de Laar 2000:214–215), the one of interest here is εἰμαρται [hēmar-tai] (see Steriade 1982:350–353ff., esp. fn. 15 (pp. 377–378)). The long vowel [ē] must be the result of compensatory lengthening after loss of *s*. The stem-initial aspiration, then, must come from a separate *s*, i.e. a reduplicant *s*: Ancient Greek *hēmartai* < Pre-Greek *s̠-e-smr̥-tai* (as per Chantraine 1980:679). This form thus requires that Pre-Greek copied C_1 for at least *sm* clusters.

The perfect stem εἴληφ- [ēlēpʰ-] could also point to the same treatment for *sl* clusters. This form is one of several perfect stems attested for the root $\sqrt{lab}$ 'take' (whose main present is λάμβάνω [lá-m-b-an-ō]). This root is usually regarded as coming from PIE $*\sqrt{sleh_2g^w}$ (e.g., Chantraine 1980:616, van de Laar 2000:201–202, Beekes & van Beek 2010:828–829; *pace* Sihler 1995:575). This etymology allows us to derive *ēlēpʰ*- from *s̠-e-slāpʰ*-, via intermediate (or virtual) *heslāpʰ*- where the expected reduplicant-initial *h* is lost (or blocked) by the operation of Grassman's Law (consult Collinge 1985:47–61, Sihler 1995:142–144/§138, 170/§170, among others).[9]

This evidence thus suggests that at some earlier stage of the language, *s*-sonorant clusters behaved just like stop-sonorant clusters for the purposes of *PCR, i.e. both displayed C_1-copying. Since there is significant dialect variation on the treatment of these forms, whereas there is complete agreement on the consonant-obstruent forms, it is likely that the change from C_1-copying to non-copying for STVX– and others pre-dates the loss of *s*. (This claim is tentative, and deserves further attention.) Therefore, there seems to be a period of the language (which we might iden-

---

[9] Steriade (1982:352, fn. 15 (pp. 377–378)) adduces several other sonorant-initial roots which display perfects with [ē-] but without an initial [h]. It is unclear if these roots should actually be reconstructed with initial *s*; see Chantraine (1980), Rix et al. (2001), Beekes & van Beek (2010).

tify as Proto-Greek) where indeed all obstruent-sonorant clusters copied C_1, while all other clusters showed non-copying. This distribution is schematized in (22).

(22)   Reduplicative behavior in Proto-Greek (tentative)

C₂ \ C₁	Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Stop	✗ (*$T_\alpha VT_\alpha T$)	✗ (*$T_\alpha VT_\alpha S$)	✓ ($T_\alpha VT_\alpha N$)	✓ ($T_\alpha VT_\alpha L$)
Fricative	✗ (*$S_\alpha VS_\alpha T$)		✓ ($S_\alpha VS_\alpha N$)	✓ ($S_\alpha VS_\alpha L$)

If this is the correct characterization of a unitary stage of the language, then this distribution can be captured with the same generalization proposed for Gothic in (14):

(23)   **Proto-Greek repetition generalization:**   [tentative]
        Obstruent-sonorant clusters license repetitions, other clusters do not.

One might interpret this to mean that the proper generalization for (Pre-)Classical Greek is indeed that repetitions are licensed for *obstruent*-sonorant clusters, not just *stop*-sonorant clusters, but the language has simply eliminated *s*-sonorant clusters independently.

Post-Classical Greek

As alluded to earlier, though, in Post-Classical Greek, there is evidence for the opposite treatment of *s*-sonorant roots. Over the intervening period, a number of *s*-sonorant roots are reintroduced into the language. A few of them, all happening to have *sm*-clusters, are verbal roots that attest perfects. These are provided in (24) (consult Liddell, Scott, & Jones 1940 for attestations). Each one exhibits non-copying. Note that (24c) $\sqrt{sma}$- 'wipe' (present σμάω [smá-ɔ̄]), which appears to be consistently written with initial <σ> [s] throughout its history, has a very late attested (2nd century AD) perfect stem spelled προεϲ̣μησμένος, which I interpret to mean [pro-_-e-zmɛ̄s-mén-os], with voicing of the root-initial /s/ to [z].

(24)   Perfects of *s*-nasal roots in Post-Classical Greek

	Root		Perfect Tense	
a.	$\sqrt{sm\bar{e}k^h}$-	'wipe'	_-e-smē̠g-	(not **s̠-e-smē̠g-)
b.	$\sqrt{sm\bar{u}k^h}$-	'burn'	_-e-smug-	(not **s̠-e-smug-)
c.	$\sqrt{sma}$-	'wipe'	_-e-zmē̠s-	(not **z̠-e-zmē̠s-, **s̠-e-zmē̠s-)

Since I have not systematically assembled evidence for the behavior of perfect reduplication in the Post-Classical period, I will refrain from making any generalization regarding the nature of *PCR at that stage.

6.3.2.4   Nasal-Liquid Clusters

Pre-Greek contained roots/stems with initial *m*-liquid clusters. These generally developed into *b*-liquid clusters in Ancient Greek: for example, PIE $\sqrt{mreg^h}$ > Ancient Greek $\sqrt{brek^h}$ 'wet' (van de Laar 2000:96); PIE $\sqrt{melh_3}$ > Ancient Greek $\sqrt{bl\bar{\mathfrak{d}}}$ (based on the zero-grade) 'go, come' (van de Laar 2000:94). In most cases, the perfects to these roots reflect the innovative initial *b*: for example, $\sqrt{brek^h}$ → perfect βέβρεγμαι [b̠-é-breg-mai]; $\sqrt{bl\bar{\mathfrak{d}}}$ → perfect βέβλωκα [b̠-é-blɔ̄-k-a]. That is to say,

the root has been restructured in its underlying representation, and the perfect reflects the productive pattern of C_1-copying for stop-sonorant roots.

Some of these roots do, however, retain archaic forms that point to C_1-copying for *m*-liquid in Pre-Greek (Steriade 1982:352, fn. 15 (pp. 377–378)): we have two perfects in *membl-*, pointing to Pre-Greek **m-e-ml...*, with later excrescence of [b] between the nasal and the liquid. First, there is μέμβλεται [membletai] < **m-e-ml-e-tai*; this is an early attested perfect to the root √*mel* 'care for', which additionally makes a more synchronically regular perfect μέμηλα [m̲-e-mēl-a] (Chantraine 1980:684, van de Laar 2000:215–216, Beekes & van Beek 2010:928–929).

We also have a perfect μέμβλωκα [memblōka] < **me-mlō-k-a* (< **me-ml(o)h₃-*).[10] This is an archaic perfect built to PIE root **√melh₃* 'go, come', the root which, as discussed above, evolves into Ancient Greek √*blō* (present βλώσκω [blō-sk-ō]). Therefore, μέμβλωκα [memblōka] exists beside the later, synchronically regular C_1-copying perfect βεβλώκα [b̲-e-blō-k-a] (Chantraine 1980:182, van de Laar 2000:94, Beekes & van Beek 2010:223). The existence of these forms strongly suggests that we should reconstruct C_1-copying also for nasal-liquid roots in the prior stages of Greek.

6.3.2.5 Apparent Exceptions, and their Ramifications for Pre-Greek

Ancient Greek attests several exceptions to the generalization in (18) that non–stop-sonorant clusters show non-copying (see Chapter 2 for more detailed discussion). These exceptions point to an earlier stage of the language where many more cluster types permitted C_1-copying.

One isolated, yet very telling exception has to do with the behavior of the root *kta* 'acquire'. Its normal perfect stem shows the expected non-copying pattern: *e-ktē-*. It also attests a perfect stem with unexpected C_1-copying: *k̲-e-ktē-*. But this stem does not actually function synchronically as the root's perfect stem. Rather, it is semantically a present with the meaning 'possess', and it attests modal bases as if it were an independent root (perfect stems do not permit further derivation). This behavior is best explained by assuming that *k̲-e-ktē-* is an archaic formation, which underwent morphological and semantic shift. This somehow licensed retention of the older reduplicative pattern, which evidently was C_1-copying.

Additional evidence in this direction comes from several C_1-copying perfect stems to non–stop-sonorant cluster-initial roots which additionally contain C_1-copying reduplicated presents in their verbal systems. These are provided in (25). In Chapter 2, I argued that the fact that the exceptions to the non-copying generalization are restricted to present-perfect pairs of this type can only be explained if the pattern they exhibit is interpreted as an archaism; that is, there is no pathway by which these could have arisen as innovations.[11]

---

[10] Note that if this form is coming from a zero-grade, then it is not evidence for the treatment of cluster-initial bases at the stage in which laryngeals are still present, because it would reflect **me-ml̥h₃-* with syllabic [l̥]. External evidence suggests that the position of the full-grade/*o*-grade vowel for this root is between the *m* and the *l* (Rix et al. 2001:433–434), so it is likely that we are dealing with a zero-grade.

[11] The exception to the exceptions is the present-perfect pair mentioned earlier: present γιγνώσκω [g̲-i-ŋnō-sk-ō] ↔ perfect ἐγνώκα [_-e-ŋnō-k-a]. This must have something to do with the properties of velar nasals. I have not yet worked out the details of the analysis of this pair.

(25)   Ancient Greek exceptional C_1-copying forms: present and perfect

Root		Present	Perfect
√*pet-*	'fall'	*p̲-í-pt-ɔ̄*	*p̲-é-ptɔ̄-k-a*
√*stē̆-*	'stand'	*h̲-í-stē̆-mi* (< **sistāmi*)	*h̲-é-stē̆-k-a* (< **sestāka*)
√*mnē̆-*	'remind'	*m̲-i-mnḗ-sk-ɔ̄*	*m̲-é-mnē̆-mai*

If we thus assume that these forms provide evidence for the productive reduplication pattern of some earlier stage, then we can conclude that there was a stage where stop-stop, fricative-stop, and nasal-nasal all permitted C_1-copying. Since this essentially covers the full range of repetitions outlawed in attested Ancient Greek (it is likely that stop-fricative roots are a relatively late development), this indicates that there was a stage where *PCR was essentially inactive. However, if we push back far enough, we find one repetition type that was indeed prohibited: repeated laryngeals before a consonant.

6.3.2.6   Attic Reduplication in Pre-Greek: Avoidance of Repeated Laryngeals

In Chapter 2, I demonstrated that Attic Reduplication — the minority treatment of vowel-initial roots in the Ancient Greek perfect — can and should be traced back to a pattern in Pre-Greek which is characterized as the avoidance of repeated "laryngeals" in pre-consonantal position. For example, Ancient Greek √*ager* 'gather' → perfect *ag̲ḗger-* comes from Pre-Greek **√$h_2$ger* → perfect **h_2əg-e-$h_2$ger-*, which specifically avoids the default C_1-copying output ***$h̲_2$-e-$h_2$ger-*.

It must be the case that laryngeal+consonant clusters were the only type to undergo an alternative copying pattern at that stage. This is because the constraint grammar necessary to generate the "Pre-Attic Reduplication" forms for laryngeals would generate either an equivalent pattern or, more probably, simple cluster-copying (equivalent to Gothic) for any other cluster type targeted for repetition avoidance. We have no evidence of any such forms. Furthermore, the persistence of the exceptional form ἵστημι [h̲-í-stē̆-mi] (< **sistāmi*), which is precisely cognate with Latin *sistō* [s̲i-st-ō] (which is also synchronically irregular), argues very strongly ST clusters were not subject to repetition at prior stages of Greek (Brugmann & Delbrück 1897–1916:40–41, Byrd 2010:103–104). Therefore, at this stage, we can assert that all clusters other than laryngeal+consonant permitted repetitions.

Unfortunately, it is not completely clear what the precise phonetics of the laryngeals were at this stage. The laryngeals are generally assumed to have been non-strident dorsal or pharyngeal fricatives (see, e.g., Fortson 2010:62, Byrd forthcoming:§3.3), at least at the point of Proto-Indo-European.[12] However, it would not be unreasonable to assume that they had undergone lenition by the stage at which they became targeted for repetition avoidance; that is to say, they could have been something more like approximants when Pre-Attic Reduplication arose (see the discussion in Chapter 2). Since we do not actually know what the precise phonetics/feature specification of the laryngeals was (at this stage of Pre-Greek, or any earlier stage), it is not appropriate to put any great weight upon this finding with respect to the precise scope of repetition avoidance effects.

6.3.2.7   Local Summary

This section has provided the full evidence for the reduplication patterns of cluster-initial roots throughout the (pre-)history of Greek. While a more careful consideration of the relative chronology

---

[12] Consult Chapter 2 of this dissertation for additional discussion of the laryngeals.

of various changes is called for, we can tentatively identify up to six stages of Greek that display slightly different scopes of repetition avoidance effects.

(26) Hypothesized stages of Greek with distinct repetition avoidance behaviors

    a. Pre-Greek I (prior to the advent of Pre-Attic Reduplication)
       All clusters permit C_1-copying (no *PCR effects).

    b. Pre-Greek II (stage when Pre-Attic Reduplication was active)
       All clusters permit C_1-copying except laryngeal+consonant.

    c. Pre-Greek III (after loss of laryngeals)
       All clusters permit C_1-copying (no *PCR effects).

    d. Pre-Greek IV (before loss of pre-sonorant *s*; $\approx$ Proto-Greek)
       i. Obstruent+sonorant clusters (including *s*-sonorant) permit C_1-copying.
       ii. Nasal+liquid and nasal+nasal clusters (probably) permit C_1-copying.
       iii. Consonant+obstruent clusters prohibit C_1-copying.

    e. **(Pre-)Classical Greek (after loss of pre-sonorant *s*)**
       i. **Stop+sonorant clusters permit C_1-copying (variation for *bl-* and *gl-*).**
       ii. **Consonant+obstruent clusters prohibit C_1-copying.**
       iii. **Nasal-nasal clusters prohibit C_1-copying.**

    f. Post-Classical Greek (after limited reintroduction of pre-sonorant *s*)
       i. *s*-sonorant clusters prohibit C_1-copying.
       ii. (Other cluster types not investigated.)

The one which is most certain is that of Pre-Classical and Classical (26e). The schematic distribution for this stage is provided in (27) below (repeated from (20)). The others are based on reconstructions, or, in the case of Post-Classical Greek (26f), limited investigation on my part. For this reason, in the remainder of this chapter, I will only consider (Pre-)Classical Greek when evaluating the scope of *PCR effects. This means that the relevant generalization is the one laid out in (21) above, repeated below, again with the caveat that the data is also consistent with *obstruent*-sonorant rather than just *stop*-sonorant.

(27) Initial clusters and reduplicative behavior in (Pre-)Classical Greek (final)

C_2 / C_1	Vcls Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Vcls Stop (T)	✗ (*$T_\alpha VT_\alpha T$)	✗ (*$T_\alpha VT_\alpha S$)	✓ ($T_\alpha VT_\alpha N$)	✓ ($T_\alpha VT_\alpha L$)
Vcd Stop (D)			✓ ($D_\alpha VD_\alpha N$)	✓~✗ ($^{(*)}D_\alpha VD_\alpha L$)
Fricative	✗ (*$S_\alpha VS_\alpha T$)			
Nasal			✗ (*$N_\alpha VN_\alpha N$)	

(28) **Ancient Greek repetition generalization:**
Stop-sonorant clusters license repetitions, other clusters do not. (Variation in voiced stop + liquid clusters.)

In the future, further investigation of the relative chronology and other details of the reconstructed stages, and more careful study of the Post-Classical Period, may allow us to more confidently include these additional stages in our understanding of *PCR.

6.3.3 A Non–Indo-European Parallel: Klamath

The sort of cluster-dependent reduplicative behavior currently under examination is not only found in the Indo-European languages. Steriade (1988:131, 136–139) puts forward Klamath (a recently extinct language isolate from the Pacific Northwest of the United States; Barker 1964, *et seq.*) as a robust typological comparandum. Similar to Gothic (and Proto-Anatolian), Klamath demonstrates an alternation between a C_1-copying pattern (CV) and a cluster-copying pattern (CCV), dependent on the type of root-initial cluster (see also Fleischhacker 2005). Klamath is a useful case, as it has a very rich inventory of root-initial cluster types, including sonorant-obstruent, a type not attested in any of the Indo-European languages under discussion (except for zero-grade clusters in Sanskrit; see below). Despite having a great many more cluster types that display reduplication than do most of the Indo-European languages, the set of clusters for which it licenses repetitions is essentially coextensive with that of Ancient Greek. In this subsection, I lay out the cluster-wise distribution of repetition avoidance effects in Klamath. I postpone the formal analysis of the system until Section 6.6.3 below.

The consonant inventory of Klamath is provided in the table in (29) below. In each cell, the left-hand representation corresponds to the IPA transcription (following Blevins 1993, 2001); the right-hand representation (enclosed in < >) corresponds to Barker's (1964) original transcription system (see Blevins 1993:237–238, fn. 2), which I will employ below (for consistency with the previous literature, especially Steriade 1988). Klamath has three consonant series: plain, aspirated, and glottalized. All three series are fully represented by both obstruents (in which the glottalized series surfaces as ejectives) and sonorants (in which the aspirated series surfaces as voiceless).

(29) Klamath consonant inventory

Obstruents							
Plain:	T	p < b >	t < d >	tʃ < j >	k < g >	q < g̣ >	s
Aspirated:	Tʰ	pʰ < p >	tʰ < t >	tʃʰ < c >	kʰ < k >	qʰ < q >	h
Glottalized (ejective):	T'	p' < ṗ >	t' < ṭ >	tʃ' < č >	k' < k̇ >	q' < q̇ >	ʔ
Sonorants							
Plain:	R	m < m >	n < n >	l < l >	w < w >	j < y >	
Aspirated (voiceless):	Rʰ	m̥ < M >	n̥ < N >	l̥ < L >	w̥ < W >	j̥ < Y >	
Glottalized (creaky):	R'	m̰ < ṁ >	n̰ < ṅ >	l̰ < l̇ >	w̰ < ẇ >	j̰ < ẏ >	

Beyond just having a large consonant inventory, Klamath also allows a wide variety of initial cluster types (see Barker 1964, Clements & Keyser 1983:118, Kingston 1985:262–264, Blevins 1993:254–255, ex. 22). Obstruent-obstruent, obstruent-sonorant, sonorant-obstruent, and sonorant-sonorant clusters are all permitted word-initially, as are obstruent-*ʔ* and sonorant-*h*. (Word-initial *hC* and *ʔC* clusters are not allowed.) This gives us a larger testing ground for the scope of the repetition avoidance effect.

Klamath marks the DISTRIBUTIVE with prefixal partial reduplication. This category attests two types of reduplicants: (i) a CV reduplicant (the C_1-copying pattern) and (ii) a CCV reduplicant (the cluster-copying pattern). Roots in initial /CV.../ always take a CV reduplicant, even if they

undergo syncope[13] to create a stem-initial cluster on the surface (i.e. $\sqrt{C_1V_2C_3...} \rightarrow$ distributive $[C_1V_2\text{-}C_1C_3...]$).[14] With the exception of stop-sonorant clusters, all root-initial clusters exclusively display cluster-copying.[15] Examples of each of these are provided in (30). The one type of cluster where the CV outcome is possible is stop-sonorant (TR). While TR roots may surface with either type, I analyze the CV outcome as the default, with CCV being a lexically restricted alternative type, generated via CONTIGUITY, not via an anti-repetition constraint (see Section 6.6.3 for details). These examples are shown in (31).

The tables provide only representative examples, not an exhaustive list of forms. Note that the LW example represents the only type of sonorant-sonorant root-initial cluster attested with reduplication. For ease of comprehension, in both of the following tables, relevant reduplicated forms are accompanied by a derivationally intermediate form which undoes late phonological rules (e.g. vowel syncope/reduction, cluster simplification, etc.) that do not directly impact the determination of reduplicant shape. All data is from Barker (1964) (abbreviated in the tables below as B).[16]

(30)   CCV reduplication with non-TR clusters (representative examples)

Cluster Type		Root	Reduplicated	(morphophonemic)	Source
ST		/stiːq̇a/	*sti-stiːq̇a*		B:84
		/scidiːla/	*sci-scdiːla*	(//sci-scidiːla//)	B:84
TT		/ktiwcṅa/	*kti-ktoːcṅa*	(//kti-ktiwcṅa//)	B:88
TS		/ksodga/	*kso-ksatga*	(//kso-ksodga//)	B:84
RR	(LW)	/lwosga/	*lwo-lwasga*	(//lwo-lwosga//)	B:89
RT	(NT)	/mbodẏdk/	*mbo-mpditk*	(//mbo-mbodẏdk//)	B:90
	(LT)	/lbogaa/	*lbo-lpga*	(//lbo-lbogaa//)	B:84
	(WT)	/w-qeːwi-'a/	*wqe-wqeːẁa*		B:121
SR	(SN)	/sṅog-y-s/	*sṅo-sngis*	(//sṅo-sṅogys//)	B:89
	(SL)	/sliṅ/	*sli-slan*	(//sli-sliṅ//)	B:92
	(SW)	/swinys/	*swi-soːnis*	(//swi-swinys//)	B:88
CʔΟ	(SʔΟ)	/sʔaba/	*sʔa-sba*	(//sʔa-sʔaba//)	B:85
	(LʔΟ)	/lʔeg-bg-m/	*lʔe-lʔakpgam*	(//lʔe-lʔegbgm//)	B:171

---

[13] Klamath has a syncope/reduction process that applies to the initial vowel of the root (if underlyingly short) when certain types of prefixation occur, one of which being reduplication (see Barker 1964, Blevins 1993:258). There are two outcomes of this process whose distributions can be approximated by syllabification (Clements & Keyser 1983, Kingston 1985, Steriade 1988): the root vowel is deleted if it would surface in an open syllable; it reduces to [ə] < a > if it would surface in a closed syllable. To minimize confusion, I have attempted to use examples without syncope to the greatest extent possible.

[14] This apparent non–surface-oriented behavior poses a problem for a fully parallel approach to repetition avoidance. See Section 6.6.3 for further discussion.

[15] I have found two exceptions in Barker (1964):

(i)   a.   /RED-kċoːsYeːni-'a/ → *ko-kċoːsYeṅa* (not *kċo-kċoːs...) (p.84)

       b.   /RED-qtana/ → *qa-qta* (not *qta-qta) (p.85)   cf. /RED-qday-'aːk̇/ → *qda-qdiːʔaːk* (p.89)

[16] I can provide a spreadsheet containing an exhaustive list of reduplicated forms upon request.

(31)   Reduplication in TR clusters (representative examples)

Cluster	Red. Type	Root	Reduplicated	Source
TN	CV	/dmesga/	_de-dmasga_	B:85
	CV	/pni-ʼaːk̓/	_pi-pnaʔaːk_	B:90
	CCV	/p̓na/	_p̓na-p̓na_	B:92
	var.	/qniy-ʼa/	_qi-qn̓ya ~ qni-qn̓ya_	B:85
TL	CV	/c̓lodga/	_c̓e-c̓latga_	B:82
	CCV	/c̓lid̲ga/	_c̓li-c̓latga_	B:82
	var.	/q̓liq-diːla/ ~ /t̓liq-diːla/	_q̓li-q̓laqdiːla ~ t̓i-t̓laqdiːla_	B:117
TW	CV	/t̓waːYa/	_t̓a-t̓waːYa_	B:85
	CCV	/tya-ʼaːk̓/	_tya-tiːʔaːk (//tya-tyaʔaːk̓//)_	B:89
	CCV	/q̓yabga/	_q̓ya-q̓yapga_	B:89

The total cluster-wise distribution of cluster-copying vs. C_1-copying in Klamath is shown in (32). What we see is that Klamath picks out TR for C_1-copying to the exclusion of all other cluster types. This is essentially the same behavior we observed for Ancient Greek in Section 6.3.2, except that now we have clear evidence for the behavior of _s_-sonorant.

(32)   Cluster-wise reduplicant shape distribution in Klamath

C_2 / C_1	Stop	_s_	_ʔ_	Sonorant (N,L,W)
Stop	Cluster: TTV	Cluster: TSV	no examples found	**C_1-copying: TV ~ Cluster: TRV**
_s_	Cluster: STV	n/a	Cluster: SʔV	Cluster: SRV
Sonorant (N,L,W)	Cluster: RTV	no examples found[17]	Cluster: RʔV	Cluster: RRV

The generalization regarding repetitions here is very clear: only stop-sonorant licenses a repetition. This is the same way the Ancient Greek generalization was framed in (18), prior to the introduction of the variation seen in voiced stop + liquid clusters. (It does not appear that there is any comparable variation in Klamath based on the laryngeal features of the stops.)

(33)   **Klamath repetition generalization:**
        Stop-sonorant clusters license repetitions, other clusters do not.

6.3.4   The Distribution of Repetition Avoidance Effects in Sanskrit

Sanskrit presents the most complicated case for determining the scope of repetition avoidance effects, both because of the system's analytical complexity and because of the difficulty in interpreting the relevant data. Sanskrit has two distinct analytical categories where repetition avoidance

---

[17] The only stem-initial sonorant-_s_ sequence is found in a cluster derived through syncope: /RED-losnc̓na/ → _lo-lsanc̓a_. As mentioned above, in all cases, such forms reduplicate as if the vowel were still present, as if //lo-losanc̓a//.

effects are felt: C_2-copying reduplication is employed to avoid particular repetitions with cluster-initial roots of the relevant shape (Section 6.2.3); but the $C_1\bar{e}C_2$ pattern is employed to avoid repetitions in the perfect weak stems of CaC roots (Section 6.2.4). In this subsection, I will explore the full cluster-wise distributions of these two patterns, paying careful attention to the nature of the data (which is, in a number of ways, frustratingly ambiguous).

The cluster-initial roots present a fairly clear and simple answer (Section 6.3.4.1), which looks minimally different than what we observed for Ancient Greek and Gothic. The distribution of the $C_1\bar{e}C_2$ pattern, on the other hand, is significantly more complex. I will first entertain a solution that understands the data as representing multiple distinct diachronic stages, with a definable mode of change, where the behavior of the *PCR constraint(s) becomes more stringent in the post–Rig-Vedic period (Section 6.3.4.2). However, the analytical system developed to account for the general distribution of C_2-copying and the $C_1\bar{e}C_2$ pattern in Chapter 5 is incapable of accommodating a distribution of this sort (Section 6.3.4.3).

Therefore, I will ultimately argue (in Section 6.3.4.4) that a slightly different interpretation of the data, which requires disregarding two grammarian-prescribed forms, allows for a unitary analysis within the post–Rig-Vedic period, with a more minor change occurring after the (earliest) Rig-Veda. I include the full assessment of the individual patterns on their own terms, rather than proceeding directly to my ultimate conclusions, so that their interpretation may be considered in the future in light of potential alternative analyses which may not be subject to the same analytical complication that motivates the reinterpretation I adopt.[18]

6.3.4.1 Sanskrit Cluster-Initial Roots

The full data illustrating the reduplicative behavior of cluster-initial roots, assembled from the forms compiled by Whitney (1885 [1988]), is provided in (34) below, sorted by cluster type.[19] Like Ancient Greek, Sanskrit displays a robust inventory of root-initial clusters. However, Sanskrit appears to license the default C_1-copying pattern in many more contexts than does Ancient Greek: the alternative pattern — C_2-copying, notated in (34) as the bolded forms — only surfaces for sibilant-stop clusters. For some cells which are robustly attested — for example, C_1-copying to stop-sonorant clusters and C_2-copying to sibilant-stop clusters — only representative examples are provided.

Notation and Transcription

In the chart in (34), empty cells are those for which no cluster-initial roots of the relevant type are attested (at least not with reduplication). Italicized forms are those where the consonant in the reduplicant differs in place and/or manner from its correspondent in the base (due primarily to velar palatalization and the *ruki* rule), as these differences render the consonants non-identical, and thus (potentially) not relevant for an investigation of repetition avoidance effects.[20] Forms marked with brackets [ ] are those which Whitney (1885 [1988]) reports as existing only in the works of the Classical Sanskrit grammarians, rather than being attested in naturally-occurring texts; they are thus

---

[18] Readers who are less concerned with the philological details of Sanskrit may wish to proceed to Section 6.3.4.4, where I lay out the cluster-wise distribution which I will assume in the subsequent exploration of *PCR.

[19] This data set, as also the one in (37) below, collapses together all of attested Sanskrit up through the Classical Period. There do not seem to be substantive differences in the treatment of cluster-initial roots across that time period. I will, however, argue in Section 6.3.4.2 below that there are real differences in the treatment of (would-be) clusters arising from zero-grade ablaut between the Rig-Veda and later periods.

[20] Aspiration is (generally) not permitted in the reduplicant (see the discussion in Chapter 5). There is no indication that a difference in aspiration renders consonants "non-identical" for the purposes of *PCR.

of less certain linguistic reality. Whitney does not report which grammarian(s) cite particular forms, and thus I do not know their exact chronology (though I believe that all of the grammatical texts drawn on by Whitney hail from the Classical period; see Whitney 1885 [1988]:vii).

Two notes on transcription and phonological identification are required. First, the consonants traditionally transcribed as <c> and <j> are undoubtedly palatal obstruents. They are frequently identified as affricates [tʃ] and [dʒ] (or perhaps [tɕ] and [dʑ]). However, the Classical grammarians' phonetic descriptions of these sounds (as reported by Whitney 1889:15–16/§42–44), seem more consistent with palatal stops [c] and [ɟ]. In all respects relevant to the current discussion, they pattern like the other stops. Therefore, I transcribe them as stops [c] and [ɟ] and include them in the stop categories for the discussion in this chapter; however, I do not believe that any of the conclusions drawn would differ if we assigned them affricate interpretations.

Second, I identify the phonetic value of the consonant traditionally transcribed as <v> to be a labiodental "narrow approximant" [ʋ̝] (in the sense of Padgett 2002) — essentially a glide with increased constriction, and possibly some frication.[21] This is because the Sanskrit <v> exhibits a sort of "mixed" behavior phonologically, in some ways patterning like a fricative, in other ways patterning like a sonorant.

Treating <v> as a fricative would make sense of differences in the phonotactic distribution of <v> and <y> (the palatal glide — IPA [j]). (Consult Whitney 1885 [1988]:151–168, Kessler 1994:esp. 37–38, and others on Sanskrit phonotactics.) Namely, <v> may freely appear word-initially before the liquids *r* and *l* (though *l* is itself rare) and before the glide *y*.[22] However, *y* may never appear word-initially before any consonant, nor medially before any consonant; it always vocalizes or coalesces with a preceding vowel. Likewise, consonantal *r* — and probably also *l*, and perhaps even the nasals — may not appear word-initially before <v>. Sanskrit generally only permits rising sonority clusters in root-initial position (with the exception of sibilant-stop), so these facts imply that <v> was less sonorous than the other glides and liquids. This would naturally lead to identifying it as a fricative.

However, there is strong evidence that it is not an obstruent. For one, it retains a number of sonorant-like behaviors reflecting its origin as Proto-Indo-European *w*. Namely, it continues to alternate with *u* in positions where sonorants vocalize, and, in coda position, it monophthongizes with a preceding /a/ to yield [ō]. This is precisely parallel to the behavior of *y*, which vocalizes to *i* and monophthongizes with /a/ to yield [ē] in equivalent positions. Second, and more convincingly, it behaves like a sonorant for the purposes of laryngeal neutralization and laryngeal assimilation (Steriade 1997, 1999). Laryngeal contrasts (voicing and aspiration) are only licensed in pre-sonorant position. Laryngeal contrasts are indeed licensed before <v>. Relatedly, obstruent sequences always agree in voicing (with regressive voicing assimilation). Voiceless obstruents are permitted before <v>. These facts show that <v> cannot be an obstruent.

Padgett (2002) proposes the category of narrow approximant to account for this for similar mixed behavior of *v* in Russian and a number of other languages (see also Barkaï & Horvath 1978, Bjorndahl 2013, 2015, among others). I thus adopt this proposal for Sanskrit as well. Since *v* is the only narrow approximant in the language, and its status as a narrow approximant could bestow upon it different properties with respect to repetition avoidance and *PCR, I give it its own row and column in (34) and the other charts that follow, placing it between the sibilant fricatives and the

---

[21] Whitney (1889:20/§57) reports that the Classical Sanskrit grammarians very distinctly describe it as a labiodental, though he does not report anything that would adjudicate between identification as a true fricative [v], a true glide [ʋ], or something in between, i.e. a "narrow approximant" [ʋ̝].

[22] There are no *v*+nasal–initial roots. There is, though, one base-initial *vn* sequence in reduplication; see below.

nasals (indicating its borderline status between obstruent and sonorant). For typographical reasons, I continue to represent <v> as *v*, but it is intended to encode phonetic/phonological [ʋ̩].[23]

(34)   Reduplicative behavior in Sanskrit cluster-initial roots

C₁ \ C₂	Stop	Sibilant	*v* (= [ʋ̩])	Nasal	Liquid	*y* (= [j])
Stop		tatsāra [papsāu] *cakṣamē* *cakṣadē*	didvēṣa dadʰvansire tatvarē [tatvakṣa] [*cakvātʰa*] jajvāla [jajvāra]	dadʰmāu jajñāu	babʰrāja dadrāu paprācʰa tatrē pupluvē *jagrabʰa* *cakranda* [jijrāya]	dadʰyāu *cakʰyāu* tatyāja didyota jijyāu cucyuvē
h			[*jahvāla*]	[*juhnuvē*] [*jahrāsa*]	*jihrāya*	
Sibilant	**tastʰāu** **tastambʰa** **tastāra** ***caskanda*** ***caskʰāla*** **pasprçē** **paspaçē** ***tiṣṭʰēva*** **[ṭiṣṭʰēva]** ***tuṣṭʰāva*** **[ṭuṣṭʰāva]** **cuccota**		sasvadē sasvaje sasvanur *suṣvāpa* çaçvāsa [çiçvāya]	sasmāra sismāya *siṣmāya* sasnur [*siṣnēha*] [çaçnātʰa]	susrāva sasransur çaçrāma çaçratʰē çaçlāgʰirē	sasyandē *siṣyanda* [çaçyē]
v (= [ʋ̩])					vavrāja	vivyāca vivyādʰa
Nasal				[mamnāu]	mumloca mamlāu	mimyakṣa

Notes on the Data

A few preliminary words are required here regarding the data. First, there is only a single data point for nasal-nasal (NN) roots: $\sqrt{mn\bar{a}}$- 'note' $\rightarrow$ _mamnāu_ (Whitney 1885 [1988]:127). This form is not attested in naturally-occurring texts, but rather only cited by grammarians. I will for the moment accept this as legitimate evidence; however, I will reassess this in Section 6.3.4.3 below.

Second, most of the stop-sibilant (TS) roots begin in a velar stop. Velar-initial bases always exhibit palatal correspondents in the reduplicant (see Whitney 1889:222–223/§590): for example, $\sqrt{k\d{s}ad}$ 'divide' $\rightarrow$ _cakṣadē_ (Whitney 1885 [1988]:27). Since our investigation is concerned with consonant identity, it should be the case that velar-initial roots can always (vacuously) satisfy the repetition avoidance constraints by non-identity. It will indeed be shown below that the facts of the $C_1\bar{e}C_2$ pattern require that non-identity for place exempts repeated consonants from a *PCR violation. Therefore, such forms are not probative for assessing the distributional properties of the repetition avoidance effect.

