UNIVERSITY OF CALIFORNIA

SANTA CRUZ

OPACITY AND SOUND CHANGE IN THE POLISH LEXICON

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Robert Nathaniel Sanders

June 2003

The dissertation of Robert Nathaniel Sanders
is approved:

____
Professor Jaye Padgett, Chair

____
Professor Junko Ito

____
Professor Armin Mester

____
Caro Struijke, PhD

____
Frank Talamantes
Vice Provost and Dean of Graduate Studies

Table of Contents

OPACITY AND SOUND CHANGE IN THE POLISH LEXICON

Abstract

OPACITY AND SOUND CHANGE IN THE POLISH LEXICON

Robert Nathaniel Sanders

The main goal of this dissertation is to provide a generative account of phonological opacity within a framework built upon direct mapping in the synchronic grammar without abstract intermediate representations, as in standard Optimality Theory (OT; Prince and Smolensky 1993/2002). Such a framework predicts that certain types of opacity cannot be synchronically productive. I take this prediction seriously and develop an analysis in which opacity is shown to arise from the interaction of sound change and a strong version of Prince and Smolensky's principle of lexicon optimization, in which the underlying lexicon is 'optimized' by becoming more faithful to the surface pronunciation. This interaction results in a progressive encoding of sound changes directly into the evolving lexicon, mirroring the stepwise effect of multistratal derivations, but diachronically rather than synchronically, preserving direct mapping.

The specific theoretical framework used in this dissertation is *Faithfulness, Dispersion, and Markedness in OT* (FDM-OT), which differs from standard OT by offering a functional account of sound change and synchronic phonology through the interaction of faithfulness (along the lines of McCarthy and Prince 1995), dispersion (generalized from Dispersion Theory (Flemming 1995, Padgett 1997, and Ní Chiosáin and Padgett 2001)), and universally ranked articulatory markedness constraints. Grounded in cognition, acoustics, and articulation, FDM-OT explains and predicts phonological patterns with fewer arbitrary or abstract stipulations than are required by competing theories.

I analyze three well-known instances of opacity from Polish, the language of focus. Additionally, I provide analyses of opacity in refined Low German, Turkish, and Tuyuca, arguing that all cases of opacity fit into the following typology: (i) *synchronically unproductive opacity*, which fails to apply to nonce forms and to lexical exceptions but is still pervasive in the lexicon due to lexicon optimization; (ii) *morphologically conditioned opacity*, which may be synchronically productive, but only at particular morphological boundaries because the relevant affixes have allomorphs created by lexicon optimization that are encoded with historically opaque alternations; and (iii) *transparent 'opacity'*, which can be reanalyzed transparently because original opaque analyses lacked sufficient phonetic detail or access to certain recent theoretical advances, such as FDM-OT's dispersion constraints.

Acknowledgements

This dissertation did not miraculously leap from my head unaided (unfortunately), so I must indulge in a bit of name-dropping to give proper recognition to the many people who have had an influence on my work.

No dissertation could exist without an advisor, and I was supremely fortunate to have Jaye Padgett as mine. He supported and encouraged me not only to succeed, but also to go beyond my goals, bettering myself by exploring ideas I would not have considered otherwise. His keen insight sliced through the fluff in my work and got to the heart of the issues at hand. I'll be forever grateful for his guiding hand. It's most definitely Miller time!

The other members of my dissertation committee have had a constant and important impact on my work, both in the present work and in my previous research. Junko Ito's no-nonsense attitude and tough questions aided me by giving my work much needed depth. More importantly, she taught me to find my own weaknesses before anyone else did. Armin Mester's intelligence and ingenuity strengthened my work in the other dimension, showing me boundaries that needed pushing, and ultimately, giving me the vision to see my work in a broader linguistic universe. Caro Struijke blended skepticism and optimism to help me question my assumptions and analysis without losing hope that it would all work out in the end. Despite being only a one year visitor, she was eager to join my committee, and I am thankful for her energy and patience.

The UCSC linguistics faculty as a whole and as individuals have all contributed to my professional and personal growth. Geoff Pullum has been my teacher, my mentor, my colleague, and my friend. When I needed a kick in the pants,

he was ready and waiting, eagerly lacing up his boot. I have been indescribably lucky to know Geoff and to have had access to his amazingly diverse experience and knowledge, his often unappreciated humor, and his uncanny ability to follow formalisms to their bitter (and sometimes unfortunate) end. Sandy Chung deserves mention for her refreshing mix of outrageous joviality and sober concern for facts. She taught me to care about linguistics in new ways and to ensure that my work is important to me and as fun as possible. I would also like to thank Judith Aissen, Donka Farkas, Jorge Hankamer, Bill Ladusaw, and Jim McCloskey for the wealth of information I gained from their classes and for the superb linguistics community they have created at UCSC. And of course, our department and my dissertation would be a shambles without Cher Bergeon, Connie Creel, Ashley Hardissy, Sylvia Zito, and especially Tanya Honig, superb people who tirelessly kept the right wheels turning behind the scenes, enabling all of us to flourish.

All of my other colleagues at UCSC have contributed to this work in some way, be it through direct commentary on a draft or just by giving me an emotional boost with a friendly greeting in the coffeeshop. Suzanne Lyon in particular was a true friend, always supportive in every way, even from hundreds of miles away. In Kazu Kurisu, I had a mock adversary whose unending skepticism pushed me harder to prove my theories to him and to myself. Outi Bat-El, Dylan Herrick, Jason Riggle, Jen Smith, Anne Sturgeon, Adam Ussishkin, and Andy Wedel have all given me much needed feedback at various stages, and my other fellow phonologists from various eras of my UCSC career, such as Dan Karvonen, Motoko Katayama, Bryn LaFollette, Vera Lee-Schoenfeld, Philip Spaelti, Emily Urban, and Rachel Walker, all contributed in various ways to my growth as a linguist and enriched my overall

research. The many other Cave-dwelling linguists that have graced the halls of Stevenson have also influenced me by being outstanding examples of dedicated researchers. In particular, I would like to thank Teal Bissell, Ryan Bush, James Darrow, Naomi Devlin, Christine Gunlogson, Rodrigo Gutierrez, Chris Kennedy, Rachel Lagunoff, Jason Merchant, Line Mikkelsen, Shira Nobles, and Chris Potts for their friendship and scholarly inspiration.

My endless appreciation goes to my Polish consultants, Anna Bandaruk, Michael Jasinski, Karolina Novicka, and Basia Torres, for their invaluable help with data judgments and for their participation in my experiments. Dziękuję bardzo! Numerous non-linguists, friends near and far, boardgame buddies, co-workers, and housemates, have supported me in a variety of ways. Many thanks to Jason, Ranjana, Elena, Brenda, Stu, Joe, Cheri, Maryanne, Liz, Chris, Tom, and Kathleen. Your friendship and support have helped keep me working towards my goals.

Finally, I must thank the very special people who have changed me in ways so profound that I can never fully quantify the full extent of their importance in my life. My best friend in every possible lifetime, Jennifer, has guided me for over fifteen years and will certainly continue to do so well into the future. My sister, Allyson, despite our frighteningly extensive similarities, has managed to do the things I never dared to do, always showing me the multitude of options that are open to me. My parents, Nathalie and George, have selflessly given me so much, whether I asked for it or not. I can never adequately repay the debt, but I hope that my love is a sufficient start. And finally, to Adam: I would be quite miserable without you, and my work would be diminished without your influence. I love you immensely!

My heartfelt thanks to each of you for everything you have given me!

Transcriptions, Abbreviations, and Symbols

All transcriptions are generally given in the 1996 version of the International Phonetic Alphabet (IPA). Significant deviations from the IPA are explained. Where it is not possible to create an accurate IPA symbol for a particular sound (because the sound is hypothetical, from an unclear source, etc.), I have chosen a symbol satisfying some mix of simplicity, convention, and apparent analytical accuracy, with appropriate discussion. The following is a list of abbreviations and special symbols used at various points in this dissertation:

grammatical

ACC	accusative
DAT	dative
FEM	feminine
GEN	genitive
IMP	imperative
INS	instrumental
LOC	locative
NEUT	neuter
NOM	nominative
NON-VIR	non-virile (FEM, NEUT, or non-human)
PAST	past
PL	plural
POSS	possessive
PRES	present
PROG	progressive
SG	singular
VIR	virile (MASC and human)
1	1st person
2	2nd person
3	3rd person

historical

PIE	Proto-Indo-European
PSl	Proto-Slavic (Common Slavic)
PWS	pre-West Slavic
*	reconstructed/unattested form
>	became through sound change
<	comes from historically

analytical

$\mathcal{D}$	dispersion (constraint)
$\mathcal{F}$	faithfulness (constraint)
$\mathcal{M}$	markedness (constraint)
□	the set of all possible words
/…/	underlying representation listed in the lexicon; also, input to the grammar
[…]	surface/output form
□…□	orthographic form
□	synchronic input-output mapping
✘	ungrammatical form; also, constraint violation in a tableau
✘^{n}	*n* constraint violations
✘!	fatal constraint violation
✓	grammatical candidate selected as optimal by current formulation of the grammar
☹	grammatical candidate not selected as optimal by current formulation of the grammar
☠	ungrammatical candidate selected as optimal by current formulation of the grammar

Chapter 1: Essentials

While this dissertation is aimed at generative phonologists, it is also designed to be accessible to non-generative Slavicists and historical linguists. Thus, it is important to first present crucial topics that may be unfamiliar to some readers. In §1.1, I give an overview of Optimality Theory (OT; Prince and Smolensky 1993/2002), a prominent framework in modern generative phonology and the foundation upon which the analyses in this dissertation are built. In §1.2, I explore the relationship between OT and *opacity*, the phenomenon in which a phonological generalization appears to be rendered false through interaction with other generalizations, yet is still analyzed as being true. Opacity presents serious problems for certain versions of OT (including the monostratal version adopted in this dissertation) and has therefore been at the center of a significant amount of recent work in OT. These two sections may be skipped by readers who are familiar with OT and opacity.

In §1.3 (crucial reading for anyone attempting to understand the remaining chapters), I define the analytical framework used in this dissertation: *Faithfulness, Dispersion, and Markedness in OT* (FDM-OT). Though based upon OT, FDM-OT differs significantly in a few respects. The most obvious point of departure can be readily seen in the inputs and candidate set. In standard OT, these are comprised solely of individual words. However, I follow recent work within Dispersion Theory (Flemming 1995) by adopting a relational version of OT, in which inputs and candidates are *sets* of words. Such a theory allows for morphologically unrelated words to influence each other's phonological shape through the overall system of the phonetic contrasts they exemplify.

§1.4 explains how sound change in the FDM-OT framework operates, beginning with the evolution of the early grammar for one generation of speakers, continuing through sound change to create the late grammar for that generation, and cycling back to the early grammar for the following generation. In addition, I show how opacity can be analyzed in FDM-OT as the interaction of sound change (represented in FDM-OT as constraint reranking) with Prince and Smolensky's (1993/2002) hypothesis of lexicon optimization, through which underlying representations in the lexicon are selected from all possible inputs in order to maximize faithfulness to the pronounced output forms. This interaction effectively encodes opaque generalizations directly into the evolving lexicon, allowing it to flourish in the lexicon without necessarily being synchronically productive. Indeed, a prediction of this account is that true opacity *cannot* be synchronically productive (for monomorphemic words; on polymorphemic words, see Chapter 4).

Since a central claim of this dissertation is that some instances of supposed synchronic opacity are in fact due to lexical storage of sound changes, it is important to lay out the history of the language of study. In §1.5, I give a tour of the history of Polish, beginning with Proto-Slavic, the reconstructed ancestral language of the Slavic languages, and ending with Modern Polish. All readers should at least skim this section to be aware of the terminology and transcriptions used in this dissertation, as the usage here often differs from the traditional literature on Slavic (which itself is not entirely consistent). Generative phonologists and historical linguists unfamiliar with the history of Polish are encouraged to read this section thoroughly, since the various historical stages of Polish play a crucial role in the analyses put forth in the remaining chapters.

1.1 Optimality Theory

In this section, I explain OT, highlighting the components which are important to the analyses presented in this dissertation.

1.1.1 Summary

OT is an analytical framework in which the pronounced form of an utterance (the *output*) is selected from multiple possible candidates that are all simultaneously evaluated and directly compared to the lexical representation (the *input*) by a ranked set of universal and violable constraints (Prince and Smolensky 1993/2002 and McCarthy and Prince 1993 are the best original sources for OT; Kager 1999a is an excellent reference which summarizes the standard theory).

The primary components of standard OT are briefly described in (1):

(1)
a. *Gen*: a function which produces the potentially infinite set of candidates for a given input
b. *Con*: the set of universal constraints
c. *markedness*: a subset of Con which enforces various well-formedness conditions on output candidates
d. *faithfulness*: a subset of Con which enforces various types of identity between the input and output candidates
e. *constraint hierarchy*: a language-particular ranking of Con; a strict, total[1] order (irreflexive, asymmetric, and transitive relation) on Con
f. *Eval*: the function which selects as the grammatical output the single most harmonic candidate given a particular input, a set of candidates, and a constraint hierarchy[2]

---

[1] This is the standard take on the constraint hierarchy. However, it does not seem crucial to prevent two different rankings from producing the same results, in which case, a strict ranking is not necessary.

[2] Because of a lack of adequate terminology, I also use Eval to refer to the function that results from restricting Eval to a specific constraint hierarchy.

The selection of a particular output in OT is graphically illustrated by means of a special table called a *tableau*. The following sample (and over-simplified) tableau for the Polish word ⟨przykład⟩ [pʂɪkwat] 'example' illustrates the layout of a typical OT tableau, along with the specific notations used in this dissertation:

(2)

/pʐɪkwad/	NO WORD-FINAL VCD. OBSTR.	OBSTR. CLUSTERS AGREE IN VOICING	FAITH [voice]
✓ a. pʂɪkwat			×^{2}
b. pʐɪkwat		×!	×
c. pʂɪkwad	×!		×
d. pʐɪkwad	×!	×!	

The input /pʐɪkwad/ occupies the upper left corner of the tableau. The set of candidates produced by Gen are listed down the first column, below the input, with the grammatical, or *optimal*, output indicated with a checkmark (✓). The constraint hierarchy is listed across the top row, to the right of the input, ordered from left to right, beginning with the highest ranked constraints.