There are only two reduplicated forms to non-velar-initial TS roots. One is _papsāu_ (Whitney 1885 [1988]:104). Like _mamnāu_, this form too is only cited by grammarians and not attested in literary texts. Furthermore, though listed by Whitney under the lemma $\sqrt{ps\bar{a}}$ 'devour', he states that the root itself is clearly derived from the more basic root $\sqrt{b^has}$ 'devour'. Therefore, it could be the case that this form ought to be interpreted as having a zero-grade base-initial cluster rather than a root-initial cluster. If there were to be differences in the behaviors of clusters that result from zero-grade and clusters that are underlying, then it might not be appropriate to include this as evidence in this category. However, I will ultimately argue that there are not (or, rather, cannot be) differences between the two categories (see Section 6.3.4.3), so this is a moot point, and it should be accepted as evidence.

The last form is _tatsāra_ $\leftarrow$ $\sqrt{tsar}$ 'approach stealthily' (Whitney 1885 [1988]:68). It is the only one whose attestation and membership in the current category cannot be questioned. It should, though, be noted that this form is only attested in the Rig-Veda, which is the earliest attested stage of the language. Therefore, if we come across conflicting diachronic evidence (as we will below), this form can only be taken as secure evidence for the scope of the repetition avoidance effect in Rig-Vedic Sanskrit, not necessarily for that of later Sanskrit.

The Distribution

With these caveats in mind, we can schematize the cluster-wise distribution of default C_1-copying vs. alternative C_2-copying for cluster-initial roots in Sanskrit as shown in (35). (Note that this table does not take into account _h_-initial bases, because they always reduplicate non-identically as _j_, vacuously satisfying the repetition avoidance constraint.) Based on the distribution seen in (35), the class of root-initial clusters for which Sanskrit permits default C_1-copying can be characterized as in (36). (This will be slightly revised later.)

(35)　　Reduplicative behavior in Sanskrit cluster-initial roots

C_1 ＼ C_2	Stop	Sibilant ([s,ç,ṣ])	_v_ ([v̥])	Nasal	Liquid ([r,l])	Glide ([y])
Stop		✓	✓	✓	✓	✓
Sibilant	✗		✓	✓	✓	✓
_v_					✓	✓
Nasal				✓	✓	✓

(36)   **Sanskrit cluster-initial roots repetition generalization:**   [to be revised]
Consonant-sonorant clusters and stop-fricative clusters (*ps*, *ts*, *kṣ*) license repetitions, other clusters do not.
(Conversely: fricative-stop clusters do not license repetitions, all other clusters do.)

This differs from each of the systems seen previously, in at least two ways. First, there is a cluster with an obstruent C_2 that licenses repetitions, namely stop-fricative. This means that drawing the dividing line at obstruent vs. sonorant as the following context (i.e. $*C_\alpha VC_\alpha$ / $\_C_{[-son]}$) is not sufficient to capture the distribution, as that would predict stop-fricative roots should show the alternative pattern.[24] Additionally, Sanskrit shows that all sonorant-sonorant clusters that are attested permit repetitions (though the evidence for nasal-nasal is shaky; see more below). This was definitively not the case for Klamath, where all sonorant-sonorant clusters were attested and all of them showed the alternative cluster-copying pattern.[25] Therefore, it seems that the *PCR constraint operative for Sanskrit will need to be somewhat different than those operative in the other languages.

6.3.4.2   Sanskrit Zero-Grade Bases (The CēC Pattern)

Now let us consider the distribution of repetition avoidance effects in the perfect weak stems, i.e. those categories of the perfect which normally undergo zero-grade ablaut, specifically for CaC roots. The chart in (37) below lays out the attested perfect forms to CaC roots which have either (i) C_1-copying with zero-grade of the root in the perfect weak stem or (ii) the $C_1 \bar{e} C_2$ pattern in the perfect weak stem.[26] I believe this table to be exhaustive, based on forms extracted from Whitney (1885 [1988]) (see also Sandell 2015b:217–220, with careful attention to period of attestation, which will turn out to be important).

Notations and the Data

The chart in (37) is formatted as follows (largely equivalent to the format of (34) above). The type of consonant represented by root-C_1 varies vertically, from least sonorous at the top to most sonorous at the bottom (with *v*, i.e. [ʋ], assumed to have a sonority value intermediate between fricatives and nasals); the type of consonant represented by root-C_2 varies horizontally, from least sonorous on the left to most sonorous on the right. The forms for each type of root-C_1 are divided into two rows: the forms in the top row are those which show C_1-copying (that is, are overtly reduplicative and result in a consonant repetition); the forms in the bottom row are those which show the $C_1 \bar{e} C_2$ pattern (and are thus assumed to disallow C_1-copying because of the type of repetition it would create). What we are trying to understand from this organization is whether the sequence $C_1 VC_1 C_2$ is permitted in each particular consonantal context, under the assumption that those forms showing the $C_1 \bar{e} C_2$ pattern do so because C_1-copying would violate the repetition avoid-

---

[24] Though keep in mind that this is dependent on the assumption that we have secure evidence for C_1-copying in stop-sibilant clusters. See again the preceding discussion for details.

[25] I claimed earlier that, in Ancient Greek, nasal-nasal clusters — the only sonorant-sonorant clusters attested in the perfect — disallowed repetitions, but the evidence was somewhat ambiguous. Gothic attests no sonorant-sonorant clusters in word-initial position.

[26] Other types of perfect weak stems, namely unexpected full grades, are not considered here. However, viewing perfect weak stems with unexpected full grades (i.e. failure to undergo ablaut) as a potential avoidance strategy for *PCR violations, and thus grouping them with the $C_1 \bar{e} C_2$ pattern for the purposes of determining the distribution of *PCR effects, would be a fruitful avenue for further research.

(37)   Perfect weak stems of CaC roots: reduplication vs. $C_1\bar{e}C_2$ (from Whitney 1885 [1988])

C_1 \ C_2	Stop	Fricative	v	Nasal	Liquid	Glide
Stop	paptur	*ɟakṣur* (bapsati)		dadʰmirē tatnē papnē *cakʰnur* ɟagmur ɟaɟɲur	babʰrē dadʰrē dadrē paprur *cakrē*	bibʰyur *cikyur* *ɟigyē* *ɟigʰyur* [cicyē]
	pētur ✓ pēdur ✓ dēbʰur ✓ tēpē ✓ ɟēpur ✗ [cētē] ✗ pēcur ✗ bʰēɟur ✗	dēhē ✗		tēnē ✓ cēmur ✗ [pʰēṇur] ✗ [pēṇē] ✗	tērur ✓ ɟērur (✗) pʰēlirē ✗ cērur ✗ cēlur ✗	
Fricative (*sib.*, *h*)	*saçcur*		*suṣvāṇa* *ɟuhvē*	(sasni) (siṣṇu)	sasrē çaçrē *ɟahrur*	çiçyē
	sēpur ✓ sēdur ✗ çēkur ✗ çēpur ✗ sēcirē ✗ sēɟur ✗	sēhur ✗		[sēnē] ✓ çēmur (✓)	[çēlē] (✗)	
v				vavnē vēmur ✗	vavrē	vivyē
Nasal				mamnātʰē	mamrur	ninyē mimyur
	nēdur ✓ mētʰur ✗ [nēbʰē] ✗	nēçur ✗ [nēhē] ✗		mēnē ✓ nēmē ✓		
Liquid						lilyē
	rēbʰē ✓ rēɟur ✓ [rēdur] ✓ [rēʈur] ✓ lēbʰē (✓) lēpur (✗)	rēsur ✓ lēṣur (✓) rēhur (✗)		rēmē ✓		
Glide	yētē ✗ yēɟē ✗			yēmur ✗		

ance constraint by creating a disallowed consonant repetition (though there is more to say on this; see below).

Following the earlier discussion, I continue to assume that <v> represents a narrow approximant [ʋ]. In (37), *h* is collapsed with the sibilants under the heading "fricative". It largely appears as though the only relevant differences between the two types of fricatives has to do with their phonotactics.

Additional notations are as follows (again largely following that of (34)). Empty cells are those for which there are no CaC roots with a perfect weak stem of the relevant sort. Italicized forms are those in which the reduplicated consonant is non-identical to root-C_1 in place and/or manner, either due to velar palatalization in the reduplicant or place assimilation of root-C_1 to root-C_2. (I again assume that reduplicated forms where the base-reduplicant corresponding consonants differ in place and/or manner satisfy the repetition avoidance constraint via non-identity, and therefore that they are potentially not relevant for determining the nature of the repetition avoidance effect. I make this reasoning more explicit below.) Forms enclosed by brackets [ ] are those which Whitney (1885 [1988]) reports as being cited only in grammatical texts and not in naturally-occurring texts.

Forms enclosed by parentheses ( ) are presents or other derivatives of CaC roots in the zero-grade which appear to have (or clearly do have) reduplication. Each $C_1\bar{e}C_2$ form is accompanied by either a ✓ or a ✗. A ✓ indicates that the cluster which would result from zero-grade would be phonotactically legal (if in parentheses, it means that it would be only marginally so). A ✗, on the other hand, indicates that the cluster would be phonotactically illegal (if in parentheses, it means that I am not completely sure of its status in this regard), assuming that the cluster were not to be repaired by any sort of feature change (other than voicing assimilation). (See Kessler 1994, Sandell 2015b:230–232, 2017:5–6 for formal discussion of Sanskrit cluster phonotactics.) The double solid vertical line represents the boundary for the repetition avoidance constraint in the cluster-initial roots: violation to the left, satisfaction to the right. The single solid vertical line separates roots where C_2 is an obstruent from roots where C_2 is a sonorant.[27]

Honing in on the Relevant Data

The abundance and heterogeneity of data in this table makes it somewhat difficult to interpret. Since we are trying to determine the distribution of the repetition avoidance effect, there are several types of data which we might reasonably be able to factor out in order to make finding the relevant generalizations easier.

First, C_1-copying forms where the reduplicant consonant and its base correspondent are non-identical are potentially also not relevant for repetition avoidance. If the *PCR constraint indeed requires total identity in order to assign a violation, then such forms would vacuously satisfy it. I will argue below that place and manner identity is required in order to satisfy the identity condition

---

[27] The class of nasal-stop roots may contain some members which attest C_1-copying forms with the string NVNT, but each possible example is highly ambiguous. To the root which Whitney (1885 [1988]:117) enters as $\sqrt{mat^h}$, *mant^h*- 'shake', we could imagine the forms with medial [n] (e.g. present *mant^{h}ati*) as being originally reduplicated: $\sqrt{mat^h} \rightarrow$ //<u>ma</u>-mt^h-// $\rightarrow$ *mant^h*-. The same situation obtains for $\sqrt{mad}$, *mand*- 'be exhilarated, exhilarate' (p. 118). (For discussion of possible interpretations of forms from this root, see Sandell 2015b:220, fn. 11.) Another root given by Whitney is $\sqrt{nand}$ 'rejoice' (pp. 87–88), with a present *nandati*. He implies that this root is to be connected with $\sqrt{nad}$ 'sound'. It might alternatively be possible to connect it with $\sqrt{mad}$, with reduplicative copying of the place-assimilated nasal.

While these could logically be viewed as reduplicated in origin, it is difficult to rule out other explanations (e.g., the medial [n] being etymological or being originally the nasal infix). In all cases, the forms with the doubled nasals are attested already in Vedic. In all cases, the NVNT sequences never represent reduplication + zero-grade of the root *as a perfect weak stem*, only (potentially) in other stems. Perfect weak stems of these roots with unexpected full-grade, though, do contain NVNT sequences *in the root*: for example, $\sqrt{nand} \rightarrow$ perfect <u>na</u>-**nand**-*ur*. How exactly to interpret this collection of data is thoroughly unclear. I will not consider it further here.

of *PCR. (The most convincing evidence is presented immediately below.) Therefore, I will factor out any form recorded with italics in (37) (except those where the difference in place is between dental and retroflex; see below).

Second, any cluster that would arise from zero-grade which would be phonotactically illicit might not provide evidence of repetition avoidance *per se*. This is because the C_1-copying form could potentially be ruled out by phonotactics independent of the operation of the repetition avoidance constraint.[28] This can be best understood by considering the potential interactions between four relevant constraint types: cluster markedness constraints, Input-Output faithfulness constraints, the *PCR constraint, and the constraint(s) militating against the $C_1\bar{e}C_2$ pattern.

We know that the $C_1\bar{e}C_2$ pattern is a means that the grammar has at its disposal for avoiding phonotactically impermissible structures, but that this is only employed under compulsion of a higher-ranked constraint violation. So far, I have only shown that this strategy is used to avoid the repetitions banned by *PCR. Using a cover constraint "NO-$C_1\bar{e}C_2$" to stand in for whatever constraint(s) militate against the $C_1\bar{e}C_2$ mapping,[29] we can schematize this ranking as follows:

(38)   Ranking for *PCR-driven $C_1\bar{e}C_2$ mappings:
        *PCR $\gg$ NO-$C_1\bar{e}C_2$

Given richness of the base, we also know that any phonotactically illegal cluster in the input maps onto some phonotactically legal sequence — either a legal cluster, or a corresponding string where that cluster has been repaired by deletion or epenthesis (even if we might not know exactly which one due to the absence of alternations). This means that, for each illegal cluster, the markedness constraint against that cluster dominates at least one faithfulness constraint, i.e. the faithfulness constraint violated by the repair mapping. We can express this in schematic terms with the following ranking:

(39)   Ranking for phonotactically illegal clusters:
        *C_1C_2 $\gg$ FAITH-IO

Now, if applying zero-grade ablaut (i.e. root-vowel deletion) to a CaC root in the perfect would bring about a phonotactically illegal cluster, there are two ways that the problem could be solved. One is to allow zero-grade to apply and repair the resulting cluster via the operation governed by the lowest ranked relevant faithfulness constraint — i.e. violate FAITH-IO. This would be concomitant with C_1-copying, which would provide evidence for *PCR *if selected as the optimal mapping*. The other option is to eschew vowel-deletion and reduplicative copying altogether, and instead use the $C_1\bar{e}C_2$ mapping — i.e. violate NO-$C_1\bar{e}C_2$. If we had the ranking FAITH-IO $\gg$ NO-$C_1\bar{e}C_2$ (as given in (40)), then the grammar would select the $C_1\bar{e}C_2$ output. Therefore, under such a ranking, the $C_1\bar{e}C_2$ pattern can come about without any influence from *PCR.

(40)   Ranking for faithfulness-driven $C_1\bar{e}C_2$ mappings:
        *C_1C_2 $\gg$ FAITH-IO $\gg$ NO-$C_1\bar{e}C_2$

The same sort of interaction can arise even with phonotactically *legal* clusters. A phonotactically legal cluster is characterized by the reverse of the ranking in (39), as given in (41). This ranking

---

[28] This point has also been articulated by Sandell (2017:6–7). See Sandell (2015b:Ch. 8) on the importance of phonotactics in the $C_1\bar{e}C_2$ pattern.

[29] In Chapter 5, I considered several possible analyses of the $C_1\bar{e}C_2$ pattern. In the one which I advocated for attested Sanskrit (the phonologically conditioned allomorphy approach), the constraint which played the role of "NO-$C_1\bar{e}C_2$" was "PRIORITY" — a constraint penalizing selection of a dispreferred allomorphy (cf. Mascaró 2007).

means that employing some repair of the relevant cluster based on any unfaithful will be suboptimal, because the markedness constraint is not ranked high enough.

(41)   Ranking for phonotactically legal clusters:
        FAITH-IO $\gg$ *C_1C_2

However, even if the cluster is legal, this entire ranking fragment still has the potential to rank above NO-C_1ēC_2. If this were the case, even though the cluster is normally tolerated, it would still be preferable to not create the cluster if it can be avoided by diverting to the C_1ēC_2 pattern. This ranking is given in (42). This sort of ranking would not be unexpected for clusters which are phonotactically "marginal", that is, clusters which are occasionally tolerated when underlying, but still fairly infrequent.

(42)   Ranking for markedness-driven C_1ēC_2 mappings:
        FAITH-IO $\gg$ *C_1C_2 $\gg$ NO-C_1ēC_2

We thus have two possible ranking conditions involving cluster markedness ((40) and (42)) that can derive C_1ēC_2 outputs without any reference to *PCR. Therefore, any roots whose zero-grade cluster would not be (robustly) phonotactically legal cannot serve as solid evidence for the scope of *PCR. That is to say, we could not be sure whether the C_1ēC_2 output for that root was driven by a *PCR violation or a cluster-related markedness or faithfulness violation. For this reason, I additionally factor out any C_1ēC_2 form in (37) with an unqualified ✗. I continue to entertain the validity of forms with qualified ✗'s, but they remain questionable.

The importance of separating out *PCR-driven C_1ēC_2 forms from markedness/faithfulness-driven C_1ēC_2 forms can be seen vividly in a distinction between the two types of perfect weak stems attested for the root $\sqrt{sac}$ 'accompany' (Whitney 1885 [1988]:182). In the Rig-Veda (the earliest period of the attested Sanskrit), there are C_1-copying forms like *saçcur*, where the root-initial *s* assimilates in place to the root-final palatal stop. In later Sanskrit (from the Atharva-Veda onward), this root instead exhibits a C_1ēC_2 perfect weak stem, as in *sēcirē*. What this is showing us is a change not in the scope of the *PCR constraint, but in the willingness of root-initial consonants to assimilate.

At all stages of the language, sibilants agree in place with a following coronal (i.e. dental vs. palatal vs. retroflex). I'll use the constraint *sc to stand in for the markedness condition that induces assimilation, since this is the relevant cluster for the present case. To ensure assimilation, we require a ranking of the type in (39) for phonotactically illegal clusters, as follows:

(43)   Ranking for sibilant place assimilation:
        *sc $\gg$ IDENT[place]-IO

This ranking remains consistent throughout the language. What changes is the relative ranking of IDENT[place]-IO and NO-C_1ēC_2. In Rig-Vedic (44a), NO-C_1ēC_2 outranks IDENT[place]-IO, such that it is preferable to undergo place assimilation than to divert to the C_1ēC_2 pattern. When the C_1-copying output with faithful realization of the underlying root-initial /s/, *sasc-* (45a), is ruled out by *sc (or indeed *PCR; see more on this below), this ranking selects the place-assimilated C_1-copying candidate *saçc-* (45b). On the other hand, in later Sanskrit (44b), the ranking is reversed, and IDENT[place]-IO has come to dominate NO-C_1ēC_2. This means that place assimilation is no longer preferable to diversion to the C_1ēC_2 pattern. Now, when the markedness constraint(s) rule out *sasc-* (46a), the ranking selects the C_1ēC_2 candidate *sēc-* (46c).

(44) a. Ranking in Rig-Vedic: No-C_1ēC_2 ≫ IDENT[place]-IO
　　　b. Ranking in later Sanskrit: IDENT[place]-IO ≫ No-C_1ēC_2

(45) C_1-copying + place assimilation in Rig-Vedic

/RED, sac-/		*PCR	*sc	No-C_1ēC_2	IDENT[place]-IO
a.	sa-sc-	*!	*!		
b. ☞	sa-çc-				*
c.	sēc-			*!	

(46) C_1ēC_2 mapping in later Sanskrit

/RED, sac-/		*PCR	*sc	IDENT[place]-IO	No-C_1ēC_2
a.	sa-sc-	*!	*!		
b.	sa-çc-			*!	
c. ☞	sēc-				*

　　　Understanding this interaction is important for our investigation of the scope of *PCR effects. For one, it shows that we do not necessarily need to explain forms of the *sēc-* sort using *PCR: it is a combination of orthogonal markedness and faithfulness constraints that generates the form. Second, it helps delineate what sorts of transvocalic consonant pairs meet *PCR's identity condition. In order to allow a form like *saçc-*, it must be the case that consonants which differ in place do not violate *PCR. There is no question that $S_\alpha VS_\alpha T$ repetitions violate *PCR; this is the core the type of repetition which is avoided in all of the languages with any *PCR effect. Therefore, if sibilants which differed in (minor) place counted as identical, *saçc-* would violate *PCR, and the derivation should be diverted to the C_1ēC_2 mapping regardless of position of faithfulness. Since this is not the case, we can conclude that non-identity for place excludes a sequence from evaluation by *PCR.

　　　This has ramifications for the cluster-initial roots. Consider the STVX– C_2-copying form *tuṣṭāva* ← √*stav/stu* 'praise' (Whitney 1885 [1988]:193). The root-initial /s/, and indeed the root-second /t/, undergo the *ruki* rule (triggered by the reduplicant *u*) and surface as retroflex. The consonant in the reduplicant is [t], matching the place of its correspondent in the underlying root, not in the output base. (Reduplicant consonants in Sanskrit tend to be more faithful to their correspondents in the underlying root than in the output base.) Given that mismatch in place between base and reduplicant is licensed in this very form, and that forms like *saçc-* show that transvocalic consonants differing only in place vacuously satisfy *PCR, why does the application of *ruki* not license a C_1-copying output ***suṣṭāva?

　　　To sustain the preceding analysis, it must be the case that *some* types of place differences are not significant enough to trigger non-identity for the purposes of *PCR. A retroflex vs. dental distinction is a reasonable candidate for this sort of non-significant non-identity, given how tenuous a contrast that is to begin with (see Steriade 1997). This is the only place difference for which there is evidence of *PCR violation (and thus sufficient identity), so I will proceed under the assumption at all other place differences do in fact render consonants non-identical for the purposes of *PCR, though this is a question that deserves further attention.

The Pared-Down Distribution

Excluding these forms significantly clarifies the distribution. The forms that remain after this pruning are provided in (47) below; in categories that continue to be well-populated, I list only repre-

sentative examples, focusing on those with the fewest uncertainties surrounding their attestation (namely, ones which are clearly phonotactically licit, and occur outside of grammatical texts). The notation in this chart is slightly different than those above. Parentheses ( ) now indicate $C_1\bar{e}C_2$ forms whose would-be zero-grade cluster is phonotactically marginal or unclear. (The ✓ vs. ✗ distinction is no longer necessary, as all forms with unqualified ✗'s have been removed.) The three remaining non-perfect forms — *bapsati* (a reduplicated present from the root $\sqrt{b^h as}$ 'devour'; Whitney 1885 [1988]:109) and *sasni* and *siṣṇu* (reduplicated derivatives from the root $\sqrt{san}$ 'gain'; Whitney 1885 [1988]:183) — are now enclosed in curly braces { }. Square brackets [ ] continue to signify grammarian forms.

(47)   Core *PCR distribution of perfect weak stems of CaC roots

C_1 \ C_2	Stop	Fricative (*sib.*, *h*)	*v*	Nasal	Liquid	*y*
Stop	paptur	{bapsati}		tatnē · dadʰmirē · papnē · ɟaɟṇur	babʰrē · dadʰrē · paprur	bibʰyur · [cicyē]
	pētur · dēbʰur · tēpē			tēnē	tērur	
Sibilant			suṣvāṇa	{sasni} · {siṣṇu}	sasrē · çaçrē	çiçyē
	sēpur			[sēnē] · (çēmur)	([çēlē])	
v				vavnē	vavrē	vivyē
Nasal				mamnātʰē	mamrur	ninyē · mimyur
	nēdur			mēnē · nēmē		
Liquid						lilyē
	rēbʰē · rēɟur · [rēdur]	rēsur · (lēsur) · (rēhur)		rēmē		

A somewhat more interpretable picture emerges from (47). For most cells, there is no variation at all; only a single outcome is attested. The generalizations which are categorical (or nearly categorical) can be stated as in (48). I will return to the exceptions to these generalizations below.

(48)   a.   All roots with glide *y* as root-C_2 show C_1-copying.

       b.   If we discount the (grammarian only) form *çēlē*, whose would-be zero-grade cluster *-çl-* is marginally licit at best, then all liquid-final roots also show C_1-copying, with the exception of *tērur* ($\leftarrow \sqrt{tar/tr}$ 'pass'; Whitney 1885 [1988]:64).

       c. C_1-copying holds for all roots of the shape stop+nasal, with the exception of *tēnē* ($\leftarrow \sqrt{tan}$ 'stretch'; Whitney 1885 [1988]:60).

       d.   The $C_1\bar{e}C_2$ pattern holds for all liquid-initial roots, except when root-C_2 is the glide *y*. Note that a glide would be the only type of consonant of greater sonority than a liquid.

       e.   The $C_1\bar{e}C_2$ pattern holds for all stop-final roots, with the exception of *paptur* ($\leftarrow \sqrt{pat}$ 'fly, fall'; Whitney 1885 [1988]:94).

       f.   The $C_1\bar{e}C_2$ pattern holds for all sonorant-fricative roots.

       g. C_1-copying holds for the one example of stop+fricative *bapsati* (though note that it is not a perfect, and that it has a mismatch in voice).

       h. C_1-copying holds for the one example of sibilant+*v* *suṣvāṇa* (though note that it has a dental~retroflex mismatch; see preceding discussion).

The two cells where there seems to be substantive variation between the two patterns are the remaining nasal-final roots, i.e. sibilant+nasal and nasal+nasal. However, when we dig deeper into the attestations of the distinct types, a further generalization emerges that may make sense of the variation. In cases of conflicting patterns for the same type, it seems to be the case that the C_1-copying form is attested in the Rig-Veda (the oldest stage of attested Sanskrit) and not in later stages, but the $C_1\bar{e}C_2$ form is attested only in later stages (often not until the Classical Period) and not in the Rig-Veda.

Most telling are the cases of doublets from the same root. For the root $\sqrt{san}$ 'gain' (Whitney 1885 [1988]:183)), the (non-perfect) derivative *siṣṇu* is attested only in the Rig-Veda and the (non-perfect) derivative *sasni* is attested in Vedic and in the Brāhmaṇa (a post-Vedic text); on the other hand, this root attests a $C_1\bar{e}C_2$ perfect weak stem *sēnē* only in Classical Sanskrit grammatical texts. Likewise, for root $\sqrt{man}$ 'think' (Whitney 1885 [1988]:118), a C_1-copying perfect weak stem *mamnāt^{h}ē* is attested only in the Rig-Veda, while a $C_1\bar{e}C_2$ perfect weak stem *mēnē* is attested in the Brāhmaṇa. (Sandell 2015b:220 cites its first attestation as the Atharva-Veda, a Vedic text that post-dates the Rig-Veda; the point remains the same.) Relevant too here are the forms of the root $\sqrt{sac}$ 'accompany' (Whitney 1885 [1988]:182) discussed earlier: Rig-Vedic *saçcur* vs. later *sēcirē* (from the Atharva-Veda onward).

Even among the non-doublets, this diachronic distribution generally holds. The C_1-copying form *vavnē* ($\leftarrow \sqrt{van}$ 'win'; Whitney 1885 [1988]:153) is attested in the Rig-Veda (but not in later periods); while the $C_1\bar{e}C_2$ form *çēmur* ($\leftarrow \sqrt{çam}$ 'be quiet'; Whitney 1885 [1988]:171) is attested from the Brāhmaṇa onward. The only counter-example to this trend within these categories is *nēmē* ($\leftarrow \sqrt{nam}$ 'bend, bow'; Whitney 1885 [1988]:88), which is attested from the Rig-Veda onward. However, there are reasons to expect an asymmetry in the treatment of coronal-labial (*nm-*) vs. labial-coronal (*mn-*) in favor of the latter (Jun 2004:63–64; see Section 6.5.3 below).

This diachronic distribution holds not only of the pre-nasal fricatives and nasals, but also generally of the exceptions within the root-initial stop classes. We again have several doublets. For the root $\sqrt{tan}$ 'stretch' (Whitney 1885 [1988]:60), the C_1-copying form *tatnē* is attested in the Rig-Veda, while the $C_1\bar{e}C_2$ form *tēnē* is attested first in the Atharva-Veda and continued in subsequent periods (Sandell 2015b:220). For the root $\sqrt{pat}$ 'fly, fall' (Whitney 1885 [1988]:94), both the C_1-copying

form *paptur* and the $C_1\bar{e}C_2$ form *pētur* are attested in the Rig-Veda; however, *paptur* occurs only in the books which are known to be older, while *pētur* occurs only in the books which are known to be younger (Sandell 2015b:220). Only *pētur* is attested outside of the Rig-Veda (i.e. in later stages). It must be noted that other $C_1\bar{e}C_2$ forms for stop+stop roots — e.g., *dēb^{h}ur* ($\leftarrow \sqrt{dab^h}$ 'harm'; Whitney 1885 [1988]:70), *tēpē* ($\leftarrow \sqrt{tap}$ 'heat'; Whitney 1885 [1988]:70) — are attested in the Rig-Veda (Sandell 2015b:218). However, it seems likely that these are explainable in the same way as *nēmē* (see Section 6.5.3).

The last, and perhaps most unexpected, exception is *tērur* (from root $\sqrt{tar/tr}$ 'pass'; Whitney 1885 [1988]:64). Stop-liquid is everywhere else a robust class for the default C_1-copying pattern (i.e. exempt from repetition avoidance effects); therefore, the existence of a $C_1\bar{e}C_2$ form rather than reduplication here is thoroughly unexpected. Nevertheless, it fits the chronological pattern seen with the other exceptions: it is attested first in Epic Sanskrit, a relatively late period of the language. In the Rig-Veda, this root attests reduplicated perfect weak stems in *titir-* and *tutur-*.[30] (Furthermore, the root also attests a reduplicated perfect weak stem with unexpected full-grade *tatarē* beginning in Classical Sanskrit.) While it is unlikely that *tērur* could receive an explanation based directly on repetition avoidance,[31] the chronology of its development as an exception fits with the broader pattern of an increasingly stringent context for reduplication in the perfect weak stem over time.

These diachronic distributions of variants and exceptions suggest that we are dealing with two slightly different versions of *PCR for different stages of the language (as concluded also by Sandell 2017).[32] This is outlined in the table in (49) below. "RED" represents a category which exhibits C_1-copying, and "CēC" represents a category which exhibits the $C_1\bar{e}C_2$ pattern.[33] Shaded cells indicate the categories where there is a consistent diachronic change between the Rig-Veda and following stages; the " > " represents the direction of this change. For stop-stop, fricative-nasal, and nasal-nasal, we see a change from reduplication in the Rig-Veda to the $C_1\bar{e}C_2$ pattern in subsequent stages. In stop-nasal and stop-liquid, we see the development of exceptional $C_1\bar{e}C_2$ forms only in the later period (indicated by "[> CēC]"), matching the direction of consistent change in the shaded cells. I record *v*-nasal and sibilant-*v* as "RED (> ?)", because the only data points in these categories — *vavnē* and *suṣvāṇa*, respectively — are attested only in the Rig-Veda (Whitney 1885 [1988]:153, 187); therefore, we can not be sure of the treatment of these cluster types in later periods.

---

[30] The alternate vowels result from the treatment of the sequence *rr < *rh_2, which arose in zero-grade forms with vowel-initial suffixes. The root in PIE terms is *$terh_2$ (Rix et al. 2001:633).

[31] It can, though, be noticed that the two most unexpected exceptions involve unaspirated coronal [t] before a coronal sonorant [r,n]. Perhaps it is something about the interaction between the identity in place and the properties of unaspirated stops vis-à-vis *PCR that leads to the prohibition on reduplication in these cases.

[32] Or perhaps it is rightly three distinct stages. Recall that the stop-stop roots already exhibit $C_1\bar{e}C_2$ forms in the Rig-Veda (at least the later Rig-Veda), the only exception being (slightly older) *paptur* beside (slightly younger) *pētur*. Unless it can be shown that these $C_1\bar{e}C_2$ forms are *all* restricted to the younger books of the Rig-Veda, it might be appropriate to identify the period of change of the stop-stop roots as some period shortly prior to the earliest Rig-Veda, and let *paptur* stand as an archaism already within the Rig-Vedic period.

[33] The sole data point in the stop-fricative cell is *bapsati*, a reduplicated present to the root $\sqrt{b^h as}$ 'devour' (Whitney 1885 [1988]:109). This form is attested already in Vedic. Although it is not a perfect weak stem, it is a zero-grade reduplicated form to a CaC root. Since it matches the behavior of the cluster-initial form *tatsāra* (see Section 6.3.4.1), I tentatively include it as evidence for the behavior of reduplication in zero-grade categories for both stages.

(49)   The distribution of perfect weak stems of CaC roots across stages

C_1 \ C_2	Stop	Fricative	*v*	Nasal	Liquid	*y*
Stop	RED > CēC	RED		RED [> CēC]	RED [> CēC]	RED
Sibilant	CēC		RED (> ?)	RED > CēC	RED	RED
v				RED (> ?)	RED	RED
Nasal	CēC			RED > CēC	RED	RED
Liquid	CēC	CēC		CēC		RED

With this established, we can now make our generalizations about which sequences were and were not permitted by the repetition avoidance constraint in the determination of C_1-copying vs. the C_1ēC$_2$ mapping for the perfect weak stem of CaC roots in the two distinct periods. The generalizations regarding repetitions in the Rig-Vedic period (the earlier stage) are provided in (50); the generalizations for the later stages of Sanskrit (beginning probably in the Atharva-Veda, or even already in the late Rig-Veda) are provided in (51).

(50)   Distribution of repetition avoidance effects for CaC roots in the Rig-Veda

    a.   Satisfies repetition avoidance constraint(s) when accompanied by C_1-copying:
        i.   All stop-initial clusters
        ii.   All sonorant-final clusters, except *liquid-nasal*
    b.   Violates repetition avoidance constraint(s) when accompanied by C_1-copying:
        i.   All stop-final clusters, except *stop-stop*
        ii.   All liquid-initial clusters, except *liquid-glide*

(51)   Distribution of repetition avoidance effects for CaC roots in later Sanskrit   [to be revised]

    a.   Satisfies repetition avoidance constraint(s) when accompanied by C_1-copying:
        i.   All non-nasal sonorant-final clusters
        ii.   All stop-initial clusters, except *stop-stop*
        iii.   (*v*-nasal clusters?)
    b.   Violates repetition avoidance constraint(s) when accompanied by C_1-copying:
        i.   All obstruent-final clusters, except *stop-fricative*
        ii.   All nasal-final clusters, except *stop-nasal* (and *v*-nasal?)

As stated, these generalizations do not seem terribly cohesive. In large part, however, these could be recast in terms of the relative sonority of the clusters: with a few exceptions (in both directions), the clusters which license repetitions have *rising sonority* (cf. Parker 2002, 2008). I will now show that a slight reinterpretation of the data, which is independently motivated by the structure of the analysis, can lead to a more parsimonious generalization in terms of relative sonority.

6.3.4.3   Reconciling the Domains

The repetition generalizations just proposed for the CaC weak stems are problematic. First, as just mentioned, there appears to be little internal structure to the generalizations when framed in terms of the manner classes involved in different cluster types. Second, the generalization for later Sanskrit is not consistent with the generalizations regarding cluster-initial roots, as stated in (36); namely, under the current formulation of the generalizations, sibilant-nasal clusters and nasal-nasal clusters

license repetitions for cluster-initial roots but not for CaC weak stems. It might not be problematic *a priori* to assume different cluster-wise behaviors in the two categories. However, the interrelated analysis of the C_2-copying and $C_1\bar{e}C_2$ patterns developed in Chapter 5 is not compatible with this sort of distinct behavior.

As illustrated by mappings like $\sqrt{sn\bar{a}}$ 'bathe' $\rightarrow$ (perfect weak stem) *sasnur* (see (34) above), which is attested in the Classical period (Whitney 1885 [1988]:195), roots whose initial cluster is sibilant-nasal exhibit C_1-copying not C_2-copying. Other examples include *sismāya* and *çiçnāt^{h}a*. This indicates that $S_\alpha VS_\alpha N$ sequences do not violate the anti-repetition constraint relevant to that mapping. This contrasts with a $C_1\bar{e}C_2$ form like *sēnē*, which would seem to indicate that the anti-repetition constraint relevant to that mapping *is* violated by $S_\alpha VS_\alpha N$ sequences.