The violations of a constraint by a candidate are listed in the cell at the intersection of the constraint's column and the candidate's row. An empty cell indicates *satisfaction* of the constraint, while *violations* are marked by × (with a superscript number to indicate more than one violation). A single solid vertical line is used to indicate a *crucial* ranking between two constraints. If a crucial ranking were reversed, the wrong candidate would win. In the sample tableau, the solid vertical line indicates that the constraint OBSTRUENT CLUSTERS AGREE IN VOICING crucially outranks FAITH-[voice]. Violation of a highly ranked constraint is often *fatal*, meaning that the candidate cannot be the optimal output under this constraint ranking; this is symbolized by putting an exclamation mark after the fatal violation marks. For example, candidate (2b) fatally violates the constraint OBSTRUENT CLUSTERS AGREE IN VOICING because it has the obstruent cluster [pz̩] which does not agree in voicing. Candidate (2c) fatally violates the constraint NO WORD-FINAL VOICED OBSTRUENTS because it has the voiced obstruent [d] at the end of the word. The winning candidate (2a) violates neither constraint, though it does violate FAITH-[voice] more than either candidate (2b) or (2c) because it differs from the input in voicing specification for both /z̩/ and /d/, while each of candidate (2b) and (2c) only differ from the input in the voicing specification of one of /z̩/ and /d/. Because the markedness constraints are crucially ranked higher than the faithfulness constraints, candidates (2b) and (2c) cannot emerge as the grammatical output.[3]

A dotted line between two constraints indicates that the output is the same regardless of the ranking of these constraints. In this example, a strict ranking between the markedness constraints is not crucial, since (2a) violates neither and thus cannot incur a fatal violation. Because the ranking is not crucial, we cannot determine which constraint candidate (2d) fatally violates, so an exclamation mark is put in both cells. Note that candidate (2d) loses to (2a) despite completely satisfying FAITH-[voice]. This is a result of the constraint ranking being a strict order: the violations of a lower constraint cannot compensate for the violations of a higher constraint. Grey shading is used to indicate that violations of a lower constraint cannot overcome a fatal violation of a higher ranked constraint. In this case, no

[3] The choice of input here is illustrative. Other inputs are possible, and in fact, the theory requires all plausible inputs to be entertained (see §1.1.4).

amount of violations of FAITH-[voice] can alter the outcome with the markedness constraints highly ranked, so the cells under FAITH-[voice] are all shaded.

In the following sections, I explore some specific fundamental issues in OT and discuss why they are important to generative phonology.

1.1.2 Constraints versus rewrite rules

Early generative phonology (as exemplified by Halle 1959, Chomsky and Halle 1968, and their derivatives) expressed phonological generalizations through rewrite rules, in which an input string *A* is changed to the output string *B* in some environment *C* ___ *D*. This can be characterized by the following schema:

(3) *A* → *B* / *C* ___ *D* ‘all instances of *A* between *C* and *D* are changed to *B* (i.e., all instances of *CAD* are changed to *CBD*), where *A*, *B*, *C*, and *D* are arbitrary strings of segments and morphological or prosodic boundaries’

Generally, theories that utilize such rewrite rules are too powerful, able to describe vast numbers of impossible languages just as easily as attested languages. To combat this problem, rule-based theories must place limits on the types of possible rules and rule interactions that are available to languages. However, such limits are often seemingly arbitrary in the context of the theory itself.

Consider the following simplified rule for nasal place assimilation, a common cross-linguistic phenomenon (see Mohanan 1993 for further discussion):

(4) [C,+nasal] → [αPLACE] / ___ [αPLACE]

This rule explains the following pattern of the pronunciation of the nasal consonant in the Latinate English prefixes ⟨in/im-⟩ and ⟨en/em-⟩ in which the nasal has the same place of articulation as the following consonant:[4]

(5)	*e*[mb]*ed*	bilabial
	i[ɱf]*ormal*	labiodental
	e[n̪θ]*use*	dental
	i[nt]*angible*	alveolar
	i[n̠d̠ʒ]*ustice*	post-alveolar
	i[ɲç]*uman*	palatal
	i[ŋg]*rain*	velar

There is no theory-internal reason why a rule like (4) is commonly found in many unrelated languages, but the subtly different rule in (6), with '–nasal' instead of '+nasal', is rarely (if ever) found in a language without (4):

(6) $[\text{C},-\text{nasal}] \rightarrow [\alpha\text{PLACE}]\ /\ \_\_\_\ [\alpha\text{PLACE}]$

The rule in (6) would ensure that all oral consonants assimilate to the place of articulation of any following consonant, but it says nothing about assimilation in nasals. Indeed, without (4), nasal place assimilation would not occur. A language with (6) but not (4) would have all consonant clusters agreeing in place unless the first consonant was a nasal. At first glance (i.e. ignoring cross-linguistic data, acoustics, and physiology), this may not seem like an entirely bad possible language. However, (4) is grounded in the physical world, while (6) is not. Ohala and Ohala (1993) argue that since the acoustic cues for place of articulation are weaker in nasals

[4] Some of these instances of assimilation (especially labiodental and palatal) only occur in faster styles of speech, but I am not concerned here with such variation.

than in oral consonants, nasals are more susceptible to articulatory reanalysis (because a sequence of two distinct places of articulation is harder to perform than a single place of articulation held over two segments). This argument would rule out (6) as a natural rule separate from (4), since (6) contrarily singles out oral consonants as undergoing place assimilation.

Nothing intrinsic to rewrite rules themselves could exclude (6) and accept (4) in any principled way. But the existence of (6) side by side with (4) as an equally possible rule would generate unattested and unexpected languages that would violate acoustic and articulatory facts. If the goal of a phonological theory is not only to describe phonological patterns, but also to *explain* and *predict* them, then (6) must be disallowed. Banning (6) outright solely because it is unattested is not an interesting or useful solution. For example, using this method of banning rules based on hindsight, we could not tell in advance which rules should be prevented from existing. In addition, there is a problem with circularity in reasoning: we create rules to explain what can and cannot exist in language, but rule (6) would be ruled out *because* it produces languages which do not exist.

The solution to this deficiency in rule-based theories lies in constraints on rules which contain information regarding universal principles that can be used to determine what counts as a valid rule and what does not. Thus, there would have to be a constraint which bans rules like (6) that yield assimilation in oral consonants preferentially over nasals. Since such a constraint can be derived from real world principles, it seems that the problem of over-generation faced by a rule-based theory can be put to rest: phonological theory should contain both (i) a set of theoretically possible rewrite rules and (ii) a set of grounded constraints that determine which subset of the theoretically possible rules are linguistically possible. However, new problems arise when rules and constraints are mixed in the same theory, and we must ask: if rules require constraints, do constraints require rules?

1.1.3 Conspiracies, universality, and violability

One of the properties of the interaction between rules and constraints is that several different rules can be shaped by the same constraint. This phenomenon is often called a *conspiracy* (Kisseberth 1970). One example of a conspiracy is the affinity between syllable weight and stress. Some languages, such as Chamorro (Chung 1983), Mohawk (Postal 1968), and Slovene (Šuštaršič, Komar, and Petek 1999), have predictably long vowels in certain stressed syllables (and in some of these languages, long vowels can *only* occur in stressed syllables). For example, in Mohawk, all vowels are short, except in open stressed syllables, where they must be long, as in [raˈgeːdas] 'scrape (3SG PRES)' (cf. [waˈhaːgedeʔ] 'scrape (3SG PAST)', with a short unstressed vowel in the same root [-ged-]). As seen in [roˈjoʔdeʔ] 'work (3SG PRES)', long vowels do not appear in closed stressed syllables. The following rule is one way to characterize the lengthening process in Mohawk:

(7) [V,+stress] → [+long] / ___]$_{\sigma}$

In other languages, such as Arabic (Kaye 1990, Thelwall and Sa'adeddin 1999), Korean (Lee 1999), and Yapese (Jensen 1977), vowel length is generally lexically contrastive, unpredictably appearing in essentially any syllable, with stress appearing on certain heavy syllables (the rightmost, the leftmost, etc.) if possible. For Korean, in which the initial syllable is stressed if it is heavy (otherwise, the second syllable is stressed), a rule like the following could be used:

(8) #□□□ □ [+stress]

Both of these rules aim for the same general goal: ensure that stress and syllable weight coincide. Thus, the multiple different rules that can be found in languages like these have their roots in a single *universal* constraint, such as STRESS□WEIGHT (S□W).[5] To explain cross-linguistic tendencies as different instances of the same universal grammar, a theoretical framework with fewer language-specific formalisms is preferable to one with many. This is the approach OT takes, by completely eliminating the need for language-specific rules, relying on universal constraints instead.

However, the universality of constraints is not sufficient to account for variation within languages. A naïve constraint-based theory would treat universal constraints as binary parameters, with each language's phonology consisting of a set of active constraints, while the remaining universal constraints are 'turned off'. This type of theory cannot achieve a satisfactory analysis of the conspiracy discussed above. Though S□W may be active in each of the languages mentioned, it can still be violated within one of the languages. For example, S□W is generally satisfied in Chamorro by lengthening stressed vowels in open syllables, as in [ɑˈliːtʊs] 'earrings'. However, long vowels may only appear in penultimate syllables, so stress on a different syllable will not trigger lengthening, and thus will violate S□W, as seen in the word [ˈigɑdu] 'liver'. This exemplifies the insight of OT that constraints like S□W can be both universal (by existing in Con, and thus being accessible to every

[5] As Caro Struijke (personal communication) points out, work in OT has used the two constraints STRESS-TO-WEIGHT and WEIGHT-TO-STRESS to capture the universal tendency expressed by S□W. With other constraints in the system, it is not completely clear that both are needed, but the main point concerning conspiracies is still the same. See Kager 1999a for another example of a conspiracy.

language) *and* violable (by being ranked lower than other constraints in a language-particular constraint hierarchy).

1.1.4 Richness of the base

The constraints in OT govern outputs or the relationship between inputs and outputs; they do not affect inputs. This is made explicit in the hypothesis of *richness of the base* (Prince and Smolensky 1993/2002), which states that inputs are unrestricted. That is, Eval should be able to produce a grammatical output regardless of what the input looks like. Richness of the base is related to conspiracies, in which different processes operating on different inputs can still lead to the same output. I adopt a version of richness of the base in FDM-OT. Specifically, I assume that the set which consists of every possible word (called □ throughout this dissertation) is evaluated during language acquisition by the evolving grammar so that when the constraint hierarchy is stabilized, Eval can take □ as the input and map it to the subset of □ which best matches the perceived adult outputs. This is further explicated in §1.3.

1.1.5 Lexicon optimization

While richness of the base maintains that inputs are unrestricted, there is a mechanism in OT which can select a particular underlying representation from the pool of possible inputs to be used as the sole input for a given output. This mechanism is known as *lexicon optimization* (named in Prince and Smolensky 1993/2002; see Kiparsky 1968 for a prescient version of the same idea). The following is a reformulation of the definition given by Prince and Smolensky (1993/2002):

(9) *lexicon optimization*

Let O be an output and let I □ □ such that I_k □ I if and only if Eval(I_k) = O. Then, the underlying representation for O will be the input in I which is most faithful to O, as determined by the ranking of the faithfulness constraints in the constraint hierarchy.

That is, for a given output O and every input that can be mapped to O, whichever input is most faithful to O is selected to be the underlying representation of O. Consider the English word [k^{h}owt] ‘coat’, which shows obligatory diphthongization of a tense vowel and obligatory lack of aspiration on a word-final voiceless stop (as well as many other processes not relevant here). There are numerous possible inputs that will result in this output. Among them are:

(10)	/k^{h}owt/	*identical to the output*
	/k^{h}ot/	*requires predictable tense vowel diphthongization to apply*
	/k^{h}owt^h/	*requires predictable final deaspiration to apply*
	/k^{h}ot^h/	*requires both diphthongization and deaspiration*

Richness of the base allows all of these as possible inputs, and Eval will ensure that all of them map to [k^{h}owt]. When it comes time to store this word in the lexicon, lexicon optimization will select /k^{h}owt/ as the underlying representation for [k^{h}owt], because out of all of the possible inputs, it is the most faithful to the output.

As Prince and Smolensky discuss, lexicon optimization is significantly complicated when words consisting of more than one morpheme are taken into consideration. The primary complication stems from *lexical minimization*, a fundamental (often implicitly assumed) hypothesis in most work in generative phonology, originating in Chomsky and Halle 1968, reformulated here:

(11) *lexical minimization*

Every morpheme has exactly one underlying representation which can be used to derive all of its surface allomorphs.

The conflict between lexicon optimization and lexical minimization can be seen with the English word [kʰowɾəd] 'coated', the past tense of [kʰowt]. As shown above, the underlying representation for [kʰowt] is /kʰowt/. By a similar computation, it would seem that the underlying representation of [kʰowɾəd] would be the analogous /kʰowɾəd/. Separating out the relevant morphemes, we see that there are two underlying strings with the same meaning of 'coat': /kʰowt/ (used in the present tense) and /kʰowɾ-/ (used in the past tense, suffixed with /-əd/). This is a clear violation of lexical minimization, since these underlying representations are not phonologically the same (they differ in the final consonant). In order to satisfy lexical minimization, one of these two underlying representations must be chosen as the *sole* underlying representation. The obvious choice is /kʰowt/, since it can be used to derive both the present tense (since it is identical to it) and the past tense (by the process of flapping, changing /t/ to [ɾ]), whereas /kʰowɾ-/ can only be used to derive the past tense (there is no attested process of 'deflapping' in English to change /ɾ/ to [t]).

However, some research on allomorphy in OT (e.g. Mester 1994, Burzio 1996, Kager 1996, 1999b, Rubach 2001) argues that lexical minimization need not be strictly adhered to (cf. non-OT analyses with similar arguments, such as Vennemann 1974, Hudson 1975, Aronoff 1976, Bybee 1988, 1995, and others). This clearly must be true for cases of obvious suppletion, in which the allomorphs cannot be derived by any reasonable purely phonological process from a single underlying representation (e.g. the oft-cited English pairs *go~went* and *be~was*). When suppletive allomorphs have utterly unpredictable distributions, most researchers are willing to allow lexical minimization to be violated, under the assumption that the lexicon should encode all information that is unpredictable. However, once any amount of regularity is established, productive or otherwise, the drive to avoid redundancy forces many to assert that a single underlying representation must be responsible for the alternations, which must be governed by the phonology. Such assertions are based solely on theoretical grounds; there is no evidence to indicate that there is any practical upper bound on the size of the human lexicon, or that language learners do not memorize redundancy. Indeed, evidence from malapropisms, so-called 'tip of the tongue' effects, and other speech errors suggests that some predictable information (such as number of syllables, stress patterns, etc.) is stored in memory (Brown and McNeill 1966, Fromkin 1973, Fay and Cutler 1977, Cutler 1986, Levelt 1989, Ullman 1993, 1999, among others).