One's initial reaction might be to index a stricter version of *PCR to the perfect weak stem (i.e. zero-grade categories), so as to rule out $\sqrt{san} \rightarrow$ ****<u>sa</u>-sn-ē* but not $\sqrt{sn\bar{a}} \rightarrow$ ****<u>sa</u>-sn-ur*. This cannot be correct, because the cluster-initial form *sasnur* itself appears in a zero-grade category. One could appeal to potential *phonetic* differences between underlying clusters and clusters arising through derivation,[34] if the repetition avoidance constraints were sensitive to phonetic properties instead of, or in addition to, phonological features; however, I do not know of any reports of such phonetic differences.

Therefore, it seems that there can only be a single constraint penalizing $S_\alpha VS_\alpha N$ sequences, positioned at one spot in the rankings. (I abbreviate this constraint below as $*S_\alpha VS_\alpha N$; this is a stand in for *PCR, focusing just on whatever aspect of it rules out this particular type of repetition.) If this constraint were to rule out such a repetition for the CaC weak stems (leading to a $C_1\bar{e}C_2$ form), it would by transitivity have to induce the C_2-copying pattern for equivalent cluster-initial roots. I illustrate this in the tableaux in (52) below with the phonologically conditioned allomorphy analysis developed in Chapter 5.[35]

(52)   Treatment of SN roots in different categories

/ {RED₁ > -ē₋₂}, san, é /	LIN-IR	$*S_\alpha VS_\alpha N$	PRIORITY	ALIGN-RT-L	ANCHOR
a. <u>sa</u>-sn-ē (1)		*!		**	
b. <u>na</u>-sn-ē (1)	*!			**	*
c. ☞ sēn-ē (2)			*		

/ {RED₁ > -ē₋₂}, smay, a /	LIN-IR	$*S_\alpha VS_\alpha N$	PRIORITY	ALIGN-RT-L	ANCHOR
a. ☺ <u>si</u>-smāy-a (1)		*!		**	
b. ☻ <u>mi</u>-smāy-a (1)				**	*
c. smēy-a (2)			*!		

Specifically, in order for $*S_\alpha VS_\alpha N$ to induce the $C_1\bar{e}C_2$ pattern for SaN weak stems, $*S_\alpha VS_\alpha N$ would have to outrank the constraint penalizing the $C_1\bar{e}C_2$ mapping; in the phonologically conditioned allomorphy analysis shown here, this constraint is PRIORITY (cf. Mascaró 2007). PRIORITY must outrank the constraints which are violated by the C_2-copying pattern — namely, ALIGN-ROOT-L and, more crucially, ANCHOR-L-BR. (See Chapter 5 for ranking arguments.) Therefore, if $*S_\alpha VS_\alpha N$ dominates PRIORITY, it also dominates ANCHOR-L-BR. This means that if $*S_\alpha VS_\alpha N$ truly does trigger the $C_1\bar{e}C_2$ pattern for SaN weak stems, it must also trigger it for SNVX– cluster-

---

[34] I will suggest that such a situation may hold of Klamath; see Section 6.6.3 below.

[35] I believe that the same problems hold of the phonologically driven allomorphy analyses considered in Chapter 5, as well.

initial roots. This is clearly not the case, meaning that there is something wrong with the current picture.

I believe that the source of the problem lies not in the formal analysis, but in the cluster-wise generalizations. As mentioned above, the two main loci of conflict between the two categories are sibilant-nasal and nasal-nasal. In both cases, the conflicting evidence is of questionable legitimacy. For sibilant-nasal, there are two $C_1\bar{e}C_2$ examples in (47) which conflict with the (robustly instantiated) C_1-copying pattern for cluster-initial roots. The first is *sēnē*, which is a grammarian cited form, and so of questionable linguistic reality; the other is *çēmur*, which is only marginally phonotactically licit, and thus could potentially be explained by phonotactics (see the earlier discussion regarding phonotactics and repetition avoidance). If these forms are thus not taken as legitimate evidence, then the evidence for the $C_1\bar{e}C_2$ pattern for sibilant-nasal disappears. It is then possible to assume that the regular treatment of sibilant-nasal CaC zero-grade clusters in later Sanskrit would indeed have been C_1-copying, but we just lack solid examples.[36]

Coming from the reverse direction, the only example for the behavior of nasal-nasal cluster-initial roots, from any period, is the grammarian cited form *mamnāu*. Among the CaC roots, we saw that there was a change from C_1-copying in the Rig-Veda — *mamnat^{h}ē* — to the $C_1\bar{e}C_2$ pattern in the later stages — *mēnē*. If we disregard *mamnāu* in the same way as *sēnē*, then it is possible to assume that this change which is evident in the actually attested weak stems would have been tracked by the behavior of cluster-initial nasal-nasal roots, which happen to not exist.[37]

6.3.4.4 The Reconciled Distribution of Repetition Avoidance Effects in Later Sanskrit

If we amend our interpretation of the data in this way (i.e. disregarding certain grammarian-only forms and phonotactically questionable forms), we arrive at new distributions across the two categories in later Sanskrit. The table in (53) shows the updated distribution for cluster-initial roots (cf. (35) above); the table in (54) shows the updated distribution for CaC roots (cf. (49) above). What we see is that the distributions are now fully compatible (that is, there are no conflicting cells). Putting these together, we get the distribution shown in (55), which can be taken as the total distribution of repetition avoidance effects in the reduplicative system of later Sanskrit.

(53)   Amended distribution of *PCR effects for cluster-initial roots in later Sanskrit

C$_1$ \ C$_2$	Stop	Fricative	v	Nasal	Liquid	Glide
Stop		✓	✓	✓	✓	✓
Fricative	✗		✓	✓	✓	✓
v					✓	✓
Nasal				✗	✓	✓
Liquid						

---

[36] The ideal situation here would be to *wug*-test speakers. Obviously, this is not possible.

[37] The "root" $\sqrt{mn\bar{a}}$ 'note' is a secondary creation from the root $\sqrt{man}$ 'think', the root of the aforementioned weak stems (Whitney 1885 [1988]:118, 127).

(54)   Amended distribution of *PCR effects for CaC roots in later Sanskrit

C₁ \ C₂	Stop	Fricative	*v*	Nasal	Liquid	Glide
Stop	✗	✓		✓	✓	✓
Fricative	✗		?	✓	✓	✓
v				?	✓	✓
Nasal	✗			✗	✓	✓
Liquid	✗	✗		✗		✓

(55)   The distribution of *PCR effects across categories in later Sanskrit

C₁ \ C₂	Stop	Fricative	*v*	Nasal	Liquid	Glide
Stop	✗	✓	✓	✓	✓	✓
Fricative	✗		✓	✓	✓	✓
v				?	✓	✓
Nasal	✗			✗	✓	✓
Liquid	✗	✗		✗		✓

    None of the emendations enacted here have effected the distributional facts of the Rig-Vedic period. When the CaC facts of this period are integrated with the evidence from cluster-initial roots (for which there is no obvious evidence of diachronic change between the earlier and later periods), we arrive at the distribution in (56). The three differences relative to (55) are boxed. These differences are: (i) there is direct evidence at this stage that *v*-nasal licenses repetitions; and (ii) stop-stop and nasal-nasal license repetitions at this stage, whereas they do not at the later stage.

(56)   The distribution of *PCR effects across categories in Rig-Vedic Sanskrit

C₁ \ C₂	Stop	Fricative	*v*	Nasal	Liquid	Glide
Stop	☑	✓	✓	✓	✓	✓
Fricative	✗		✓	✓	✓	✓
v				☑	✓	✓
Nasal	✗			☑	✓	✓
Liquid	✗	✗		✗		✓

    Whereas our previous attempts at generalizations for the various distributions led to fairly disjoint statements, these updated distributions allow for parsimonious generalizations based on sonority. If we assume the sonority scale in (57) (cf. Parker 2002, 2008), then the repetition behavior of the two stages is describable in terms of the relative sonority of the two consonants in the cluster. In the Rig-Vedic period, repetitions are allowed for all clusters with *rising or level* sonority (58a). In later Sanskrit, repetition is permitted only in *rising* sonority clusters (58b). The difference

between the two is thus whether or not level sonority clusters, namely stop-stop and nasal-nasal, permit C_1-copying.

(57) Sonority scale for Sanskrit (least sonorous to most sonorous)

1		2		3		4		5		6
Stop	<	Fricative	<	*v*	<	Nasal	<	Liquid	<	*y*

(58) **Sanskrit repetition generalizations:**

   a. **Rig-Vedic:** Clusters with rising or level sonority license repetitions; clusters with falling sonority do not.
   b. **Later Sanskrit:** Clusters with rising sonority license repetitions; clusters with level or falling sonority do not.

6.3.5 Repetition Licensing and Sonority

6.3.5.1 Local Summary

This section has explored the cluster-wise distribution and scope of repetition avoidance effects in the reduplicative systems of the Indo-European languages, and also the non–Indo-European language Klamath. The generalizations for the five distinct systems identified are repeated below. What this has revealed is that a simple division between pre-sonorant repetitions (permitted) and pre-obstruent repetitions (banned), enforced by a single anti-repetition constraint $*C_\alpha VC_\alpha$ / _$C_{[-son]}$, is not sufficient to capture the full distributions in most of the languages (it is fully sufficient only for Gothic).

(59) **Gothic repetition generalization:**
      Obstruent-sonorant clusters license repetitions, other clusters do not.

(60) **Ancient Greek repetition generalization:**
      Stop-sonorant clusters license repetitions, other clusters do not. (Variation in voiced stop + liquid clusters.)

(61) **Klamath repetition generalization:**
      Stop-sonorant clusters license repetitions, other clusters do not.

(62) **Later Sanskrit repetition generalization:**
      Clusters with rising sonority license repetitions; clusters with level or falling sonority don't.

(63) **Rig-Vedic repetition generalization:**
      Clusters with rising or level sonority license repetitions; clusters with falling sonority don't.

6.3.5.2 Minimum Sonority Distance

What these generalizations appear to suggest is that repetition licensing is sensitive to the relative sonority of the repeated consonant and what immediately follows. This was made explicit in the generalizations for the two stages of Sanskrit: in both cases, all clusters with rising sonority license repetitions; and, for Rig-Vedic Sanskrit, level sonority clusters also license repetitions. Gothic can also be cast in terms of relative sonority, as the division between pre-obstruent and pre-sonorant repetitions lines up with falling vs. rising sonority clusters. Greek and Klamath, though, are not

explicable simply in terms of rising vs. non-rising sonority clusters; however, they may be understandable in terms of the degree of sonority rise (or "minumum sonority distance") between the first and second members of the clusters, as proposed by Steriade (1982).

For Greek, while stop-fricative is of rising sonority (see Parker 2002, 2008), that rise in sonority is significantly smaller than the one between stop and sonorant. Therefore, if repetitions are only licensed when the root-initial cluster is separated by a sonority distance exceeding some threshold, there could potentially be a dividing line between stop-fricative and stop-sonorant. This might begin to make sense of the variation in voiced stop + liquid clusters: since voiced stops are more sonorous than voiceless stops (see again Parker 2002, 2008), the sonority distance is decreased when C_1 is a voiced stop rather than a voiceless one. If this difference in sonority brings the distance down towards the threshold for repetition licensing, then we might expect variation in such cases. The problem with this approach is that we should expect equivalent variation in voiced stop + nasal clusters. We do not seem to observe this, though it is true that there are very few data points, and therefore the lack of variation could be an accidental gap.

Klamath is somewhat similar, but poses more of a challenge to a minimum sonority distance approach. Like Greek, Klamath disallows repetitions for all level and falling sonority clusters, and also for stop-fricative clusters (which have a very small sonority rise). Crucially, however, unlike (Classical and Pre-Classical) Greek, Klamath attests fricative-sonorant clusters, and these also disallow repetitions. If we were to employ a sonority scale with regular, small increments — as shown in (64) (roughly equivalent to the one used for Sanskrit in (57)) — then we might, for example, expect the sonority distance in stop-nasal clusters to be equivalent to (or even less than) fricative-glide clusters. This is not the way Klamath works: all stop-sonorant clusters license repetitions, all fricative-sonorant clusters do not. It might be possible to adjust the spacing between values on the scale in just such a way that stop-sonorant and fricative-sonorant could be distinguished in the desired manner, but this would be an *ad hoc* solution. This suggests that sonority distance is not quite the right approach to this problem.

(64)   Hypothetical sonority scale for Klamath

1		2		3		4		5
Stop	<	Fricative	<	Nasal	<	Liquid	<	Glide

Putting aside for the moment the problem posed by Klamath, let us consider how this sort of minimum sonority distance approach could be encoded in a constraint. This could be done by adapting the $*C_\alpha VC_\alpha \: / \: \_C_{[-son]}$ schema in the following way:

(65)   *Sonority-distance anti-repetition constraint schema*:
        $*C_\alpha VC_\alpha C_\beta$; sonority($C_\alpha$) – sonority($C_\beta$) $\leq n$
        For each sequence of repeated consonants ($C_\alpha VC_\alpha$) which precedes another consonant (C_β), assign a violation mark * if the sonority distance between C_α and C_β does not exceed some threshold n.

For Rig-Vedic Sanskrit, where all rising and level sonority clusters license repetitions, $n = 0$. For later Sanskrit, where all rising sonority clusters license repetitions, n is some number just greater than 0. For the other systems, which require greater sonority distance, n is some larger positive number.

This line or reasoning could easily be extended to explain why all the languages license repetitions for all CV-initial roots. If we were to replace C_β in the constraint schema with just *any*

segment (X), so as to include both consonants and vowels, it will always be the case that if the language licenses a repetition with any consonants in that position, it will necessarily also license the repetition pre-vocalically, because all vowels are of greater sonority than all consonants.

6.3.5.3 Beyond Sonority

Modulo the problem with Klamath regarding the stop-sonorant vs. fricative-sonorant distinction (and perhaps the variation in Greek voiced stop + liquid clusters), this approach appears to be sufficiently flexible to capture the set of data explored in this chapter. However, it is not obviously imbued with any explanatory power. That is to say, why are repeated consonants preferentially licensed when they are followed by segments of greater and greater sonority? In the following section, I develop a hypothesis that these distributions are the result of the way that repetition interacts with acoustic/auditory cues to certain consonantal contrasts (Steriade 1994, 1997, 1999 *et seq.*, Flemming 1995/2002 *et seq.*), specifically the contrast between a consonant and its absence, i.e. C∼Ø (cf. Côté 2000, among others).

Under this proposal, sonority rise — whose acoustic correlate is *intensity rise* (Parker 2002, 2008) — serves as a strong cue to the presence of the repeated consonant in such a position, because a rise in intensity in the interlude between a vowel and a consonant (i.e. V_C_β) can only result from the presence of a consonant in that interlude. I will suggest that other cues which are normally robust cues for identifying the presence of a consonant, especially frication noise, are less robust under repetition because they are more difficult to anchor at a particular location in the speech stream (cf. Blevins & Garrett 2004). That is to say, something about repetition alters the effectiveness of cue perception in some way. These ideas together comprise the POORLY-CUED REPETITIONS HYPOTHESIS, which will serve as the basis for the final version of the *PCR constraint.

6.4 Acoustic/Auditory Cues to Consonantal Contrasts and the Definition of *PCR

The "Licensing by Cue" framework, developed by Steriade (1994, 1997) *et seq.*, is built upon the notion that phonological contrasts are dispreferred in contexts where they are less perceptible (see also Flemming 1995/2002 *et seq.* on the "Dispersion Theory of Contrast"). The ease or difficulty of accurate contrast perception is largely determined by the availability and robustness of various acoustic/auditory "cues" to those contrasts. Wright (2004:36) defines a cue as a particular piece of "information in the acoustic signal that allows the listener to apprehend the existence of a phonological contrast". Different sorts of cues vary in their inherent strength and relevance for contrast perception, and in the types of segments and contexts in which they are produced. Steriade (1994, 1997, 1999) and others argue that phonological contrasts which are poorly-cued are actively avoided in language (due to the difficulty of their accurate perception), and that the distribution of cues is crucial in shaping the phonological typology (see generally the contributions in Hayes, Kirchner, & Steriade 2004 for contemporary views of the role of phonetics in phonology).

In the previous section, I showed that a repetition avoidance constraint (or constraint schema) based on minimum sonority distance between the repeated consonant and the following segment was nearly sufficient to capture the observed distributions. In this section, I show that defining the repetition avoidance constraint in terms of acoustic/auditory cues to particular consonantal contrasts — focusing primarily on *intensity rise* as a cue to C∼Ø contrasts — is also possible, and I argue that this approach provides a more successful and explanatory analysis of the various distributions

under consideration. I call this approach the NO POORLY-CUED REPETITIONS constraint approach, abbreviated *PCR. This is the repetition avoidance constraint I advocate throughout the remainder of this dissertation.

6.4.1 Acoustic/Auditory Cues and Consonantal Contrasts

There are a variety of acoustic/auditory cues (henceforth just "cues") which play a role in the perception of consonantal contrasts. By consonantal contrasts, I mean one of two types. First, there is the contrast between a particular consonant and another particular consonant: i.e. $C_\alpha \sim C_\beta$, the contrast distinguishing [...$XC_\alpha Y$...] from [...$XC_\beta Y$...] (where X or Y may be null). For example, English has a contrast in terms of place for stops between *p* (labial), *t* (coronal), and *k* (dorsal). Among other cues, this contrast is cued by the properties of the *release burst*, which differ based on point of articulation, and also by the effects of the stop's constriction on the formant structure of following segments (especially of F2 of a following vowel), referred to as *transitions* (see below for discussion of these cues).

The other type of consonantal contrast — which, according to the hypothesis advanced below, is more significant for the analysis of repetition avoidance effects — is that holding between a particular consonant and its absence: i.e. $C_\alpha \sim \emptyset$, the contrast distinguishing [...$XC_\alpha Y$...] from [...XY...]. Take, for example, the English minimal pair *stop* and *top*. These two differ in the presence vs. absence of *s* before *t*. Therefore, we can say that English licenses a contrast between *s* and $\emptyset$ in the word-initial pre-stop context (e.g. #_*t*). The primary cue that is involved in licensing this contrast is the *high-intensity frication noise* (in the higher frequencies) which is an inherent property of *s*. English does not license other types of fricatives in this position — for example, *f* cannot appear word-initially before a stop (***ftop) — because their *frication noise* is of *low* intensity, which makes it a less robust cue for the sake of contrast perception. While the difference in the licensing of contrast in this case can alternatively be stated in terms of phonological features (namely, [±strident]), these features simply recapitulate the phonetic properties of the acoustic/auditory cues (that is to say, [+strident] refers to segments with *high-amplitude* frication noise).

One way to conceptualize the repetition avoidance effect is in terms of the (non-)licensing of $C \sim \emptyset$ contrasts in the repetition context. What a repetition avoidance effect amounts to is the following. A language normally licenses a $C \sim \emptyset$ contrast in some particular position. This is exemplified in (66), where a $p \sim \emptyset$ contrast is licensed in the position between a vowel and a (pre-vocalic) stop (V_TV) when it is preceded (transvocalically) by no consonant (66a) or by a non-identical consonant (66b). However, when the position of the contrast is preceded by an identical consonant (66c), the contrast is no longer licensed. The effect of this is that the more marked member of the potentially contrasting pair — the one with the repeated *p*, [papta] — is not allowed to surface; only the less marked member — the one without the repeated *p*, [pata] — is a possible form.[38]

(66) $C \sim \emptyset$ contrast and the repetition context

Preceding Context		(C)V{$p \sim \emptyset$}TV			$p \sim \emptyset$ *contrast licensed?*
a.	No C	a_ta	∼	ap̲ta	✓
b.	Different C	ka_ta	∼	kap̲ta	✓
c.	Same C	pa_ta	∼	**pap̲ta**	✗

---

[38] The *p* forms are more marked by virtue of the fact that, in the general sense, *something* is always more marked than *nothing*, and, more specifically, by virtue of the problem that repetition creates. This is discussed below.

What this indicates is that languages may impose special licensing requirements for C~Ø contrasts under repetition, relative to the baseline non-repetition context. I now turn to discussion of the POORLY-CUED REPETITIONS HYPOTHESIS, which spells out a potential way of understanding why and how languages impose these special requirements.

6.4.2 The Poorly-Cued Repetitions Hypothesis

It seems likely that repetition places an extra burden on the perceptual system that is absent when repetition is not at stake. In all cases, given a speech signal, the perceptual system is tasked with (i) dividing the signal into its discrete segments, (ii) identifying the relative order of those segments, and (iii) determining the specific identity of those segments. Under repetition — that is, when the target segment is preceded by an identical segment — the perceptual system is burdened by an additional responsibility: to determine if the acoustic imprint created by the second segment's articulation is actually the result of a second instance of the segment, as opposed to some sort of delayed effect of the first segment's articulation.

If this is a correct characterization of the problem, then it would be reasonable to assume that C~Ø contrasts may become more difficult to perceive under repetition; that is to say, the distance between C and Ø in perceptual space shrinks when there is a preceding identical C. This could be because there is some amount of pressure towards interpreting the cues which properly signal the presence of the second consonant instead as cues to the presence of the preceding identical consonant. That would suggest that cues which are relatively more difficult to locate at a particular point in the speech signal (or rather, the acoustic events signaled by such cues; cf. Stevens 2002's "acoustic landmarks") will be even less effective in the repetition context than in the general context, since they may be more easily interpretable as originating from a different part of the speech stream. I gather these notions together under the heading of the POORLY-CUED REPETITIONS HYPOTHESIS, as formulated in (67).

(67) **THE POORLY-CUED REPETITIONS HYPOTHESIS**
There is some property of the perceptual system which degrades listeners' ability to apprehend the presence of a consonant (i.e. the contrast between that consonant and its absence) when that consonant is adjacent to an identical consonant.

   i. This property diminishes the effectiveness of some or all acoustic/auditory cues to C~Ø contrasts, such that some cues which are normally sufficient to license those C~Ø contrasts (in otherwise equivalent positions) are no longer sufficient to license those contrasts under repetition.

   ii. This property diminishes the effectiveness of different cues to different extents: the effectiveness of cues to acoustic events which are more difficult to anchor at a particular point in the speech stream and/or tend to extend across multiple segments are diminished to a greater degree than cues to acoustic events which are more reliably located at their correct position in the speech stream.

This should be taken as a working hypothesis, generated based on the empirical generalizations deduced from the reduplication patterns examined in this chapter and throughout this dissertation. Experimental evidence will need to be brought to bear to probe the details of this hypothesis.[39] I will

---

[39] See Zukoff & Storme (2017) for a very preliminary attempt in this direction.

not here speculate about the specific physical or cognitive underpinnings of this effect, as nothing about the evidence adduced in this dissertation directly bears on that question.[40]

6.4.3 The No Poorly-Cued Repetitions Constraint (*PCR)

Whatever the specifics, the POORLY-CUED REPETITIONS HYPOTHESIS provides a basis for a repetition avoidance constraint. If accurate perception of a contrast is more difficult in a particular context, the grammar may impose special conditions to avoid that contrast in that position (cf. Steriade 1994, 1997, 1999, Flemming 1995/2002, Hayes & Steriade 2004). Since there is, by hypothesis, increased difficulty of C∼Ø contrast perception under repetition, the grammar should include a mechanism to penalize repetition. The empirical evidence shows that not all repetition types are treated equally, and that different repetition types may be treated differently across languages. Therefore, we need to formulate the constraint in such a way as to be sensitive to the relevant differences.

I propose the formulation in (68) as a working definition f the *PCR constraint schema. It penalizes repeated consonants which lack the cues required to license the presence of a consonant under repetition, and it allows for the licensing conditions to vary by language.

(68)   ***PCR [*constraint schema — full version*]**
         Languages may set stricter conditions (in terms of cues) for the licensing of C∼Ø contrasts (i.e. the presence of C) when that C would be the second member of a transvocalic consonant repetition (i.e. C_α^2 in a $C_\alpha^1 V C_\alpha^2$ sequence) than in other contexts. Assign a violation mark * for each C_α^2 (i.e. each C∼Ø contrast where C is a C_α^2) which is not cued to the level required by the language-specific repetition licensing conditions.

There are at least two formal aspects of this definition that require some exposition. One is the notion that *transvocalic* consonant repetition should be significant, as opposed to, say, repetition at a greater distance. Recall that the in Gothic (and also Klamath and Proto-Anatolian), *PCR violations are avoided by "cluster-copying". That is to say, if default C_1-copying would create a $C_1 V C_1 C_2$ sequence that would violate *PCR, reduplication instead results in a $C_1 \mathbf{C_2} V C_1 C_2$ sequence; for example, Gothic $\sqrt{stald} \rightarrow \underline{ste}$-*stald* not ***$\underline{se}$-*stald*. In a cluster-copying form, root-C_1 is still preceded by an identical copy, but it is now separated not just by a vowel but also by a consonant. (Note that root-C_2 is now also repeated in roughly that same configuration.) It is not at all obvious to me why the presence of an intervening consonant should make such a significant difference. Nevertheless, it is a necessary consequence of the proposal that this be the case.

What I can say is that the restriction to transvocalic repetitions is not unprecedented in work relating to repetition avoidance; in fact, far from it. The transvocalic relationship of the two consonants matches that which Rose (2000:95) defines as *consonant adjacency*: "two consonants in sequence are adjacent irrespective of intervening vowels". Rose argues that the proper formulation of the Obligatory Contour Principle (OCP; cf. Goldsmith 1976, McCarthy 1986) — the notion that identical adjacent elements are dispreferred — must crucially allow reference to the surface domain defined by consonant adjacency, at least for segmental effects. That is to say, not only does the string $C_\alpha C_\alpha$ induce an OCP violation when it contains (distinct) identical elements, but so does

---

[40] For biases against repetition in phonological processing, see Frisch (2004); see also Kanwisher (1987) on a general bias against repetition across other human perceptual systems. One other possibility to consider is a gestural underpinning. Walter (2007) shows that gestures may be reduced under repetition. Reduced gestures may degrade the number and/or robustness of acoustic cues, and thus make repeated segments less perceptually salient in that way.

the string $C_\alpha V C_\alpha$.[41] Indeed, many of the now-classic OCP effects discussed in McCarthy (1986) and subsequent work involve the $C_\alpha V C_\alpha$ environment. In a similar vein, a great deal of work has demonstrated that gradient effects of similarity avoidance between consonants in this configuration (though also, in some cases, at a greater distance) are evident in the lexicons of virtually all of the languages which have been examined in this way (Greenberg 1950, McCarthy 1991, 1994, Berkley 2000, Frisch, Pierrehumbert, & Broe 2004, Pozdniakov & Segerer 2007, among others).[42] Therefore, there is significant precedent for identifying transvocalic position as a locus for effects relating to consonantal identity.

The second aspect of the above definition of *PCR that requires further discussion is the identity condition itself, i.e. the α subscripts. Consider the facts from Ancient Greek provided in (69). When aspirated stops have a correspondent in the reduplicant, that consonant is unaspirated (69a). Stop-stop clusters do not tolerate repetitions (69b). Yet, when a root-initial cluster consists of aspirated stops, the repetition is still not tolerated (69c), despite the fact that the ban on aspirates in the reduplicant would render the "repeated" consonants distinct with respect to aspiration.

(69)   Repetition and aspiration in Ancient Greek
  a. $\sqrt{p^h eug}$ 'flee' $\rightarrow$ perfect *p-e-p^{h}eug-*, not ***p^h-e-p^{h}eug-* (van de Laar 2000:306–307)
  b. $\sqrt{kten}$ 'kill' $\rightarrow$ perfect *e-kton-*, not ***k-e-kton-* (van de Laar 2000:197–198)
  c. $\sqrt{p^h t^h er}$ 'destroy' $\rightarrow$ perfect *e-p^{h}t^{h}or-*, not ***p-e-p^{h}t^{h}or-* (van de Laar 2000:308–309)

Similar evidence comes from the Sanskrit $C_1\bar{e}C_2$ form *pēd-ur* ($\leftarrow$ $\sqrt{pad}$ 'go'; Whitney 1885 [1988]:94–95), which is preferred to **pa-bd-ur, a C_1-copying output where the root-initial /p/ assimilates in voicing, rendering the base and reduplicant consonants distinct in voice. The fact that we see a $C_1\bar{e}C_2$ form here shows that a difference in only voicing is also not sufficient to count as non-identical.[43] This suggests that there are certain *phonological* differences that do not nullify the identity condition for the purposes of the *PCR constraint.

It is a significant question what counts as identical and what does not, for the purposes of the identity condition in $C_\alpha V C_\alpha$ sequences. I will not be able to resolve this definitively here. Since laryngeal features like aspiration and voicing are the only features for which there is clear evidence that their difference does not nullify identity, I will assume that identity is calculated based on place and manner features.[44]

Returning to the definition of the constraint (schema), since we are dealing with C~Ø contrasts (rather than C_α~C_β contrasts), there is a useful simplification we can make regarding the formulation of the constraint, at least as a shorthand. Since C~Ø contrasts which are deemed too indistinct by the grammar will (presumably) always be resolved in favor of Ø (because that avoids the problem created by repetition, and because something is always more marked than nothing), penalizing a C~Ø contrast is the same as penalizing the presence of C in the output. Therefore, this constraint can be thought of in much the same way as a standard issue markedness constraint, penalizing a particular output structure. The abbreviated constraint definition in (70) is framed in terms of

---

[41] Part of Rose's (2000) argument is that absolute-adjacent identical consonants (e.g. *kk*) are automatically treated as geminates (i.e. *kː*), and that geminates violate not the OCP but rather a distinct constraint against geminate consonants (No-Gem). For Rose, the CC environment yields an OCP violation only if the consonants are in fact *not* fully identical, yet both contain the feature targeted by the particular OCP constraint (e.g. *xk* violates OCP-Place).

[42] See Cooper (2009), Sandell (2015a) on similarity avoidance in Proto-Indo-European.

[43] This cannot be a base-reduplicant faithfulness effect, as voicing differences are tolerated if the resulting repetition is licensed: for example, $\sqrt{b^h as}$ 'devour' $\rightarrow$ reduplicated present *bapsati* (Whitney 1885 [1988]:109).

[44] This also accords completely with the conditions for root co-occurrence restrictions in Proto-Indo-European (Sandell 2015a).

this sort. For clarity's sake, this shorthand formulation will be used for the language-specific definitions of *PCR below.

(70)   ***PCR [*constraint schema — abbreviated*]**
         Assign a violation mark * for each output C_α^2 which lacks the requisite cues.
         ▶ **REQUISITE CUES:** ...

This is not necessarily the typical approach for integrating cues to contrast into the phonological grammar (cf. Flemming 1995/2002, 2008, among others). I leave it as an open question whether *PCR should be thought of in these (quasi-)markedness terms, or should instead be formulated in some other way, perhaps located in a somewhat different module of the grammar.

There is an implicit assumption behind these formulations that deserves to be made explicit. These constraint definitions treat all C~Ø contrasts in the repetition context as being subject to the same licensing requirements. (Though, as will be shown below, it has to allow for disjunctive statements within the licensing conditions, meaning that one could construct a set of possible licensing conditions which are tailored to respective sorts of C~Ø contrasts. In practice, this is not how I will frame the language-specific licensing conditions.) I believe this is justified given the empirical evidence, and the type of cue which will be central to the explanation, namely, *intensity rise*.

6.4.4   Intensity Rise as the Central Cue for *PCR

6.4.4.1   Intensity and Sonority

As discussed in Section 6.3.5, sonority-based generalizations may be fairly successful in describing which clusters do or do not license repetitions in the various languages currently under consideration. For example, the repetition generalization for Later Sanskrit was that all and only rising sonority clusters licensed repetitions. Parker (2002, 2008) demonstrates that *intensity* — as measured by sound level extrema in the low frequency bands (in decibels) — is a reliable acoustic correlate of the phonological notion of sonority. Low sonority sounds have low intensity values (low sound level minima), while high sonority sounds have high intensity values (high sound level minima in the case of consonants; high sound level maxima in the case of vowels). Therefore, the generalization that clusters with a *sonority rise* license repetitions in Later Sanskrit can be recast in cue-based terms by saying that it is clusters with an *intensity rise* that license repetitions.

What I will here call *intensity rise* refers to an abrupt change in the overall intensity (increasing) and attenuation (decreasing) of the acoustic signal (Yun 2016:21–38; cf. Wright 2004:39). The high-amplitude periodic noise of a vowel has the highest possible acoustic intensity, whereas the complete silence created by a voiceless stop closure has virtually no intensity (Parker 2002, 2008). We will see, however, that what is relevant for repetition licensing is only whether an intensity rise is present, not the degree of that intensity rise.

To outline how intensity rise captures sonority-based generalizations, consider what cluster types contain an intensity rise at the cluster-internal juncture. The table in (71) presents the sonority scale assumed by Parker (2008).[45] Parker's experimental results show that this sonority scale is generally matched by the relative intensity values of the different consonant classes. Using the

---

[45] Parker shows that there is a consistent difference between voiced obstruents (relatively higher intensity) and voiceless obstruents (relatively lower intensity). Since all of the languages under discussion have voicing assimilation in obstruent sequences, and I will be treating intensity rise as a binary property rather than a scalar property, this distinction is not relevant for the purposes of the present proposal.

sonority ranks as a proxy for relative intensity values,[46] the table in (72) indicates which clusters contain an intensity rise — i.e., where C_1 has lower intensity than C_2.[47] Clusters with an intensity rise are marked with a ✓; clusters which lack an intensity rise are marked with a ✗. Given that here we are simply equating intensity values with sonority values, it is obviously the case that all and only rising sonority clusters have an intensity rise.

(71)   Sonority Scale (Parker 2008:60, ex. 4)

Consonant Type	Sonority Rank ≈ Intensity Rank
Glides	12
Rhotic approximants (ɹ)	11
Flaps	10
Laterals	9
Trills	8
Nasals	7
Voiced fricatives	6
Voiced affricates	5
Voiced stops	4
Voiceless fricatives	3
Voiceless affricates	2
Voiceless stops	1

(72)   Cluster-wise distribution of the intensity rise cue

C_1 \\ C_2	Stop	Fricative	Nasal	Liquid	Glide
Stop	✗	✓	✓	✓	✓
Fricative	✗	✗	✓	✓	✓
Nasal	✗	✗	✗	✓	✓
Liquid	✗	✗	✗	✗	✓
Glide	✗	✗	✗	✗	✗

6.4.4.2   Intensity Rise and Repetition

Why should intensity rise be especially important in the repetition context? Abrupt change from a low intensity signal to a high intensity signal is a strong indicator of the presence of a consonant preceding the high-intensity portion. This is especially true in post-vocalic position. Given a produced sequence $V_1C_1C_2V_2$, a percept of rising intensity going into C_2 will be an unambiguous

---

[46] The numbers in (71) refer to the ranking of different consonant types on the sonority scale. The numbers are not intended to indicate that the different classes are evenly spaced with respect to their sonority or intensity differences. Yet, since all we will be concerned with is whether or not there is *any* rise in intensity, rankings on a scale will be fully equivalent to more realistic relative values.

[47] See Yun (2016:21–38) for discussion of cases where this direct mapping between sonority and intensity is not completely justified.

cue that there was a consonant between V_1 and C_2. This is because vowels have the highest possible intensity. If there had been no consonant intervening between V_1 and C_2, there necessarily would have been a *fall* in intensity going into C_2. Therefore, a rise in intensity in that position can only have come about by the presence of an intervening consonant.[48]

The Poorly-Cued Repetitions Hypothesis, as formulated in (67) above, hypothesized that the problem with repetition is that it leads to increased confusion about whether or not the second acoustic imprint signals the presence of a distinct consonant in the speech stream; that is, it decreases the effective perceptual distance between the repeated C and Ø. If this is correct, then cues which unambiguously anchor that second acoustic imprint at the right point in the speech stream should be particularly valuable in the repetition context. The intensity rise cue does exactly this. Therefore, it is entirely reasonable to pursue an explanation of repetition licensing which centers around intensity rise.

To return to the question raised at the end of Section 6.4.3, the fact that it is intensity rise which is the central cue for repetition licensing may justify using unified licensing conditions across different types of C∼Ø contrasts. This is because intensity rise is a cue that is in some sense independent of consonant type. Intensity is a continuous property found throughout a speech stream, and so all segments must have some intensity value. Therefore, the listener's need to identify an intensity value is not specific to any particular type of C∼Ø contrast; it is applicable to all C∼Ø contrasts. Given that a rising intensity contour is the most unambiguous type of contour for apprehending the presence of consonant, and this can be assessed across all types of consonants, there is no reason to think that the repetition licensing conditions would have to be relativized to particular types of C∼Ø contrasts. As long as intensity rise is the central cue for repetition licensing, it is therefore reasonable to posit unified licensing conditions across different possible C∼Ø contrasts under repetition.

6.4.4.3 Intensity Rise at Release

The notion of intensity rise as just introduced will not on its own capture all of the repetition distributions examined in this chapter. Consider Rig-Vedic Sanskrit (see Section 6.3.4.4). Like Later Sanskrit, this stage of the language licenses repetitions for all of the clusters with an intensity rise. However, it also licenses repetitions for the attested level sonority clusters, namely, stop-stop and nasal-nasal (at least for certain place combinations; see Section 6.5.3 below). These lack the sort of intensity rise found in the other clusters which license repetitions, since segments of the same manner category have equivalent intensity values (Parker 2002, 2008). However, there may be a minor extension of the notion of intensity rise that could bring these particular level sonority clusters into the fold.

Yun (2016) demonstrates that the presence of an intensity rise is central for the determination of licit epenthesis sites. Specifically, epenthesis into a consonant cluster is only licensed when that cluster would contain an intensity rise from the first consonant to the second, because that intensity rise is similar to the consonant-to-vowel intensity rise that will result from epenthesizing a vowel in that position. Most of the clusters which contain the necessary sort of intensity rise are those which were identified as having an intensity rise in (72), based on Parker's (2002, 2008) sonority-like definition of intensity. However, there are additional clusters which Yun (2016) claims to have an intensity rise of a slightly different sort.

This second type of intensity rise is the rise in intensity at the release of the oral constriction of a stop consonant, both obstruent stops and nasal stops. This is found in any instance where

---

[48] A rise in intensity into C_2 would also be generated by a pause between V_1 and C_2.

an obstruent stop or a nasal stop is produced with an audible release. Relevant for our purposes is that stop-stop clusters and nasal-nasal clusters will have this type of intensity rise cuing the first consonant in cases where that consonant is actually released. (If the stop or nasal is produced without an audible release in that position, this type of intensity rise will be absent.) Clusters where C_1 is a released stop or nasal and where C_2 is of greater overall intensity than C_1 have both types of intensity rise.

I will refer to this type of intensity rise as *intensity rise at release* — abbreviated *IntRise:Rel*. This is to be contrasted with Parker's (2002, 2008) sonority-like type of intensity rise based on the relative sound level extrema. For clarity, I will now refer to this type as *intensity rise between sound level extrema* — abbreviated *IntRise:SLE*. The way to characterize the difference between Rig-Vedic Sanskrit and Later Sanskrit is that Rig-Vedic licenses repetitions with either *IntRise:SLE* or *IntRise:Rel*, but Later Sanskrit licenses them only with *IntRise:SLE*.[49]

6.4.4.4 Preview of the Typology

The table in (73) lists the types of intensity rises found in different cluster types. Stops and nasals — the two types of consonants which can have a release, and thus an intensity rise at release — are here separated into released and unreleased in C_1 position. It is definitional that unreleased consonants cannot have an *intensity rise at release*.

(73)   Cluster-wise distribution of the intensity rise cue

C_1 \ C_2	Stop	Fricative	Nasal	Liquid	Glide
Released Stop	Rel	SLE & Rel	SLE & Rel	SLE & Rel	SLE & Rel
Unreleased Stop		SLE	SLE	SLE	SLE
Fricative			SLE	SLE	SLE
Released Nasal	Rel	Rel	Rel	SLE & Rel	SLE & Rel
Unreleased Nasal				SLE	SLE
Liquid					SLE
Glide					

If we allow for both conjunction and disjunction of the two types of intensity rise cues as licensing conditions for *PCR, then we predict four possible languages. These possibilities are provided in (74). I will demonstrate in Section 6.5 that three of these possibilities are instantiated. Later Sanskrit and Gothic are characterized by *IntRise:SLE* as the sole licensor (74a). Rig-Vedic licenses repetitions with either *IntRise:SLE* or *IntRise:Rel* (or both) (74c). And Klamath and Ancient Greek require both *IntRise:SLE* **and** *IntRise:Rel* to license a repetition (74d).[50]

---

[49] It could alternatively be the case that there was a change in the language-specific phonetics of the relevant clusters between the two periods. Specifically, the difference would be explained without positing a change in licensing conditions if C_1 was released in these clusters in Rig-Vedic Sanskrit, but came to be produced without audible release in Later Sanskrit. I know of no evidence which bears on this question.

[50] Because of the strictness of their main licensing condition, Klamath and Ancient Greek require an additional disjunctive condition to account for CV-initial roots.

(74)   Types of intensity rise and *PCR licensing conditions

	Cues that license repetition	Language
a.	IntRise:SLE	Later Sanskrit, Gothic
b.	IntRise:Rel	none
c.	IntRise:SLE or IntRise:Rel	Rig-Vedic Sanskrit
d.	IntRise:SLE & IntRise:Rel	Klamath, Ancient Greek

The one case which is not attested among the languages examined in this dissertation (and thus not attested among the languages known to me that exhibit *PCR effects of this type) is that where *IntRise:Rel* is sufficient on its own to license repetition, but *IntRise:SLE* is not (74b). This would be a language that, for example, licensed repetitions for TT clusters (where the first stop is released) but not for fricative-sonorant clusters or liquid-glide clusters. Having just five attested systems is certainly not sufficient to claim that the entire typology is accurately represented. Nevertheless, if it could be shown that the absence of languages like (74b) is a systematic gap and not an accidental one, this typology could be explained by claiming there to be an implicational relation between *IntRise:SLE* and *IntRise:Rel*. If *IntRise:SLE* is inherently a stronger/better cue to the presence of a consonant in the repetition context than *IntRise:Rel*, then we would expect any system that licenses repetitions with *IntRise:Rel* alone to also license repetitions with *IntRise:SLE*. This is consistent with the available data.

6.5   Cues and the Language-Specific Definitions of *PCR

In this section, I go through the facts of the individual languages and show which cues are sufficient to license repetitions in each language, i.e., which cues are required to avoid a *PCR violation in that language. I first focus on what cues *are* sufficient to license repetitions in each of the languages. I will then discuss several other possible cues which turn out not to be good repetition licensors and consider why this might be the case. Note that the languages are presented in a different order in this section than in Section 6.3; they are here organized by the types of cues which comprise their *PCR constraints.

6.5.1   Later Sanskrit

As alluded to in Section 6.4.4, the *PCR constraint for Later Sanskrit can be characterized simply in terms of the presence vs. absence of the *intensity rise between sound level extrema* cue (*IntRise:SLE*). The table in (75) (repeated from (55) above) shows the distribution of *PCR effects in later Sanskrit, collapsing together the evidence from cluster-initial roots and CaC zero-grade clusters. One way to economically describe this distribution is in terms of the sonority contour of the base-initial cluster, as provided in (76) (repeated from (58b)/(62) above).

(75)   The distribution of *PCR effects in later Sanskrit

C₁ \ C₂	Stop	Fricative	*v*	Nasal	Liquid	Glide
Stop	✗	✓	✓	✓	✓	✓
Fricative	✗		✓	✓	✓	✓
v					✓	✓
Nasal	✗			✗	✓	✓
Liquid	✗	✗		✗		✓

(76)   **Later Sanskrit repetition generalization:**
Clusters with rising sonority license repetitions; clusters with level or falling sonority don't.

From the preceding discussion, we know that *IntRise:SLE* directly recapitulates the notion of sonority rise. This is confirmed in the table in (77), where all clusters which license repetitions in (75) have *IntRise:SLE* as a cue to C_1 (marked with ✓) while all clusters which fail to license repetitions in (75) lack *IntRise:SLE* as a cue to C_1 (marked with ✗). Note that this requires that Sanskrit *v* (= "narrow approximant" [ʋ̝]; see Section 6.3.4.1) has an intensity value intermediate between nasals and fricatives.[51] Therefore, the generalization in (76) can equally well be stated in cue-based terms, as given in (78).

(77)   Cluster-wise distribution of *IntRise:SLE* cue in Later Sanskrit

C₁ \ C₂	Stop	Fricative	*v*	Nasal	Liquid	Glide
Stop	✗	✓	✓	✓	✓	✓
Fricative	✗	✗	✓	✓	✓	✓
v	✗	✗	✗	✓	✓	✓
Nasal	✗	✗	✗	✗	✓	✓
Liquid	✗	✗	✗	✗	✗	✓

(78)   **Later Sanskrit cue-based repetition generalization:**
Repeated consonants are permitted only when cued by *IntRise:SLE*.

This generalization can easily be translated into the *PCR constraint format introduced in Section 6.4.3. (I here use the abbreviated format from (70), but recall that this is a shorthand for the more explicit formulation in (68).) The "requisite cues" clause is comprised simply of *IntRise:SLE*. This is the definition adopted for the Later Sanskrit *PCR constraint in (79). In addition to capturing the repetition facts for the cluster-initial roots/bases, this definition also suffices to explain the fact that all CV-initial roots/bases license repetitions as well. This is because vowels have greater intensity than any consonant, so there will necessarily be an intensity rise cuing the repeated consonant for CV-initial roots.

---

[51] While we obviously cannot know the intensity value of Sanskrit *v*, this question could be experimentally investigated for modern languages which have narrow approximants of this sort. Note that Parker's (2008) experimental results indicate that approximants might be subject to greater variation in intensity across languages than other segment types.

(79)   ***PCR [*for Later Sanskrit*]**
         Assign a violation mark * for each output C_α^2 which lacks the requisite cues.
         ► **REQUISITE CUES:** *IntRise:SLE*

6.5.2 Gothic

Gothic has many fewer types of clusters attested with reduplication than does Sanskrit (and indeed all the other languages discussed in this section). This leaves it open to a number of possible interpretations. Nevertheless, one of the most straightforward interpretations is that it works in exactly the same way as Later Sanskrit; namely, only the clusters where C_1 is cued by *IntRise:SLE* license repetitions.

The distribution of default C_1-copying (notated with ✓) vs. alternative cluster-copying (notated with ✗) in Gothic is repeated in (80). The cells with C_1-copying are those that license a repetition; the cells with cluster-copying are those that don't. The only attested cluster that does not license a repetition is ST. The attested clusters that do license a repetition are TL and FL (where F = {f,s}). This distribution was captured in terms of phonological features and natural classes by the generalization in (81) (repeated from (14)/(59) above).

(80)   Distribution of *PCR effects in Gothic by root-initial cluster (repeated from (13))

C_1 \ C_2	Stop (T)	Liquid (L)
Stop (T)		✓ ($T_\alpha VT_\alpha L$)
Fricative (F/S)	✗ ($*S_\alpha VS_\alpha T$)	✓ ($S_\alpha VS_\alpha L, F_\alpha VF_\alpha L$)

(81)   **Gothic repetition generalization:**
         Obstruent-sonorant clusters license repetitions, other clusters do not.

Since it is the case that obstruent-sonorant clusters cue C_1 with *IntRise:SLE* but fricative-stop clusters do not, this generalization too can be recast simply in terms of *IntRise:SLE*, just as for Later Sanskrit. This leads us to the cue-based generalization regarding the distribution of repetitions in Gothic in (82), and thus the cue-based definition of *PCR provided in (83).

(82)   **Gothic cue-based repetition generalization:**   (= Later Sanskrit (78))
         Repeated consonants are permitted only when cued by *IntRise:SLE*.

(83)   ***PCR [*for Gothic*]**   (= Later Sanskrit (79))
         Assign a violation mark * for each output C_α^2 which lacks the requisite cues.
         ► **REQUISITE CUES:** *IntRise:SLE*

While *IntRise:SLE* properly captures the distribution, it is indeed possible to use a different cue to make the same division, namely, *CR transitions*. *Transitions* are rapid changes in the formant values — especially of the second formant (F2) — of a high-amplitude segment, induced by the articulatory characteristics of an adjacent constriction, that is, by the specific shape of the vocal tract created by the constriction gestures of an adjacent consonant (see Wright 2004:37). Transitions may be thought of as the automatic result of coarticulation between adjacent segments.

Since there is no formant structure in a stop (certainly not in a voiceless stop), *s*-stop clusters cannot cue C_1 with transitional information in C_2. On the other hand, since liquids do have

internal formant structure, and stops and fricatives have greater constriction than liquids, both stop-liquid and fricative-liquid do cue C_1 with transitional information in C_2. I refer to this cue as *CR transitions*. *CR transitions* are primarily a cue to place contrasts (Flemming 1995/2002:22–23). However, they might also be viewed as a cue to the $C\sim\emptyset$ contrast. If a segment bears transitional formant structure at its onset, this can only be the result of a preceding segment with greater constriction. Since vowels have minimal constriction, identifying formant transitions in C_2 necessarily implies the presence of C_1 immediately preceding it.

Therefore, the Gothic distribution could be explained in at least two ways: *IntRise:SLE* is the sole repetition licensor or *CR transitions* is the sole repetition licensor. This is actually somewhat of a moot point. The underlying logic of *PCR is that repetitions will only be tolerated in the best positions for apprehending $C\sim\emptyset$ contrasts. The clusters for which Gothic does license repetition are those which have two potentially robust cues, namely, *IntRise:SLE* and *CR transitions*. The cluster type that fails to license repetition lacks both of theses. Therefore, whether either cue could license repetitions independently, it is certainly the case that they can do so jointly.

6.5.3   Rig-Vedic Sanskrit

The distribution of *PCR effects in Rig-Vedic Sanskrit differs minimally from that of Later Sanskrit, but in a systematic way. As discussed in Section 6.4.4, the Rig-Vedic distribution can be captured by positing that the other type of intensity rise — *intensity rise at release* (*IntRise:Rel*) — is also sufficient to license repetitions, even in the absence of *IntRise:SLE*.

Rig-Vedic licenses repetitions in two additional contexts beyond those which are licensed in Later Sanskrit: stop-stop (i.e. $T_\alpha V T_\alpha T_\beta$) and nasal-nasal (i.e. $N_\alpha V N_\alpha N_\beta$). This is shown in (84) (repeated from (56) above), with these cells notated with frames around the new ✓'s. Like for Later Sanskrit, this distribution can be described in terms of the sonority contour of the base-initial cluster. In these terms, the difference between the two is that, at this stage, level sonority clusters *do* license repetitions. This is stated in (85) (repeated from (58a)/(63) above).

(84)   The distribution of *PCR effects across categories in Rig-Vedic Sanskrit

C_1 \ C_2	Stop	Fricative	*v*	Nasal	Liquid	Glide
Stop	☑	✓	✓	✓	✓	✓
Fricative	✗		✓	✓	✓	✓
v				✓	✓	✓
Nasal	✗			☑	✓	✓
Liquid	✗	✗		✗		✓

(85)   **Rig-Vedic Sanskrit repetition generalizations:**
Clusters with rising or level sonority license repetitions; clusters with falling sonority don't.

Notice, however, that the level sonority clusters we are dealing are limited to oral stops and nasal stops. These particular level sonority clusters have the unique property of (potentially) being articulated with an audible release of a complete oral closure. If the cluster-initial stops are indeed released, they will be cued release-related cues not available for other types of segments. The crucial additional cue for our purposes is the intensity rise that results from audible release — the *IntRise:Rel* cue.

270

As long as we assume that C_1 was released in Rig-Vedic stops-stop and nasal-nasal clusters, then the sonority-based generalization in (85) can be recast in terms of the two types of intensity rises, as shown in (86). This leads to the definition of *PCR provided in (87).

(86)   **Rig-Vedic Sanskrit cue-based repetition generalization:**
Repeated consonants are permitted only when cued by *IntRise:SLE* or *IntRise:Rel* (or both).

(87)   ***PCR [for Rig-Vedic Sanskrit]***
Assign a violation mark * for each output C^2_α which lacks the requisite cues.
▶ REQUISITE CUES:
     i.   *IntRise:SLE*, and/or
     ii.   *IntRise:Rel*

The use of *intensity rise at release* as a repetition licensor for Rig-Vedic Sanskrit might also make sense of an otherwise puzzling asymmetry in the treatment of different types of nasal-nasal and stop-stop clusters at that stage. As mentioned in Section 6.3.4.2, there is a disagreement in the treatment of labial-coronal (*mn-*, *pt-*, *etc.*) and coronal-labial (*nm-*, *tp-*, *etc.*) clusters at the Rig-Vedic stage. Labial-coronal clusters (88a) exhibit C_1-copying perfect weak stems, and thus license repetitions. On the other hand, coronal-labial clusters (88b) exhibit $C_1\bar{e}C_2$ perfect weak stems, and thus ban repetitions.

(88)   Labial-coronal vs. coronal-labial in Rig-Vedic
     a.   Labial-coronal licenses repetition
        i.   √*man* → <u>*mam*</u>*nāt^{h}ē*
        ii.   √*pat*   → <u>*pap*</u>*tur*
     b.   Coronal-labial bans repetition
        i.   √*nam* → *nēmē*   (**<u>*nan*</u>*mē*)
        ii.   √*tap*   → *tēpē*   (**<u>*tat*</u>*pē*)
        iii.   √*dab^h* → *dēb^{h}ur*   (**<u>*dad*</u>*b^{h}ur*)

Jun (2004:63–64) (and citations therein) claims that there is much greater overlap in the articulation of coronal-initial clusters than non-coronal–initial clusters, due to the more rapid articulation of the coronal constriction. If Rig-Vedic Sanskrit also had this gestural timing, it could well have been the case that this overlap masked the coronal release altogether in coronal-labial clusters. This would mean coronal-labial nasal-nasal and stop-stop clusters would have lacked the *intensity rise at release* cue, because there was no audible release. Since these clusters lack *IntRise:SLE* as well (since they are composed of consonants of the same manner class, and thus of the same overall intensity), depriving them of *intensity rise at release* would cause them to lack all the possible cues which could license a repetition, and thus they would be expected to violate the *PCR constraint. I see no equivalent explanation available to an account that makes reference only to sonority.

The inclusion of *intensity rise at release* as a repetition licensor might though also predict the licensing of repetitions for nasal-stop clusters. This would depend on whether nasals were released in pre-stop position. Given that Sanskrit displays place assimilation for pre-stop nasals, it is reasonable to assume that nasals were not released in this position, and thus should not be expected to have had their repetition licensed in pre-stop position. Nevertheless, there is a good deal of highly ambiguous evidence pointing towards nasal repetitions in pre-stop position, as laid out in footnote 27 above. If some or all of those forms do actually represent *bona fide* evidence of C_1-copying

at an earlier stage, the inclusion of *intensity rise at release* as a repetition licensor would explain that data, if we posit a nasal release in pre-stop position at that stage.

6.5.4 Klamath

The evidence from Klamath indicates that *IntRise:SLE* is again central to the licensing of repetitions. However, unlike Sanskrit (both stages) and Gothic, *IntRise:SLE* turns out not to be sufficient on its own to license repetitions.

The table in (89) — adapted from (32) above, but now with the different types of sonorants separated out — shows the cluster-wise distribution of *PCR effects in Klamath: ✓ represents clusters with a CV reduplicant (repetition licensed); ✗ represents clusters with a CCV reduplicant (repetition not licensed); grayed out cells are not attested with reduplication.[52] The only cluster type that licenses a repetition is stop-sonorant (as also in Ancient Greek; see below).

(89)   Cluster-wise reduplicant shape distribution in Klamath

C$_1$ \ C$_2$	Stop	*s*	*ʔ*	Nasal	Liquid	Glide
Stop	✗	✗		✓	✓	✓
s	✗		✗	✗	✗	✗
Nasal	✗					
Liquid	✗		✗			✗
Glide	✗					

The fact that *s*-sonorant clusters fail to license repetitions while stop-sonorant clusters do shows that *IntRise:SLE* is not on its own a sufficient condition for licensing repetitions, because both types contain a clear rise in overall intensity between C$_1$ and C$_2$. Similarly, stop-*s* cues C$_1$ with *IntRise:SLE*, and this too fails to license repetition. Nevertheless, the absence of *IntRise:SLE* does succeed in ruling out all of the other clusters that don't license repetition. So, it seems that the way to conceive of the *PCR licensing condition for Klamath is that *IntRise:SLE* and *some other cue* are simultaneously necessary in order to license repetition.

That other cue could be *IntRise:Rel*. If we assume that stop-*s* is produced without an audible release of the stop, then the only clusters with both *IntRise:SLE* and *IntRise:Rel* are the stop-sonorant clusters. The lack of *IntRise:Rel* would rule out stop-*s* and *s*-sonorant. However, this would predict that nasal-liquid and nasal-glide clusters would also license repetitions, if the nasal was produced with an audible release in these positions. No such root-initial clusters are attested with reduplication; therefore, this prediction cannot be tested.

To avoid the potential issue with nasals, we might instead say that *stop release burst* is the additional cue, rather than (or in addition to) *IntRise:Rel*. *Stop release burst* is a property of oral stops (see Flemming 1995/2002:23–24, Wright 2004:38). It results from the sudden expulsion of built-up air pressure in the oral cavity after the release of the complete oral occlusion which characterizes oral stops. The stop release burst consists of a short duration of high-amplitude aperiodic noise. This distinguishes oral stop releases from nasal stop releases.

Stop release burst is primarily a cue to stop place contrasts (see Flemming 1995/2002:23–24 and many others). However, since a stop release burst can only be generated by the production of

---

[52] Note that the stop-sonorant category does attest cluster-copying forms in addition to C$_1$-copying forms. I analyze these as the result of a (lexically-indexed) CONTIGUITY effect; see Section 6.6.3 below.

an oral stop, this will also be an unambiguous cue to the presence of a stop (i.e. the C∼Ø contrast). Therefore, it is reasonable to assume that *stop release burst* could also participate in the licensing of repetitions

The Klamath distribution in (89) can thus be captured by positing that *IntRise:SLE*, coupled with either *IntRise:Rel* or *stop release burst*, is sufficient to license repetition. Note, however, that this exclusively picks out stops (and possibly nasals) as the only type of segment which can be locally repeated. This is not correct. All CV-initial roots show CV-reduplication (i.e. [C_1V-C_1V...]), which yields a transvocalic repetition. This would incorrectly predict that only stop-initial CV roots could reduplicate in this way, and that CV roots with all other initial consonants would have to employ some other pattern (for example, copying a coda consonant or copying an additional syllable).[53]

In order to circumvent this problem, we must posit that there is an alternative licensing condition for the Klamath *PCR constraint. Namely, *CV transitions* can license repetition. *CV transitions* are the type of *CR transitions* (see Section 6.5.2) that surface on a vowel. Since vowels have higher amplitude then sonorant consonants, it is reasonable to assume that formant transitions affecting vowels serve as a more robust cue to relevant contrasts than equivalent formant transitions affecting sonorant consonants. Note that *CR transitions* which are not of the *CV* variety cannot be a sufficient repetition licensor in Klamath because this would predict that *s*-sonorant clusters would license repetitions, contrary to fact.

The inclusion of *CV transitions* among the repetition licensing conditions for Klamath now captures the full distribution. The Klamath *PCR constraint is formalized in (90). It consists of disjunctive licensing conditions. A repetition is licensed when the repeated consonant is cued by both *IntRise:SLE* and *IntRise:Rel* (or perhaps *IntRise:SLE* and *stop release burst*). Alternatively, a repetition can also be licensed when the repeated consonant is cued by *CV transitions*. A repetition that meets both conditions will obviously be licensed as well.

(90) ***PCR [*for Klamath*]**
      Assign a violation mark * for each output C_α^2 which lacks the requisite cues.
      ▶ **REQUISITE CUES:**
        i. *IntRise:SLE* + *IntRise:Rel* (or + *stop release burst*?), or
        ii. *CV transitions*

6.5.5 Ancient Greek

Ancient Greek (by which I mean Classical and Pre-Classical Greek; see Section 6.3.2) can be accounted for in basically the same way as Klamath. It's distribution is essentially a subset of what is observed for Klamath; each cluster type attested in Ancient Greek (with the exception of nasal-nasal) is also attested in Klamath and shows the equivalent treatment. The one significant new point raised by Ancient Greek comes from the variation observed in voiced stop + liquid clusters.

Ancient Greek permits C_1-copying, and thus the formation of a repetition, for stop-sonorant clusters. It disallows repetitions, exhibiting instead the "non-copying" pattern, for all other cluster types attested in the perfect — namely, stop-stop, stop-fricative, fricative-stop, and (probably)

---

[53] It may be possible to construct a consistent analysis where these roots actually do violate *PCR. If *PCR is ranked below the constraints which would be violated by the possible alternative reduplication strategies available to CV-initial roots, then CV-reduplication could be optimal even when it resulted in a *PCR violation. This would simplify the *PCR constraint. However, since I have not yet worked out the details of such an analysis, I proceed under the assumption that *PCR is not violated under these circumstances.

nasal-nasal. Additionally, there is variation between these two types of treatments for voiced stop + liquid clusters. This distribution is schematized in (91), and the resulting generalization regarding the licensing of repetitions is provided in (92) (repeated from (21)/(60) above). Like each of the other languages, all CV-initial roots reduplicate in CV, and thus all consonant repetitions are licensed in pre-vocalic position.

(91)   Initial clusters and reduplicative behavior in Ancient Greek (repeated from (20)/(27) above)

C$_1$ \ C$_2$	Vcls Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Vcls Stop (T)	✗ (*T$_\alpha$VT$_\alpha$T)	✗ (*T$_\alpha$VT$_\alpha$S)	✓ (T$_\alpha$VT$_\alpha$N)	✓ (T$_\alpha$VT$_\alpha$L)
Vcd Stop (D)			✓ (D$_\alpha$VD$_\alpha$N)	✓∼✗ ($^{(*)}$D$_\alpha$VD$_\alpha$L)
Fricative	✗ (*S$_\alpha$VS$_\alpha$T)			
Nasal			✗ (*N$_\alpha$VN$_\alpha$N)	

(92)   **Ancient Greek repetition generalization:**
Stop-sonorant clusters license repetitions, other clusters do not. (Variation in voiced stop + liquid clusters.)

Putting aside for the moment the variation in voiced stop + liquid clusters, the Klamath *PCR constraint will be completely sufficient to account for this distribution, under the assumption that stops are not released before fricatives (as assumed for Klamath as well). All the clusters which Greek has in common with Klamath pattern in the same way; therefore, the Klamath *PCR constraint will obviously account for those in Greek as well. The only Greek cluster not attested in Klamath is nasal-nasal. This is comprised of two segments of the same type, and so it does not have *IntRise:SLE*. Therefore, this is correctly ruled out by the Klamath licensing conditions as well. Given this, we can simply import the Klamath *PCR constraint for Ancient Greek:

(93)   ***PCR [*for Ancient Greek*]**   (= Klamath (90))
Assign a violation mark * for each output C$_\alpha^2$ which lacks the requisite cues.
   ▶ REQUISITE CUES:
      i.   *IntRise:SLE + IntRise:Rel* (or + *stop release burst*?), or
      ii.   *CV transitions*

The only wrinkle here is indeed the variation observed in the voiced stop + liquid clusters. I propose that this effect results not specifically from confusion between C and Ø, but place confusion between voiced stops. Flemming (2007) shows that voiced stops are more confusable with one another before *l* than before *r*, because of the way the lateral articulation of *l* affects the stop burst properties and formant transitions out of a preceding voiced stop. If the perception of stop place contrasts before *l* is significantly worse in voiced stops than voiceless stops, if place confusion can somehow be rolled into or overlaid on the *PCR C∼Ø confusion problem, then we might expect something of this sort. In general, more thought is required about how place and manner contrasts may play into repetition licensing vis-à-vis *PCR. But this fact of Greek suggests that it is somehow implicated in solution.

274

6.5.6 Discussion

6.5.6.1 Assessing the (Micro-)Typology of *PCR Effects

We have now identified three distinct versions of the *PCR constraint, i.e. three different sets of "requisite cues" that functioned as sufficient conditions for licensing repeated consonants in one language or another. These are collected in (94). In the cases of Gothic, Klamath, and Ancient Greek, there were alternative interpretations employing slightly different cues. I focus here on the interpretations based on the two types of intensity rise cues.

(94)   *PCR licensing conditions (most permissive to least permissive)
      a.   Rig-Vedic Sanskrit   (from (87))
          i.   Intensity rise between sound level extrema
          ii.   Intensity rise at release
      b.   Later Sanskrit, Gothic   (from (79), (83))
          i.   Intensity rise between sound level extrema
      c.   Klamath, Ancient Greek   (from (90), (93))
          i.   Intensity rise between sound level extrema and Intensity rise at release
          ii.   CV transitions

The various definitions appear to show a cline in terms of the strictness of the repetition licensing conditions. In Rig-Vedic Sanskrit (94a), either sort of intensity rise is sufficient to license a repetition. Later Sanskrit (94b) differs from this minimally: it has removed *intensity rise at release* as a sufficient licensor on its own, but still allows an *intensity rise between sound level extrema* to do the job.

Although it has many fewer cluster types with which to determine the definition of *PCR, Gothic (94b) is consistent with this definition as well. It is alternatively consistent with a definition based just on *CR transitions*. However, since the other systems appear to be organized around the intensity rise cues, it seems appropriate to stick with *intensity rise between sound level extrema* as the repetition licensor for Gothic as well.

Klamath and Ancient Greek are decidedly more stringent than all of these other systems. Whereas Gothic and the two periods of Sanskrit all licensed repetitions with at least an *intensity rise between sound level extrema*, Klamath and Ancient Greek license repetitions with an *intensity rise between sound level extrema* only if it is accompanied by another cue — this might be either an *intensity rise at release* or a *stop release burst*. Otherwise, they require *CV transitions*.

Based on this admittedly very small sample, the trend that appears to be emerging is that intensity rise is the backbone of repetition licensing, and that languages differ with respect to what type of intensity rise is sufficient, and whether any other cue can or must be involved.

6.5.6.2 Other Cues and the Repetition Context

I have now argued that intensity rise is the crucial cue for repetition licensing. I have also suggested that *CR/CV transitions* and *stop release burst* could play an ancillary role. Why is it these cues which seem to be relevant here and not others?

Blevins & Garrett (2004:121–125) argue that phonetic properties which tend to be perceived over an extended domain are the same ones which most frequently take part in harmony/spreading processes and most frequently undergo metathesis. According to Blevins & Garrett (p. 123), these include nasalization, rhoticity, and laterality. All of these are properties which would be present in the respective sonorant-initial rising intensity clusters. They argue that uncertainty in identifying the

original source of extended phonetic/phonological properties of this sort is the origin of metathesis processes; that is to say, metathesis results from learners anchoring a property at a historically incorrect point in the phonological string.

This logic suggests that extended properties of this sort might be unhelpful, or maybe even detrimental, for the perception of C∼Ø contrasts under repetition, in the following way. Listeners perceive the acoustic event contributed by the articulation of the first member of the repetition. That acoustic event has knock-on effects for some additional period of time after the completion of that consonant's articulation. At some time distance away from the articulation of that first consonant, an identical acoustic event is produced by the second member of the repetition. If that acoustic event is within the domain in which the effects of the first acoustic event are normally still perceptible, listeners will not know whether that acoustic imprint derives from a new articulation, or instead is the continued result of the first articulation.[54]

Frication noise, especially the high-intensity frication noise (in the higher frequency regions) of sibilants — which is the main cue to the s∼Ø contrast under normal circumstances — is likely to suffer from a similar problem. Blevins & Garrett (2004:128) state that:

> "...in consonant clusters containing sibilants, the sibilant noise somehow distracts the listener, leading to high confusion rates with respect to the linear order of segments (Bregman 1990). Specifically, there is a tendency to decouple sibilant noise from the rest of the speech stream, and this decoupling can result in dramatic misperceptions."

See also Ladefoged (2001:174–175) (as cited by Blevins & Garrett 2004:128, n. 8). If listeners have difficulty in this way with identifying the precise location of the sibilant noise event relative to the other events/segments in the speech stream, then it is reasonable to assume that, when two such events are located at a close proximity in time, listeners may have difficulty in determining that they are in fact separate events, and not one continuous event.

These concepts suggest that cues like frication noise, nasal resonance, and liquid-induced formant structure are at least *less effective* at cuing the presence of a repeated segment than more localized/localizable cues like intensity rise, CR transitions, and stop release burst, and that they may even be *detrimental* to cuing the presence of a repeated segment. This is supported by the empirical evidence: in all the languages which impose some special restriction on repetition (i.e. have an active *PCR constraint), these cues seem never to be able to license repetitions on their own.

6.5.6.3 Interim Conclusions

This section has shown that the range of repetition avoidance effects observed in the reduplicative patterns of the Indo-European languages, and Klamath as well, can be explained using a constraint that enforces licensing conditions for repeated consonants based on acoustic/auditory cues to potential C∼Ø contrasts. I have named this the NO POORLY-CUED REPETITIONS constraint (*PCR). *PCR economically captures the various reduplicative distributions by specifying which cues — or combinations of cues — count as sufficient repetition licensors on a language-by-language basis. The primary cue that seems to be relevant throughout is intensity rise (cf. Parker 2002, 2008, Yun 2016).

This cue-based approach leads to a more economical analysis, and one with clearer explanatory power, than does an approach based on sonority and/or phonological features alone. This is strong evidence that acoustic/auditory cues to consonantal contrasts play an important role in the

---

[54] It is conceivable that the confusion could be going in the opposite direction, with the accurate apprehension of the second member masking identification of the first member after the fact.

phonological grammar, since a constraint referencing cues is strong enough to affect the behavior of reduplicative copying, a quintessential morphophonological process.

6.6   Additional Empirical Evidence for *PCR from Reduplication

The evidence for *PCR effects is not limited to those discussed thus far. This section adduces and analyzes several more reduplication patterns that can, or indeed must, be analyzed using *PCR. The first is a fairly straightforward though limited case of infixal reduplication from Latin (Section 6.6.1), triggered by a desire to avoid *s* repetitions before a stop, i.e. *PCR. A more complex and revealing infixal reduplication pattern is found in the Sanskrit desiderative with vowel-initial roots (Section 6.6.2). Employing *PCR allows for a simple explanation of a distinction in the position of the reduplicative infix for roots with different types of post-vocalic clusters. An important subsidiary point made by this pattern is that a templatic syllable-alignment account, which otherwise could eschew the use of *PCR, appears to be inconsistent with the syllabification facts of the language; this argues for the *PCR approach. Lastly, in Section 6.6.3, I formalize the analysis of the alternation between C_1-copying and cluster-copying in Klamath, and show how *PCR is again crucial.