Since predictability does not seem to be a determining factor, it is not clear where to draw the line to determine which information is stored in the lexicon and which is not. Rather than draw an arbitrary line, I propose that the principle of lexical minimization (11) should be abandoned altogether and that lexicon optimization (9) should be adhered to strongly, resulting in *all* information being stored in the lexicon, predictable or otherwise. Since lexicon optimization is generally assumed to work in conjunction with lexical minimization, I will use the term *strong lexicon optimization* to refer to lexicon optimization without lexical minimization. Recall that for the English word 'coat', strong lexicon optimization selects /k^{h}owt/ and /k^{h}owɾ-/ as the underlying representations for the shape of the stem in the present and past tenses, respectively. Without lexical minimization, both of these underlying representations exist in the lexicon, marked for use in the environments they are supposed to be used in. The formal representation of this lexical allomorphy is discussed further in Chapter 4, where morphology plays a crucial role. The analyses in Chapters 2 and 3 do not depend on morphology, so in these chapters, I assume a simplified version of strong lexicon optimization that stores entire words, rather than individual morphemes.

At this point, one may wonder how lexicon optimization without lexical minimization can be a part of phonology in the generative sense. After all, one of the goals of generative phonology was to move away from a 'memorized list' of words to explain and predict patterns in a principled way. It is therefore important to emphasize that abandoning lexical minimization does not entail an abandonment of generative phonology as a whole. Recall that richness of the base allows every possible word to be a possible input to the grammar. Thus, the grammar must still function to filter the set □ of all possible words by mapping it to the actual set of words in the language. This is achieved through the constraint hierarchy. Sometime after the constraint hierarchy has been fixed (during the language acquisition process), strong lexicon optimization takes effect, and the outputs are stored in the lexicon as underlying representations. Once strong lexicon optimization fixes the lexicon, the underlying representations will pass through the constraint hierarchy unchanged, but crucially, *the constraint hierarchy is still there!* This means that the phonology is still active, no matter what the lexicon looks like. Hypothetical inputs can still be tested and mapped to actual words because phonological generalizations are still expressed by the grammar through a principled mechanism (cf. Pinker's

(1999) 'words-and-rules' theory, which is also based on the concept that regularity in language is both stored and derived).

1.1.6 Monostratality

One of the strongest claims of standard OT is that the mapping from input to output is governed solely by constraints on the output and its relationship to the input. In other words, the mapping is direct, or *monostratal*. Monostratality in standard OT does not allow any sort of abstract intermediate representations to influence the selection of the optimal candidate. Indeed, such representations simply do not exist in standard OT. In this dissertation, I maintain that monostratal OT is sufficient to account for the range of phenomena multistratal theories are designed for, including opacity, the primary topic of this dissertation and of the next section.[6] (See §5.5.3 of Chapter 5 for further discussion of multistratal models.)

1.2 Opacity

It is possible for a generalization describing the phonological shape of actual and potential words of a language to be *transparent* (that is, always true in the language).[7] An example of a transparent generalization in English is 'a word-initial oral stop must be voiceless, but can be aspirated or unaspirated'.[8] That is, there are English words

---

[6] I should note that I do not consider faithfulness between morphologically related outputs (as in Benua 1995, 1997, Steriade 1996, Kenstowicz 1996, Burzio 1998, etc.) to be multistratal in the same sense as used in this section. This type of output-output faithfulness can be encoded with a single constant constraint hierarchy that requires Eval to consider an input only once. The important difference between output-output faithfulness and multistratal OT lies in the issue of abstract intermediate representations. None need to be calculated for output-output faithfulness but they are a crucial component to multistratal models. Output-output faithfulness is discussed in more detail in §5.5.3 of Chapter 5.

[7] Hooper's (1976) True Generalization Condition claims that this is not just possible, but required.

[8] This formulation is overly simplified for the sake of discussion and should not be taken as a complete analysis of this phenomenon, which warrants a more complex analysis.

such as [**p**æt] ‘bat’ (the initial stop is voiceless and unaspirated) and [**p**^{h}æt] ‘pat’ (voiceless and aspirated), but there are no words like ×[**b**æt] (fully voiced and unaspirated) or ×[**b**^{h}æt] (voiced and aspirated). This generalization is always true of English; it has no lexical exceptions, and nonce forms and foreign loanwords invariably conform to it. However, there are numerous examples in the literature of generalizations which are assumed to always be true for a language, despite apparently being false for some words or in some environments. Of specific interest to this dissertation are instances of *opacity* (Kiparsky 1971), in which some opaque generalization *G* appears to be false (but is still assumed to be true) precisely when *G* interacts with some other generalization that obscures *G*’s truth.

The inability of monostratal OT to adequately account for opacity in general is not a new subject and is well-documented (see McCarthy and Prince 1993, Prince and Smolensky 1993/2002, Chomsky 1995, Goldsmith 1996, numerous papers in Roca 1997, Idsardi 1998, Kager 1999a, etc.). Various proposals have come forth to modify OT to allow it to account for opacity, either generally or only in a handful of cases. Among such proposals are the multistratal models discussed in the previous section, as well as other extensions to OT such as constraint conjunction (Smolensky 1993/2000, Kirchner 1996) and output-output faithfulness (Benua 1995, etc.). As stated before, I adhere to monostratality, rejecting the multistratal frameworks. The other attempts to analyze opacity in monostratal OT are not sufficiently powerful to account for all types of opacity, including the cases of opacity in Polish analyzed in this dissertation. Thus, the burden is upon this author to provide a suitable framework which can handle opacity.

1.3 Formalization of FDM-OT

The analytical framework adopted in this dissertation is *Faithfulness, Dispersion, and Markedness in OT*, or *FDM-OT* for short. FDM-OT is a monostratal version of OT, following the principles of richness of the base and strong lexicon optimization (thus rejecting lexical minimization). FDM-OT is distinguished from standard OT by having three constraint families: faithfulness, dispersion, and markedness.

1.3.1 Faithfulness constraints

The first family of constraints in FDM-OT consists of *faithfulness* ($\mathcal{F}$) constraints of the general sort discussed in Prince and Smolensky 1993/2002 and McCarthy and Prince 1995. At a first level of approximation, this family ensures that various properties of the input are preserved in the output. For example, if a particular segment in the input is nasalized, it should be nasalized in the output as well. If not, a violation of the constraint $\mathcal{F}$-nasal is incurred, indicating that the output differs from the input with respect to the property of nasality. The $\mathcal{F}$ family can be formalized as follows:

(14) $\mathcal{F}$-*P* 'If *x* and *y* are segments in a correspondence relationship with each other, then their specifications for property *P* must be the same. Violations are counted gradiently: the more different *x* and *y* are, the worse the violation.'

The effects of $\mathcal{F}$-constraints will depend on the nature of the various *P*s in the system. I assume that properties are generally scalar. That is, they may have multiple possible specifications along a single scale, with a suitable metric for evaluating closeness. For example, the property of vowel height has at least three possible specifications (and likely more): high, mid, and low. High vowels are closer to mid vowels than they are to low vowels. Thus, a change from input /i/ to output [a] will be worse with respect to $\mathcal{F}$-height than changing /i/ to [e] (which is, of course, worse than faithfully keeping /i/ the same from input to output):

(15)

/i/	$\mathcal{F}$ height
a. i	
b. e	✖
c. a	✖^{2}

Note that the notion of 'property' that I am adopting is somewhat different from traditional distinctive phonological features. In the case of vowel height, I am conflating the features [high] and [low]. Many properties, such as nasality and rounding seem to only have two values, but many others, such as vowel height, vowel color, voicing, tone, duration, etc. seem to operate in a more gradient fashion, with more than two viable values over a range. While these multi-valued properties can often be decomposed into binary features, there is a large body of work which rejects binary features for some properties (for example, Williamson 1977, Lindau 1978, Selkirk 1984, Clements 1991; see Gnanadesikan 1997 for further references and discussion of this issue). I assume no inherent requirement for binary properties and adopt multi-valued properties where appropriate.

In FDM-OT, as in Dispersion Theory (Flemming 1995, Pad get 1997, Ní Chiosáin and Padgett 2001), inputs to Eval and candidates created by Gen are *sets* of words, rather than single words as in standard OT. Following the OT hypothesis of richness of the base, I assume that the input to an FDM-OT grammar is [illegible], the set of all possible words. Obviously, I cannot show such a set in full within a tableau, so it is necessary to abstract away from all possibilities to just those that are relevant to the analysis at hand. This is not a novel concept in any theory, though it is rarely mentioned explicitly in the literature (but see Ní Chiosáin and Padgett 2001 for an excellent discussion of this issue). As an example, consider vowel height in American English. Most dialects of American English allow three levels of height in the front lax vowels before nasal codas, as in *t*[ɪ]*n*, *t*[ɛ]*n*, *t*[æ]*n*. However, in my dialect (Appalachian Georgian) and many other southern dialects, the high and mid vowels neutralize before nasal codas to [ɪ] (more or less; I am ignoring the characteristic 'drawl' of southern dialects in which some amount of diphthongization also occurs, resulting in a pronunciation more like [ɪə]).

To analyze this neutralization, it is not important to consider the nature of the onset, the place of articulation of the nasal coda, or other syllables, as these have no effect on the neutralization. Thus, it is sufficient to use as the input a set of words that differ only in vowel height and coda nasality, such as the input in (16):

(16)

$\text{p}^{\text{h}}\text{ɪt}_1$ $\text{p}^{\text{h}}\text{ɪn}_4$ $\text{p}^{\text{h}}\text{ɛt}_2$ $\text{p}^{\text{h}}\text{ɛn}_5$ $\text{p}^{\text{h}}\text{æt}_3$ $\text{p}^{\text{h}}\text{æn}_6$	$\mathcal{F}$ height
a. $\text{p}^{\text{h}}\text{ɪt}_1$ $\text{p}^{\text{h}}\text{ɪn}_4$ $\text{p}^{\text{h}}\text{ɛt}_2$ $\text{p}^{\text{h}}\text{ɛn}_5$ $\text{p}^{\text{h}}\text{æt}_3$ $\text{p}^{\text{h}}\text{æn}_6$	
b. $\text{p}^{\text{h}}\text{ɪt}_1$ $\text{p}^{\text{h}}\text{ɪn}_{4,5}$ $\text{p}^{\text{h}}\text{ɛt}_2$ $\text{p}^{\text{h}}\text{æt}_3$ $\text{p}^{\text{h}}\text{æn}_6$	×

Each word in the input is indexed by a number, shown as a subscript. In the output candidates, the same indices appear, but sometimes on different words. In this case, the word [p^hɪn] in candidate (16b) is indexed with the numbers 4 and 5, which means it corresponds to the input words with those same numbers, /p^hɪn/ and /p^hɛn/. Thus, candidate (16b) represents those dialects with neutralization (i.e. homophony), while (16a) represents standard American English with a three-way height distinction.

The input set could be extended to further probe the properties at work (how does neutralization interact with diphthongs? with vowel color?), but that is not necessary for the purpose of the current discussion, which is merely to give an example of an $\mathcal{F}$-constraint acting upon a limited subset of □.

1.3.2 Dispersion constraints

Dispersion, or $\mathcal{D}$, constraints were first proposed for OT by Flemming (1995) in his formulation of Dispersion Theory. The principle idea behind $\mathcal{D}$-constraints is that the perceived acoustic differences between words play a role in phonology, an idea that significantly predates OT (see for example de Saussure 1959, Martinet 1964, Lindblom 1986, 1990) and has recently been adapted to OT (Flemming 1995, Steriade 1995, Padgett 1997, to appear a, to appear b, Boersma 1998, and Ní Chiosáin and Padgett 2001, to name a few). I follow the general model put forth by Flemming (1995) and expanded by Padgett (1997) and Ní Chiosáin and Padgett (2001) in which a universal hierarchy of $\mathcal{D}$-constraints penalizes candidates that contain words which are 'too close together' with respect to some perceived acoustic dimension. These constraints are defined generally as follows:

(17) $\mathcal{D}_n$-P ‘Every pair of words x and y in the output which contrast for property P must be at least as far apart as the nth from smallest allowable perceptual distance for P.’

The properties referenced in $\mathcal{D}$-constraints are based purely on perception of acoustics. Thus, they are potentially different from those used in $\mathcal{F}$-constraints, since the latter are cognitive properties which may involve some combination of articulation and perception (cf. Boersma 1998). However, the phonetic reality captured by these two properties are often identical, and except where it is confusing or important, I ignore the differences between perceptual $\mathcal{D}$-properties and cognitive $\mathcal{F}$-properties.

For every property P there is an inherently ranked family of $\mathcal{D}$-constraints:

(18) $\mathcal{D}_0\text{-}P \gg \mathcal{D}_1\text{-}P \gg \mathcal{D}_2\text{-}P \gg \ldots \gg \mathcal{D}_{N-1}\text{-}P \gg \mathcal{D}_N\text{-}P \gg \mathcal{D}_{N+1}\text{-}P$

The constraints $\mathcal{D}_0$-P, $\mathcal{D}_N$-P, and $\mathcal{D}_{N+1}$-P can be rewritten as the following, to better highlight their functions:

(19) $\mathcal{D}_0$-P ‘Every pair of words x and y in the output which contrast for property P must be at least as far apart as the *smallest* allowable perceptual distance for P.’

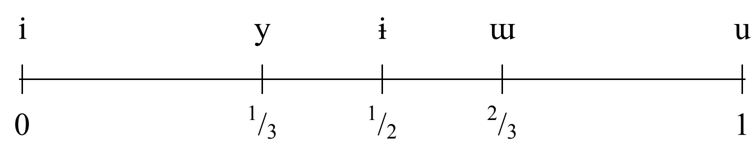
(20) $\mathcal{D}_N$-P ‘Every pair of words x and y in the output which contrast for property P must be at least as far apart as the *largest* allowable perceptual distance for P.’

(21) $\mathcal{D}_{N+1}$-P ‘Every pair of words x and y in the output which contrast for property P must be farther apart than the largest allowable perceptual distance for P.’

These constraints deal with the lower and upper bounds on contrast for a given property. I assume (non-crucially) that $\mathcal{D}_0$-P is undominated (perhaps a part of Gen), preventing any candidate from having a pair of words which are closer than the smallest perceptual distance. In comparison, only the most contrastive pair of words can satisfy $\mathcal{D}_N$-P, since it requires all contrasts to be as good as the maximum contrast.[9] The constraint $\mathcal{D}_{N+1}$-P completely prevents any contrast for P, since it requires all pairs to be better than the maximum possible contrast, a clear impossibility.

I now give an example of $\mathcal{D}$-constraints for the property of vowel color. The perceptual property of vowel color involves some weighted function of vowel formants, with particular emphasis on F2 and F3 (cf. Carlson, Granström, and Fant 1970). Without mentioning a particular function, I simply assume that [i] and [u] are at the extremes (the minimum and maximum on an arbitrarily labeled scale from 0 to 1), with [y] and [ɯ] equally spaced between them (at $^1/_3$ and $^2/_3$ respectively), and [ɨ] occupying the center at the $^1/_2$ mark:

(22)



I ignore further subdivisions of the color space, though in theory, an infinite number is possible. The following shows the relative perceptual distance within various pairs

[9] It may seem odd at first to talk about a maximum allowable perceptual distance, since the acoustic properties we judge sounds on (e.g. formant values) need not necessarily be limited to those which produce human speech sounds. It is likely that during the language acquisition process, language learners weed out useless constraints based perceptual distances that cannot be created by human mouths. For example, it is impossible to produce a vowel fronter than [i], so there is no need for a constraint that bans any perceptual distances larger than that between such a vowel and [u].

of vowels from the smallest (and impossible) to the largest, with the $\mathcal{D}$-constraints that rule out each pair listed below:

(23) [y ɨ],[ɨ ɯ] < [i y],[y ɯ],[ɯ u] < [i ɨ],[ɨ u] < [i ɯ],[y u] < [i u]

ruled out by: $\mathcal{D}_0$ $\mathcal{D}_1$ $\mathcal{D}_2$ $\mathcal{D}_3$ $\mathcal{D}_4$

The first vowel pairs, [y ɨ] and [ɨ ɯ], are unattested and thus represent impossible pairs. Their members are too close together perceptually to be accurately and consistently distinguished. Presumably, this is a fact that the language learner acquires through observation, filtered by some innate minimum threshold of confusability. I am not concerned with the exact mechanism by which this threshold is obtained. All that matters is that $\mathcal{D}_0$-color is a measure of it and that candidate languages which contain contrasting vowels that are closer than this lower bound will violate $\mathcal{D}_0$-color. Similarly, increasingly larger tolerances for closeness are governed by the higher $\mathcal{D}$-constraints. I am being deliberately vague on the exact phonetic value associated with these constraints. The kind of experimental work required to determine these values has not been conducted and can (for now) only be inferred from cross-linguistic typology. Thus, I abstract away from any specific numerical values and merely examine inventories.

In the following tableau, various candidate vowel inventories are subjected to this family in order to show how constraint violations for $\mathcal{D}$-constraints are to be counted (no input language needs to be specified since $\mathcal{D}$-constraints only refer to outputs):

(24)

	$\mathcal{D}_0$ color	$\mathcal{D}_1$ color	$\mathcal{D}_2$ color	$\mathcal{D}_3$ color	$\mathcal{D}_4$ color
a. i y ɨ ɯ u	✖[2]	✖[5]	✖[7]	✖[9]	✖[10]
b. i y ɯ u		✖[3]	✖[3]	✖[5]	✖[6]
c. i y u		✖	✖	✖[2]	✖[3]
d. i ɯ u		✖	✖	✖[2]	✖[3]
e. i ɨ u			✖[2]	✖[2]	✖[3]
f. i ɯ				✖	✖
g. i u					✖
h. ɨ					

Violations for $\mathcal{D}$-constraints are counted for every pair of words which are too close in the relevant perceptual dimension. Note that the number of violations a particular candidate incurs does not decrease through successive constraints (the violation count is monotonically increasing, either equal or larger at each step). This is because the required minimal distance increases while the candidate's actual spacing remains constant. Candidate (24a) violates the constraint $\mathcal{D}_0$-color by having two pairs of vowels, [y ɨ] and [ɨ ɯ], which are closer than the smallest allowable perceptual distance. This candidate can never surface in any language because $\mathcal{D}_0$-color is undominated in all languages.

The $\mathcal{D}$-constraints described above are generalized versions of their predecessors: Flemming's (1995) MINIMAL DISTANCE constraints and Padgett's (1997) SPACE constraints. MINIMAL DISTANCE constraints mark violations based on a finite set of acoustic features (such as [high F2], [lowest F2], etc.). Padgett (1997) argues that the categorical nature of MINIMAL DISTANCE constraints does not adequately capture the gradient properties of perception and contrast dispersion, especially in regards to vowel quality. He proposes SPACE constraints instead, which mark violations based on fractional divisions of the perceptual space, in a fixed format of $^{1}/_{n}$, for integral values of n. However, Padgett's SPACE constraints specify particular divisions of the perceptual space ($^{1}/_{2}$, $^{1}/_{3}$, etc.) which do not necessarily represent the perceptual spacing speakers actually use (the spacing could in reality be something as seemingly arbitrary as $^{43}/_{97}$ or even $^{3}/_{\square}$). Only precise phonetic experimentation on speaker perception could determine the true values (or an approximation of them). The more abstract $\mathcal{D}$-constraints used here avoid a priori specification of any exact perceptual distances, as they are compatible with any monotonically increasing set of distances (including those defined by MINIMAL DISTANCE and SPACE constraints).

By interspersing $\mathcal{F}$-color among the $\mathcal{D}$-color constraints, different winning candidates can be selected. For clarity of exposition, I do not put subscripts on the inputs or candidates in the following tableaux, instead simply assuming that each input vowel that does not faithfully surface merges with its nearest neighbor in the surface, with $\mathcal{F}$ violations counted as the number of input vowels that do not surface faithfully. In (25), $\mathcal{F}$-color is ranked between $\mathcal{D}_0$-color and $\mathcal{D}_1$-color (recall that $\mathcal{D}_0$-color must be undominated; not all of the lower-ranked $\mathcal{D}$-constraints are shown):

(25)	i y ɨ ɯ u	$\mathcal{D}_0$ color	$\mathcal{F}$ color	$\mathcal{D}_1$ color	$\mathcal{D}_2$ color	$\mathcal{D}_3$ color
	a. i y ɨ ɯ u	×²!		×⁵	×⁷	×⁹
✓	b. i y ɯ u		×	×³	×³	×⁵
	c. i y u		×²!	×	×	×²
	d. i ɯ u		×²!	×	×	×²
	e. i ɨ u		×²!		×²	×²
	f. i ɯ		×³!			×
	g. i u		×³!			
	i. ɨ		×⁴!			

Candidate (25a) is ruled out immediately because it is impossible: despite being maximally faithful to the input, it violates the undominated constraint $\mathcal{D}_0$-color, which bans vowels from being too close perceptually. The remaining candidates violate $\mathcal{F}$-color to some extent. Because candidate (25b) has the fewest violations of this highly ranked constraint, it is selected as the output. This candidate represents the color inventory of languages like Turkish (Zimmer and Orgun 1999).

In the next tableau, $\mathcal{F}$-color is ranked between $\mathcal{D}_1$-color and $\mathcal{D}_2$-color, lower than in (25). To simplify the tableau, I ignore the topmost $\mathcal{D}_0$-color constraint and the candidates that fatally violate it:

(26)

i y ɨ ɯ u	$\mathcal{D}_1$ color	$\mathcal{F}$ color	$\mathcal{D}_2$ color	$\mathcal{D}_3$ color
a. i y ɯ u	×[3]!	×	×[3]	×[5]
b. i y u	×!	×[2]	×	×[2]
c. i ɯ u	×!	×[2]	×	×[2]
✓ d. i ɨ u		×[2]	×[2]	×[2]
e. i ɯ		×[3]!		×
f. i u		×[3]!		
g. ɨ		×[4]!		

Highly ranked $\mathcal{D}_1$-color rules out candidates (26a–c), leaving (26d) as the most faithful candidate and thus the selected output since $\mathcal{F}$-color is the next ranked constraint. This candidate represents the vowel color system in Amharic (Hayward and Hayward 1999).

The next tableau shows the result of ranking $\mathcal{F}$-color even lower, below $\mathcal{D}_2$-color. For simplicity, $\mathcal{D}_1$-color and the candidates which fatally violate it are ignored:

(27)

i y ɨ ɯ u	$\mathcal{D}_2$ color	$\mathcal{F}$ color	$\mathcal{D}_3$ color
a. i ɨ u	✗²!	✗²	✗²
b. i ɯ		✗³	✗!
c. y u		✗³	✗!
✓ d. i u		✗³	
e. ɨ		✗⁴!	

With $\mathcal{D}_2$-color ranked high, three-color systems like (27a) are ruled out. Candidates (27b–d) are more faithful than (27e), which fatally violates $\mathcal{F}$-color. Because they are equally faithful, they must be compared with $\mathcal{D}_3$-color, which selects (27d) as the optimal candidate. This output represents the vowel systems of languages like Arabic (Thelwall and Sa'adeddin 1999) and Catalan (Carbonell and Llisterri 1999).[10] $\mathcal{F}$-color can be ranked lower, producing smaller vowel inventories, until eventually $\mathcal{F}$-color is ranked below all of the relevant $\mathcal{D}$ constraints, which results in a system containing only one segment to exemplify the relevant property.

In the next section, I show that in addition to $\mathcal{F}$-constraints and $\mathcal{D}$-constraints, a third type of constraint is also required.

[10] In order to get a vowel system like Japanese (27b), the grammar needs markedness constraints to rule out round vowels (since these involve more complex articulations than unrounded vowels). Thus, with $\mathcal{M}$-u ranked over $\mathcal{D}_3$-color, candidate (27b) can be obtained. I leave it as an exercise to the reader to prove that an ill-formed system such as (27c) cannot be derived with these constraints and a markedness hierarchy in which round vowels are universally more difficult than their unrounded counterparts.

1.3.3 Markedness constraints

It is relatively uncontroversial that the limitations of the oral tract must play a role in phonology. Numerous phonological generalizations are motivated solely or primarily by articulatory difficulty rather than any sort of drive to be faithful or to be more perceptually distinct. For example, there seem to be no non-African languages which make linguistic use of clicks, sounds made with an ingressive velaric airstream mechanic. Since linguistically meaningful clicks are possible (and common) sounds in several south African languages (as in the Khoisan language !Xóõ, in which the large majority of the words listed in Traill's (1994) !Xóõ dictionary begin a click), so it is not the case that they simply do not exist (Ladefoged and Maddieson 1996:246). By the richness of the base hypothesis, the lack of clicks in languages like English cannot be due to $\mathcal{F}$-constraints, which would require identity with whatever input is used, including those containing clicks. $\mathcal{D}$-constraints also seem useless for explaining the absence of clicks in English; indeed, clicks are in fact quite likely the most distinctive linguistic sounds, as they are generally louder than other sounds, are often separated from surrounding sounds by brief silence to accommodate the articulations required to make them, and are rarely confused with other sounds (Ladefoged and Maddieson 1996:259, Ladefoged and Traill 1994). Thus, if neither $\mathcal{F}$- nor $\mathcal{D}$-constraints are able to motivate the lack of clicks in English, a third constraint family is required. I propose that this family of *articulatory markedness* constraints punishes difficult maneuvers in the vocal tract, in this case, the double articulations and ingressive airflow required to make the characteristic popping sound of a click. Articulatory markedness, or $\mathcal{M}$, constraints, are defined generally as follows:

(28) $\mathcal{M}$-*A* ‘No output word can contain the articulation *A*.’

The relevant $\mathcal{M}$-constraints for clicks are $\mathcal{M}$-ʘ, $\mathcal{M}$-ǃ, $\mathcal{M}$-ǂ, etc. which would each be universally ranked higher than their non-click counterparts $\mathcal{M}$-p, $\mathcal{M}$-t, $\mathcal{M}$-c, etc. since singly-articulated pulmonic egressive stops are easier to produce than clicks.

$\mathcal{M}$-constraints in FDM-OT are similar, but not identical, to the markedness or well-formedness constraints from standard OT. Very often, ‘markedness’ has been used to describe many different concepts, sometimes at the same time: articulatory difficulty, order of acquisition, implications in inventories, cross-linguistic frequency, etc. The $\mathcal{M}$-constraints in FDM-OT target structures that are marked *solely* because of the physical articulations involved. $\mathcal{M}$-constraints thus never refer to acquisitional or typological ‘markedness’. There are two reasons why articulatory $\mathcal{M}$-constraints are preferable to the more standard typological markedness constraints. First, $\mathcal{M}$-constraints are based on information that all speakers of every language have access to at some (early) point in their lifetime: an awareness of the physiology of the human mouth. Speakers do not have access to the cross-linguistic distribution of sounds. For example, a random language learner could not be expected to treat [ʔ] as more marked than [t], despite [t]’s predominance in the languages of the world, because the learner can only rely on data from his or her own language (which could possibly be Hawai’ian, which in fact has [ʔ] but not [t], contrary to typological markedness). Having the typological distribution of speech sounds dictate the universal ranking of markedness constraints is circular and has no explanatory power.

In addition, traditional typological markedness constraints cannot always derive the correct empirical results. Padgett (1997) discusses this with respect to vowel height. By typological standards, [ə] is one of the most marked vowels, since it occurs as a contrastive vowel in fewer languages than [i] or [a] do, suggesting a universal ranking in standard OT such as ✖ə ≫ ✖i,✖a. However, [ə] is very often chosen as the default vowel for epenthesis or in unstressed (i.e. relaxed) positions. Without reference to articulatory difficulty, this discrepancy must be ignored or given an ad hoc explanation. But [ə] should be treated as the easiest vowel in terms of articulatory effort; all other vowels are deviations from the neutral position which characterizes [ə]. In other words, the actual markedness hierarchy, using FDM-OT constraints, is $\mathcal{M}$-i, $\mathcal{M}$-a ≫ $\mathcal{M}$-ə, the opposite of what is expected in standard OT This means the typological distribution of sounds must come from some factor other than conventional markedness constraints. Adapting concepts from Flemming 1995, I propose that typological facts are derived from the interaction of $\mathcal{D}$-constraints with $\mathcal{M}$- and $\mathcal{F}$-constraints, rather than from classical markedness constraints alone. Since articulatory markedness is still required for phonology, the FDM-OT formulation of $\mathcal{M}$-constraints expresses a necessary property of language without being conflated with the derivable property of typology.

1.4 Sound change in FDM-OT

In this section, I provide details for how the early grammar, sound change, and opacity are analyzed in the FDM-OT framework. This analysis derives from work in Sanders 2001 and 2002b, which utilize the same core concepts, but in standard OT rather than FDM-OT.