6.6.1   Latin: Infixing Perfect Reduplication in STVX– Roots

Among the Indo-European languages, Latin displays a unique reduplicative behavior for its STVX– roots in the perfect (see Weiss 2009:410): *PCR violations are avoided by *infixing* the reduplicant (cf. Fleischhacker 2005, DeLisi 2015). In this pattern, rather than the reduplicant deviating from its target shape *CV*, it deviates from its target position, namely at the left edge. It does this by placing the reduplicant *after* the root-initial *s*:[55]

(95)   Latin infixing perfect reduplication to ST roots (forms from Weiss 2009:410)

Root		Perfect		
$\sqrt{st}$	'stand/stop'	*s-te-t-ī*	not **_se_-st-ī*	(but present *si-st-ō*)
$\sqrt{spond}$	'promise'	*s-po-pond-ī*	not **_so_-spond-ī*	
$\sqrt{scid}$	'cut'	*s-ci-cid-ī*	not **_si_-scid-ī*	

This is predicted if ALIGN-RED-L, a constraint preferring the reduplicant to surface at the left edge of the word (defined in (96a)), is subordinated to *PCR and the other constraints whose violation could obviate *PCR — for example, ANCHOR-L-BR and *CC.[56] CONTIGUITY-IO,

---

[55] It may be possible to describe the pattern in slightly different terms. For example, one could identify the reduplicant as the *second* repeated string rather than the first. Such a segmentation might more transparently explain the lack of a second stem vowel in *stetī*. However, it would make deriving the vowel of the reduplicant significantly more complicated. See Riggle (2006) for an analysis of infixing reduplication in Pima, which touches on questions of this sort.

[56] Note that the size minimizer constraint must here be *CC and not ALIGN-ROOT-L, as the ranking ALIGN-ROOT-L ≫ CONTIGUITY-IO, ALIGN-RED-L would predict infixation also for CVX– roots, since the abiding factor would be placing the root at the left edge. While many of the reduplicated CVX– roots could be consistent with an infixing analysis (because the vowel is frequently identical, or in part determined by context; cf. McIntyre 1992), I find at least one example where it is clear that the reduplicant is the first syllable rather than the second: *ce-cīd-ī* ← $\sqrt{caed}$ 'cut down'. The long vowel in the second syllable must relate to the root's diphthong.

which prohibits disrupting input adjacency relations in the output (defined in (96b)), must also be subordinated to these constraints. This alignment approach also correctly predicts that infixation is minimal: candidate (97d) ≻ candidate (97e). (In (97), the base of reduplication is the string to the right of the reduplicant.)

(96) a. **ALIGN-RED-L**
          Assign one violation mark * for each segment intervening between the left edge of the reduplicant and the left edge of the word.
      b. **CONTIGUITY-IO**
          Assign one violation mark * for each pair of segments which are adjacent in the input that have non-adjacent correspondents in the output.

(97) Infixing reduplication in Latin to avoid *PCR violation

/RED, scid, ī/		*PCR	ANCHOR-L-BR	*CC	CONTIG-IO	ALIGN-RED-L
a.	s<u>i</u>-scid-ī	*!		*		
b.	c<u>i</u>-scid-ī		*!	*		
c.	sc<u>i</u>-scid-ī			**!		
d. ☞	s-c<u>i</u>-cid-ī			*	*	*
e.	sc-<u>id</u>-id-ī			*	*	**!

This analysis predicts that TRVX– roots should exhibit the standard Indo-European C_1-copying pattern — for example, hypothetical $\sqrt{plen\text{-}} \rightarrow pe\text{-}plen\text{-}$, not $p\text{-}le\text{-}len\text{-}$ — because infixation would be triggered only by *PCR-violating repetitions (of which $T_\alpha VT_\alpha R$ would not be one). Latin unfortunately does not have any reduplicated forms to TRVX– roots (see Cser 2009), so we cannot test whether this behavior is truly triggered by *PCR or whether infixation would apply to cluster-initial roots across-the-board. I do not believe that any of the constraints involved in the analysis of the other Indo-European systems could generate across-the-board infixation for cluster-initial roots without also predicting it for singleton-initial roots.

6.6.2 Sanskrit: Infixing Desiderative Reduplication in Vowel-Initial Roots

In addition to the perfect reduplication pattern discussed in Section 6.2.3 above (see also Chapter 5), Sanskrit also shows reduplication in a number of other verbal categories (consult Kulikov 2005). One such category is the desiderative (see Whitney 1889:372–374/§1026–1031), which is marked by prefixal reduplication (with a fixed [+high] vowel, matching the base vowel in [±round]) and a suffix -(i)ṣa-, which attaches immediately after the root. For consonant-initial roots, the distribution of reduplicant shape is the same as in the perfect: C_1-copying to TRVX– roots (98a), C_2-copying to STVX– roots (98b).

(98) a. √*tvar* 'hasten' → desiderative <u>ti</u>-tvar-iṣa-,   perfect *ta-tvar-*
      b. √*stambʰ* 'prop' → desiderative <u>ti</u>-stambʰ-iṣa-, perfect <u>ta</u>-stambʰ-

For vowel-initial roots, however, it displays something different. In Vedic Sanskrit, the earliest attested stage of the language, truly vowel-initial roots are relatively infrequent, and there does not seem to be a particularly consistent pattern for their reduplication. In the perfect, at least, they generally display vowel-lengthening, which is understandable based on these roots' diachronic origins as laryngeal-initial roots (compare the origin of vowel-lengthening perfects in Ancient Greek

in Chapter 2). In Classical Sanskrit, though, vowel-initial roots in the desiderative show consistent behavior, namely infixation. A good number of these forms are not attested in literary texts, but rather only reported by the grammarians. We can thus wonder about their linguistic reality; but, given their internal consistency, and the way in which they are consistent with the rest of the grammar, they at the very least reflect the active phonological grammar of the grammarians themselves. The pattern is illustrated in (99) below.

(99) Classical Sanskrit infixing desiderative reduplication to vowel-initial roots

(forms from Whitney 1885 [1988])

	Root shape	Root		Desiderative	
a.	VC	√aɟ	'drive'	a-ɟi-ɟ-iṣa-	not **aɟ-aɟ-iṣa-
		√īḍ	'praise'	ī-ḍi-ḍ-iṣa-	not **īḍ-īḍ-iṣa-
		√ēdʰ	'thrive'	ē-di-dʰ-iṣa-	not **ēd-ēdʰ-iṣa-
		√ēɟ/īɟ	'stir'	ī-ɟi-ɟ-iṣa-	not **īɟ-īɟ-iṣa-
b.	VCC	√arc	'praise'	ar-ci-c-iṣa-	not **a-ri-rc-iṣa-
		√ard	'stir'	ar-di-d-iṣa-	not **a-ri-rd-iṣa-
		√ubɟ	'force'	ub-ɟi-ɟ-iṣa-	not **u-bi-bɟ-iṣa-
		√aɲɟ	'anoint'	aɲ-ɟi-ɟ-iṣa-	not **a-ɲi-ɲɟ-iṣa-
		√uɲcʰ	'glean'	uɲ-ci-cʰ-iṣa-	not **u-ɲi-ɲcʰ-iṣa-
c.	Vkṣ	√akṣ	'attain'	ā-ci-kṣ-iṣa-	not **āk-ṣi-ṣ-iṣa-
		√īkṣ	'see'	ī-ci-kṣ-iṣa-	not **īk-ṣi-ṣ-iṣa-

Modulo the special behavior of Vkṣ roots, an identical pattern is found in a number of other languages (McCarthy & Prince 1993b:132), most famously Timugon Murut (Prentice 1971). For an analysis of the Timugon Murut pattern, see McCarthy & Prince (1993b:132–135) (also McCarthy & Prince 1986 [1996], 1993a, *et seq.*). Following Kiparsky (1986), McCarthy & Prince already recognized Sanskrit's membership in this group.

6.6.2.1 The Position of Infixation and *PCR

When the root is of the shape VC_1 (99a), the reduplicant is infixed after the root-vowel, and takes the shape -C_1i-, where the i is (presumably) a copy of the suffix vowel. The fact that it is the first vowel in the VC_1VC_1 string and not the second that matches the root vowel — i.e. √aɟ → aɟiɟ-, not **iɟaɟ- — shows that the reduplicant is indeed infixal and not prefixal. While the infixal nature of this pattern is on its own of significant interest (and I return to the motivation for infixation extensively below), the reason why this pattern is relevant to a discussion of *PCR is the way that it treats roots of the shape VC_1C_2 (99b), and its sub-type Vkṣ (99c).[57]

Given that the reduplicant is prefixed in consonant-initial forms, we should expect infixation to be minimal; that is to say, the reduplicant should surface as close to the left edge as possible, so as to minimize violations of ALIGN-RED-L, just as for Latin in Section 6.6.1 above. Once infixation is itself motivated (again see below), this incorrectly predicts that VC_1C_2 roots should have redupli-

---

[57] At the end of this subsection, I will entertain a templatic syllable-alignment analysis of infixation. Within that analysis, *PCR does not need to play a crucial role, though there is nothing about that analysis which contradicts *PCR. This suggests that such an analysis may be missing an important generalization.

cated forms of the shape $**V\text{-}C_1i\text{-}C_1C_2\text{-}i\!\!\:sa\text{-}$, where the reduplicant surfaces between the root-vowel and the first consonant. Instead, the reduplicant surfaces after the first consonant: $VC_1\text{-}C_2i\text{-}C_2\text{-}i\!\!\:sa\text{-}$. The reason: for the type of clusters that are attested among these roots (with one crucial exception), infixing after the vowel would cause a *PCR violation. With the exception of roots in Vkṣ (99c), all of the VC_1C_2 roots have post-vocalic clusters which would not be expected to license a consonant repetition under the definition of the *PCR constraint for Later Sanskrit as defined in (79) above — which requires that the repeated consonant be cued by an intensity rise (reckoned by comparing the relative sound level extrema of the repeated consonant and the segment that follows). By moving the reduplicant one position to the right, it is able to copy a consonant which is pre-vocalic — i.e. root-C_2, which is followed by the suffix vowel i. $C_\alpha VC_\alpha$ sequences are always licensed pre-vocalically, so this satisfies *PCR. Put another way, the need to satisfy *PCR compels extra violations of ALIGN-RED-L.

This case of *PCR-driven non-minimal infixation is demonstrated in the tableau in (100). (I continue to temporarily assume that only those candidates with an infixal reduplicant will be well-formed.) The minimal infixation candidate (100a) is prevented from surfacing because its -*bibɟ*-sequence violates *PCR. The candidate that pushes the reduplicant all the way past the root (100c) incurs an additional and unnecessary violation of ALIGN-RED-L. Therefore, despite candidate (100b)'s non-minimal number of ALIGN-RED-L violations (two compared to candidate (100a)'s one), it is selected as optimal.

(100)   Non-minimal infixation to *PCR-violating cluster: √*ubɟ* 'force' → desiderative *ub-ɟi-ɟ-iṣa-*

/RED, ubɟ, -iṣa-/	*PCR	ALIGN-RED-L
a. u-<u>bi</u>-bɟ-iṣa-	*!	*
b. ☞ ub-<u>ɟi</u>-ɟ-iṣa-		**
c. ubɟ-i<u>ṣ</u>-iṣa-		***!

What confirms this analysis, though, is the behavior of the Vkṣ roots in (99c). When a root velar is copied in reduplication, its reduplicant correspondent is palatal. The sequence -*cVkṣ*- does not violate *PCR, both because stop-fricative clusters meet the repetition licensing conditions (since they contain an intensity rise at the cluster-internal juncture) and because the base-reduplicant correspondents are non-identical. It is in just this case, where a *PCR violation is not at stake, that a root with two post-vocalic consonants exhibits minimal infixation, as shown in the tableau in (101). With no *PCR violation to rule out candidate (101a), candidate (101b)'s extra ALIGN-RED-L violation is now fatal.[58]

(101)   Minimal infixation to Vkṣ roots: √*akṣ* 'attain' → desiderative *ā-ci-kṣ-iṣa-*

/RED, ākṣ, -iṣa-/	*PCR	ALIGN-RED-L
a. ☞ ā-<u>ci</u>-kṣ-iṣa-		*
b. āk-<u>ṣi</u>-ṣ-iṣa-		**!
c. ākṣ-i<u>ṣ</u>-iṣa-		**!*

---

[58] Donca Steriade suggests to me that candidate (101b) might be dispreferred because it has three retroflex sibilants in sequence. This does not affect the arguments being made here regarding *PCR, because the ranking *PCR ≫ ALIGN-RED-L is still necessary in order to account for (100), and the activity of ALIGN-RED-L would still be required to select (101a) over (101c).

This analysis predicts that, if roots of the shape VTR existed, that they would reduplicate like Vkṣ roots, showing minimal infixation: for example, hypothetical √*atr* → desiderative *a-tï-tr-iṣa-*. No such roots are attested.

6.6.2.2 Infixation as Cue-Based Faithfulness

To complete the analysis, we need to establish what constraint is driving infixation in the first place. I propose that we should understand the motivation for infixation as a desire to leave (root) vowels that were non–post-consonantal in the input, or in an output base, non–post-consonantal in the reduplicated form. I formalize this in terms of a cue-based approach to faithfulness. Subsequently, I show that a "templatic" analysis based on syllable alignment is not workable, as it either runs into a ranking paradox, has to posit syllabification facts that don't comport with independent syllabification evidence, or requires a type of templatic constraint which is without parallel in the literature.

The intuition behind the cue-based approach is that the effects on a vowel of being post-consonantal or not are properties that can be referenced by faithfulness constraints. Specifically, any vowel which is post-consonantal in the output will host CV transitions (see, e.g., Wright 2004:37). On the other hand, if a vowel is non–post-consonantal in the output, it will surface in its "pristine" state, unburdened by the coarticulatory effects of a preceding consonant. I propose that this difference, i.e. the presence vs. absence of CV transitions in a vowel, is a property that can be tracked by faithfulness constraints.

The way to deploy this type of faithfulness for the present problem is to have a constraint that militates against giving CV transitions to a vowel that previously lacked CV transitions. This can be viewed as a DEP constraint. As for which correspondence dimension over which this faithfulness constraint holds, the simplest approach will be to define it over the Base-Derivative correspondence relation. Since all output forms will necessarily contain fully detailed cue information, correspondence between output forms will easily have access to cues in the calculation of faithfulness. This is the faithfulness dimension I will employ here.[59]

Therefore, the constraint that triggers infixation is DEP[CVtransitions]/V-BD, which is defined in (102). I assume that the base is some verbal category which realizes the root faithfully, for example the present. Note that this constraint only requires that vowels not come to be associated with new CV transitions (hence the "/V" in the constraint name); it does not prohibit consonants from being associated with CV transition.

(102)   **DEP[CVtransitions]/V-BD**   (DEP[CVtrans]-BD)
        Assign one violation mark * for each vowel in the derivation which hosts CV transitions that has a corresponding vowel in the base which does not host CV transitions.

The basic case of infixation is the VC root, for example √*aɟ* 'drive' → desiderative *a-ɟi-ɟ-iṣa-*. This is shown in (103). Note that, in this and the following tableaux, the root portion of the output string is in bold; violations of DEP[CVtrans]-BD can be reckoned by whether the bolded vowel is

---

[59] Depending on our conception of the input, it may be possible to define these constraints over the Input-Output correspondence dimension instead or in addition. Given that the distribution of cues is predictable in the output, cues ought to be viewed as derived properties rather than properties of the input. If, though, we adopted a serial model of the phonetics-phonology interface along the lines of Flemming's (2008) "Realized Input" — where the input to the core phonological derivation is based upon the phonetically-detailed realization of the faithful mapping of the input — the phonology could have access to an input representation that contains cues without requiring that that information be stored in the lexicon. Therefore, it is possible that we could use Input-Output cue-based faithfulness constraints to capture the pattern. A conceptual difficulty with this approach as regards reduplication is how to understand the "faithful" candidate, and, more generally, how this system would interact with morphologically complex derivations. These are questions for another day.

preceded by a consonant. I assume that the base of reduplication is everything to the right of the reduplicant (see below for a substantive argument on this point); therefore, ANCHOR-L-BR violations are reckoned by whether the leftmost segment of the reduplicant corresponds to the segment which immediately follows the reduplicant.[60]

(103)　　Infixation for VC roots

INPUT: /RED, aɟ, -iṣa-/　BASE: [aɟ-]	DEP[CVtrans]-BD	ONSET	ALIGN-RED-L	ALIGN-ROOT-L	ANCHOR-L-BR
a.　　aɟ-**aɟ**-iṣa-	*!	*		**	
b.　☞　**a-ɟi-ɟ**-iṣa-		*	*		
c.　　**aɟ**-iṣ-iṣa-		*	**!		
a.　　a-**aɟ**-iṣa-		**!		*	
e.　　ɟ-**aɟ**-iṣa-	*!			*	*

In (103), aligning a VC reduplicant to the left edge (candidate (103a)) causes a violation of DEP[CVtrans]-BD, because the consonant that has been copied into the reduplicant induces CV transitions on the root vowel, whose correspondent in the output base (i.e. the [a] of BASE [aɟ-]) lacked CV transitions. The same holds for candidate (103e), which has copied just the root consonant into the reduplicant. (This alleviates an ONSET violation, but as long as DEP[CVtrans]-BD ≫ ONSET, this will not be optimal. This candidate also violates ANCHOR-L-BR, but this constraint must be low-ranked in order to allow for the C_2-copying for STVX– roots; see Chapter 5 and the summary in Section 6.2.3.) Candidate (103d) copies just the vowel and places it at the left edge. This succeeds in maximizing left-alignment without creating CV transitions for the root vowel, but at the expense of an extra ONSET violation, and is thus suboptimal. The remaining candidates (103b) and (103c) are those which displace the reduplicant from the left edge. This allows the root /a/ to remain in absolute initial position, and thus not acquire new CV transitions. The two differ in the extent to which they displace the reduplicant, and thus the number of ALIGN-RED-L violations they incur. Candidate (103b) places the reduplicant just after the root-initial vowel, incurring one ALIGN-RED-L violation, while candidate (103c) pushes the reduplicant past the root-final consonant as well, incurring an additional ALIGN-RED-L violation, which is unnecessary and fatal. The ranking DEP[CVtrans]-BD ≫ ONSET ≫ ALIGN-RED-L thus derives minimal infixation, resulting in a CV reduplicant after the root-initial vowel.

　　The derivation for VCC roots, illustrated with √*ubɟ* 'force' → desiderative *ub-ɟi-ɟ-iṣa-* in (104), is identical to that for the VC case, but with the added complication of *PCR (as demonstrated in Section 6.6.2.1), and the additional possible candidates furnished by having an extra root consonant. In comparison to the tableau in (100) above (which only considered infixal candidates), the noteworthy addition is that DEP[CVtrans]-BD is responsible for ruling out left-aligned reduplicant candidates like (104a–b) and (104g–h), because each places a consonant before the root-initial vowel. Among the infixal candidates (104c–e), the normally optimal minimal infixation candidate (104c) is blocked by *PCR (because it creates a pre-obstruent consonant repetition). Therefore, the

---

[60] This approach to ANCHOR-L-BR is reminiscent of Nelson's (2003) LOCALITY constraint. The two are equivalent for the analysis of Sanskrit.

next most minimal infixation candidate (104d), which places the reduplicant after the root-initial VC sequence, is selected as optimal.

(104)   Infixation for VCC roots

INPUT: /RED, ubɟ, -iṣa-/ BASE: [ubɟ-]	*PCR	DEP[CVtrans]-BD	ONSET	ALIGN-RED-L	ALIGN-ROOT-L	ANCHOR-L-BR
a. u͟b-**ubɟ**-iṣa-	*!	*!	*		**	
b. ub͟ɟ-**ubɟ**-iṣa-		*!	*		***	
c. **u**-b͟i-**bɟ**-iṣa-	*!		*	*		
d. ☞ **ub**-ɟ͟i-**ɟ**-iṣa-			*	**		
e. **ubɟ**-i͟ṣ-iṣa-			*	***!		
f. u͟-**ubɟ**-iṣa-			**!		*	
g. b͟-**ubɟ**-iṣa-	*!	*!			*	*
h. ɟ͟-**ubɟ**-iṣa-		*!			*	*

There are two more candidates worth mentioning here. The first is **ub**-*i*-**ɟ**-*iṣa*- (105b). This is a candidate which copies just the vowel, and places it in the position where the optimal CV reduplicant goes, between the first and second root consonants. This candidate fares equally well as the winner on *PCR and DEP[CVtrans]-BD, as well as the alignment constraints ALIGN-RED-L and ALIGN-ROOT-L. What distinguishes it from the winner is its ANCHOR-L-BR violation: it copies a segment which is not at the beginning of the string immediately following the reduplicant.

(105)   Infixation for VCC roots

INPUT: /RED, ubɟ, -iṣa-/ BASE: [ubɟ-]	*PCR	DEP[CVtrans]-BD	ONSET	ALIGN-RED-L	ALIGN-ROOT-L	ANCHOR-L-BR	*CC
a. ☞ **ub**-ɟ͟i-**ɟ**-iṣa-			*	**			*
b. **ub**-i͟-**ɟ**-iṣa-			*	**		*!	

This demonstrates that we must use ALIGN-ROOT-L as the general size minimizer constraint in the fuller analysis (i.e. the constraint which disallows copying an entire root-initial consonant cluster into the reduplicant), for the following reason. In Chapter 5 (reviewed in Section 6.2.3), I showed that a size minimizer must dominate ANCHOR-L-BR in order to select C_2-copying, and not cluster-copying, as the optimal *PCR avoidance strategy for STVX– roots. The cluster-initial roots data would permit either ALIGN-ROOT-L or *CC to be that size minimizer constraint. Now, losing candidate (105b) does succeed in avoiding a consonant cluster by copying just a vowel; the winning candidate (105a) exhibits this cluster (i.e. the root *b* + the reduplicant *ɟ*). Since candidate (105b) is ruled out by ANCHOR-L-BR, we know that ANCHOR-L-BR ≫ *CC. This leaves ALIGN-ROOT-L as the only size minimizer that can dominate ANCHOR-L-BR.

The other additional candidate worth mentioning is the following. There is in theory a candidate **ub**-u͟-**ɟ**-*iṣa*-, which has a structure identical to (105a) but copies the root vowel to its left rather than the suffix vowel to its right. If the string to the left of the reduplicant could equally well serve as the base for reduplication as the string to its right, then this would actually seem to satisfy

ANCHOR-L-BR, because the root vowel is leftmost in that string. Therefore, we either need to reify the condition that the reduplicant copies only from its right, or we need to switch to LOCALITY (Nelson 2003). LOCALITY, which penalizes intervention between the reduplicant and the string containing the segments it copies, will equally well penalize non-local copying from the left and from the right. This seems like the superior option in this case.

The tableau in (106) simply confirms that this analysis continues to predict that infixation will once again be minimal for roots of the shape Vkṣ — i.e. √ākṣ 'attain' → desiderative ā-ci-kṣ-iṣa- — because they are not subject to *PCR violations. All candidates are equivalent to those in (104).

(106) Infixation for Vkṣ roots

INPUT: /RED, ākṣ, -iṣa-/ BASE: [ākṣ-]		*PCR	DEP[CVtrans]-BD	ONSET	ALIGN-RED-L	ALIGN-ROOT-L	ANCHOR-L-BR
a.	āc-**ākṣ**-iṣa-	*!		*		**	
b.	**ākṣ**-**ākṣ**-iṣa-	*!		*		***	
c. ☞	**ā**-ci-**kṣ**-iṣa-			*	*		
d.	**āk**-si-**ṣ**-iṣa-			*	**!		
e.	**ākṣ**-iṣ-iṣa-			*	***!		
f.	ā-**ākṣ**-iṣa-			**!		*	
g.	c-**ākṣ**-iṣa-	*!				*	*
h.	ṣ-**ākṣ**-iṣa-	*!				*	*

Lastly, the tableau in (107) confirms that there will be no infixation for STVX– roots, e.g. √*stan* 'thunder' → desiderative *ti-stan-iṣa-*. Only a violation of DEP[CVtrans]-BD is sufficient to trigger infixation. DEP[CVtrans]-BD will never be in danger of violation for STVX– roots, because all of their vowels are post-consonantal in the base, and thus already have CV transitions. Without the threat of a DEP[CVtrans]-BD violation, any ALIGN-RED-L violation is fatal, and thus the reduplicant will surface at the left edge.

(107) No infixation for STVX– roots

INPUT: /RED, stan, -iṣa-/ BASE: [stan-]		*PCR	DEP[CVtrans]-BD	ONSET	ALIGN-RED-L	ALIGN-ROOT-L	ANCHOR-L-BR
a.	si-**stan**-iṣa-	*!				**	
b. ☞	ti-**stan**-iṣa-					**	*
c.	sti-**stan**-iṣa-					***!	
d.	s-ti-**tan**-iṣa-				*!		
e.	**sta**-ni-**n**-iṣa-				*!**		

This form additionally provides the ranking argument for ALIGN-RED-L ≫ ALIGN-ROOT-L. Aligning the reduplicant to the left edge of the word necessarily displaces the root from the left edge, and thus induces violations of ALIGN-ROOT-L in candidates (107a–c). Since these violations

284

would be absent in the infixal candidates (107d–e), which instead incur ALIGN-RED-L violations, we require the ranking ALIGN-RED-L ≫ ALIGN-ROOT-L.

The full DEP[CVtrans]-BD–based analysis of the Sanskrit desiderative, integrated with the results from the Sanskrit perfect in Chapter 5, is summarized in (108).

(108)   Ranking for Sanskrit desiderative reduplication

ONSET   DEP[CVtrans]-BD   *PCR

ALIGN-RED-L

ALIGN-ROOT-L

ANCHOR-L-BR

CONTIGUITY-BR   *CC
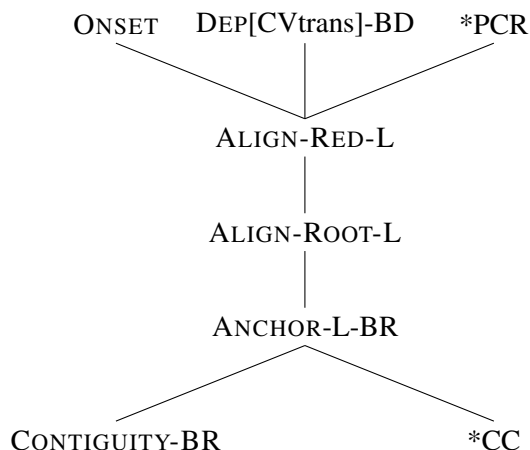
6.6.2.3 Infixation as Templatic Syllable-Alignment

The traditional templatic analysis of infixing patterns like the Sanskrit desiderative is that the reduplicant seeks to be aligned to syllable edges (see McCarthy & Prince 1993b:132–135 for this analysis of the equivalent pattern in Timugon Murut). However, the totality of the Sanskrit patterns, and (one interpretation of) the syllabification facts of the language, make this analysis problematic.

If we are to construct an analysis based on syllable-alignment, it will be crucial to understand the syllabification facts of the language. For Vedic, the earliest period of the language, the preponderance of evidence — specifically, the patterns of poetic metrics and (morpho)phonological lengthening processes and alternations — argues that *all* medial two-consonant clusters, regardless of their sonority profile, were heterosyllabic, i.e. parsed as simplex coda + onset (see Vaux 1992:294–296, Cooper 2014:Ch. 2, and references therein; consult also Hermann 1923, Byrd 2015). However, for later periods of the language (Epic Sanskrit, Classical Sanskrit, etc.), Vaux (1992:esp. 296–298) argues, based on claims about gemination processes and Classical Sanskrit meter (citing Varma 1929), that many clusters (mainly rising sonority clusters — including *-kṣ-*, though the evidence here is minimal) came to have tautosyllabic (i.e. complex onset) parses. Since many of the infixing desiderative forms (including all of the ones to VCC roots except *ī̆cikṣiṣa-* ← √*ī̆kṣ* 'see'; Whitney 1889:373/§1029) are not attested until the Classical period (or rather, first cited by Classical grammarians), if the syllabification had changed in this way, the syllable-alignment analysis presented below could be made to work.

I will first show that an analysis which treats medial *-kṣ-* as heterosyllabic completely fails. I will then show that the same analysis appears to succeed if we assume a tautosyllabic parse for *-kṣ-*. Therefore, the syllable-alignment analysis rests entirely on the claim of tautosyllabicity for *-kṣ-* in Classical Sanskrit. I present evidence from length alternations in the reduplicated aorist that seem argue against a tautosyllabic parse.

Failure of the Syllable-Alignment Analysis with the Heterosyllabic Parse

Let us start by considering the ramifications of across-the-board heteroyllabic parsing of medial two-consonant clusters in Classical Sanskrit. Under this assumption, in, for example, reduplicated desideratives of cluster-initial roots, which reduplicate with prefixal CV (where which root consonant C corresponds with depends on the type of cluster; see again Section 6.2.3), the root-initial consonant would always syllabified as the coda of the syllable headed by the reduplicant vowel: i.e. [.t̲i̲-t.va.r-i.ṣa-], [.t̲i̲-s.ta.n-i.ṣa-], etc. (" . " represents syllable boundaries). The consistency of heterosyllabic parsing is even more important with respect to the two distinct treatments of VCC roots — infixation after the first root consonant in the general case (as also for VC roots), but infixation before the first root consonant (i.e. after the root vowel) in Vkṣ roots.

In the general treatment, the CV infix surfaces before a pre-*vocalic* consonant: e.g., *ubɟi̲ɟiṣa-*. Therefore, in such a form, the reduplicant is perfectly aligned to (i.e. coextensive with) a syllable: [ub.ɟi.ɟi.ṣa-]. On the other hand, in the treatment for Vkṣ roots, the reduplicant surfaces before a pre-*consonantal* consonant: e.g., *āc̲i̲kṣiṣa-*. If medial two-consonant clusters are indeed always heterosyllabic, then such a form would have to have the syllabification [ā.c̲i̲k.ṣi.ṣa-], where the reduplicant's right edge is not aligned to a syllable boundary, i.e. *k* intervenes between the rightmost reduplicant segment *i* and that syllable's right boundary.

If we were to try to employ a constraint requiring that the right edge of the reduplicant be aligned to the right edge of a syllable — ALIGN(RED, R; σ, R), defined in (109a) — to trigger infixation, this could derive the general pattern, as shown in (110.i), as long as this constraint dominated the constraint which requires the reduplicant to surface at the left edge of the word — formerly ALIGN-RED-L, now (for consistency and explicitness) ALIGN(RED, L; WD, L), defined in (109b). However, this makes exactly the wrong prediction for the Vkṣ roots, as demonstrated in (110.ii).

(109)   a.   **ALIGN(RED, R; σ, R)**
            Assign one violation mark * if the rightmost segment of the reduplicant is not the rightmost segment of a syllable.
        b.   **ALIGN(RED, L; WD, L)**   [= ALIGN-RED-L]
            Assign one violation mark * if the leftmost segment of the reduplicant is not the leftmost segment of a word.

(110)   VCC vs. Vkṣ roots in the desiderative

i. /RED, ubɟ, -iṣa-/		ALIGN(RED, R; σ, R)	ALIGN(RED, L; WD, L)
a.	u̲.b̲-ub.ɟ-i.ṣa-	*!	
b.	u.-b̲i̲-b.ɟ-i.ṣa-	*!	*
c. ☞	ub.-ɟi̲-ɟ-i.ṣa-		*
ii. /RED, ākṣ, -iṣa-/		ALIGN(RED, R; σ, R)	ALIGN(RED, L; WD, L)
a.	ā̲.c̲-āk.ṣ-i.ṣa-	*!	
b. ☺	ā.-c̲i̲-k.ṣ-i.ṣa-	*!	*
c. 💣	āk.-s̲i̲-ṣ-i.ṣa-		*

The ranking ALIGN(RED, R; σ, R) ≫ ALIGN(RED, L; WD, L) will treat the two root shapes in exactly the same way (again assuming that their clusters have the same syllabification properties), and thus incorrectly selects candidate (110.iic) **āk.-s̲i̲.-ṣ-i.ṣa-*. But not only does this particular ranking not select the right output, the desired candidate (110.iib) is doubly harmonically bounded.

If we assume categorical alignment constraints as in the definitions in (109), then the (b) candidates have one violation for each constraint, while the other two candidates have only a single violation for one of the two constraints. If we used gradient definitions, the (a) candidates would get an extra violation of ALIGN(RED, R; σ, R) (since there are two segments between its reduplicant *c* and the next syllable edge to its right), and the (c) candidates would get an extra violation of ALIGN(RED, L; WD, L) (since there are two segments between its reduplicant ṣ and the left edge to the word). However, since it does not add any violations to the otherwise empty cells in the (a) and (c) candidates, this does not change the harmonic bounding relationship. Therefore, under the assumption that medial *-kṣ-* was heterosyllabic in Classical Sanskrit, the alignment-based templatic analysis not only requires a ranking that does not generate the behavior of the Vkṣ roots, it requires constraints that leave its desired candidate harmonically bounded on both ends.

Success of the Syllable-Alignment Analysis with the Tautosyllabic Parse

There are obviously two possible responses to this result: (i) admit the failure of the templatic syllable-alignment analysis; or (ii) deny the underlying assumption that causes the analytical failure, namely that medial *-kṣ-* was parsed heterosyllabically at the relevant period. If (110.iib) were parsed as [ā.ci.kṣi.ṣa-] (with tautosyllabic *kṣ*) rather than [ā.cik.ṣi.ṣa-] (with heterosyllabic *kṣ*), then (110.iib) would no longer violate ALIGN(RED, R; σ, R). Candidate (110.iia), now parsed [ā.cā.kṣi.ṣa-], still would violate this constraint, and thus be eliminated. If we adopted the gradient formulation of ALIGN(RED, L; WD, L), then candidate (110.iic), now parsed [ā.kṣi.ṣi.ṣa-], would incur two violations compared to candidate (110.iib)'s, and thus be eliminated. (It would also violate a constraint requiring that the reduplicant be left-aligned with a syllable, though this constraint is not needed for any other reason.) Therefore, if we were to follow Vaux's (1992) claim that Classical Sanskrit parsed medial *-kṣ-* as a complex onset, a syllable-alignment account would lead to a convergent analysis.

The tableaux in (111) below summarize the syllable-alignment analysis under the assumption of a tautosyllabic parse for medial medial *-kṣ-*. ALIGN(RED, R; σ, R) has to rank below *COMPLEX in order to prevent infixation for STVX– roots roots (111.iii), whose reduplication patterns violate ALIGN(RED, R; σ, R). Given that *COMPLEX must be ranked above the alignment constraints, we need additional high-ranked syllabification constraints to generate the assumed syllabification — where all rising sonority clusters (crucially including *stop-fricative*, namely *kṣ*) form complex onsets while flat and falling sonority clusters are parsed as coda + simplex onset.[61] Using the SONORITY SEQUENCING PRINCIPLE constraint (SSP; cf. Clements 1990) to prevent complex onsets for flat and falling sonority clusters, the constraint ranking that would generate this distribution is: SSP ≫ NOCODA ≫ *COMPLEX. (These additional constraints are not shown in the tableaux.)

*PCR, ANCHOR-L-BR, and CONTIGUITY-BR are still relevant in the determination of which cluster to copy for cluster-initial roots. (I leave determining the ranking arguments as an exercise for the reader.) However, under this analysis, *PCR is *not* needed in order to determine the site of infixation in vowel-initial roots; alignment can handle it on its own. Nevertheless, the infixation patterns generated by alignment are ultimately in compliance with *PCR. Given that *PCR is crucial elsewhere in the system (i.e. the cluster-initial roots), and it is capable of determining infixation site in the vowel-initial roots (when coupled with DEP[CVtransitions]/V-BD, as shown above), one could argue that the syllable-alignment analysis is missing an important generalization. However, given the novelty of DEP[CVtransitions]/V-BD, this is not a strong argument against the syllable-alignment approach.