1.4.1 The early grammar

While I do not provide a full analysis of language acquisition, it does play a role in this dissertation. I assume that through some learning mechanism (which I am not concerned with formalizing), a language learner who hears some set of adult outputs *A* creates an early constraint ranking and Eval function designed to map □ (the set of all possible words) to *A*, so that the language learner can reproduce a possible adult output from every possible input.[11] The following diagram shows the relationships involved in language acquisition:

(29) *language acquisition (LA) and the early grammar*

child grammar

□ (all possible inputs)

A —**LA**→ Eval □

adult output *A* (identical to adult output)

The adult output is the set *A*. Based on *A*, the language learning child creates Eval (or more precisely, a constraint hierarchy that Eval makes use of)), which maps □ to *A*. I call the grammar created in this way the *early grammar*.

1.4.2 Sound change

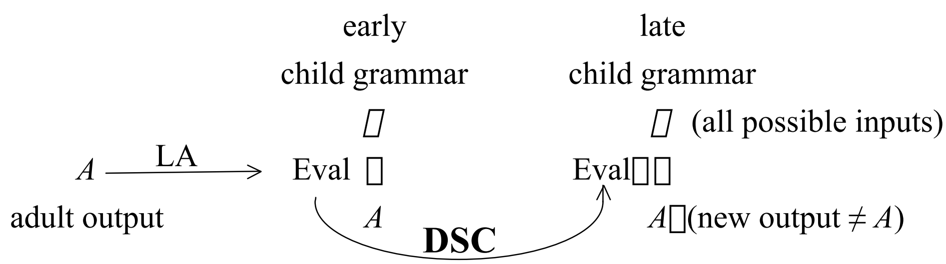
Sound changes occur when phonological generalizations are added, lost, or altered over the course of time. In FDM-OT, sound change is represented by reranking constraints in the constraint hierarchy (cf. similar proposals in Jacobs 1995, Gess 1996, Holt 1997, Zubritskaya 1997, etc.), which results in a new Eval function that is different from that derived by regular language acquisition. This new Eval□maps □

---

[11] This is a very simplistic, and ultimately inaccurate model of acquisition. The adult output set *A* is finite (since the language learner only hears a finite number of words when acquiring a language), but Eval must map □ to an infinite set *A*□, which contains *A* as a subset. Importantly, *A*□is not any arbitrary infinite superset of *A*, but a well-defined (in the phonological sense) superset, with the extra words representing possible words which happen not to exist (or were not heard during the acquisition process). I ignore the complexity here and simply assume that *A* is actually the infinite set which represents all possible words in the adult grammar.

to a set of words *A*′ that is similar to *A* but shows evidence of the sound change caused by the constraint reranking which distinguishes Eval from Eval′:

(30) *diachronic sound change (DSC)*



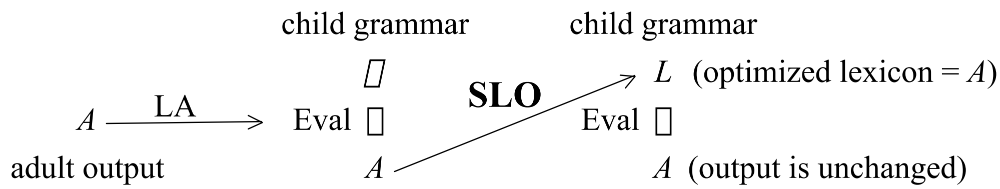
When discussing a particular sound change, I refer to the grammar after the sound change as the *late grammar*. Tableaux for late grammars are visually distinguished from early grammar tableaux by shadows behind the late grammar tableaux.

Since each generation’s Eval depends on the previous generation’s outputs, it is clear that any sound change will be passed along to the next generation when it attempts to recreate the adult outputs and will continue to be a part of the evolving language as each generation learns the sound change. Later sound changes can interact with earlier sound changes in numerous ways. The most interesting interaction is opacity, which I claim crucially relies on strong lexicon optimization.

1.4.3 Strong lexicon optimization and opacity

As stated in §1.1.5, strong lexicon optimization (sans lexical minimization) will eventually cause the outputs of Eval to be stored in the lexicon directly, so that the speaker does not have to continually re-evaluate every possible input (□) every time a word is spoken. While it is probably the case that strong lexicon optimization applies to different words at different times (likely due to frequency), I assume a simpler model here in which strong lexicon optimization applies to the entire lexicon all at once, resulting in the creation of a lexicon *L* that is identical to the set of outputs generated by Eval:

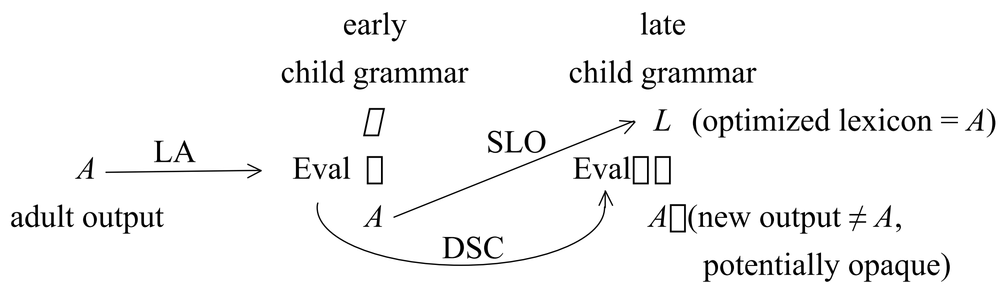
(31) *strong lexicon optimization (SLO) in the child grammar*



Once strong lexicon optimization has occurred, the new lexicon *L* is the input to Eval, which maps *L* completely faithfully to *A* (since *L* is merely a copy of *A*, which is already the output generated by Eval when applied to ▯). This is the key to allowing opacity to occur in a grammar that is synchronically monostratal.

Opacity requires intermediate representations, which are banned by monostratal direct mapping. Thus, whatever input *I* is passed to Eval, the output cannot be opaque with respect to *I*. However, it could easily be opaque with respect to previous generations' outputs or lexicons, since there is no requirement of direct mapping between generations. The newly optimized lexicon acts as a set of *diachronic* intermediate representations. Suppose that diachronic sound change occurs after strong lexicon optimization (not a requirement of FDM-OT, but a crucial possibility). Then opacity can occur if the new sound change opaquely obscures sound changes from previous generations, simulating multistratal synchronic derivations with diachronic serialism:

(32) *opacity as strong lexicon optimization followed by diachronic sound change*



This is the general schema which will be used to analyze a variety of cases of opacity in Polish that have arisen through diachronically ordered sound changes. In the following section, I give a tour of the history of Polish, summarizing the major features of each stage of the language, as well as documenting the types of evidence used to determine these features.

1.5 The history of the Polish language

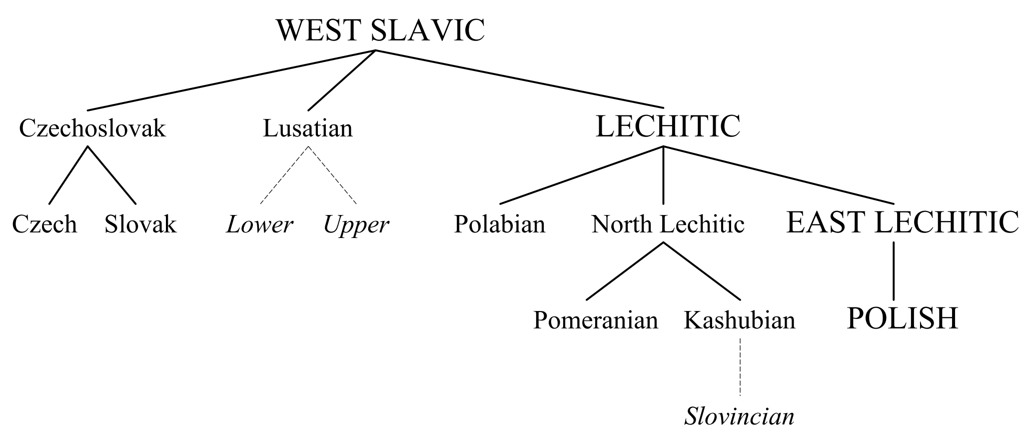
1.5.1 Polish's genetic relationship to other languages

As a Slavic (or 'Slavonic') language, Polish is a descendant of the reconstructed language generally known as Proto-Slavic (PSl) or Common Slavic, itself a descendent of Proto-Indo-European (PIE), the proposed ancestral language of numerous language families (Germanic, Romance, Indic, etc.). Since PSl stopped being a unified language before any Slavic language was put into written form, everything that is known or assumed about PSl is based on historical reconstruction from the modern Slavic languages and various ancient literary languages (Carlton 1991:9). In most analyses of PSl, the 'fall of the jers' (deletion or change of the short high vowels *[i] and *[u] in certain contexts; see §1.5.3 for further discussion) is singled out as the last sound change which affected all of PSl, marking the end of PSl and the true beginning of three major dialects of PSl as independent languages

(Carlton 1991:171). The largest PSl group by number of modern speakers is the *East Slavic* group, which includes Russian, Ukrainian (sometimes called Ruthenian), and Belorussian (or 'White Russian'). The oldest attested group of PSl is *South Slavic*, which includes Serbo-Croatian, Bulgarian, Slovene, and Macedonian. South Slavic is preserved in an ancient literary language known as Old Church Slavonic (sometimes also referred to as Old Bulgarian) which is often used to add validity to reconstruction of PSl due to the significant similarity between attested Old Church Slavonic forms and hypothesized PSl forms (Entwistle and Morison 1949:50–52, de Bray 1969:1, Carlton 1991:13).

Finally, of most importance to the present discussion, *West Slavic* is the group of PSl which includes Polish and numerous other languages and dialects, as the following diagram shows (dialects are in italics, attached to their parent languages by dotted lines; nodes in Polish's direct lineage are in large capitals):

(33)



The modern languages of Czech and Slovak constitute the southernmost subgroup of the West Slavic languages. The western subgroup of West Slavic is represented only by Lusatian (also known as Sorbian or Wendish). Lusatian, divided into the two dialects Lower Lusatian and Upper Lusatian, is spoken in the eastern corner of

Germany near the Polish and Czech borders. The diverse northeastern Lechitic subgroup is the largest subgroup of West Slavic (Carlton 1991:13–14).

Polabian, the most western Lechitic language, became extinct in the mid-18th century, with less than 3,000 written words left behind. Pomeranian (‘West Pomeranian’ in Stieber 1973), also extinct, is a blanket term for the coastal variety of Lechitic spoken in the north. Very little remains of Pomeranian (only some proper names in 12th–14th century German documents), making it difficult to accurately reconstruct more complex relationships among the westernmost Lechitic languages. Kashubian (spelled ‘Cassubian’ in some sources), is often classified as a dialect of Polish, despite the fact that it is mostly unintelligible by Polish speakers, even those who border the Kashubian region (Stieber (1973:139) classifies Kashubian as an eastern dialect of North Lechitic, his ‘Pomeranian’). Slovincian is an extinct western dialect of Kashubian which died out circa 1940 (Entwistle and Morison 1949, Carlton 1991:13–17).

1.5.2 Proto-Slavic

PSl is often taken as the starting point in any historical analysis of a particular Slavic language, so I begin this excursus into the history of Polish with PSl. The standard analysis of PSl (e.g. in Entwistle and Morison 1949, Bidwell 1963, Carlton 1991, Townsend and Janda 1996) posits a vowel system with two heights, two colors, and two lengths. Following Carlton (1991), I use the following symbolization for this reconstructed vowel system (long vowels are indicated with the IPA symbol for length [ː], while short vowels are unadorned):

(34)

	front	back
high	i,iː	u,uː
low	æ,æː	ɑ,ɑː

PSl is also analyzed to have the four diphthongs with all four possible low-to-high contours: *[æj], *[æw], *[ɑj], and *[ɑw].

Most researchers assume that PSl was a pitch accent language, with two possible tones: rising (traditionally called 'acute'; symbolized here with the IPA symbol [ˇ]) and falling ('circumflex'; IPA [ˆ]). These tones could only occur on an accented (stressed) syllable. The falling tone was further restricted in that it could only occur on the initial syllable of a form with mobile accent in the paradigm (i.e. different forms in the paradigm may have different syllables with accent, as in modern Russian [goloˈva]/[ˈgolovu] 'head (NOM/ACC SG)') (Entwistle and Morison 1949:71–75, Stieber 1973:13, Carlton 1991:186–187).

PSl also distinguished three places of articulation (labial, coronal, and velar) and two glottal states (voiced and voiceless) in oral stops. The fricatives were more limited: there were no labial fricatives at this stage, and there was only one velar fricative, voiceless *[x]. PSl had one rhotic, one lateral, two nasal stops (labial and coronal), and two glides corresponding to the two colors of high vowels:

(35)

	labial	coronal	palatal	velar
oral stops	p b	t d		k g
fricatives		s z		x
nasal stops	m	n		
rhotic		r		
lateral		l		
glides	w		j	w

With a few quibbles (the phonemic status of glides, for example), the sound system described in this section is generally accepted as the single language from which the various Slavic dialects emerged.

1.5.3 The period of disintegration: pre-West Slavic

Over the course of the next few centuries, a period of time sometimes referred to as *the period of disintegration*, the Slavic-speaking region expanded, and the speech of eastern, southern, and western Slavs increasingly diverged as various sound changes crisscrossed throughout the area. While these sound changes were roughly universal, applying in most dialects identically or similarly, they were often shifted in time due to the delay inherent to spreading a change throughout such a large area. Even the relative chronology of two sound changes could differ from area to area, if the two sound changes originated in different parts of the Slavic region. The end result was the three distinct languages East Slavic, South Slavic, and West Slavic. When referring to the reconstructed language during the period of disintegration, I use the term *pre-West Slavic* (PWS), rather than following the more common practice of using PSl or Common Slavic to refer to everything from the beginning of PSl up to and including the period of disintegration. The use of the term PWS emphasizes the focus of the discussion on the sound changes during the period of disintegration as they applied to create West Slavic (and not East or South Slavic) during the period of disintegration. The other dialects during this period, pre-East Slavic and pre-South Slavic, are referred to for comparison when necessary.

During the period of disintegration, many significant and distinctive sound changes occurred, including the creation of the *jers*. Emerging as the result of a general shift from a vowel system based on length to one based on quality distinctions, the jers (often symbolized in work on Slavic with the Cyrillic letters ⟨ь⟩ and ⟨ъ⟩) began as the short high vowels *[i] and *[u] in PSl, and then lowered and laxed during the period of disintegration to *[ɪ] and *[ʊ]. By the end of the period of disintegration, the jers were ultimately either deleted or mutated by a combination of sound changes which are often grouped together as the so-called *fall of the jers*.