---

[61] Vaux (1992:298) implies that, for at least some varieties of Classical Sanskrit, *s-stop* clusters also form medial complex onsets. If so, the definition of the SSP constraint at work in the language would have to be adjusted accordingly.

(111)　Syllable-alignment analysis of the desiderative (with tautosyllabic -*kṣ*-)

i.	/RED, ubɟ, -iṣa-/	*COMPLEX	ALIGN (RED, R; σ, R)	ALIGN (RED, L; WD, L)
a.	u̲.b̲-ub.ɟ-i.ṣa-		*!	
b.	u.-b̲i̲-b.ɟ-i.ṣa-		*!	*
c. ☞	ub.-ɟ̲i̲.-ɟ-i.ṣa-			**

ii.	/RED, ākṣ, -iṣa-/	*COMPLEX	ALIGN (RED, R; σ, R)	ALIGN (RED, L; WD, L)
a.	ā̲.c̲-ā.kṣ-i.ṣa-	*	*!	
b. ☞	ā.-c̲i̲-.kṣ-i.ṣa-	*		*
c.	ā.k-s̲i̲.-ṣ-i.ṣa-	*		**!

iii.	/RED, stan, -iṣa-/	*COMPLEX	ALIGN (RED, R; σ, R)	ALIGN (RED, L; WD, L)
a. ☞	t̲i̲-s.ta.n-i.ṣa-		*	
b.	s̲t̲i̲-s.ta.n-i.ṣa-	*!	*	*
c.	sta.-n̲i̲.-n-i.ṣa-	*!		***

iv.	/RED, tvar, -iṣa-/	*COMPLEX	ALIGN (RED, R; σ, R)	ALIGN (RED, L; WD, L)
a. ☞	t̲i̲.-tva.r-i.ṣa-	*		
b.	t̲v̲i̲.-tva.r-i.ṣa-	**!		
c.	tva.-r̲i̲.-r-i.ṣa-	*		*!**

In sum, if it can be maintained that there was a difference in syllabification between medial *kṣ* and other medial clusters (specifically, falling sonority clusters) — i.e. a tautosyllabic parse for the former, heterosyllabic parses for the latter — at the stage of Sanskrit where the infixal desideratives to vowel-initial roots first develop (presumably Classical Sanskrit), then the templatic syllable-alignment analysis stands as a viable alternative to the *PCR + DEP[CVtransitions]/V-BD analysis proposed in Section 6.6.2.2. I now present evidence from the reduplicated aorist that argues against a tautosyllabic parse in Classical Sanskrit.[62]

---

[62] If one were to deny the tautosyllabic analysis of *kṣ*, another way to rescue the templatic analysis could be to appeal to a very different notion of template, one which does not employ Generalized Alignment (McCarthy & Prince 1993a) in its standard form. The constraint could be stated in one of two ways: (i) the reduplicant may not contain segments belonging to different syllables; or, roughly equivalently, (ii) a syllable boundary may not be contained within the reduplicant.

I know of no precedent for either type of formulation. Furthermore and more importantly, such a constraint would be incompatible with the facts of the intensive, where syllable boundaries can indeed be contained within the reduplicant: for example, √*krand* 'cry out' → intensive *ka̲.n̲(-)i-k.rand-* (Whitney 1885 [1988]:24; see Steriade 1988 on the analysis of the intensive). Since I have not constructed an analysis for the intensive, it is potentially possible that such a constraint could be ranked low enough to not interfere with this mapping, but it means that it is at least not a surface true condition across all reduplicative categories in the language. With no external support for a constraint of this type, and potential language-internal counter-evidence, this approach does not seem appealing.

Evidence for the Heterosyllabic Parse from the Reduplicated Aorist

The reduplicated aorist (Whitney 1889:308–312/§856–868) allows us to determine the syllabification of a following cluster (Cooper 2014:41–42). Much like the perfect and the desiderative (and also the reduplicated present), the reduplicated aorist has a CV reduplicant, normally with fixed [i] vowel quality. Unlike the other reduplicated categories, the aorist places a condition on the weight of its reduplicated syllable: it must always be heavy.[63] This gives us a direct window into syllabification. (The same condition holds of the reduplicated intensive; Whitney 1889:362–365/§1000–1004, Steriade 1997. However, this category is moribund by the later language, so it does not provide sufficient evidence for the current question.)

When the root begins in a single consonant, the reduplicant vowel is long: for example, $\sqrt{car}$ 'move' → aorist <u>cī</u>-car-, $\sqrt{vāç}$ 'bellow' → aorist <u>vī</u>-vaç- (notable for the shortening of the root vowel), and many others. For *s*-stop roots (which show C_2-copying in this category as well), the reduplicant vowel surfaces as short: for example, $\sqrt{sparç}$ 'touch' → aorist <u>pi</u>-spṛç-.[64] Under a syllabic view of weight, this clearly means that *s*-stop clusters had a heterosyllabic parse. In Vedic, all cluster-initial roots with reduplicated aorists have *short* reduplicant vowels; for example, *cikrada*, *titrasa*, *piplava*, *siṣvada*, *çiçnaya*, *vivyat^ha*, and, most notably for the present discussion, *cikṣipa*. This is strong evidence that *all* medial two-consonant clusters had a heterosyllabic parse in Vedic.

This vowel length alternation is maintained throughout the later language. Therefore, we can ascertain the syllabification facts of later Sanskrit by applying this same test to the reduplicated aorists attested at that stage. Whitney (1885 [1988]:224–225) compiles a list of all the reduplicated aorist stems attested in Sanskrit. He divides these into three categories (see Whitney 1885 [1988]:211): those which appear only in the earlier language, those which appear only in the later language (i.e. Epic Sanskrit and/or Classical Sanskrit), and those found throughout the history of the language. Since we are concerned with the syllabification only of the later language, it is the last two of these categories which will be of interest to us. All of the cluster-initial roots from among these are reproduced in (112).[65] The single cluster-initial reduplicated intensive that has a linking -*i*-, which is subject to the same length alternation as the reduplicant vowel of the aorist, that is attested in the later language displays a short vowel as well: <u>dar**i**</u>-drā- (Whitney 1885 [1988]:233).

(112)   Reduplicated aorists of cluster-initial roots attested in Later Sanskrit

    a.   Earlier and later Language
         i.   *dudruva*
         ii. *çiçriya*
    b.   Later language only
         i.   *jigraha*
         ii. *titvara*
         iii. *bib^hrama*

Since these cluster-initial roots show short reduplicant vowels in later Sanskrit, it must be the case that these clusters were heterosyllabified at this period. This includes stop-sonorant, fricative-sonorant, and stop-*v* (which unquestionably was a fricative by the later language). While there are no *kṣ*-initial roots to directly test its syllabification, it seems almost certain that if all these clusters

---

[63] Whitney (1889:310/§860) states that it remains light when the root syllable is heavy.

[64] Examples of reduplicated aorists for *s*-stop roots are only attested in the earlier language; see below.

[65] The Classical Sanskrit aorist stems *cac^hada* and *cic^hada* show short reduplicant vowels. This likely indicates that aspirated affricates, at least by the time of Classical Sanskrit, functioned as clusters; see Whitney (1889:16/§42).

were heterosyllabic, then *-kṣ-* was too. If this is correct, then the syllable-alignment analysis of the infixing desiderative cannot be sustained.[66]

6.6.2.4 Local Summary

A templatic syllable-alignment approach to the infixing reduplicated desiderative in Sanskrit is contingent upon a particular interpretation of the syllabification facts of the language. While I have not yet been able to assess the metrical evidence, which Vaux (1992) (citing Varma 1929) claims supports tautosyllabic parsing of *-kṣ-*, the evidence from the length alternation of the reduplicated aorist points very strongly to heterosyllabic parsing of *-kṣ-*. As a tautosyllabic parse is a prerequisite for the syllable-alignment analysis, this conclusion would argue strongly against such an analysis.

On the other hand, an approach based on cue-based faithfulness produces a straightforward and consistent analysis which is not dependent on syllabification facts.[67] Furthermore, it allows the alternation in the position of the reduplicative infix in vowel-initial roots to be viewed as yet another *PCR effect, which is evident elsewhere in the language's reduplicative system (including the treatment of cluster-initial roots in the desiderative itself). The syllable-alignment analysis does not employ *PCR as an active factor in this process, and thus could be said to be missing an important generalization.

6.6.3 Analysis of the Klamath Distributive

As demonstrated in Section 6.3.3 above, Klamath (Barker 1964) exhibits the same sort of cluster-dependent copying behavior, driven by *PCR, as do the Indo-European languages under discussion in this dissertation. Parallel to Gothic (and Proto-Anatolian), Klamath demonstrates an alternation between a CV C_1-copying pattern and a CCV cluster-copying pattern, dependent on the type of root-initial cluster (see Steriade 1988, Fleischhacker 2005). In this section, I flesh out the formal analysis of the Klamath reduplication pattern. For reference, in (113), I reproduce Klamath's consonant inventory (repeated from (29) above).

(113)   Klamath consonant inventory

Obstruents

Plain:	T	p < b >	t < d >	ʧ < j >	k < g >	q < g̣ >	s
Aspirated:	Tʰ	pʰ < p >	tʰ < t >	ʧʰ < c >	kʰ < k >	qʰ < q >	h
Glottalized (ejective):	T'	p' < p̓ >	t' < t̓ >	ʧ' < c̓ >	k' < k̓ >	q' < q̓ >	ʔ

Sonorants

Plain:	R	m < m >	n < n >	l < l >	w < w >	j < y >
Aspirated (voiceless):	Rʰ	m̥ < M >	n̥ < N >	l̥ < L >	w̥ < W >	j̥ < Y >
Glottalized (creaky):	R'	m̰ < m̓ >	n̰ < n̓ >	l̰ < l̓ >	w̰ < w̓ >	j̰ < y̓ >

Klamath marks the DISTRIBUTIVE with prefixal partial reduplication. This category attests two types of reduplicants: (i) a CV reduplicant (the C_1-copying pattern, when applied to cluster-

---

[66] Vaux (1992:298), citing Benveniste (1937), alludes to the idea that *kṣ* (and likewise *ɟɲ*) had actually become a single segment by this period (though he provides no explicit evidence). If this were correct, then we would expect Vkṣ roots to pattern with VC roots in their infixation pattern in the desiderative, as they do.

[67] Insofar as the difference between "tautosyllabic" and "heterosyllabic" parses could be said to correlate with the presence vs. absence of (robust) CX transitions (as controlled by a language's phonetic grammar), then perhaps some of these questions might carry over to the cue-based analysis to a greater extent than meets the eye.

initial roots) and (ii) a CCV reduplicant (the cluster-copying pattern). Roots in initial /CV.../ always take a CV reduplicant, even if they undergo syncope (see footnote 13) to create a stem-initial cluster on the surface (i.e. $\sqrt{C_1V_2C_3...} \rightarrow$ distributive $[\underline{C_1V_2}\text{-}C_1C_3...]$).[68] With the exception of stop-sonorant clusters, all root-initial clusters exclusively display cluster copying.[69] Examples of each of these are provided in (114). The one type of cluster where the CV outcome is possible is stop-sonorant (TR). While TR roots may surface with either type, it is usually categorical by root (though Barker reports two cases of free variation). These examples are shown in (115). For ease of comprehension, in both of the following tables, relevant reduplicated forms are accompanied by a derivationally intermediate form which undoes late phonological rules (e.g. vowel syncope/reduction, cluster simplification, etc.) that do not directly impact the determination of reduplicant shape. All data is from Barker (1964) (abbreviated in tables as B).

(114)   CCV reduplication with non-TR clusters (repeated from (30) above)

Cluster Type		Root	Reduplicated	(morphophonemic)	Source
ST		/stiːq̓a/	<u>sti</u>-stiːq̓a		B:84
		/scidiːla/	<u>sci</u>-scdiːla	(//<u>sci</u>-scidiːla//)	B:84
TT		/ktiwcṅa/	<u>kti</u>-ktoːcṅa	(//<u>kti</u>-ktiwcṅa//)	B:88
TS		/ksodga/	<u>kso</u>-ksatga	(//<u>kso</u>-ksodga//)	B:84
RR	(LW)	/lwosga/	<u>lwo</u>-lwasga	(//<u>lwo</u>-lwosga//)	B:89
RT	(NT)	/mbod̓ẏdk/	<u>mbo</u>-mpditk	(//<u>mbo</u>-mbod̓ẏdk//)	B:90
	(LT)	/lbogaa/	<u>lbo</u>-lpga	(//<u>lbo</u>-lbogaa//)	B:84
	(WT)	/w-qeːwi-'a/	<u>wqe</u>-wqeːẁa		B:121
SR	(SN)	/sṅog-y-s/	<u>sṅo</u>-sṅog̣is	(//<u>sṅo</u>-sṅogys//)	B:89
	(SL)	/sliṅ̓/	<u>sli</u>-slan	(//<u>sli</u>-sliṅ̓//)	B:92
	(SW)	/swinys/	<u>swi</u>-soːnis	(//<u>swi</u>-swinys//)	B:88
Cʔ	(Sʔ)	/sʔaba/	<u>sʔa</u>-sba	(//<u>sʔa</u>-sʔaba//)	B:85
	(Lʔ)	/lʔeg-bg-m/	<u>lʔe</u>-lʔakpgam	(//<u>lʔe</u>-lʔegbgm//)	B:171

---

[68] This poses a problem for a fully parallel *PCR-based approach, as it would seem that *PCR would have to be evaluated at an intermediate level of representation. One could hope to maintain a fully parallel analysis by claiming that the syncope process does not actually result in complete deletion (i.e. it is extreme vowel reduction, not deletion), and/or that the acoustic/articulatory properties of consonants that precede a deletion site are different than those of consonants that are underlyingly pre-consonantal (i.e. underlyingly pre-vocalic consonants are produced with at least some of the robust cues normally restricted to surface pre-vocalic position). I know of no evidence which bears on this question, so I leave it as a question for future consideration.

[69] I have found two exceptions in Barker (1964):

(i)   a.   /RED-k̓coːsYeːni-'a/ → <u>ko</u>-k̓coːsYeːṅa (not *<u>k̓co</u>-k̓coːs...) (p.84)

       b.   /RED-qtana/ → <u>qa</u>-qta (not *<u>qta</u>-qta) (p.85)   cf. /RED-qday-'aːk̓/ → <u>qda</u>-qdiːʔaːk (p.89)

(115) Reduplication in TR clusters (repeated from (31) above)

Cluster	Red. Type	Root	Reduplicated	Source
TN	CV	/pni-'aːǩ/	*pi-pnaʔaːk*	B:90
	var.	/qniy-'a/	*qi-qnẏa ~ qni-qnẏa*	B:85
TL	CV	/c̓lodga/	*c̓e-c̓latga*	B:82
	CCV	/c̓lidga/	*c̓li-c̓latg̣a*	B:82
	var.	/q̓liq-diːla/ ~ /t̓liq-diːla/	*q̓li-q̓laqdiːla ~ t̓i-t̓laqdiːla*	B:117
TW	CV	/t̓waːYa/	*t̓a-t̓waːYa*	B:85
	CCV	/tya-'aːǩ/	*tya-tiːʔaːk (//tya-tyaʔaːǩ//)*	B:89

Given that the outline of this system is identical to Gothic, we can simply important the basics of the Gothic analysis. Putting aside for the moment the variation seen in the TR roots, we can describe the system as one where the desired reduplicant shape is CV (shown in (116)), but this preference is blocked when this would create a poorly-cued repetition, via the ranking *PCR ≫ *CC, as demonstrated in (117).

(116) CV C_1-copying reduplication in TR roots

/RED, pni-'aːǩ/		ANCHOR-L-BR	*PCR	*CC
a. ☞	pi-pnaʔaːk			*
b.	pni-pnaʔaːk			**!
c.	ni-pnaʔaːk	*!		*

(117) CCV cluster reduplication in non-TR roots

/RED, stiːq̓a/		ANCHOR-L-BR	*PCR	*CC
a.	si-stiːq̓a		*!	*
b. ☞	sti-stiːq̓a			**
c.	ti-stiːq̓a	*!		*

If *PCR is the motivation for cluster-copying in ST roots, the question then becomes how to account for the cases of cluster-copying in TR roots, as these should not be subject to a *PCR violation. As we saw for Hittite in Chapter 3, the way to impose cluster-copying reduplication in general is to use a CONTIGUITY constraint (Kenstowicz 1994, McCarthy & Prince 1995). In Hittite, the CONTIGUITY constraint could be defined over the Base-Reduplicant correspondence relation, because the roots vowels (generally) surfaced in the output base. This cannot be the case for Klamath, as the vowels which crucially induce the violations are frequently deleted in the output base. This would vacuously satisfy the constraint under conditions where we need to leverage its violation. Therefore, the CONTIGUITY constraint for Klamath must be defined over the Input-Reduplicant correspondence relation (see McCarthy & Prince 1995, 1999; see also the analysis of Sanskrit in Chapter 5), as in (118).

(118)   **CONTIGUITY-IR**
   Assign one violation * for each pair of adjacent segments in the reduplicant who have correspondents in the input root which are not adjacent.

This penalizes the C_1-copying candidate. If it were ranked above *CC, it would select cluster-copying. However, this constraint cannot in general outrank *CC, or else we would never observe C_1-copying. Therefore, this CONTIGUITY constraint must be lexically-indexed (see Kraska-Szlenk 1997, 1999, Fukazawa 1999, Itô & Mester 1999, 2001, Pater 2000, 2009; see also Chapter 2 for a similar application in Greek) to the TR roots that take the cluster-copying pattern, as illustrated with the tableau in (119). The non-indexed version of CONTIGUITY then ranks below *CC. As long as all TR roots that do not show the cluster-copying pattern lack the index, this ranking allows *CC to determine the evaluation for these roots in favor of C_1-copying. This is demonstrated in (120).

(119)   CCV cluster reduplication to lexically-indexed TR roots

/RED, c̓lidga̠*lex*/	ANCHOR-L-BR	*PCR	CONTIG-IR*lex*	*CC	CONTIG-IR
a. c̓i-c̓latga̠			*!	*	*
b. ☞ c̓li-c̓latga̠				**	
c. li-c̓latga̠	*!			*	

(120)   CV C_1-copying reduplication in non-lexically-indexed TR roots

/RED, pni-ʾaːk̓/	ANCHOR-L-BR	*PCR	CONTIG-IR*lex*	*CC	CONTIG-IR
a. ☞ pi-pnaʔaːk				*	*
b. pni-pnaʔaːk				**!	
c. ni-pnaʔaːk	*!			*	

This approach yields a tidy explanation for one of the free variation cases: /RED-q̓liq-diːla/ → *q̓li-q̓laqdiːla* ∼ /RED-t̓liq-diːla/ → *t̓i-t̓laqdiːla* 'peek underneath-DIST'. For this word, there are two variants of the root, one with initial /q̓/ and one with initial /t̓/. If the q̓-root variant is lexically-indexed (i.e. /q̓liq*lex*/) but the t̓-root variant is not (i.e. /t̓liq/), then we generate the distribution with no further machinery. The other case of free variation is harder to explain, as there do not seem to be distinct root variants: /RED-qniy-ʾa/ → *qi-qnˊya* ∼ *qni-qnˊya* 'have an erection-DIST'. If these forms are in free variation among the population but categorical for individual speakers, then we could say that some speakers have indexed the root while others have not. If it is truly in free variation for individual speakers, then we would have to resort to some sort of variable ranking. As this is well beyond the main focus of this discussion, I leave this as a question for future investigation.

6.7   Empirical Evidence for *PCR outside of Reduplication

Up until this point, all of the empirical evidence presented in support of *PCR has come from reduplication. If *PCR is indeed a constraint in the phonological grammar, then we should expect it to exhibit effects outside of reduplication. In this section, I adduce and analyze two such examples: a case of suffix allomorphy in Latin (Section 6.7.1) and an exceptional behavior of aspiration in

Sanskrit (Section 6.7.2).[70] Additionally, I lay out some of the basic facts of the *sCVC* constraint in English (Fudge 1969; Section 6.7.3), and suggest that this too might find an explanation in the *PCR constraint or something similar .

These cases are important evidence in favor of the *PCR approach to the problems addressed in this dissertation. Not only does external evidence of differing sorts in general bolster the proposal, but the fact that non-reduplicative patterns are explained by the same constraint that explains facts about reduplication sets this analysis apart from alternatives that rely on reduplication-specific machinery to derive the reduplication facts.

6.7.1 Allomorphy in Latin Suffixes in *–is...*

Cser (2015:13) documents unexpected allomorphic alternations involving the perfect suffixes that begin in /-is.../ for stems ending in *s*.[71] When the suffixes listed in (121) attach to consonant-final perfect stems (and also *u*-final perfect stems), they surface faithfully: for example, √*nov-* 'know' → *nōv-isse*, √*ta(n)g-* 'touch' → *tetig-isse*, √*mon-* 'warn' → *monu-isse*, etc. When attached to vowel-final stems, these suffixes generally lose their initial vowel: for example, √*amā-* 'love' → *amā-sse*, √*comple(v)-* 'complete' → *complē-sse*, √*abi-* 'leave' → *abī-sse*, etc. This is a general property of most vowel-initial suffixes, and can be explained as a simple hiatus avoidance effect, where root faithfulness trumps suffix faithfulness (see Cser 2015).

(121)   Perfect endings in *–is...* (Allen & Greenough 1903 [2006]:79/§166)

  a.   *-istī*   (2SG perfect active indicative)
  b.   *-istis*   (2PL perfect active indicative)
  c.   *-isse*   (perfect infinitive)
  d.   *-isse-*   (pluperfect subjunctive)

However, this distribution is partially disrupted for perfect stems ending in *s*. When the perfect stem ends in *s*, the full variant is generally permitted; however, there are also a number of forms which attest the vowel-less variant as well (or instead).[72] Leumann (1977:598/§438) illustrates this with the attested perfect forms of the root √*dīk* 'say' (present *<dīcō>*). This root has a perfect stem with final /s/: *dīks-* (e.g. 1SG pluperfect *<dīx-eram>*). (Orthographic *<x>* represents the sequence [ks].) As shown in (122), the /-is.../ suffixes may surface without their initial vowel /i/, which triggers subsequent deletion (and/or degemination) of one of the underlying /s/'s. The same holds for all the examples in (123), cited also by Leumann (1977:598/§438).

(122)   Perfects of *dīcō* with reduced suffix variants (forms from Leumann 1977:598/§438)

	Reduced suffix		Full suffix		
a.	*dīxtī* [dīkstī]	∼	*dīxistī*	[dīksistī]	(2SG perfect active indicative)
b.	*dīxem* [dīksem]	∼	*dīxissem*	[dīksissem]	(1SG pluperfect subjunctive)
c.	*dīxe* [dīkse]	∼	*dīxisse*	[dīksisse]	(perfect infinitive)

---

[70] The Sanskrit case does involve reduplication, but it does not seem as though reduplication is central to the behavior being addressed.

[71] I am indebted to Donca Steriade for bringing this example to my attention.

[72] In the forms provided below, I am not certain whether each of the corresponding forms with the full variants are actually attested. The forms with the reduced variants are drawn directly from Leumann (1977:598/§438), where they are accompanied with the source of their attestation.

(123)   Additional examples (Leumann 1977)

a. *dētraxe*   (∼ *dētrax-isse*)
b. *prōmīstī*   (∼ *prōmīs-istī*)
c. *admīsse*   (∼ *admīs-isse*)
d. *scrīpstis*   (∼ *scrīps-istis*)
e. *cōnsūmpse*   (∼ *cōnsūmps-isse*)
f. *lūxtī*   (∼ *lūx-istī*)
g. *dērēxtī*   (∼ *dērēx-istī*)
h. *traxe*   (∼ *trax-isse*)

i. *exstīnxem*   (∼ *exstīnx-issem*)
j. *accestis*   (∼ *access-istis*)
k. *surrēxe*   (∼ *surrēx-isse*)
l. *ērēpsēmus*   (∼ *ērēps-issē-mus*)
m. *percustī*   (∼ *percuss-istī*)
n. *ēvāstī*   (∼ *ēvās-istī*)
o. *dīvīsse*   (∼ *dīvīs-isse*)

The use of the reduced variant should be viewed as a *PCR effect. When these suffixes are attached to an *s*-final stem, maintenance of the full form would violate *PCR. This is clearest in the 2SG and 2PL suffixes /istī/ and /istis/ (on the suffixes in *iss...*, see below): concatenation of the full form creates a repetition of *s* in pre-*t* position, as can be seen in the full form *dīxistī* [dīk**sistī**]. This is a *PCR violation, of exactly the sort which is avoided in reduplication through infixation (see Section 6.6.1). The reduced forms of the suffix will avoid this *PCR violation, because they eliminate the suffix's /s/, and thus never create the problematic repetition.

Cser (2015) advocates for a phonologically conditioned allomorphy approach, where each of these affixes has stored allomorphs with and without the vowel. *PCR would then (directly or indirectly) variably select the vowel-less variant for *s*-final stems. This is probably an unnecessary complication, as it is straightforward to derive the variation directly in the phonology.

*PCR can motivate phonological deletion of the entire suffix-initial /is/ sequence from a single underlying representation that contains the full version of the suffix.[73] A variable ranking between *PCR (defined in shorthand in (124a)) and MAX-AFFIX (defined in (124b)) can generate the variable realization of these suffixes following *s*-final stems. When MAX-AFFIX outranks *PCR (125.i), the suffix will be realized faithfully. When *PCR outranks MAX-AFFIX (125.ii), there is deletion. Two other requirements make deletion of just a single segment suboptimal. First, geminates are only allowed in inter-sonorant (cf. Weiss 2009:157): this is encoded with *GEM//OBS (defined in (124c)). Second, there is a requirement for contiguity between affix segments to be maintained: this is encoded with CONTIGUITY-AFFIX (defined in (124d)).

(124)   a. ***PCR [*for Latin*]**
         Assign one violation mark * for each repeated *s* before an obstruent (*SVST).
         b. **MAX-AFFIX**
         Assign one violation mark * for each affix segment in the input that lacks a correspondent in the output.
         c. ***GEM//OBS**
         Assign one violation mark * for each geminate in the output that is adjacent to an obstruent.
         d. **CONTIGUITY-AFFIX**
         For any pair of affix segments x_i, y_i in the input with correspondents in the output x_o, y_o, assign one violation mark * if x_o and y_o are adjacent but x_i and y_i are not, and vice versa.

---

[73] The *-is-* sequence likely originated as a separate morph (Weiss 2009:392), which somehow spread throughout the perfect inflectional system. (The *-er-* sequence in many of the other perfect endings is the diachronically expected outcome of *is / _V, due to rhotacism and pre-*r* lowering; see Weiss 2009:392, Cser 2015:14, fn. 14.) If this status as a separate morph were still somehow maintained in Latin, then perhaps a morphological analysis along the lines proposed by Cser (2015) would become more attractive. However, the allomorphy evident from vowel-final suffixes would suggest a morph boundary between the *-i-* and the *-s...*, which does not accord with the diachronically justified division.

(125) Variable ranking between *PCR and MAX-AFFIX generates variable realization

i. When MAX-AFFIX ≫ *PCR: faithful realization of suffix (*dīx-istī*)

/dīks, -istī/	*GEM//OBS	CONTIG-AFX	MAX-AFX	*PCR
a. ☞ dīks-istī				* (sist)
b. dīks-stī	*!		* (i)	
c. dīks-itī		*! (i↔t)	* (s)	
d. dīks-tī			*!* (is)	

ii. When *PCR ≫ MAX-AFFIX: deletion of suffix-initial /is/ (*dīx-tī*)

/dīks, -istī/	*GEM//OBS	CONTIG-AFX	*PCR	MAX-AFX
a. ☞ dīks-istī			*! (sist)	
b. dīks-stī	*!			* (i)
c. dīks-itī		*! (i↔t)		* (s)
d. dīks-tī				** (is)

The one minor complication here is the treatment of the suffixes in /-iss.../. These behave exactly like the suffixes in /-ist.../. The difference, though, is that, if <ss> represents a true geminate, which would just be a single long consonant, we should not expect it to meet the conditions for *PCR; that is to say, if the <ss> were a single long consonant [sː], the repeated consonants (the root-final [s] and the suffix-medial [sː]) would both be in pre-vocalic position, and would thus be expected to satisfy any version of *PCR.

Therefore, if we are to explain the behavior of the two types of *-is...*-initial suffixes in the same way (i.e. via *PCR), then we must deny that the <ss> string is treated as a single segment, at least for the purposes of *PCR's identity calculation. That is to say, this <ss> string must be a "fake geminate" (cf. Hayes 1986). This might actually be a reasonable assumption, because, as mentioned in footnote 73, the *-is-* sequence in these affixes was originally a distinct morph, meaning that the <ss> string arose historically from morpheme concatenation. As long as the <ss> string is treated phonologically as distinct segments (whether or not it was phonetically equivalent to a true geminate), *PCR will identify the first portion as a repeated segment, and successfully penalize candidates with the sequence [siss] resulting from concatenation of these affixes to an *s*-final stem.[74]

6.7.2 Aspiration in Sanskrit

The distribution of aspiration in Sanskrit is fairly complex and involves the interaction of several processes and alternations. There is at least one case, however, that seems to run counter to the normal rules governing this distribution. This exception can be explained by *PCR.[75]

---

[74] As discussed in Section 6.4, transitions appear to be the most important cue for the purposes of licensing repetitions. The repeated *s* in this position (and in *sVst*) lacks any sort of transitions. Since we don't have any obvious evidence for the behavior of repetition types other than those involving *s* before an obstruent, it is unclear what the precise repetition licensing conditions are for Latin. But the fact that it is these which show clear evidence of repetition avoidance comports well with the overall *PCR approach, since these are among the most poorly-cued repetition types.

[75] The relevance of this phenomenon to *PCR was first noticed by Donca Steriade. The outline of the analysis proposed below follows Steriade (2015).

In Sanskrit, voiced aspiration (probably murmur or breathy voice) is a contrastive feature for stops /bʰ, dʰ, ɖʰ, gʰ/ (= Dʰ). However, this is only licensed in pre-sonorant position; all laryngeal contrasts — voicing, aspiration, breathy voice (≈ voicing + aspiration) — are neutralized in pre-obstruent and word-final position (see Whitney 1889:53/§1553). When an underlying Dʰ would surface in a position where the laryngeal contrast is not licensed, the breathy voicing can migrate to a nearby stop, subject to certain restrictions. If there is no stop on which it can land, the breathy voice is lost. There are two places the breathy voice can migrate to.

If the *preceding* consonant (or a member of the preceding consonant cluster) is an *unaspirated voiced stop* (/b, d, ɖ, g/), breathy voice has the potential to migrate to that stop if laryngeal contrasts are not licensed in its underlying position. For example, the root √*budʰ* 'know'[76] makes a root noun. In the locative plural, it has the underlying representation /budʰ-su/, with aspiration on the second root consonant. However, since aspiration is not permitted to surface immediately preceding the *s*, this form surfaces instead as [bʰutsu], with aspiration on the first root consonant. This is often referred to as Aspiration Throw Back (ATB).

Similarly, if the immediately *following* consonant is an *unaspirated voiceless stop* (/p, t, ʈ, k/) or an *unaspirated voiced stop* (/b, d, ɖ, g/), breathy voice also has the potential to migrate to that stop if laryngeal contrasts are not licensed in its underlying position (and it indeed never will be licensed because of said following stop). That stop also becomes voiced, regardless of its original voicing specification. For example, the root √*rudʰ* 'obstruct' builds a nasal-infix present. In the 3rd singular, this formation has the underlying representation /ru-na-dʰ-ti/, with aspiration on the root-final consonant. Aspiration is not licensed in root-final position here because of the suffix-initial /t/, forcing the aspiration to migrate. It surfaces on the suffix-initial /t/: [runaddʰi]. This process is known as Bartholomae's Law (BL).

When both ATB and BL are in principle available, BL is preferred. This can be seen in the past participle of √*budʰ* 'know', which is underlyingly /budʰ-ta-/. It surfaces as [buddʰa-], with aspiration migrating rightwards via the application of BL, even though it could have migrated leftwards to surface as **[bʰutta-] by applying ATB.

However, there is at least one instance where both options are available but the usually dispreferred ATB option surfaces. This occurs in several cells of the paradigm for the reduplicated present of √*dʰā* 'place', as shown in (126).[77] In this table, grey cells indicate morpheme concatenations where aspiration is licensed in its underlying position, and we observe no application of either ATB or BL, as expected. White cells, on the other hand, are those in which aspiration would not be licensed in its underlying position. In all such cases, ATB is available, and it applies. The bolded cells — the active 3rd dual and the middle 3rd singular — however, in addition to having the conditions to support ATB also have the conditions to support BL. Even though BL is normally applied in such circumstances, these forms show ATB.

---

[76] I adopt the position that the Indo-European "diaspirate" roots are stored synchronically as /DVDʰ/ not /DʰVDʰ/ in the Sanskrit lexicon.

[77] I assume that there is no independent reason why aspiration would surface on the initial *d*. Reduplicants in virtually all other cases reject aspiration, the opposite behavior of what seems to be happening here. Otherwise, I abstract away from the fact that this is a reduplicated paradigm. While I cannot rule out that the presence of reduplication plays some role in the unusual distribution of aspiration, there is no obvious way in which simply appealing to the presence of reduplication would explain these facts.

(126) Reduplicated present of $\sqrt{d^{ɦ}\bar{a}}$ 'place' (forms are RED-ROOT-ENDING)

	Active			Middle		
	Sing	Dual	Plural	Sing	Dual	Plural
1st	dá-dʰā-mi	da-dʰ-vás	da-dʰ-más	da-dʰ-é	dá-dʰ-vahe	dá-dʰ-mahe
2nd	dá-dʰā-si	dʰa-t-tʰás	dʰa-t-tʰá	dʰa-t-sé	da-dʰ-áte	dʰá-d-dʰve
3rd	dá-dʰā-ti	**dʰa-t-tás**	dá-dʰ-ati	**dʰa-t-té**	da-dʰ-áte	dá-dʰ-ate

I argue that this is a *PCR effect. This can be seen when we consider what the BL form would have been: MID.3SG //da-dʰ-té// → **[**da-d**-dʰé] (BL). BL would create a sequence of two identical stops in a position where the second of the repeated stops is not cued by an intensity rise, the minimal licensing condition for repetitions in Sanskrit (see Section 6.5.1). This would be a violation of the Sanskrit *PCR constraint. It is thus avoidance of the *PCR violation that results in applying ATB instead of BL, contrary to expectation.

6.7.3 The *sCVC Constraint in English

It has long been noted (Fudge 1969, Clements & Keyser 1983, Davis 1984, 1989, 1991, Coetzee 2005, among others) that there are context-sensitive restrictions on (near-)identical consonants in English. Specifically, there is a massive asymmetry in the English lexicon between, on the one hand, $C_{\alpha}VC_{\alpha}$ words — which are extremely frequent, and, on the other hand, $sC_{\alpha}VC_{\alpha}$ words — which are virtually non-existent (with the exception of coronal stops). Compare the abundance of words like *kick* [kɪk], *cake* [keɪk], *pup* [pʌp], *pipe* [paɪp], etc., with the complete non-existence (and indeed ungrammaticality) of words like *[skɪk], *[skeɪk], *[spʌp], *[spaɪp], etc. The ungrammatical post-*s* repetitions can also be compared with post-*s* non-repetitions. That is, beside ungrammatical *[spaɪp] and *[skaɪk], there is grammatical *spike* [spaɪk] and *Skype* [skaɪp].[78] It thus seems that the presence of a preceding *s* converts a perfectly licit repetition into one that is completely disallowed, and that repetition is banned in a position where non-identical consonants of the same sort are allowed.

Furthermore, Davis (1991) shows that this effect holds medially as well, regardless of the additional surrounding context. Since the following context in particular could lead to different syllabifications — that is, the second member of the repetition could be a coda (as it is in word-final position) when followed by certain consonants, but it could also be an onset when followed by a vowel — it is not possible to define the constraint in syllabic terms (Davis 1991); this has left the problem without an adequate solution.

This distribution has very similar contours to the *PCR effects discussed throughout this chapter: some contrast is normally licensed, but the contrast is suddenly disallowed when it is part of a repetition. And furthermore that this distribution is crucially determined by what sorts of segments are or are not adjacent to the repetition. There are, though, several ways in which this distribution is significantly different than the other ones considered above; namely, the changing context is before rather than after the repetition, and the way that that context affects the repeated consonant would seem only to be with respect to laryngeal cues — which were explicitly unimportant in the reduplicative *PCR effects in Ancient Greek and Sanskrit.

At present, I do not have an analysis of the English *sCVC constraint either. This is simply to say that there are distinct parallels between it and the other *PCR effects, and that further consideration of the two types of cases in tandem is likely to lead to a better understanding of both.

---

[78] Coppe van Urk (personal communication) informs me that the same constellation of facts holds in Dutch as well.

6.7.4 Interim Conclusions

In this section, I have shown that *PCR effects hold outside of reduplication as well. These appeared in the form of conditions on suffix allomorphy in Latin and aspiration-related repairs in Sanskrit. There is also a lexical constraint of English against repetitions of particular sorts in particular contexts which seems like it could be a related phenomenon. This provides evidence that *PCR effects are not specific to reduplication, but rather a more general phonological problem.

One other place to look for *PCR effects is in gradient statistical tendencies in the lexicon. Per Zuraw (2000) and many others, phonotactic constraints that lead to categorical phonological patterns in some languages should show gradient effects in other languages. And indeed if the categorical effects it exerts in a language are restricted in their scope, in the way that the *PCR effects documented for the ancient Indo-European languages are limited to reduplication and other minor patterns arising under affixation, it is extremely likely that that same language will show statistical evidence of that constraint in its lexicon. A great deal of work has demonstrated that gradient effects of similarity avoidance between consonants are evident in the lexicons of virtually all languages that have been studied in this way (Greenberg 1950, McCarthy 1991, 1994, Berkley 2000, Frisch, Pierrehumbert, & Broe 2004, Pozdniakov & Segerer 2007, among others). If it could be shown that some of these similarity avoidance effects could be better quantified by taking cue-based context into account, this would be very strong evidence in favor of *PCR as a phonological constraint.

6.8   Conclusion

In this chapter, I have motivated and developed the NO POORLY-CUED REPETITIONS (*PCR) approach to repetition avoidance. This approach is based around the notion that languages may establish particular acoustic/auditory cues as special licensing conditions for consonantal repetitions (in terms of the contrast between the repeated consonant and Ø), over and above the normal contrast licensing conditions that hold in the language. Specifically, languages set well-cuedness thresholds for repeated consonants based primarily on the presence vs. absence of an intensity rise as a cue to the presence of the repeated consonant. When these special repetition licensing conditions are not met, *PCR can induce phonological repairs and avoidance strategies.

The *PCR approach economically and consistently explains the distributions of consonant repetitions in the reduplication patterns of the ancient Indo-European languages. In these languages, when defined and ranked appropriately, it contributes to generating all of the following alternative reduplication patterns: non-copying in Ancient Greek; cluster-copying in Gothic; the principled distribution of C_2-copying, the $C_1\bar{e}C_2$ pattern, and infixal reduplication in Sanskrit; and infixal reduplication in Latin. The same contributing factors, ranked in slightly different ways, are also sufficient to account for the reduplicative distributions in the non–Indo-European language Klamath; namely, Klamath combines the *PCR licensing conditions of Ancient Greek with the ranking of phonological constraints found in Gothic. *PCR also explains non-reduplicative effects in several languages, namely, a case of suffixal allomorphy in Latin and an unexpected pattern of aspiration mobility in Sanskrit. It may also provide an avenue for future explanation of the *sCVC constraint in English.

This approach provides an alternative to Fleischhacker's (2005) similarity-based approach to partial reduplication, designed to cover much of the same data. (Time unfortunately prevents me from laying out a full-fledged review and comparison of her proposal here.) A major virtue of Fleischhacker's proposal is that it relates the behavior of cluster reduction in reduplication to

cluster repair via epenthesis, as well as certain facts about alliteration, rhyming, and punning. The *PCR approach cannot be extended to capture the latter types of facts in any obvious way; conversely, Fleischhacker's account would have nothing to say about the non-reduplicative *PCR effects discussed in Section 6.7. However, Fleischhacker's insight that the reduplication patterns are connected to the epenthesis facts can be maintained in the *PCR approach. Yun (2016) argues that epenthesis site is determined by the position of intensity rises in clusters. I have proposed that the licensing of repetitions is crucially tied to the presence of intensity rises in clusters. Therefore, it appears that the two types of patterns are indeed linked insofar as they both depend on the same property, namely, intensity rise. Additional work — both experimental and analytical — is required to fully evaluate the relative merits of the two proposals.

Chapter 7

Conclusion

7.1 Summary of Dissertation

This dissertation has examined the reduplicative systems of the ancient Indo-European languages from both a synchronic and diachronic perspective. The main chapters of the dissertation have dealt with Ancient Greek (Chapter 2), the Anatolian languages Hittite and Luwian (Chapter 3), Gothic (Chapter 4), and Sanskrit (Chapter 5). Each of these languages — either in their synchronic systems and/or in reconstructed prior stages — exhibit evidence of an alternation in reduplication pattern based on the composition of the reduplicative base. Throughout these chapters, this alternative was characterized as a difference between stop-sonorant cluster-initial roots/bases (TRVX–), which consistently showed the default Indo-European C_1-copying pattern ($\underline{TV}$-TRVX–), and s-stop cluster-initial roots/bases (STVX–), which routinely displayed some alternative strategy, differing by language.

The final chapter (Chapter 6) collected this evidence, and expanded the scope of the inquiry to look at the full distribution of cluster types involved in these languages (as well as in an equivalent pattern in Klamath). To explain these more complicated distributions, I proposed the NO POORLY-CUED REPETITIONS constraint (*PCR). This constraint posits that languages may impose special licensing conditions, in terms of acoustic/auditory cues to the contrast between a consonant and its absence (the C∼Ø contrast), for repeated identical consonants. The languages examined employ different combinations of such cues, centering around the *intensity rise* cue, such that the precise distributions vary between languages, but in understandable ways.

In the remainder of this section, I summarize the main analytical conclusions of the individual chapters. I then, in Section 7.2, discuss some of the themes which emerge from these analyses, focusing on the methodology employed across these chapters whereby synchronic analyses across diachronic stages are linked via models of (morpho)phonological learning. Finally, in Section 7.3, I collect the evidence from the dissertation and use it to argue for a particular reconstruction of the reduplicative system of Proto-Indo-European, namely, across-the-board C_1-copying.

7.1.1 Greek

In Chapter 1, I developed a comprehensive account of the reduplicative system of Ancient Greek at multiple diachronic stages, paying special attention to the historical development of the Attic Reduplication pattern. First, I constructed a detailed synchronic analysis of the reduplicative system of attested Ancient Greek which simultaneously generates the productive patterns displayed by consonant-initial roots — the default C_1-copying pattern for singleton-initial roots and for stop-

sonorant–initial roots, in opposition to the alternative non-copying pattern for obstruent-obstruent–initial roots (motivated by the *PCR constraint) — and the productive vowel-lengthening pattern for vowel-initial roots. The Attic Reduplication pattern, and also the previously unrecognized sub-regularity of the exceptional C_1-copying cluster-initial perfects associated with reduplicated presents, had to be accounted for through lexical indexation of a special constraint requiring copying (named there REDUP(RED)$_{lex}$) in cases where the phonotactics or alignment would normally divert the derivation to the non-copying mapping. The existence of the Attic Reduplication pattern along-side the productive vowel-lengthening pattern, and the way in which Attic Reduplication has to be encoded in the synchronic grammar, raised questions about its diachronic origin.

Based on the clear etymological connection between Attic Reduplication and the laryngeals, I argued that laryngeal-specific phonotactics operative in Pre-Greek — including a constraint against locally repeated laryngeals, of the same sort as the more general *PCR constraint active in later Greek and many other Indo-European languages — spawned the precursor of Attic Reduplication (Pre-AR). Pre-AR was then shown to be consistent with the interaction between another laryngeal-specific phonotactic repair (laryngeal vocalization) and the general reduplicative grammar as still evidenced in attested Ancient Greek.

In an attempt to retain the reflex of the pattern subsequent to the loss of the laryngeals (and thus the loss of the pattern's conditioning factors), speakers innovated a new constraint system based on lexical indexation, as is necessary to account for Attic Reduplication synchronically in attested Ancient Greek. This was a natural result of the learning process given the inconsistency posed by the data regarding vowel-initial perfects at the stage immediately following the loss of the laryngeals. I formalized this using a version of Becker's (2009) model for Inconsistency Detection and Constraint Cloning in Recursive Constraint Demotion (Tesar 1995, Tesar & Smolensky 1998, 2000). This same system can be used to derive indexation of the exact same sort for the exceptional C_1-copying perfects to roots with reduplicated presents. This demonstrates that both patterns are not simply frozen, archaic forms which have arbitrarily persisted in the language, but rather synchronically generable minority patterns which are subject to the normal demands of the grammar.

This chapter in particular illustrated how synchrony and diachrony can be used in tandem to help explain systematic irregularities. This approach, to one degree or another, characterizes each of the case studies in the Indo-European languages undertaken in this dissertation.

7.1.2   Anatolian

Chapter 2 examined the synchronic reduplicative systems of the Anatolian daughter languages Hittite and Luwian. Both of these languages differ in a crucial respect from most of the other languages examined in this dissertation: their productive reduplication patterns show no direct evidence of a distinction in copying behavior based on cluster type. That is to say, neither language has a synchronically active *PCR constraint. In fact, both of them show evidence of violation of that constraint in reduplication in the form of the _VC_-_VCX_– pattern.

Interestingly, however, internal and comparative reconstruction points very strongly in favor of reconstructing their proximate common ancestor, Proto-Anatolian, with a reduplicative alternation driven by an active *PCR constraint: C_1-copying reduplication for TRVX– roots vs. cluster-copying reduplication for STVX– roots (as attested also in Gothic). Insofar as the reduplicative systems of Hittite and Luwian represent clear divergences from the system reconstructible for Proto-Anatolian, this is indicative of *grammar change*. Like the other Indo-European languages examined in this dissertation, Proto-Anatolian reduplication shows effects of *PCR. Yet, by the period of attested Hittite and Luwian, *PCR has been demoted to the bottom of the grammar, as shown by the inno-

vative _VC_-_VCX_– reduplicative pattern. I argued that these developments can be accounted for with the following diachronic scenario.

First, independent phonological changes in both Hittite and Luwian affected the reduplication of cluster-initial roots: the development of across-the-board cluster copying in Hittite; the deletion of word-initial *_s_ in Luwian. This produced ambiguities in the learning data with respect to *PCR. By adopting a slightly revised version of Recursive Constraint Demotion which favors installation of maximally informative winner-preferring constraints over less informative winner-preferring constraints — which I termed Maximally Informative Recursive Constraint Demotion (MIRCD) — this ambiguity leads directly to the demotion of *PCR to the bottom of the rankings, in both languages. Second, the _VC_-_VCX_– reduplicative pattern emerged independently in each language after the post-Proto-Anatolian loss of pre-vocalic word-initial *h_1, when newly vowel-initial roots were input into the innovative synchronic grammar for the first time. This innovation was made possible (and indeed perhaps necessary) by MIRCD's total demotion of *PCR.

7.1.3   Germanic

The analysis developed in Chapter 4 showed that the pattern of morphophonological markings of the preterite stem(s) found among the "strong" verbs in Gothic (or, rather, its historical precursor, Pre-Proto-Germanic) — which includes, but is not limited to, reduplication — can be generated through the interaction of markedness and faithfulness constraints with REALIZE MORPHEME constraints (cf. Kurisu 2001) that require phonological contrast between morphologically related stems; specifically, a contrast between the present (or default) verbal stem and the preterite stem, and a contrast between the default preterite stem (which surfaces in the plural and elsewhere) and the preterite singular indicative stem. This system of verbal morphology thus provides compelling evidence that specific morphosyntactic features are visible to the phonology, and that the extent to which such morphosyntactically distinct forms remain formally distinct can be modeled in terms of language-specific constraint grammars. Furthermore, the morphophonological system that controls the patterns of stem formation is fully compatible with the grammar that controls the shape of reduplication, in the one corner of the language where it is retained.

7.1.4   Sanskrit

In Chapter 5, I developed an account of the reduplicative system of the Sanskrit perfect. Sanskrit is perhaps unique among the attested Indo-European languages in the way it deals with potential *PCR violations, insofar as it employs distinct strategies based on the derivational history of the offending base-initial sequence. For cluster-initial roots which would violate *PCR when accompanied by the default C_1-copying pattern (namely STVX– roots), Sanskrit employs C_2-copying as its alternative reduplication strategy (i.e. _TV_-_STVX_–). However, for CaC roots in morphological categories that normally trigger root-vowel deletion (the "perfect weak stem"), *PCR violations are dealt with in a completely different way. Rather than showing C_2-copying as the *PCR avoidance strategy, these forms exhibit the apparently non-reduplicative "$C_1\bar{e}C_2$ pattern".

I demonstrated that the $C_1\bar{e}C_2$ pattern can be derived in one of several ways, consistent with the same synchronic grammar necessary to generate the C_2-copying for STVX– roots. First, it could be analyzed as a case of phonologically _conditioned_ allomorphy, where the $C_1\bar{e}C_2$ forms are synchronically derived from a non-reduplicative morph of PERFECT — namely, a floating feature bundle that converts underlying root /a/ to surface [ē]. This is likely the most appropriate analysis for synchronic Sanskrit. Alternatively, the $C_1\bar{e}C_2$ pattern can be treated as phonologically _driven_ allomorphy, where the $C_1\bar{e}C_2$ forms actually represent the optimal mapping from an underlying

representation with /RED/. While the phonological approach is perhaps simplest when cast in a derivational/serial phonological framework, a parallel OT analysis is possible if we allow Input-Reduplicant faithfulness and faithfulness to timing slots. It undoubtedly must be the case that some version of the phonological approach was responsible for the origin of this pattern, as the allomorphy approach needs a diachronic source for the /-$\bar{e}$-/ morph that is used to derive the $C_1\bar{e}C_2$ forms.

The $C_1\bar{e}C_2$ pattern in Sanskrit is matched almost exactly by the $C_1\bar{e}C_2$ pattern in the Germanic Strong Class V preterite plurals (and perhaps also several minor patterns attested elsewhere in Indo-European). The origin of the Germanic pattern can be explained in the very same way as that of Sanskrit. I proposed that it is exactly the $C_1\bar{e}C_2$ forms that drive the development of the contrast-based stem-formation system of Pre-Proto-Germanic proposed in Chapter 4. Specifically, the processes of zero-grade vowel-deletion and *PCR-driven consonant-deletion, which stood in a feeding order at the prior stage, are rendered opaque by the Germanic accent shift. This undermined learners' ability to construct a consistent grammar based on a reduplicated underlying representation of the preterite, leading the language ultimately down the path towards the dramatically different contrast-based system it eventually developed. The ongoing work represented by Sandell & Zukoff (2017) aims to show that this change is effectively pre-destined after the accent shift, using a computational model based on multi-generational learning.

7.1.5   *PCR

Chapter 6 tied together the common thread running through the preceding chapters, developing a concrete proposal regarding the nature of the *PCR constraint. This approach is based around the notion that languages may establish particular acoustic/auditory cues as special licensing conditions for consonantal repetitions (in terms of the contrast between the repeated consonant and Ø), over and above the normal contrast licensing conditions that hold in the language. Specifically, languages set well-cuedness thresholds for repeated consonants based primarily on the presence vs. absence of an *intensity rise* as a cue to the presence of the repeated consonant. When these special repetition licensing conditions are not met, *PCR can and does induce phonological repairs and avoidance strategies.

The *PCR approach economically and consistently explains the distributions of consonant repetitions in the reduplication patterns of the ancient Indo-European languages. In these languages, when defined and ranked appropriately, it contributes to generating all of the following alternative reduplication patterns: non-copying in Ancient Greek; cluster-copying in Gothic; the principled distribution of C_2-copying, the $C_1\bar{e}C_2$ pattern, and infixal reduplication in Sanskrit; and infixal reduplication in Latin. The same contributing factors, ranked in slightly different ways, are also sufficient to account for the reduplicative distributions in the non–Indo-European language Klamath; namely, Klamath combines the *PCR licensing conditions of Ancient Greek with the ranking of phonological constraints found in Gothic. *PCR also explains non-reduplicative effects in several languages, namely, a case of suffixal allomorphy in Latin and an unexpected pattern of aspiration mobility in Sanskrit. It may also provide an avenue for future explanation of the *sCVC* constraint in English.

The *PCR approach provides an alternative to Fleischhacker's (2005) similarity-based analysis of cluster reduction in partial reduplication. While, unlike Fleischhacker's account, *PCR does not provide an explanation for certain facts about alliteration, rhyming, and punning, the *PCR approach does capture Fleischhacker's insight that these patterns of cluster simplification in reduplication are linked to the distribution of different types of epenthesis, in that we now know that both of these processes can be explained with reference to the presence vs. absence of an intensity rise within a cluster (see Yun 2016 on intensity rise and epenthesis). Additional work — both experimental and analytical — is required to fully evaluate the relative merits of the two proposals.

7.2   Discussion of the Methodology

Beside the contributions in the realm of the more detailed and thorough analytical treatments of the reduplicative systems of these languages, and the novel *PCR proposal, the main contribution of this dissertation is methodological in nature. While certainly not without precedent, this dissertation represents one of the most thoroughgoing exercises in the integration of synchrony and diachrony in theoretical phonological explanation.

The case studies represented by Chapters 2–5 in large part have followed a consistent expository structure.[1] First, a detailed synchronic theoretical analysis was constructed to account for a complex morphophonological system in an attested language. This analysis helped reveal certain aspects of the system which stand apart from the general productive properties of that system, pointing to their status as archaisms. Standard methods of historical reconstruction, both internal and comparative, could then be brought to bear to identify the specific earlier stage of the language where (the pre-cursor of) that pattern would have been productive, or at least regular.

The structure described thus far is fairly standard for research in historical (morpho)phonology. But the approach to historical explanation adopted in this dissertation does not stop here. In each of the cases, the stage identified via historical reconstruction was then itself subjected to full-fledged synchronic theoretical analysis, undertaken with the same level of detail (or the greatest level of detail possible, given the limits of reconstruction) as the synchronic system of the attested language. The grammars of the diachronically separated stages were then compared, such that the differences between the stages could be characterized not just phenomenologically (as is typical in traditional historical linguistic approaches), but in terms of the systemic differences in constraint rankings and/or underlying representations that comprise the distinct grammars.

This type of comparison allowed for considerations of (morpho)phonological learning to be employed as explanatory devices linking one stage to the next. That is to say, the relationship between the grammar of an earlier stage of a language and the grammar of a later stage of that language are not related solely through the application of random and spontaneous instances of language change in the intervening period; rather, at least some aspects of the systematic differences between stages can be derived from the way that particular other changes affect the learning input and how those effects drive the learning process in a particular direction that is predictable (at least after the fact).

It is, though, certainly the case that many aspects of language change are not deterministically predictable before the fact, at least given our current levels of understanding of these matters. I have not in general tried to account for instances of change that we might characterize as "simple" sound changes, where some particular segment or phonological string changes or is deleted. Rather, I have for the most part taken these sound changes as given, and tried to derive the systemic morphophonological changes observable within these systems from the learning conditions which result from said sound changes. This division of labor is artificial and not entirely appropriate, as both types of change (presumably) derive from the same language learning process. The overall approach advanced in this dissertation will therefore be significantly improved when it can be more explicitly integrated with a concrete theory of sound change and learning.

Nevertheless, this interplay of synchrony, diachrony, learning, and morphophonological theory achieves a greater degree of explanatory power than most previous ways of approaching questions of the same sorts. Specifically, this dissertation has demonstrated that, given the right sorts of evidence

---

[1] Chapter 4 on Gothic covers much the same ground, though it turns this structure somewhat on its head. That chapter would have benefited from a presentation of the sort described here.

and tools, an approach which integrates these often disparate linguistic methodologies can lead to true *explanation* of synchronic irregularities/exceptions, rather than just description.

7.3 Reconstruction of Proto-Indo-European Reduplication

This dissertation has assembled the vast majority of the evidence regarding (verbal) reduplication in the Indo-European languages,[2] and analyzed it in a thorough and consistent manner. This puts us in a better position to evaluate the reconstruction of the reduplicative system of Proto-Indo-European (PIE) itself.[3] Throughout the dissertation, I have focused on the analysis and reconstruction of reduplicant *shape*, and paid relatively less attention to questions regarding reduplicant *vocalism*. I therefore limit myself here to discussion of reduplicant shape in PIE, specifically the question of whether or not PIE made a distinction in reduplicant shape between TRVX– roots and STVX– roots. This question can be broken up into three parts:

(1)   a.   Did PIE exhibit *PCR effects in reduplication? If so, then:
        b.   What was the alternative reduplication pattern induced by *PCR? and
        c.   Which cluster types patterned with STVX– in undergoing the alternative treatment?

The answer to (1a) that I will argue for here is *no*: PIE is to be reconstructed as having across-the-board C_1-copying. This renders questions (1b) and (1c) moot.

    The Indo-Europeanist literature contains two main views on the reconstruction of PIE reduplication with respect to the potential TRVX– vs. STVX– distinction. The more traditional view reconstructs for PIE the distribution found in Gothic (and now Proto-Anatolian): i.e., C_1-copying for TRVX– roots (<u>TV</u>-TRVX–) but cluster-copying for STVX– roots (<u>STV</u>-STVX–). This allows for the productive STVX– treatments of Ancient Greek (<u>V</u>-STVX–), Sanskrit (<u>TV</u>-STVX–), and Latin (S-<u>TV</u>-TVX–) to be seen as different kinds of reductions (often referred to as "dissimlation") from the original type. The precise means by which these reductions are meant to have happened are rarely (if ever) made explicit, and none of them are matched by independently observed types of dissimilation in those languages.

    I argue for an alternative reconstruction for PIE, one which is present in the Indo-Europeanist literature (for example, Byrd 2010:100–105), though not the most commonly accepted view: in PIE, both TRVX– and STVX– roots, and indeed all types of root-initial clusters, exhibited C_1-copying; i.e., <u>TV</u>-TRVX– and <u>SV</u>-STVX–. The strongest evidence in favor of this alternative reconstruction comes from archaisms in Greek and Latin (Brugmann & Delbrück 1897–1916:40–41, Byrd 2010:103–104), and their agreement with Iranian (Byrd 2010:103), which all points to <u>SV</u>-STVX–. Specifically, there is exact correspondence between the archaic reduplicated present forms of the PIE root √*$steh_2$ 'stand' in Ancient Greek ἵστημι [hí-stɛ̄-mi] (< Proto-Greek *<u>si</u>-stā-mi) and Latin *sistō*

---

[2]   There are several Indo-European languages/branches which I have not been able to (fully) treat in this dissertation. There are substantial reduplicative patterns in Avestan (Iranian) and Tocharian which have here not been examined at all, and there is a good deal more to say about Old Irish (and perhaps Celtic more generally) than was presented in the brief discussions in Chapters 1 and 3. Furthermore, there are additional reduplication facts in Sanskrit not discussed in either Chapter 5 or 6, including the behavior of vowel-initial roots in the perfect, the perfects with long vowel reduplicants (which probably only in part can be explained with reference to laryngeals), and the Sanskrit intensive. A comprehensive reconstruction of reduplication in Proto-Indo-European will need to incorporate these additional languages and patterns. Another aspect which requires further consideration is how exactly the $C_1\bar{e}C_2$ patterns fit into our reconstruction. These questions all must be left for future work.

[3]   See McIntyre (1992), Niepokuj (1997), Keydana (2006, 2012) for recent work on the reconstruction of reduplication in Proto-Indo-European. (Consult Katz 2000 for a critique of some of the ideas advanced in Niepokuj 1997.)

([s̲i̲-st-ō]), despite neither belonging to the productive pattern of the language (compare Ancient Greek perfect ἔσταλκα [é-stal-k-a], Latin perfect *stetī* [s-t̲e̲-t-ī]). Avestan and Old Persian appear to match the archaic treatment shown by Greek and Latin: Avestan h̲i̲-štaiti, vi-š̲a̲-starᵒ, Old Persian a-h̲i̲-štatā (Byrd 2010:103). The fact that the Latin and Greek forms agree with the Iranian forms can only be explained if that pattern is reconstructed to Proto-Indo-European, the last common ancestor of those three branches.

Note that this negates the evidence of C_2-copying from Sanskrit for the purposes of reconstructing PIE. If the claim is that C_1-copying in Iranian reflects the PIE situation, then it must be the case that the same situation held at the level of Proto-Indo-Iranian. This means that the Sanskrit C_2-copying pattern has to be an innovation against Proto-Indo-Iranian, and therefore cannot be telling us anything about PIE.

One additional piece of evidence in favor of reconstructing C_1-copying to PIE comes from Old Irish. The sole STVX– root attested with reduplication in Old Irish also exhibits C_1-copying: √*scenn* → s̲e̲-scann- (Thurneysen 1946 [1980]:424/§687; Byrd 2010:103). Therefore, we actually have fairly substantial evidence for positing C_1-copying reduplication for PIE STVX– roots.

The main reason why traditional Indo-Europeanists were keen on the cluster-copying reconstruction was that it was difficult to motivate the set of changes that would take you from this state of affairs to the respective daughter languages. Understanding the changes in terms of grammatical reorganization rather than sound change alleviates this worry. Each of the changes can be characterized in a unified way: the change results from the promotion of *PCR.

Prior to such a change, *PCR would have been ranked *below* all the constraints whose violation would have resulted in some pattern other than C_1-copying. The change in pattern is the result of the re-ranking of *PCR *above* one such constraint. The reason why the different languages develop different patterns for STVX– roots is because they each "choose" different constraints to promote *PCR above. Proto-Anatolian (2a) and Gothic[4] (2b) independently promote *PCR above *CC (and/or ALIGN-ROOT-L), leading to cluster-copying. Ancient Greek (really Proto-Greek; (2c)) promotes *PCR above ONSET (subsequent to developing a morphologically fixed reduplicative vowel; see Chapter 2), yielding non-copying. After keeping *PCR relatively low-ranked through Proto-Indo-Iranian, Sanskrit (2d) promotes *PCR above ANCHOR-L-BR, which creates C_2-copying. And Latin (2e) promotes *PCR above both ALIGN-RED-L and CONTIGUITY-IO, allowing for the infixal pattern to arise.

(2)   a.   PIE C_1-copying → Proto-Anatolian cluster-copying | ***PCR ≫ *CC** |
     b.   PIE C_1-copying → Gothic cluster-copying | ***PCR ≫ *CC** |
     c.   PIE C_1-copying → Ancient Greek non-copying | ***PCR ≫ ONSET** |
     d.   Proto-Indo-Iranian C_1-copying → Sanskrit C_2-copying | ***PCR ≫ ANCHOR-L-BR** |
     e.   PIE C_1-copying → Latin infixation | ***PCR ≫ ALIGN-RED-L, CONTIGUITY-IO** |

We now see that it actually is straightforward to characterize the set of changes that derive the daughter languages from C_1-copying for STVX– roots in PIE: they each result from the promotion of *PCR. In a certain sense, then, the changes in overt reduplication pattern indeed all arise from the *same* underlying change: increased sensitivity to the repetition avoidance constraint. Assuming that *PCR was independently promoted in the various languages, rather than being retained as an active constraint from PIE, is logically consistent with the fact that the languages differ somewhat in exactly which repetition types are targeted by *PCR (see Chapter 6 for extensive discussion).

---

[4]   This change may well have happened already in (Pre-)Proto-Germanic; see Chapter 4. This question is not relevant for the current discussion.

While all of these languages make a consistent distinction between stop-sonorant–initial roots (TRVX– roots) and *s*-stop–initial roots (STVX– roots), they show substantial differences in the treatment of the other cluster types. This seems a likely state of affairs if the *PCR effects represent a parallel development driven by similar inherited conditions. If this scenario regarding *PCR and the general reconstruction of reduplication in Proto-Indo-European is correct, then this presents a further argument that reconstructions of dynamic properties/systems like reduplication need to be based on fully articulated theoretical analyses, not just correspondences in surface forms.

References

Albright, Adam. 2002a. The Identification of Bases in Morphological Paradigms. UCLA, PhD Dissertation.

———. 2002b. Islands of Reliability for Regular Morphology: Evidence from Italian. *Language* 78(4):684–709.

Albright, Adam & Bruce Hayes. 1999. An Automated Learner for Phonology and Morphology. Ms., UCLA.

———. 2003. Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study. *Cognition* 90(2):119–161.

Alderete, John. 2001. Dominance Effects as Transderivational Anti-Faithfulness. *Phonology* 18(2):201–253.

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy & Suzanne Urbanczyk. 1999. Reduplication with Fixed Segmentism. *Linguistic Inquiry* 30(3):327–364.

Allen, Joseph Henry & James Bradstreet Greenough. 1903 [2006]. *New Latin Grammar*. Mineola, New York: Dover Publications, Inc.

Arregi, Karlos & Andrew Nevins. 2012. *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Dordrecht: Springer.

Baal, Berit Anne Bals, David Odden & Curt Rice. 2012. An Analysis of North Saami Gradation. *Phonology* 29(2):165–212.

Barkaï, Malachi & Julia Horvath. 1978. Voicing assimilation and the Sonority Hierarchy: Evidence from Russian, Hebrew and Hungarian. *Linguistics* 16(212):77–88.

Barker, M. A. R. 1964. *Klamath Grammar*. Berkeley: University of California Press.

Becker, Michael. 2009. Phonological Trends in the Lexicon: The Role of Constraints. University of Massachusetts, Amherst, PhD Dissertation.

Beckman, Jill N. 1998. Positional Faithfulness. University of Massachusetts, Amherst, PhD Dissertation.

Beckwith, Miles Christopher. 1996. The Greek Reduplicated Aorist. Yale University, PhD Dissertation.

Beekes, Robert S. P. 1969. *The Development of the Proto-Indo-European Laryngeals in Greek*. The Hague: Mouton.

Beekes, Robert S. P. & Lucien van Beek. 2010. *Etymological Dictionary of Greek*. Leiden; Boston: Brill.

Benua, Laura. 1997. Transderivational Identity: Phonological Relations Between Words. University of Massachusetts, Amherst, PhD Dissertation.

Benveniste, Emile. 1937. Le problème du *þ* indo-européen. *Bulletin de la Société de Linguistique* 38:139.

Berkley, Deborah Milam. 2000. Gradient Obligatory Contour Principle Effects. Northwestern University, PhD Dissertation.

Bermúdez-Otero, Ricardo. 2013. Amphichronic Explanation and the Life Cycle of Phonological Processes. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford Handbook of Historical Phonology*, Oxford: Oxford University Press.

Bjorndahl, Christina. 2013. Phonetic properties of [v] in Russian, Serbian and Greek. In *Proceedings of Meetings on Acoustics ICA2013*, vol. 19, Acoustical Society of America. http://asa.scitation.org/doi/abs/10.1121/1.4800950.

———. 2015. The Phonetics and Phonology of Segment Classification: A Case Study of /v/. In Charles E. Cairns & Eric Raimy (eds.), *The Segment in Phonetics and Phonology*, 236–250. Wiley-Blackwell.

Blevins, Juliette. 1993. Klamath Laryngeal Phonology. *International Journal of American Linguistics* 59(3):237–279.

———. 2001. Klamath /sʔ/ Clusters. *International Journal of American Linguistics* 67(4):467–474.

Blevins, Juliette & Andrew Garrett. 2004. The Evolution of Metathesis. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, 117–156. Cambridge: Cambridge University Press.

Boersma, Paul. 1997. How We Learn Variation, Optionality, and Probability. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam 21*, 43–58. ROA 221.

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró. 2007. Allomorph Selection and Lexical Preferences: Two Case Studies. *Lingua* 117:903–927.

Bowers, Dustin Andrew. 2015. A System for Morphophonological Learning and its Consequences for Language Change. UCLA, PhD Dissertation.

Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Brugmann, Karl & Berthold Delbrück. 1897–1916. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Strassburg: K.J. Trübner.

Byrd, Andrew Miles. 2010. Reconstructing Indo-European Syllabification. UCLA, PhD Dissertation.

———. 2011. Laryngeal Vocalization as Schwa Epenthesis: Why does Compensatory Lengthening not Occur? Paper presented at the 30th East Coast Indo-European Conference, Cambridge, MA. June 10, 2011.

———. 2015. *The Indo-European Syllable*. Leiden: Brill.

———. forthcoming. The Phonology of Proto-Indo-European. In Jared S. Klein & Brian Joseph (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*, Berlin: Mouton de Gruyter.

Carstairs, Andrew. 1988. Some Implications of Phonologically Conditioned Suppletion. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1988*, 67–94. Dordrecht: Kluwer.

Casali, Roderic F. 1996. Resolving Hiatus. UCLA, PhD Dissertation. ROA-215.

Chantraine, Pierre. 1980. *Dictionnaire étymologique de la langue grecque: Histoire des mots*. Paris: Klincksieck.

Clackson, James. 1994. *The Linguistic Relationship Between Armenian and Greek*. Oxford: Blackwell.

Clements, George N. 1990. The Role of the Sonority Cycle in Core Syllabification. In John Kingston & Mary Beckman (eds.), *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, 283–333. Cambridge: Cambridge University Press.

Clements, George N. & Samuel Jay Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Coetzee, Andries W. 2005. The Obligatory Contour Principle in the Perception of English. In Sónia Frota, Marina Vigário & Maria João Freitas (eds.), *Prosodies: With Special Reference to Iberian Languages*, 223–245. Berlin/New York: Mouton de Gruyter.

Collinge, N. E. 1985. *The Laws of Indo-European*. Amsterdam: John Benjamins Publishing Company.

Cooper, Adam I. 2009. Similarity Avoidance in the Proto-Indo-European Root. *University of Pennsylvania Working Papers in Linguistics* 15(1):55–64. http://repository.upenn.edu/pwpl/vol15/iss1/8/.

———. 2014. *Reconciling Indo-European Syllabification*. Leiden: Brill.

Côté, Marie-Hélène. 2000. Consonant Cluster Phonotactics: A Perception-Based Approach. MIT, PhD Dissertation.

Cowgill, Warren. 1965. Evidence in Greek. In Werner Winter (ed.), *Evidence for Laryngeals*, 142–180. The Hague: Mouton.

Crosswhite, Katherine. 1999. Intra-Paradigmatic Homophony Avoidance in Two Dialects of Slavic. In Matthew Gordon (ed.), *UCLA Working Papers in Linguistics 1: Papers in Phonology*, 48–67. Los Angeles: UCLA Department of Linguistics.

Cser, András. 2009. The History of Perfective Reduplication and Stem-Initial Patterns in Latin. *Acta Antiqua Academiae Scientiarum Hungaricae* 49(2):107–115.

———. 2015. The Nature of Phonological Conditioning in Latin Inflectional Allomorphy. *Acta Linguistica Hungarica* 62(1):1–35.