The fall of the jers began around the 10th century in the southern and western Slavic regions and swept north and east into modern-day Russia around the 13th century (Carlton 1991:76). The West Slavic languages are generally analyzed as having undergone the fall of the jers around the 11th century (Stieber 1968:4, Carlton 1991:171–172). Jers were deleted in numerous places, including at the end of words (Stieber 1968:5–7). The fall of the jers is regarded as the last sound change which affected all of the Slavic languages, and thus, it is often used as the dividing point between the period of disintegration and the major groupings of West Slavic, South Slavic, and East Slavic.

1.5.4 After the period of disintegration

I divide the development of Polish after the period of disintegration into the following approximate periods:

(36)	1000–1150	West Slavic	very little direct philological evidence exists; analyses based primarily on traditional historical reconstruction via comparative method (Čiževskij 1971:33, de Bray 1980:33,131)
	1150–1350	(East) Lechitic	first written words in Pope Innocent II's Papal Bull to the Archbishop of Gniezno (1136); evidence comes from religious texts, such as *Kazania Świętokrzyskie* (13th c.), and *Psałterz Floriański* and *Kazania Gnieźnieńskie* (both from the late 14th c.); some sound changes attributed to Lechitic might have only affected eastern dialects, but paucity of evidence makes determining Lechitic subdivisions difficult (de Bray 1980:230–231)
	1350–1500	Old Polish	evidence comes from mixture of secular documents such as Jakub Parkoszowic's *Trakut o ortografii polskiej [Treatise on Polish orthography]* (circa 1440) and religious documents, such as *Psałterz Puławski* (circa 1450) (de Bray 1980:231–232)
	1500–1750	Middle Polish	written record of Polish is greatly increased due to Renaissance and Age of Enlightenment; exemplified by works of Łukasz Górnicki, Jan Kochanowski, Mikołej Rej, Samuel Twardowski, among many others (Čiževskij 1971:73–104)
	1750–present	Modern Polish	revival of Polish after various wars; exemplified by Adam Mickiewicz, Cyprian Norwid, Bolesław Prus, Juliusz Słowacki, Saint Tremecki, etc. (Čiževskij 1971:105–197, Gotteri 1998:362)

1.5.5 Modern Polish

There are five main vowels in modern Polish, filling out a common triangular vowel space with three heights and two colors, except at the lowest height, where there is only one (central, neutral) vowel color:

(37)

	front-unround	central	back-round
high	i		u
mid	ɛ		ɔ
low		a	

Some examples of these vowels are given below:[12]

(38)				
	kɔɲ̟**i**	'horses (GEN)'	kɔɲ̟**u**	'horse (LOC)'
	tɕ**i**	'these (VIR)'	t**u**	'here'
	kɔɲ̟**ɛ**	'horses'	kɔɲ̟**ɔ**m	'horses (DAT)'
	t**ɛ**	'these (NON-VIR)'	t**ɔ**	'this (NEUT)'
	kɔɲ̟**a**	'horse (ACC)'		
	t**a**	'this (FEM)'		

In addition to the five oral vowels, Polish has two marginal nasal vowels (see Chapter 3 for discussion and analysis) and another high oral vowel, spelled ⟨y⟩. Modern Polish ⟨y⟩ is pronounced more or less as [ɪ], slightly lower and back than [i] (though some sources incorrectly claim that this vowel is pronounced centrally as [ɨ]). The two front high vowels essentially occur in complementary distribution: [i] can appear at the beginning of a word, after vowels, and after palatalized consonants (see below), while [ɪ] can only occur after plain (non-palatalized) consonants.

Modern Polish also has the phonemic consonant inventory in (39):

---

[12] Unless indicated otherwise, Polish forms cited in this dissertation are based on base forms from Jaślan and Stanisławski 1993 with inflections and declensions from Janecki 2000 (verbs) and Bielec 1998 (nouns and adjectives). Unmarked forms are given in the infinitive (verbs), the nominative singular (nouns), and the masculine nominative singular (adjectives). Listed pronunciations are standard and have been confirmed by native speakers of Polish.

(39)

	bilabial	labiodental	dental	alveolar	post-alveolar	alveolo-palatal	palatal	velar
oral stops	p b p^j b^j		t d					k g
fricatives		f v f^j v^j	s z		ʂ ʐ	 ɕ ʑ		x
affricates			ts dz		tʂ dʐ	 tɕ dʑ		
nasal stops	m m^j		n			 ɲ̟		ŋ
trill				r				
lateral				l				
glides	w						 j	w
nasal glide	w̃							w̃

For many manners of articulation, there are two rows of consonants. The upper series is generally called *plain* or *hard*, while the lower series is called *palatalized* or *soft*. Some sounds undergo a shift in the primary place of articulation and/or manner in palatalizing environments, so they are listed in a different column and/or row. For example, the palatalized version of dental [n] is an alveolo-palatal (or pre-palatal) [ɲ̟], not a palatalized dental [n^j]. There is also a variety of allophonic variation, especially before [i], where many sounds are pronounced more towards the palatal region. In particular, the velars are noticeably fronter before [i], pronounced as [k̟ g̟] (often transcribed as [c ɟ] in some sources).

For typographic simplicity, I do not mark the dental consonants with the IPA bridge diacritic [̪]. While most of the remaining consonant symbols above are standard, there are at least four different sets of symbols used in the literature for the post-alveolar fricatives (and the fricative half of the corresponding affricates), representing orthographic ⟨sz⟩ and ⟨rz/ż⟩. Most sources use non-IPA [š] and [ž] (Bidwell 1963, Stieber 1973, Rubach 1984a, Czaykowska-Higgins 1988, Bethin 1992) or IPA [ʃ] and [ʒ] (Spencer 1986, Dogil 1990, Jaślan and Stanisławski 1993), sometimes without remarking that the Polish post-alveolars are different from the post-alveolars in English that are often transcribed with the same symbols. However, the English post-alveolars are domed (the tongue is slightly raised behind the point of constriction), whereas the Polish post-alveolars are flat (Ladefoged and Maddieson 1996). Following more careful sources (e.g. Nowak 2001, and Hamann 2002a,b), I use the IPA retroflex symbols [ʂ] and [ʐ] to distinguish the Polish post-alveolars from [ʃ] and [ʒ], with the understanding that while the Polish fricatives are articulated in the same post-alveolar region of the roof of the mouth as other retroflex sounds, there is no curling of the tongue tip as seen in the retroflex stops of languages like Tamil (Ladefoged and Maddieson 1996). (Švarný and Zvelebil's (1955) 'cacuminal' could be used to describe non-curled retroflexes, but since the distinction is not necessary to the present work, I avoid this rather non-standard term.)

In the remainder of this dissertation, I give FDM-OT analyses of various instances of opacity, beginning in Chapter 2 with the alternation in Polish between [ɔ] and [u], which is triggered by underlying word-final oral voicing and is rendered opaque due to devoicing of word-final obstruents.

Chapter 2: The Story of ⟨ó⟩

The Modern Polish graphemes ⟨ó⟩ and ⟨o⟩ are the descendants of the same pre-West Slavic (PWS) vowel,*[ɔ]. In some environments, PWS *[ɔ] ultimately raised to modern [u] (so that it is perceptually indistinguishable from the historically unrelated vowel ⟨u⟩), while it retained its original quality of [ɔ] elsewhere. With the advent of printing in Poland in the 16th century, the acute accent ⟨ ́⟩ was borrowed from the Czech writing system to distinguish the two pronunciations of this vowel (Entwistle and Morison 1949:294, de Bray 1980:231):

(1)

pre-West Slavic pronunciation	*[ɔ]		*[u]
	↙ ↘		↓
Modern Polish grapheme	⟨o⟩	⟨ó⟩	⟨u⟩
	↓	↘	↙
Modern Polish pronunciation	[ɔ]	[u]	

Because some instance of [ɔ] and [u] derive from the same historical vowel, they alternate in various environments. What makes this alternation in back vowel height interesting is that the historical triggers for it have generally disappeared in the modern language, resulting in historical opacity. Most generative analyses of this alternation (Kenstowicz and Kisseberth 1979, Bethin 1992, Kenstowicz 1994, among others) have assumed that the back vowel alternation necessitates a synchronic explanation, requiring abstract intermediate forms that appear neither on the surface nor in the underlying form (see Chapter 1 for discussion of synchronic opacity). I contend that the alternation is not synchronically productive, but some explanation must still be given for why it is so pervasive in the lexicon.

In this chapter, I review the relevant synchronic generalizations concerning one instantiation of the back vowel alternation before word-final consonants (§2.1), bringing in further evidence from lexical exceptions in the modern vocabulary (§2.2), from my own recent experimental work on nonce forms (§2.3), and from the historical development of Polish phonology (§2.4), all of which suggest an alternative analysis with intermediate stages that are diachronic, but not synchronic. In §2.5, I provide a new analysis of the back vowel alternation within the FDM-OT framework developed in Chapter 1, which shows how sound changes are stored in the lexicon, allowing opacity to flourish without being synchronically productive. In §2.6, I discuss a recent proposal by Łubowicz (2001) which provides a synchronic analysis of the same alternation in back vowels. Finally, I conclude in §2.7 with a summary and discussion of the major results and outstanding problems of the FDM-OT analysis presented in this chapter.

2.1 Opacity in Modern Polish

In this section, I provide details of two phenomena in Modern Polish: the ban on voiced obstruents in word-final position and the back vowel alternation. I further discuss the opaque interaction of these two phenomena and the consequences of analyzing this type of opacity as a synchronically productive interaction.

2.1.1 Voicing in word-final codas

Polish, like many languages such as German, Hungarian, Russian, etc., does not allow word-final voiced obstruents (Kenstowicz and Kisseberth 1979, Rubach 1984a, Bethin 1992, etc.).[1] This generalization holds true of all Polish words, regardless of language of origin, morphological features, or grammatical category, as seen in (2):

(2)	*word-medial*	*word-final*	
	(voicing allowed)	*(voiceless only)*	*glosses*
	klu**b**ɪ	klu**p**	'club (PL/SG)'
	tɕɛka**v**ɪ	tɕɛka**f**	'ready (regular/short form)'
	kɔlɛ̃n**d**a	kɔlɛ̃n**t**	'Christmas carol (NOM SG/GEN PL)'
	dva ra**z**ɪ	ra**s**	'twice/once'
	talɛ**ʐ**ɛ	talɛ**ʂ**	'plate (PL/SG)'
	grɪ**ʑ**ɛʂ	grɪ**ɕ**	'bite (2SG/IMP)'
	bʐɛ**g**u	bʐɛ**k**	'edge (GEN/ACC)'

The restriction in (2) is obligatory (Rubach 1984a:208) and is so uncontroversial that it often receives little to no discussion on its own merits in research on Polish phonology (though its interaction with voicing assimilation has received a great deal of attention, as in Bethin 1992 and Lombardi 1991). For example, Gussmann (1980) does not include a rule of devoicing in his summary of phonological rules for Polish (pp. 133–135), though there is a passing mention of a rule of Final Unvoicing elsewhere in his analysis (p. 114). It seems clear that (2) is a completely robust and fully productive generalization which is synchronically active in Modern Polish. This is important to note, since the back vowel alternation is not so well-behaved.

---

[1] The actual environment varies somewhat by dialect. The southern and western dialects (such as the dialect spoken in Kraków) require word-final obstruents before sonorant- and vowel-initial words to be voiced. In contrast, the northern and central dialects (exemplified by the Warsaw dialect) have only voiceless obstruents in that position (Stieber 1973:116, Kenstowicz and Kisseberth 1979:418–419, Rubach 1984a:213 fn.6, Bethin 1992:183). I refer here simply to the more general word-final (more properly, pre-pausal) environment, abstracting away from dialectal variance.

2.1.2 The back vowel alternation

There are various paradigms in Polish which exhibit the back vowel alternation in which /ɔ/ surfaces as [u] before word-final, underlyingly voiced, oral consonants. I am primarily concerned with the masculine nominal paradigm (other paradigms, such as the feminine and neuter nominal paradigms, are discussed briefly at the end of this chapter). In many masculine nouns, the back vowel alternation is transparent in stems that end in oral sonorants (the liquids [r] and [l] and the glides [j] and [w]). When the sonorant is word-final, as it generally is in the nominative singular, a preceding [ɔ] is banned, emerging as [u] instead (3). In other forms in the paradigm, such as the nominative plural, typically formed by suffixing [-ɪ], [-i], or [-ɛ] to the nominative singular, the stem-final sonorant is no longer word-final, so [ɔ] can appear:

(3)

stem UR	NOM SG	NOM PL	gloss
/dv**ɔ**r/	dv**u**r	dv**ɔ**rɪ	'mansion'
/b**ɔ**l/	b**u**l	b**ɔ**lɛ	'ache'
/pɔk**ɔ**j/	pɔk**u**j	pɔk**ɔ**jɛ	'room'
/st**ɔ**w/	st**u**w	st**ɔ**wɪ	'table'

The data below show that this alternation involves raising of /ɔ/ to [u] rather than lowering of /u/ (4a) and that it is not triggered by voiceless consonants (4b) or nasal stops (4c):

(4)

	stem UR	NOM SG	NOM PL	gloss
a.	/z̦**u**r/	z̦**u**r	z̦**u**rɪ	'a kind of sour soup'
	/**u**l/	**u**l	**u**lɛ	'beehive'
	/v**u**j/	v**u**j	v**u**jɛ	'uncle'
	/m**u**w/	m**u**w	m**u**wɪ	'mule'

b.	/tʂ̹ɔp/	tʂ̹ɔp	tʂ̹ɔpɪ	‘peg’
	/kɔt/	kɔt	kɔtɪ	‘cat’
	/vwɔs/	vwɔs	vwɔsɪ	‘hair’
	/wɔç/	wɔç	wɔçɛ	‘elk’
	/sɔk/	sɔk	sɔk̟i	‘juice’
	/grɔx/	grɔx	grɔxɪ	‘pea’
c.	/dɔm/	dɔm	dɔmɪ	‘house’
	/tʂ̹ɔn/	tʂ̹ɔn	tʂ̹ɔnɪ	‘trunk’
	/kɔɲ̟/	kɔɲ̟	kɔɲ̟ɛ	‘horse’

Since most analyses of the back vowel alternation claim that the environment for raising is word-final oral voicing, and since voiced obstruents cannot appear word-finally (as seen in §2.1.1), there are two theoretical possibilities for the output of an underlying stem like /bɔb/ ‘bean’, depending on the interaction between the two generalizations seen in (2) and (3). In the *transparent* interaction, the output is [bɔp], with the back vowel alternation transparently blocked because there is no word-final oral voicing in the output to trigger it. In the *opaque* interaction, the output is [bup], with the back vowel alternation triggered by underlying word-final oral voicing that ultimately does not appear on the surface.

2.1.3 Opaque interaction

In many words, the back vowel alternation interacts opaquely with the ban on word-final obstruent voicing, with /bɔb/ ‘bean’ emerging as [bup]. The underlying word-final oral voicing (which appears unchanged in the plural [bɔbɪ]), seems to be triggering the back vowel alternation, despite not appearing on the surface:

(5)

	stem UR	*NOM SG*	*NOM PL*	*gloss*
	/b**ɔb**/	b**up**	b**ɔb**ɪ	‘bean’
	/r**ɔv**/	**ruf**	r**ɔv**ɪ	‘ditch’
	/l**ɔd**/	l**ut**	l**ɔd**ɪ	‘ice’
	/dɔv**ɔz**/	dɔv**us**	dɔv**ɔz**ɪ	‘supply’
	/n**ɔʐ**/	n**uʂ**	n**ɔʐ**ɛ	‘knife’
	/r**ɔg**/	r**uk**	r**ɔǧ**i	‘horn’

In a serial rule-based theory, this type of opaque interaction is easily constructed using ordered rules. The rules in (6) and (7) can account for the generalizations seen in the data in (2) and (3)–(5), respectively:

(6) Devoicing: [–sonorant] → [–voice] / ___ #

(7) ɔ-Raising: ɔ → u / ___ [C,–nasal,+voice] #

These rules must be serially ordered so that ɔ-Raising applies before Devoicing, as shown in (8), in order to produce the correct output (in this case, [bup] ‘bean’). If the rule order is reversed, as in (9), an incorrect output, ✗[bɔp], with no application of ɔ-Raising, is produced:

(8)			(9)		
	UR	/bɔb/		UR	/bɔb/
	ɔ-Raising	bub		Devoicing	bɔp
	Devoicing	bup		ɔ-Raising	—
	output	[bup]		output	✗[bɔp]

This is the type of opacity which proves to be problematic for monostratal OT (see §5.5 of Chapter 5 for discussion of some of the proposals designed to solve this problem in general; see §2.6 for a response to a recent proposal by Łubowicz (2001) which analyzes the same data in this chapter). In the following sections, I present evidence which suggests that the opaque back vowel alternation is not in fact synchronically productive and thus is not a problem for a monostratal OT framework such as FDM-OT.

2.2 Lexical exceptions

In both the native and loanword vocabulary of Modern Polish, there are many lexical exceptions to the back vowel alternation in both the transparent and the opaque case. The following words have [ɔ] instead of [u] in the nominative singular, despite having the proper triggering environment for the back vowel alternation (word-final oral voicing). In (9), the final consonants are all sonorants, while in (10), the final consonants are obstruents that are voiced in the plural. As shown in §2.1, all of these words should have [u] in the singular (as in the ungrammatical forms in the third column), but these data unquestionably have the seemingly anomalous [ɔ]:

(9)	*stem UR*	*NOM SG*	×*NOM SG*	*NOM PL*	*gloss*
	/p**ɔ**r/	p**ɔ**r	×p**u**r	p**ɔ**rɪ	'leek'
	/kɔl**ɔ**r/	kɔl**ɔ**r	×kɔl**u**r	kɔl**ɔ**rɪ	'card suit'
	/x**ɔ**l/	x**ɔ**l	×x**u**l	x**ɔ**lɛ	'lobby'
	/paras**ɔ**l/	paras**ɔ**l	×paras**u**l	paras**ɔ**lɛ	'umbrella'
	/kɔvb**ɔ**j/	kɔvb**ɔ**j	×kɔvb**u**j	kɔvb**ɔ**jɛ	'cowboy'
	/ɔɕ**ɔ**w/	ɔɕ**ɔ**w	×ɔɕ**u**w	ɔɕ**ɔ**wɪ	'donkey'
	/grutʂ̪**ɔ**w/	grutʂ̪**ɔ**w	×grutʂ̪**u**w	grutʂ̪**ɔ**wɪ	'gland'

(10)	/gl**ɔ**b/	gl**ɔ**p	×gl**u**p	gl**ɔ**bɪ	‘globe’
	/sn**ɔ**b/	sn**ɔ**p	×sn**u**p	sn**ɔ**bɪ	‘snob’
	/ɛpʲiz**ɔ**d/	ɛpʲiz**ɔ**t	×ɛpʲiz**u**t	ɛpʲiz**ɔ**dɪ	‘episode’
	/k**ɔ**d/	k**ɔ**t	×k**u**t	k**ɔ**dɪ	‘code’
	/nɛkrɔl**ɔ**g/	nɛkrɔl**ɔ**k	×nɛkrɔl**u**k	nɛkrɔl**ɔ**ɡ̟i	‘obituary’
	/prɔl**ɔ**g/	prɔl**ɔ**k	×prɔl**u**k	prɔl**ɔ**ɡ̟i	‘prologue’
	/rɛk**ɔ**rd/	rɛk**ɔ**rt	×rɛk**u**rt	rɛk**ɔ**rdɪ	‘record’
	/fʲ**ɔ**rd/	fʲ**ɔ**rt	×fʲ**u**rt	fʲ**ɔ**rdɪ	‘fjord’
	/x**ɔ**wd/	x**ɔ**wt	×x**u**wt	x**ɔ**wdɪ	‘homage’
	/tʂ̟**ɔ**wg/	tʂ̟**ɔ**wk	×tʂ̟**u**wk	tʂ̟**ɔ**wɡ̟i	‘tank’

Such examples are not hard to find.[2] The number of lexical exceptions, especially in recent loanwords, strongly suggests that the back vowel alternation is not synchronically productive, even in transparent environments (before word-final oral sonorants).

2.3 Experimental evidence against synchronic opacity

To further explore the synchronic productivity of the back vowel alternation, I conducted an experiment in which native speakers were asked to produce singular forms from nonce plurals, similar in form to the plurals in (10). The singular forms should be opaque if the back vowel alternation is synchronically productive, having [u] instead of [ɔ] despite the lack of final oral voicing on the surface. However, the results of the experiment, given in (11), are similar to the lexical data in (10), with no

[2] It is a potentially interesting fact that there are more lexical exceptions for stems ending in sonorants than stems ending in voiced obstruents. If the number of lexical exceptions is related to the length of time a generalization has been unproductive, then this suggests that the back vowel alternation before voiced obstruents persisted longer than before sonorants. Further research is required, and if this hypothesis holds, then the analysis presented in this chapter would have to be modified to allow for the back vowel alternation to be lost in two stages rather than one.

back vowel alternation in the expected environment (before a stem-final obstruent which has oral voicing in the plural):

(11)

	response	*unobserved opaque*	*stimulus*
	(NOM SG)	*response (✖NOM SG)*	*(NOM PL)*
	ʑnab**ɔ**t	✖ʑnab**u**t	ʑnab**ɔ**dı
	pʂak**ɔ**t	✖pʂak**u**t	pʂak**ɔ**dı
	ʂtap**ɔ**t	✖ʂtap**u**t	ʂtap**ɔ**dı
	ʂlap**ɔ**k	✖ʂlap**u**k	ʂlap**ɔ**ği
	ɕrab**ɔ**k	✖ɕrab**u**k	ɕrab**ɔ**ği
	smat**ɔ**k	✖smat**u**k	smat**ɔ**ği

The two subjects who took part in this experiment are: MJ, a male in his mid-thirties from Gliwice (in southern Silesia), who had been in the United States for over ten years at the time of the experiment; and KN, a female teenager from Warsaw, who had been in the United States for three months. The subjects were presented with the following types of sentences as aural stimuli, in which the underlined word, a nonce noun in the masculine nominative plural, was the only variable from sentence to sentence. The subjects were told in advance that the nonce words were to be thought of as imaginary forest creatures like elves or fairies who were helping poor John, who was lost and thirsty:

(12) Bardzo ładne źnabody dały Jankowi kawę, nie herbatę.
Bardzo ładne szlapogy dały Jankowi kawę, nie herbatę.
(etc.)
'The very pretty źnabods (szlapogs, ...) gave John coffee, not tea.'

The stimulus sentences were read by a 23-year-old female native speaker from Będzin (in western Małopolska, near Katowice), who had been in the United States for four years, recorded on a Sony Professional Walkman, and converted to WAV format for aural presentation to the subjects. The nonce words used in the stimuli consisted of the six plurals in (11), as well as the 18 similar plurals which differed by having a voiceless stem-final consonant ([ʐnabɔtɪ], etc.), by having [u] instead of [ɔ] ([ʐnabudɪ]), and by having both differences ([ʐnabutɪ]). These stimuli were each used twice in the experiment. In addition, there were 25 filler stimuli with different vowels and stem-final consonants to mask the purpose of the experiment, for a total of 73 stimulus sentences, which were presented in random order in four groups (of sizes 20, 20, 17, and 16), with a very short break between groups. Each of the 73 stimuli were presented three times in succession before the subjects responded.

The subjects were asked to say the sentence in (13) three times after the third repetition of a stimulus sentence, using the appropriate form of the nonce word in the blank. Responses were recorded on the Sony Professional Walkman and converted to WAV format for analysis.

(13) Jeden bardzo ładny ______ pożyczył Jankowi i pieniądze, i koszulę.
'One very pretty ______ lent John both money and a shirt.'

The only syntactically valid form that can go in the blank is the masculine nominative singular, which drops the vowel ending of the plural and creates the environment for both the back vowel alternation and for the ban on final voiced obstruents. Thus, these forms should be opaque if the back vowel alternation is productive: [ʐnabut], [ʂlapuk], etc. Impressionistically, no alternation is apparent in the responses; to my ears, the acoustic height of the response vowels (singular) is identical to that of the stimulus vowels (plural). However, precise phonetic measurements are warranted, as it is possible that some amount of statistically significant raising did in fact occur but could not be perceived.

F1, the strongest phonetic correlate of phonological vowel height, was measured for the final vowel for all of MJ's tokens of the relevant nonce words, using Boersma and Weenink's (1992/2003) Praat program. Measurements were taken as averages over a 10 ms range centrally selected from the F1 peak. These measurements were grouped into four rime families based on the final vowel and consonant of the stem, with the following means:

(14)	*family*	*description*	*mean F1*
	ɔtks	/ɔ/ followed by a voiceless obstruent	514.8 Hz
	ɔgdz	/ɔ/ followed by a voiced obstruent	511.4 Hz
	utks	/u/ followed by a voiceless obstruent	404.9 Hz
	udgz	/u/ followed by a voiced obstruent	418.4 Hz

The data from the four families were subjected to the Tukey method of multiple comparison, as described in Glass and Hopkins 1996, to test whether the differences between the means are statistically significant. The relevant statistic, the Studentized range statistic *q*, for each pairwise comparison of families is given in the table below:[3]

[3] The dashes indicate vacuous comparisons of a family with itself, which obviously results in $q = 0$. Note that multiple comparison is symmetric: if *X* is statistically different from *Y*, then *Y* is statistically different from *X*. Thus, the lower left cells are a mirror image of the upper right cells.

(15)

	ɔtks	ɔdgz	udgz	utks
ɔtks	—	0.367	10.278	11.717
ɔdgz	0.367	—	9.910	11.350
udgz	10.278	9.910	—	1.439
utks	11.717	11.350	1.439	—

The critical value of q is approximately 3.7, based on 73 data points, 4 families, and a confidence interval of $\alpha = 0.05$. If the value of q for a pair of families is greater than the critical value, as in the shaded boxes above, the F1 values for families are statistically different (that is, we can reject the null hypothesis H_0 = 'the two families in the pair have statistically identical F1 values') with 95% confidence (in fact, the null hypothesis can be rejected with 99% confidence, because the relevant critical value of q is about 4.6). It is clear that the ɔtks family and (crucially) the would-be opaque ɔdgz family both have statistically different F1 values from those in the utks and udgz families (all four of the relevant q values are 9.910 or greater, which is much larger than the critical value of 3.7). In addition, the F1 values for the ɔtks and ɔdgz families are statistically similar (q = 0.367, which is less than 3.7). Thus, the height of [ɔ] does not seem to change significantly in the expected environment.

Combined with the lexical evidence in the previous section, it would seem clear that the back vowel alternation is not productive in Modern Polish, and thus, there is no need to analyze this case of opacity with a synchronic grammar.

2.4 The relevance of historical change

In this section, I discuss the crucial sound changes in the history of Polish which led to the modern back vowel alternation. This section is important to the FDM-OT analysis in §2.5, since I treat the back vowel alternation as an instance of diachronic opacity which has been lexicalized. The following chart summarizes the sound changes described in this section, with a step-by-step derivation of the early Lechitic word *[bɔb] 'bean' given to help show how these changes operated in sequence (listed forms are from after the sound change applied).

(16)	Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
	Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
	Step 2b	Old Polish	1350–1500	first vowel raising	boːp
	Step 3	Middle Polish	1500–1750	vowel shortening	bop
	Step 4	Modern Polish	1750–present	second vowel raising	bup

This chart will be repeated in each section, with the relevant sound change under consideration highlighted.

2.4.1 Step 1: Lechitic vowel lengthening

After the fall of the jers (Proto-Slavic (PSl) short high vowels *[i] and *[u]) at the end of the period of disintegration (Chapter 1, §1.5.3), environments were created in which a variety of new sound changes could occur. In many Slavic languages, new long vowels were created in the closed syllables that resulted from the loss of a following jer. Because of the seemingly compensatory nature of this sound change, it is generally referred to as 'compensatory lengthening' (CL) in the literature (e.g. in Stieber 1968 and Carlton 1991). CL is usually described as a general process across the Slavic languages, with language to language variation in the conditioning environment. However, the nature of the lengthenings are so diverse that lumping them together as one process blurs the unique properties of each and runs the risk of causing confusion in determining the relative chronology of Slavic sound changes. When each language's lengthening is considered separately, a more accurate picture of CL is obtained. In the following discussion, I explain how instances of CL in various languages are different from each other, and I adopt a novel analysis of Lechitic 'CL' which actually has nothing to do with compensation for the loss of a jer, suggesting that it not be classified as part of some general CL process in Slavic.