Davis, Stuart. 1984. Some Implications of Onset-Coda Constraints for Syllable Phonology. In *CLS 20: Papers from the Twentieth Regional Meeting of the Chicago Linguistics Society*, 46–51. Chicago: Chicago Linguistic Society.

———. 1989. Cross-vowel Phonotactic Constraints. *Computational Linguistics* 15(2):109–110.

———. 1991. Coronals and the Phonotactics of Non-Consonants in English. In Carole Paradis & Jean-François Prunet (eds.), *The Special Status of Coronals: Internal and External Evidence* (Phonetics and Phonology 2), 49–60. San Diego: Academic Press.

de Haas, Wim. 1988. *A Formal Theory of Vowel Coalescence: A Case Study of Ancient Greek*. Dordrecht, Holland; Providence, RI, U.S.A.: Foris.

DeLisi, Jessica. 2015. Sonority Sequencing Violations and Prosodic Structure in Latin and Other Indo-European Languages. *Indo-European Linguistics* 3(1):1–23.

Dempsey, Timothy Richard. 2015. Verbal Reduplication in Anatolian. UCLA, PhD Dissertation.

Fleischhacker, Heidi Anne. 2005. Similarity in Phonology: Evidence from Reduplication and Loan Adaptation. UCLA, PhD Dissertation.

Flemming, Edward. 1995/2002. *Auditory Representations in Phonology*. New York: Routledge.

———. 2007. Stop Place Contrasts Before Liquids. In *Proceedings of the 16th International Congress on Phonetic Sciences*, 233–236. http://icphs2007.de/conference/Papers/1764/1764.pdf.

———. 2008. The Realized Input. Ms., MIT. http://web.mit.edu/flemming/www/paper/RI2.pdf.

Fortson, Benjamin W. 2010. *Indo-European Language and Culture: An Introduction*. 2nd edn. Malden, MA: Wiley-Blackwell Publishing.

Fox, Anthony. 1995. *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford & New York: Oxford University Press.

Frampton, John. 2009. *Distributed Reduplication*. Cambridge, MA: MIT Press.

Frisch, Stefan A. 2004. Language Processing and Segmental OCP Effects. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, 346–371. Cambridge, UK: Cambridge University Press.

Frisch, Stefan A., Janet B. Pierrehumbert & Michael B. Broe. 2004. Similarity Avoidance and the OCP. *Natural Language & Linguistic Theory* 22:179–228.

Fudge, Erik C. 1969. Syllables. *Journal of Linguistics* 5(2):253–286.

Fukazawa, Haruka. 1999. Theoretical Implications of OCP Effects on Features in Optimality Theory. University of Maryland, PhD Dissertation. ROA-307.

Fulk, Robert D. 1987. Reduplicating Verbs and their Development in Northwest Germanic. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 109:159–79.

Fullerton, G. Lee. 1991. Reduplication and the Prosody of Ancient Germanic. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 113:1–21.

Gafos, Diamandis. 1998. A-templatic Reduplication. *Linguistic Inquiry* 29(3):515–527.

Giannakis, Giorgios. 1992. Studies in the Syntax and Semantics of the Reduplicated Presents in Homeric Greek and Indo-European. UCLA, PhD Dissertation.

Goldsmith, John A. 1976. Autosegmental Phonology. MIT, PhD Dissertation.

Golston, Chris. 2014. Hiatus. In Giorgios K. Giannakis, Vit Bubenik, Emilio Crespo, Chris Golston & Alexandra Lianeri (eds.), *Encyclopedia of Ancient Greek Language and Linguistics*, vol. 2: G–O, 160–161. Leiden; Boston: Brill.

Greenberg, Joseph H. 1950. The Patterning of Root Morphemes in Semitic. *Word* 6(2):162–181.

Güterbock, Hans G. & Harry A. Hoffner, Jr. (eds.). 1989. *The Hittite Dictionary of the Oriental Institute of Chicago*, vol. L–N. Chicago: The Oriental Institute of the University of Chicago.

Halle, Morris. 1997. On Stress and Accent in Indo-European. *Language* 73(2):275–313.

———. 2008. Reduplication. In Robert Freidin, Carlos P. Otero, Maria Luisa Zubizarreta & Samuel Jay Keyser (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 325–357. Cambridge, MA: MIT Press.

Hayes, Bruce. 1986. Inalterability in CV Phonology. *Language* 62(2):321–351.

———. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20(2):253–306.

———. 2004. Phonological Acquisition in Optimality Theory: The Early Stages. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, 158–203. Cambridge: Cambridge University Press.

Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.). 2004. *Phonetically Based Phonology*. Cambridge: Cambridge University Press.

Hayes, Bruce & Zsuzsa Cziráky Londe. 2006. Stochastic Phonological Knowledge: The Case of Hungarian Vowel Harmony. *Phonology* 23(1):59–104.

Hayes, Bruce & Donca Steriade. 2004. Introduction: The Phonetic Bases of Phonological Markedness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, 1–33. Cambridge: Cambridge University Press.

Hayes, Bruce, Bruce Tesar & Kie Zuraw. 2013. OTSoft 2.5, software program, http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Hayes, Bruce & James White. 2015. Saltation and the P-map. *Phonology* 32(2):267–302.

Hendricks, Sean Q. 1999. Reduplication without Template Constraints: A Study in Bare-Consonant Reduplication. University of Arizona, PhD Dissertation.

Hermann, Eduard. 1923. *Silbenbildung im Griechen und in den andern indogermanischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht.

Hoffner, Harry A., Jr. & H. Craig Melchert. 2008. *A Grammar of the Hittite Language (Part I: Reference Grammar)*. Winona Lake, Indiana: Eisenbrauns.

Idsardi, William & Eric Raimy. 2008. Reduplicative Economy. In Bert Vaux & Andrew Nevins (eds.), *Rules, Constraints, and Phonological Phenomena*, 149–185. Oxford: Oxford University Press.

Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: Doubling in Morphology*. Cambridge, UK: Cambridge University Press.

Insler, Stanley. 1997. Vedic Denominatives to Thematic *a*-Stems. In Alexander Lubotsky (ed.), *Sound Law and Analogy: Papers in Honor of Robert S. P. Beekes on the Occasion of his 60th Birthday*, 103–110. Amsterdam: Rodopi.

Itô, Junko. 1989. A Prosodic Theory of Epenthesis. *Natural Language & Linguistic Theory* 7(2):217–259.

Itô, Junko & Armin Mester. 1999. The Phonological Lexicon. In Natsuko Tsujimura (ed.), *The Handbook of Japanese Linguistics*, 62–100. Oxford: Blackwell.

———. 2001. Covert Generalizations in Optimality Theory: The Role of Stratal Faithfulness Constraints. *Studies in Phonetics, Phonology, and Morphology* 7(2):273–299.

Jasanoff, Jay H. 2002. The Nom. Sg. of Germanic *n*-Stems. In Alfred Wedel & Hans-Jörg Busch (eds.), *Verba et Litterre: Explorations in Germanic Languages and German Literature: Essays in Honor of Albert L. Lloyd*, 31–46. Newark, Delaware: Linguatext.

———. 2007. From Reduplication to Ablaut: The Class VII Strong Verbs of Northwest Germanic. *Historische Sprachforschung* 120(1):241–284.

———. 2012. Long-Vowel Preterites in Indo-European. In H. Craig Melchert (ed.), *The Indo-European Verb. Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles 13-15 September 2010*, 127–135. Wiesbaden: Reichert.

Jun, Jongho. 2004. Place Assimlation. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, 58–86. Cambridge: Cambridge University Press.

Kanwisher, Nancy G. 1987. Repetition Blindness: Type Recognition Without Token Individuation. *Cognition* 27(2):117–143.

Katz, Joshua T. 2000. Notes on Indo-European Reduplication. *Indo-European Studies Bulletin* 9(1):7–13.

Kavitskaya, Darya. 2001. Hittite Vowel Epenthesis and the Sonority Hierarchy. *Diachronica* 18(2):267–299.

Kawahara, Shigeto. 2007. Copying and Spreading in Phonological Theory: Evidence from Echo Epenthesis. In Leah Bateman, Michael O'Keefe, Ehren Reilly & Adam Werle (eds.), *Papers in Optimality Theory III* (University of Massachusetts Occasional Papers in Linguistics 32), 111–144. Amherst, MA: University of Massachusetts Graduate Linguistic Student Association.

Kennedy, Robert. 2011. Reduplication in Sanskrit. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, vol. 5, 2855–2878. Wiley-Blackwell.

Kenstowicz, Michael. 1994. Syllabification in Chukchee: A Constraints-Based Analysis. In Alice Davison, Nicole Maier, Glaucia Silva & Wan Su Yan (eds.), *Proceedings of the Formal Linguistics Society of Mid-America 4*, 160–181. Iowa City: Department of Linguistics, University of Iowa. ROA-30.

Kessler, Brett. 1994. Sandhi and Syllables in Classical Sanskrit. In Erin Duncan, Donka Farkas & Philip Spaelti (eds.), *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 35–50. Stanford: CSLI.

Keydana, Götz. 2006. Die indogermanische Perfektreduplikation. *Folia Linguistica Historica* 27(1–2):61–116.

———. 2012. Evidence for Non-Linear Phonological Structure in Indo-European: The Case of Fricative Clusters. In Benedicte Nielsen Whitehead, Thomas Olander & Birgit Anette Olsen (eds.), *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics*, 99–117. Denmark: Museum Tusculanum Press.

Kimball, Sara. 1999. *Hittite Historical Phonology*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Kingston, John. 1985. The Phonetics and Phonology of the Timing of Oral and Glottal Events. University of California, Berkeley, PhD Dissertation.

Kiparsky, Paul. 1982. Lexical Morphology and Phonology. In I.-S. Yang (ed.), *Linguistics in the Morning Calm*, 3–91. Seoul: Hanshin.

———. 1986. The Phonology of Reduplication. Ms., Stanford.

———. 2010a. Compositional vs. Paradigmatic Approaches to Accent and Ablaut. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 21st Annual UCLA Indo-European Conference*, 137–182. Bremen: Hempen. [Extended version available at http://www.stanford.edu/~kiparsky/Papers/ucla_IE_09.submitted.new.pdf].

———. 2010b. Reduplication in Stratal OT. In Linda Uyechi & Lian Hee Wee (eds.), *Reality Exploration and Discovery: Pattern Interaction in Language & Life*, 125–142. Stanford: CSLI. http://www.stanford.edu/~kiparsky/Papers/reduplication.pdf.

———. 2015. Stratal OT: A Synopsis and FAQs. In Yuchau E. Hsiao & Lian-Hee Wee (eds.), *Capturing Phonological Shades Within and Across Languages*, 2–44. Cambridge Scholars Publishing. https://web.stanford.edu/~kiparsky/Papers/taipei.2014.pdf.

———. 2016. Cyclic and Non-Cyclic Ablaut and Accent. Paper presented at the 28th Annual UCLA Indo-European Conference, UCLA. November 11–12, 2016.

———. forthcoming. Accent and Ablaut. In Michael Weiss & Andrew Garrett (eds.), *Handbook of Indo-European Studies*, Oxford: Oxford University Press.

Kitto, Catherine & Paul de Lacy. 1999. Correspondence and Epenthetic Quality. In Catherine Kitto (ed.), *Proceedings of AFLA VI*, 181–200. University of Toronto: Toronto Working Papers in Linguistics.

Klein, Jared S. 2012. Ling 8320 Gothic lecture notes, University of Georgia, Spring 2012.

Kloekhorst, Alwin. 2004. The Preservation of *h_1 in Hieroglyphic Luwian: Two Separate *a*-Signs. *Historische Sprachforschung* 117(1):26–49.

———. 2006. Initial Laryngeals in Anatolian. *Historische Sprachforschung* 119:77–108.

———. 2008. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden; Boston: Brill.

Kraska-Szlenk, Iwona. 1997. Exceptions in Phonological Theory. In *Proceedings of the 16th International Congress of Linguists*, Paper number 0173. Oxford: Pergamon.

———. 1999. Syllable Structure Constraints in Exceptions. In John Rennison & Klaus Kühnhammer (eds.), *Phonologica 1996: Syllables!?*, 113–131. The Hague: Thesus.

Kulikov, Leonid. 2005. Reduplication in the Vedic Verb: Indo-European Inheritance, Analogy, and Iconicity. In Bernard Hurch (ed.), *Studies on Reduplication*, 431–454. New York: Mouton de Gruyter.

Kümmel, Martin Joachim. 2000. *Das Perfekt im Indoiranischen*. Wiesbaden: Reichert.

Kurisu, Kazutaka. 2001. The Phonology of Morpheme Realization. University of California, Santa Cruz, PhD Dissertation.

Kuryłowicz, Jerzy. 1927. Origine indoeuropéenne du redoublement attique. *EOS; Commentarii Societatis Philologae Polonorum* 30:206–210.

van de Laar, Henri M. F. M. 2000. *Description of the Greek Individual Verbal Systems*. Amsterdam: Rodopi.

Ladefoged, Peter. 2001. *Vowels and Consonants: An Introduction to the Sounds of Languages*. Malden, MA: Blackwell.

Lambdin, Thomas O. 2006. *An Introduction to the Gothic Language*. Eugene, Oregon: Wipf & Stock Publishers.

Leumann, Manu. 1977. *Lateinische Laut- und Formenlehre*. München: C. H. Beck.

Liddell, Henry George & Robert Scott. 1889. *An Intermediate Greek-English Lexicon*. Oxford: Oxford University Press.

Liddell, Henry George, Robert Scott & Henry Stuart Jones. 1940. *A Greek-English Lexicon*. Oxford: Clarendon Press. http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3atext%3a1999.04.0057.

Lundquist, Jesse & Anthony D. Yates. forthcoming. The Morphology of Proto-Indo-European. In Jared S. Klein & Brian Joseph (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*, Berlin: Mouton de Gruyter.

MacBride, Alexander Ian. 2004. A Constraint-Based Approach to Morphology. UCLA, PhD Dissertation.

Magri, Giorgio. 2012. Convergence of Error-Driven Ranking Algorithms. *Phonology* 29(2):213–269.

Marantz, Alec. 1982. Re Reduplication. *Linguistic Inquiry* 13(3):435–482.

Marchand, James W. 1973. *The Sounds and Phonemes of Wulfila's Gothic*. The Hague / Paris: Mouton.

Mascaró, Joan. 2007. External Allomorphy and Lexical Representation. *Linguistic Inquiry* 38(4):715–735.

Mayrhofer, Manfred. 1986. Lautlehre (Segmentale Phonologie des Indogermanischen). In Manfred Mayrhofer (ed.), *Indogermanische Grammatik*, vol. 1, 73–181. Heidelberg: Carl Winter.

McCarthy, John J. 1979. Formal Problems in Semitic Phonology and Morphology. MIT, PhD Dissertation.

———. 1986. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* 17(2):207–263.

———. 1991. Semitic Gutturals and Distinctive Feature Theory. In Bernard Comrie & Mushira Eid (eds.), *Perspectives on Arabic Linguistics III: Papers from the Third Annual Symposium on Arabic Linguistics*, 63–91. Amsterdam/Philadelphia: John Benjamins Publishing Company. http://works.bepress.com/john_j_mccarthy/69.

———. 1994. The Phonetics and Phonology of Semitic Pharyngeals. In Patricia A. Keating (ed.), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*, 191–233. Cambridge, UK: Cambridge University Press.

———. 1998. Morpheme Structure Constraints and Paradigm Occultation. In M. Catherine Gruber, Derrick Higgins, Kenneth Olsen & Tamra Wysocki (eds.), *Proceedings from the Annual Meeting of the Chicago Linguistic Society 32, Part II: The Panels*, 123–150. Chicago: Chicago Linguistic Society.

———. 2000. Harmonic Serialism and Parallelism. In Masako Hirotani, Andries Coetzee, Nancy Hall & Ji-yung Kim (eds.), *NELS 30: Proceedings of the North East Linguistic Society*, 501–524. Amherst, MA: University of Massachusetts Graduate Linguistic Student Association. https://works.bepress.com/john_j_mccarthy/79/.

———. 2003. OT Constraints are Categorical. *Phonology* 20(1):75–138.

———. 2008. The Gradual Path to Cluster Simplification. *Phonology* 25(2):271–319.

———. 2010. An Introduction to Harmonic Serialism. *Language and Linguistics Compass* 4(10):1001–1018. http://works.bepress.com/john_j_mccarthy/103.

———. 2011. Perceptually Grounded Faithfulness in Harmonic Serialism. *Linguistic Inquiry* 42(1):171–183.

McCarthy, John J., Wendell Kimper & Kevin Mullin. 2012. Reduplication in Harmonic Serialism. *Morphology* 22(2):173–232.

McCarthy, John J. & Alan Prince. 1986 [1996]. Prosodic Morphology. *Linguistics Department Faculty Publication Series* 13. http://scholarworks.umass.edu/linguist_faculty_pubs/13.

———. 1993a. Generalized Alignment. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1993*, 79–153. Kluwer.

———. 1993b. Prosodic Morphology I: Constraint Interaction and Satisfaction. *Linguistics Department Faculty Publication Series* 14 (2001 version). http://scholarworks.umass.edu/linguist_faculty_pubs/14.

———. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology. In Mercè Gonzàlez (ed.), *NELS 24: Proceedings of the North East Linguistic Society*, 333–379. Amherst, MA: University of Massachusetts Graduate Linguistic Student Association. http://works.bepress.com/john_j_mccarthy/43.

———. 1995. Faithfulness and Reduplicative Identity. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh Dickey (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers in Linguistics 18), 249–384. Amherst, MA: University of Massachusetts Graduate Linguistic Student Association. http://works.bepress.com/john_j_mccarthy/44.

———. 1999. Faithfulness and Identity in Prosodic Morphology. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The Prosody-Morphology Interface*, 218–309. Cambridge: Cambridge University Press. http://works.bepress.com/john_j_mccarthy/77.

McIntyre, Linda Lee. 1992. Classical Reduplication. University of North Carolina, Chapel Hill, PhD Dissertation.

Meid, Wolfgang. 1971. *Das germanische Praeteritum: Indogermanische Grundlagen und Ausbreitung im Germanischen*. Innsbruck: Institut für Sprachwissenschaftliche der Universität Innsbruck.

Melchert, H. Craig. 1983. A 'New' PIE *men Suffix. *Die Sprache* 29(1):1–26.

———. 1984. *Studies in Hittite Historical Phonology*. Göttingen: Vandenhöck & Ruprecht.

———. 1993. *Cuneiform Luvian Lexicon*. Chapel Hill, NC: H.C. Melchert.

———. 1994. *Anatolian Historical Phonology*. Amsterdam/Atlanta: Rodopi.

———. 2011. The PIE Verb for 'to pour'and Medial *h_3 in Anatolian. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 22nd Annual UCLA Indo-European Conference*, 127–132. Bremen: Hempen.

———. 2012. Hittite ḫi-Verbs of the Type *-āC₁i*, *-aC₁C₁anzi*. *Indogermanische Forschungen* 117:173–186.

———. 2015. Hittite Historical Phonology After 100 Years (and After 20 Years). Paper presented at Hrozný and Hittite: The First 100 Years, Prague.

———. 2016. Initial *sp- in Hittite and *šip(p)and-* 'to libate'. *Journal of Language Relationship* 14(3):187–205. http://rggu.ru/upload/main/vestnik/pmorv/vestnik_fvir3(14)_2016.pdf#page=41.

Mohanan, K.P. 1982. Lexical Phonology. MIT, PhD Dissertation.

Nelson, Nicole Alice. 2003. Asymmetric Anchoring. Rutgers, PhD Dissertation.

Nevins, Andrew & Bert Vaux. 2003. Metalinguistic, Shmetalinguistic: The Phonology of *shm*-reduplication. In J. Cihlar, A. Franklin, I. Kimbara & D. Kaiser (eds.), *CLS 39: Proceedings from the Annual Meeting of the Chicago Linguistic Society 39*, vol. 1, 702–721. Chicago: Chicago Linguistic Society.

Niepokuj, Mary. 1997. *The Development of Verbal Reduplication in Indo-European* (Journal of Indo-European Studies, Monograph Series 24). Washington, D.C.: Institute for the Study of Man.

Oettinger, Norbert. 1979. *Die Stammbildung des hethitischen Verbums*. Nürnberg: Hans Carl.

Padgett, Jaye. 2002. Russian Voicing Assimilation, Final Devoicing, and the Problem of [v]. Ms., University of California, Santa Cruz. http://roa.rutgers.edu/article/view/538.

Parker, Steve. 2002. Quantifying the Sonority Hierarchy. University of Massachusetts, Amherst, PhD Dissertation.

———. 2008. Sound Level Protrusions as Physical Correlates of Sonority. *Journal of Phonetics* 36(1):55–90.

Pater, Joe. 2000. Non-Uniformity in English Secondary Stress: The Role of Ranked and Lexically Specific Constraints. *Phonology* 17(2):237–274.

———. 2009. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Stephen Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*, 123–154. London: Equinox. https://rucore.libraries.rutgers.edu/rutgers-lib/41017/.

Pesetsky, David. 1979. Russian Morphology and Lexical Theory. Ms., MIT. http://web.mit.edu/linguistics/people/faculty/pesetsky/%0020russmorph.pdf.

Pozdniakov, Konstantin & Guillaume Segerer. 2007. Similar Place Avoidance: A Statistical Universal. *Linguistic Typology* 11(2):307–348.

Prentice, David John. 1971. *The Murut Languages of Sabah* (Pacific Linguistics, Series C 18). Canberra: Australian National University.

Prince, Alan. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry* 11(3):511–562.

———. 2002. Arguing Optimality. In Angela C. Carpenter, Andries W. Coetzee & Paul de Lacy (eds.), *Papers in Optimality Theory II*, vol. 1 (University of Massachusetts Occasional Papers in Linguistics 26), 269–304. Amherst, MA: University of Massachusetts Graduate Linguistic Student Association. ROA 562.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell Publishing.

Prince, Alan & Bruce Tesar. 2004. Learning Phonotactic Distributions. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, 245–291. Cambridge: Cambridge University Press.

Prokosch, Eduard. 1939. *A Comparative Germanic Grammar*. Philadelphia: LSA.

Raimy, Eric. 2000. *The Phonology and Morphology of Reduplication*. Berlin: Mouton de Gruyter.

———. 2009. Deriving Reduplicative Templates in a Modular Fashion. In Eric Raimy & Charles E. Cairns (eds.), *Contemporary Views on Architecture and Representations in Phonology* (Current Studies in Linguistics 48), 383–404. Cambridge, MA: MIT Press.

Reichardt, Konstantin. 1953. The Inscription on Helmet B of Negau. *Language* 29(3):306–316.

Rieken, Elisabeth. 2010. Das Zeichen <sà> im Hieroglyphen-Luwischen. In Aygül Süel (ed.), *Acts of the VIIth International Congress of Hittitology*, vol. 2, 651–660. Ankara: Anıt.

Riggle, Jason. 2006. Infixing Reduplication in Pima and its Theoretical Consequences. *Natural Language & Linguistic Theory* 24(3):857–891.

Ringe, Don. 2006. *From Proto-Indo-European to Proto-Germanic*. 1st edn. Oxford & New York: Oxford University Press.

Rix, Helmut. 1992. *Historische Grammatik des Griechischen: Laut- und Formenlehre*. 2nd edn. Darmstadt: Wissenschaftliche Buchgesellschaft.

Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp & Brigitte Schirmer. 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen*. 2nd edn. Wiesbaden: Ludwig Reichert.

Rose, Sharon. 2000. Rethinking Geminates, Long-Distance Geminates, and the OCP. *Linguistic Inquiry* 31(1):85–122.

Saba Kirchner, Jesse. 2010. Minimal Reduplication. University of California, Santa Cruz, PhD Dissertation.

———. 2013. Minimal Reduplication and Reduplicative Exponence. *Morphology* 23(2):227–243.

Sandell, Ryan. 2014a. Compensatory Lengthening in Vedic and the Outcomes of Proto-Indo-Iranian *[az] and *[až]. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 25th Annual UCLA Indo-European Conference*, 183–202. Bremen: Hempen.

———. 2014b. The Phonological Origins of Indo-European Long-Vowel ("Narten") Presents. Paper presented at the 26th Annual UCLA Indo-European Conference, Los Angeles. October 24–25, 2014. https://www.academia.edu/8858713.

———. 2015a. Obligatory Contour Principle Effects in Indo-European Phonology: Statistical Evidence and the Morphology-Phonology Interface. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 26th Annual UCLA Indo-European Conference*, 141–60. Bremen: Hempen. http://www.academia.edu/download/38741234/Sandell_UCLA26Proceedings.pdf.

———. 2015b. Productivity in Historical Linguistics: Computational Perspectives on Word-Formation in Ancient Greek and Sanskrit. UCLA, PhD Dissertation.

———. 2017. Allomorph Selection in Vedic Sanskrit Perfects of the Form $C_i e$:C_j. In Karen Jesney, Charlie O'Hara, Caitlin Smith & Rachel Walker (eds.), *Supplemental Proceedings of the 2016 Annual Meetings on Phonology*, Washington, DC: Linguistic Society of America. http://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/3990.

Sandell, Ryan & Andrew Miles Byrd. 2014. In Defense of Szemerényi's Law. Paper presented at the 33rd East Coast Indo-European Conference, Blacksburg, Virginia. June 6–8, 2014. https://www.academia.edu/8382487.

———. 2015. Extrametricality and Non-Local Compensatory Lengthening: The Case of Szemerényi's Law. Paper presented at the LSA Annual Meeting, Portland, Oregon. January 8–11, 2015. https://www.academia.edu/10119139.

Sandell, Ryan & Sam Zukoff. 2017. The Development of the Germanic Preterite System: Learnability and the Modeling of Diachronic Morphophonological Change. Paper presented at the LSA Annual Meeting, Austin, TX. January 5–8, 2017. http://web.mit.edu/szukoff/www/pdfs/sandell_zukoff_LSA2017.pdf.

de Saussure, Ferdinand. 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipzig: Teubner.

Schumacher, Stefan. 2005. 'Langvokalische Perfekta' in indogermanischen Einzelsprachen und ihr grundsprachlicher Hintergrund. In Gerhard Meiser & Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17–23. September 2000, Halle an der Salle*, 591–626. Wiesbaden: Reichert.

Schwyzer, Eduard. 1939. *Griechische Grammatik*. München: C.H. Beck.

Sherer, Tim D. 1994. Prosodic Phonotactics. University of Massachusetts, Amherst, PhD Dissertation.

Sihler, Andrew L. 1995. *New Comparative Grammar of Greek and Latin*. New York: Oxford University Press.

Simon, Zsolt. 2013. Once Again on the Hieroglyphic Luwian Sign *19 <á>. *Indogermanische Forschungen* 118:1–22.

Smyth, Herbert Weir. 1920 [1984]. *Greek Grammar*. Harvard University Press.

Sommerstein, Alan H. 1973. *The Sound Pattern of Ancient Greek*. Basil: Blackwell.

Spaelti, Philip. 1997. Dimensions of Variation in Multi-Pattern Reduplication. University of California, Santa Cruz, PhD Dissertation.

Stanton, Juliet & Sam Zukoff. 2016a. Prosodic Effects of Segmental Correspondence. In Ksenia Ershova, Joshua Falk & Jeffrey Geiger (eds.), *CLS 51: Proceedings of the Fifty-first Annual Meeting of the Chicago Linguistic Society*, 501–515. Chicago: Chicago Linguistic Society.

———. 2016b. Prosodic Identity in Copy Epenthesis and Reduplication: Towards a Unified Model of Transitive Correspondence. Ms., MIT. http://web.mit.edu/szukoff/www/pdfs/stantonzukoff_manuscript.pdf.

———. to appear. Prosodic Identity in Copy Epenthesis: Evidence for a Correspondence-Based Approach. *Natural Language & Linguistic Theory.* [Pre-publication version available at https://ling.auf.net/lingbuzz/003522].

Steriade, Donca. 1982. Greek Prosodies and the Nature of Syllabification. MIT, PhD Dissertation.

———. 1988. Reduplication and Syllable Transfer in Sanskrit and Elsewhere. *Phonology* 5(1):73–155.

———. 1994. Positional Neutralization and the Expression of Contrast. Ms., UCLA. http://lingphil.mit.edu/papers/steriade/contrastive-gesture.pdf.

———. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. Ms., UCLA.

———. 1999. Alternatives to the Syllabic Interpretation of Consonantal Phonotactics. In O. Fujimura, B.D. Joseph & B. Palek (eds.), *Proceedings of Linguistics and Phonetics 1998*, 205–245. Prague: Karolinum Press.

———. 2009. The Phonology of Perceptibility Effects: The P-Map and its Consequences for Constraint Organization. In Kristin Hanson & Sharon Inkelas (eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, 151–179. Cambridge, MA: MIT Press.

———. 2015. 24.962 Advanced Phonology lecture notes, MIT, January 30, 2015.

Stevens, Kenneth N. 2002. Toward a Model for Lexical Access Based on Acoustic Landmarks and Distinctive Features. *The Journal of the Acoustical Society of America* 111(4):1872–1891. http://asa.scitation.org/doi/abs/10.1121/1.1458026.

Struijke, Carolina Maria. 2000. Reduplication, Feature Displacement, and Existential Faithfulness. University of Maryland, PhD Dissertation.

Suzuki, Seiichi. 1994. Accounting for Attic Reduplication: A Synthesis. *Journal of Indo-European Studies* 22(3&4):399–415.

Tesar, Bruce. 1995. Computational Optimality Theory. University of Colorado, Boulder, PhD Dissertation.

Tesar, Bruce & Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29(2):229–268.

———. 2000. *Learnability in Optimality Theory.* Cambridge, MA: MIT Press.

Thurneysen, Rudolf. 1946 [1980]. *A Grammar of Old Irish.* Dublin: Dublin Institue for Advanced Studies.

Varma, Siddheshwar. 1929. *Critical Studies in the Phonetic Observations of Indian Grammarians.* Delhi: Munshi Ram Manohar Lal. Reprinted 1961.

Vaux, Bert. 1992. Gemination and Syllabic Integrity in Sanskrit. *Journal of Indo-European Studies* 20(3–4):283–303.

Vine, Brent. 1998. The Etymology of Greek κώμη and Related Problems. In Jay Jasanoff, H. Craig Melchert & Lisi Olivier (eds.), *Mír Curad: Studies in Honor of Calvert Watkins*, 685–702. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

———. 2011. On Dissimilatory *r*-Loss in Greek. In Thomas Krisch & Thomas Lindner (eds.), *Indogermanistik und Linguistik im Dialog*, 1–17. Wiesbaden: Reichert.

Wackernagel, Jacob. 1896. *Altindische Grammatik*, vol. I (Lautlehre). Göttingen: Vandenhoeck and Ruprecht.

Walter, Mary Ann. 2007. Repetition Avoidance in Human Language. MIT, PhD Dissertation.

Weeden, Mark. 2011. Spelling, Phonology and Etymology in Hittite Historical Linguistics. *Bulletin of the School of Oriental and African Studies* 74(1):59–76.

Weiss, Michael. 2009. *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave Press.

Whitney, William Dwight. 1885 [1988]. *The Roots, Verb-Forms, and Primary Derivatives of the Sanskrit Language: A Supplement to his Sanskrit Grammar*. New Haven: American Oriental Society. Reprinted 1945 as American Oriental Series 30.

———. 1889. *Sanskrit Grammar*. 2nd edn. Cambridge, MA: Harvard University Press.

Winter, Werner. 1950. On the Origin of the Saṃprasaraṇa Reduplication in Sanskrit. *Language* 26(3):365–370.

Wolf, Matthew Adam. 2008. Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model. University of Massachusetts, Amherst, PhD Dissertation.

Wright, Richard. 2004. A Review of Perceptual Cues and Cue Robustness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, 34–57. Cambridge: Cambridge University Press.

Yates, Anthony D. 2014. Acrostatic Neuter **s*-Stems and Sonority-Driven Epenthesis in Proto-Anatolian. Paper presented at the 33rd East Coast Indo-European Conference, Blacksburg, Virginia. June 6–8, 2014. https://www.academia.edu/7219133.

———. 2016. Left But Not Leftmost? Interactions between Epenthesis and Ictus Assignment in Anatolian. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 26th Annual UCLA Indo-European Conference*, 161–178. Bremen: Hempen. https://www.academia.edu/15608531.

———. 2017. Lexical Accent and its Diachrony: Studies in Hittite and Cupeño Prosody. UCLA, PhD Dissertation.

Yates, Anthony D. & Sam Zukoff. 2016a. The Phonology of Anatolian Reduplication. Paper presented at the 35th East Coast Indo-European Conference, Athens, GA. June 6–8, 2016. http://web.mit.edu/szukoff/www/pdfs/ECIEC2016.pdf.

———. 2016b. Variation and Change in Anatolian Reduplication. Paper presented at the 28th Annual UCLA Indo-European Conference, UCLA. November 11–12, 2016. http://web.mit.edu/szukoff/www/pdfs/yateszukoff_WECIEC2016.pdf.

Yu, Alan C. L. 2005. Toward a Typology of Compensatory Reduplication. In John Alderete, Chung-hye Han & Alexei Kochetov (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 397–405. Somerville, MA: Cascadilla Proceedings Project.

———. 2007. *A Natural History of Infixation*. Oxford: Oxford University Press.

Yun, Suyeon. 2016. A Theory of Consonant Cluster Perception and Vowel Epenthesis. MIT, PhD Dissertation.

Zec, Draga. 1995. Sonority Constraints on Syllable Structure. *Phonology* 12(1):85–129.

Zukoff, Sam. 2014. On the Origins of Attic Reduplication. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 25th Annual UCLA Indo-European Conference*, 257–278. Bremen: Hempen.

———. 2015. Poorly-Cued Repetition Avoidance in Indo-European Reduplication. Paper presented at the LSA Annual Meeting, Portland, Oregon. January 8–11, 2015. http://web.mit.edu/szukoff/www/pdfs/LSA_2015.pdf.

———. 2017a. Arabic Nonconcatenative Morphology and the Syntax-Phonology Interface. In *Proceedings of the 47th Annual Meeting of the North East Linguistic Society*, Amherst, MA: University of Massachusetts Graduate Linguistic Student Association. http://web.mit.edu/szukoff/www/pdfs/NELSpaper2017.pdf.

———. 2017b. The Mirror Alignment Principle: Morpheme Ordering at the Morphosyntax-Phonology Interface. In Snejana Iovtcheva & Benjamin Storme (eds.), *Papers on Morphology:*

MIT Working Papers in Linguistics 81, 105–124. Cambridge, MA: MITWPL. http://web.mit.edu/szukoff/www/pdfs/MITWPL.pdf.

———. 2017c. The Reduplicative System of Ancient Greek and a New Analysis of Attic Reduplication. *Linguistic Inquiry* 48(3):459–497.

Zukoff, Sam & Ryan Sandell. 2015. The Phonology of Morpheme Realization in the Germanic Strong Preterites. In Thuy Bui & Deniz Özyıldız (eds.), *Proceedings of the 45th Annual Meeting of the North East Linguistic Society*, vol. 3, 39–49. Amherst, MA: University of Massachusetts Graduate Linguistic Student Association.

Zukoff, Sam & Benjamin Storme. 2017. Reduplication and the Perception of Consonants under Repetition. Poster presented at Acoustics '17 Boston, Boston, MA. June 25-29, 2017. http://web.mit.edu/szukoff/www/pdfs/Acoustics17Poster.pdf.

Zuraw, Kie. 2000. Patterned Exceptions in Phonology. UCLA, PhD Dissertation.

———. 2002. Aggressive Reduplication. *Phonology* 19(3):395–439.