In the South Slavic languages, only Serbo-Croatian and Slovene show signs of CL. In Serbo-Croatian, PSl vowels with a falling tone lengthened in newly closed syllables, while a rising tone on the vowel prevented CL. Eventually, the rising tone turned into a falling tone in some environments, opaquely masking the historical environment which conditioned the change. In the following derivations of the PSl words *[rǎːku] 'crab' and *[nâsu] 'nose' (adapted from Carlton 1991), the PSl contrast in tone between falling [ˆ] and rising [ˇ] evolves into a contrast in length in modern Serbo-Croatian. For comparison, the genitive singular forms, with no jer loss, and thus no opportunity for CL, are also given:

(17)

		'crab'		'nose'	
		NOM SG	*GEN SG*	*NOM SG*	*GEN SG*
a.	Proto-Slavic	*rǎːku	*rǎːkɑː	*nâsu	*nâsɑː
b.	⋮ (PSS)	*rǎkʊ	*rǎka	*nɔ̂sʊ	*nɔ̂sa
c.		*rǎk	*rǎka	*nɔ̂ːs	*nɔ̂sa
d.	Serbo-Croatian	râk	râka	nɔ̂ːs	nɔ̂sa

Early in pre-South Slavic (PSS), during the period of disintegration (17b), PSl length contrasts were lost, though later, new longs vowels were created based on the old PSl tones (17c). Where PSl had a falling tone, Serbo-Croatian has a long vowel. Note that even the PSl long vowel in *[rǎːku] emerges as a short vowel in Serbo-Croatian, since it shortened during PSS, and its historical rising tone resisted CL. Slovene seems to have undergone the same instance of CL, but other lengthenings make it difficult to tell whether true CL took place. However, as in Serbo-Croatian, vowels with a rising tone did not lengthen, suggesting that both languages underwent the same sound change (Carlton 1991:216).

In the East Slavic languages, only Ukrainian and southwest Belorussian have any type of CL, with the pre-East Slavic (PES) mid vowels *[ɛ] and *[ɔ] (from PSl *[æ] and *[ɑ]) diphthongizing to [jɛ] and [wɔ] respectively (in north Ukrainian) or both raising to [i] (in south Ukrainian and southwest Belorussian) when a following jer was lost. However, Carlton (1991:218) argues that this may not be an instance of CL at all, since there is evidence that CL needed to have occurred (or started to occur) prior to the fall of the jers in PES. He suggests that the mere presence of the jers in the following syllable raised the vowels to *[e] and *[o] through some sort of vowel harmony or metaphony, with a later process of diphthongization applying to these newly created mid vowels.

Similar lengthenings that occurred in many West Slavic languages have also been traditionally analyzed as instances of CL. As I show below, this is probably not an accurate classification for any West Slavic language except Upper Lusatian:

(18)

		'honey'		'nose'	
		NOM SG	*GEN SG*	*NOM SG*	*GEN SG*
a.	Proto-Slavic	*mædu	*mædɑw	*nɑsu	*nɑsɑː
b.	⋮ (PWS)	*mʲɛdʊ	*mʲɛdu	*nɔsʊ	*nɔsa
c.		*mʲɛːd	*mʲɛdu	*nɔːs	*nɔsa
d.	Upper Lusatian	mʲi̯e̯d	mʲɛdu	nʊs	nɔsa

In Upper Lusatian, PWS *[ɛ] or *[ɔ] in an initial syllable lengthened when a following jer was lost (18c), with subsequent sound changes modifying the newly long vowels to [i̯e̯] and [ʊ] (18d). This seems to be a case of true CL triggered solely by loss of a vowel (Carlton 1991:217–218).

In the remaining cases of purported CL in West Slavic, the nature of the consonant between the lengthened vowel and the lost jer played a crucial role that was not a part of CL in Upper Lusatian. In Czech, lengthening occurred in the development PSl *[ɑ] > PWS *[ɔ] > *[ɔː] > [u], only when the consonant between the vowel and the lost jer was a sonorant or voiced fricative (though lengthening before [v] is seen only in the genitive ending *[-ɑvu] > [-uv]), and only if the resulting word was monosyllabic. The actual environment is more complex, and there are many exceptions (Carlton 1991:219). In Slovak, which also has significant complications in defining the lengthening environment, the crucial consonant was similar to that in Czech (it had to be a sonorant, [z], or [z̦]), and only PWS *[ɛ] and *[ɔ] lengthened (Carlton 1991:220–221).

In Lechitic, an ancestral stage of Polish, there was lengthening of vowels in final closed syllables that had been created by the fall of the jers, but only when the final consonant was voiced (Stieber 1973:28, Carlton 1991:216–217):

(19)

	'table'		'house'	
	NOM SG	*NOM PL*	*NOM SG*	*NOM PL*
Proto-Slavic	*stɑlu	*stɑluː	*dɑmu	*dɑmuː
⋮ (PWS)	*stɔlʊ	*stɔlɨ	*dɔmʊ	*dɔmɨ
Lechitic	*stɔːl	*stɔli	*dɔːm	*dɔmi

	'bean'		'nose'	
	NOM SG	*NOM PL*	*NOM SG*	*NOM PL*
Proto-Slavic	*bɑbu	*bɑbuː	*nɑsu	*nɑsuː
⋮ (PWS)	*bɔbʊ	*bɔbɨ	*nɔsʊ	*nɔsɨ
Lechitic	*bɔːb	*bɔbi	*nɔs	*nɔsi

At a later stage, vowel length was lost before nasals, so in Modern Polish, the evidence for lengthening can only be seen before oral sonorants and underlyingly voiced obstruents. This connects directly to the back vowel alternation described in §2.1, since long Lechitic [ɔː] eventually raised to Modern Polish [u], while short [ɔ] remained unchanged.

Rather than attributing the Lechitic lengthening to 'compensation' for the loss of jers, I analyze it as the same type of vowel lengthening that occurs in English and many other languages with word-final voiced codas:

(20)	mɑp	'mop'	bɛt	'bet'	pɪk	'pick'
	mɑːb	'mob'	bɛːd	'bed'	pɪːg	'pig'

Once final jers had deleted in PSl, voiced consonants could appear in word-final codas, triggering the same lengthening in Lechitic as in English.[4]

Some vowels had also already been lengthened during the period of disintegration for other reasons (the so-called 'neo-acute' accent and intervocalic glide deletion), creating a general contrast between long and short vowels in most environments. As evidenced by Jakub Parkoszowic's *Trakut o ortografii polskiej [Treatise on Polish orthography]* (written circa 1440), in which double letters were used to represent long vowels, a general vowel length contrast persisted at least into the mid-15th century (Old Polish), with the stipulation that only long vowels could appear before word-final voiced codas (de Bray 1980:231).

[4] Given the similarity of the environments for the lengthenings in Czech, Slovak, and Polish, it is possible that West Slavic had a lengthening process affecting vowels before sonorants and (some) voiced fricatives, whereas Lechitic later lengthened vowels before all word-final voiced consonants. For simplicity, I only analyze and refer to Lechitic lengthening in this section, ignoring a possible previous West Slavic lengthening.

(21)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.4.2 Step 2a: Old Polish word-final obstruent devoicing

In the *Kazania Gnieźnieńskie [Gniezno Sermons]* of the late 14th century, occasional misspellings of the type ⟨bok⟩ for ⟨Bóg⟩ 'God' can be found, though correct spelling of final voiced obstruents was the general rule. By the 15th century, such misspellings were much more frequent (though of course, correct spellings were still often used, since writers likely tried to spell morphemes consistently across derived forms). These misspellings suggest that early in Old Polish and continuing through the 15th century, a sound change emerged which required word-final obstruents to be voiceless (Stieber 1968:77):

(22)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.4.3 Step 2b: Old Polish long mid vowel raising

About the same time that word-final obstruents were losing their voicing, the vowel qualities of long vowels were also changing. In particular, the long mid vowels *[ɛː] and *[ɔː] seem to have been pronounced somewhat higher (likely *[eː] and *[oː]), as evidenced by misspellings from the 14th century *Psałterz Floriański [Florian Psalter]* and other documents from the 14th and 15th centuries. These misspellings include the use of ⟨uo⟩ to represent long [oː], as in ⟨guor⟩ (cf. Modern Polish ⟨gór⟩ 'mountain'), and the use of ⟨y⟩ (which was a high vowel) to represent long [eː], as in ⟨gnywacz⟩ (Modern Polish ⟨gniewać⟩ 'to anger') (Stieber 1968:23). The evolution of the oral vowels can be seen in the following charts, with the changing long mid vowels shown in bold:

(23) *early Old Polish*

i iː		u uː
ɛ **ɛː**		ɔ **ɔː**
	a aː	

late Old Polish

i iː		u uː
eː		**oː**
ɛ		ɔ
	a aː	

It is possible that the short mid vowels also changed slightly, to *[ɛ̞] and *[ɔ̞], to make better use of the perceptual height space. However, since these vowels were able to reoccupy the mid space once the long vowels merged in quality with other vowels, we cannot see any effects of a hypothetical lowering of the short mid vowels. Thus, I will not use different symbols for the early and late Old Polish short mid vowels, and I will treat them as unchanged in height. The resulting pronunciation of 'bean' due to raising is [boːp]:

(24)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.4.4 Step 3: Middle Polish vowel shortening

The marking of vowel length slowly faded away during the late 15th century. A tiny half-page fragment of a 1516 Kraków printing of the Gospel of St. John contains the last known use of double letters to mark length. By the end of the 16th century, old vowel length was certainly completely eradicated, having been replaced in some instances by changes in vowel quality (Stieber 1968:20–21). The loss of the length contrast in Middle Polish oral vowels can be seen in the following charts:

(25)

early Old Polish			*late Old Polish*		
i iː		u uː	i		u
eː		oː	e		o
ɛ		ɔ	ɛ		ɔ
	a aː			a	

Thus, in late Middle Polish, there were at least seven oral vowels (more in some dialects which had lowered short *[i] and *[u] and/or backed long *[aː]), with no distinction in length. By this point, 'bean' was now pronounced [bop]:

(26)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.4.5 Step 4: Modern Polish higher mid vowel raising

In Kashubian and various Polish dialects, the reflex of Middle Polish *[o] is still pronounced differently from both [ɔ] and [u] (e.g. as [ʊ] in Kashubian). However, in the standard modern dialect, this vowel has fully merged with [u], giving us the modern pronunciation [bup] for 'bean':

(27)	Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
	Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
	Step 2b	Old Polish	1350–1500	first vowel raising	boːp
	Step 3	Middle Polish	1500–1750	vowel shortening	bop
	Step 4	Modern Polish	1750–present	second vowel raising	bup

Poets did not systematically rhyme Middle Polish *[o] with *[u] until the late 1800s, but the lack of rhymes prior to that may be simply due to traditional spelling. Thus, it is unclear exactly when *[o] raised to *[u], but it is most assuredly a modern sound change (Stieber 1968:41).

2.5 FDM-OT analysis of the back vowel alternation

2.5.1 Step 1: Lechitic vowel lengthening

Recall that due to the neo-acute accent and intervocalic glide deletion, long vowels could contrast with short vowels in West Slavic. Additionally, there was a voicing contrast in obstruent onsets and codas. Thus, the early grammar of Lechitic must be able to produce outputs with long and short vowels and with voiced and voiceless obstruents. I only consider the input containing CVC(V) words that begin with [b], followed by [ɔ] or [ɔː], followed by [p] or [b], and followed by an optional [i]. That is, the input is the subset of ▯ characterized by the expression [bɔ(ː){p,b}(i)], which includes words such as [bɔb] 'bean' and [bɔbi] 'beans'. These words will be examples throughout this section. Other factors like word length, consonantal place and manner, vowel quality, etc., do not matter for the later process of Lechitic lengthening, so they are not included in the input.

The following tableau shows how early Lechitic is derived, with the input emerging faithfully:

(28) *Step 1.1: early Lechitic (prior to lengthening)*

bɔp$_1$ bɔb$_5$ bɔːp$_2$ bɔːb$_6$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$	𝓕 voi	𝓕 dur	𝓜 V̆C̬#	𝓜 C̬#	𝓜 Vː
✓ a. bɔp$_1$ bɔb$_5$ bɔːp$_2$ bɔːb$_6$ bɔpi$_3$ bɔbɨ$_7$ bɔːpi$_4$ bɔːbɨ$_8$			×	×^{2}	×^{4}
b. bɔp$_1$ bɔːp$_2$ bɔːb$_{5,6}$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$		×!		×	×^{4}
c. bɔp$_{1,5}$ bɔːp$_{2,6}$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$	×^{2}!				×^{3}

Clearly, as long as all 𝓕-constraints are the highest ranked constraints, no input word changes in the output (28a), so there is a full contrast in vowel length and obstruent voicing. Candidate (28b), with lengthening of a vowel before word-final voiced consonants, violates high ranking 𝓕-duration in an attempt to satisfy the constraint 𝓜-V̆C̬# (which punishes short vowels before word-final voiced consonants).[5] Both 𝓜-V̆C̬# and 𝓜-C̬# (which punishes word-final voiced obstruents) are satisfied by (28c), which devoices all word-final obstruents, at the expense of 𝓕-voicing.

[5] It is not entirely clear exactly why vowels tend to be longer before voiced codas, especially in final syllables. However, this is a well-known cross-linguistic phenomenon, so I do not attempt to justify it here, relying on the cover constraint 𝓜-V̆C̬# to account for it.

Within the FDM-OT framework, strong lexicon optimization eventually occurs, so a speaker of early Lechitic creates underlying representations that are faithful to the outputs. Thus, the underlying representation for [bɔb] 'bean' will be /bɔb/, while for [bɔbi] 'beans' it is /bɔbi/. The effects of strong lexicon optimization are not crucial for this sound change because late Lechitic vowel lengthening is transparent. Sound change in FDM-OT is achieved through reranking of constraints in the constraint hierarchy. To obtain late Lechitic lengthening, represented by candidate (28b) with /bɔb/ → [bɔːb], the speaker's grammar must change so that ℳ-V̆C̬# is promoted over ℱ-duration (or equivalently, ℱ-duration is demoted under ℳ-V̆C̬#; I make no claims about the relative merits of constraint promotion and constraint demotion). Since this is still the grammar of a single speaker, constraint rankings form the early grammar do not change unless specifically required by a sound change. Thus, ℱ-voicing continues to dominate ℳ-V̆C̬#, and ℱ-duration continues to dominate ℳ-C̬# and ℳ-Vː in late Lechitic. The following tableau shows the results late Lechitic lengthening, with the revised constraint hierarchy selecting candidate (29b) as the output (note the visual convention used throughout this dissertation: the tableau for a late grammar has a 'shadow' to distinguish it from the unadorned tableau used for the early grammar):

(29) *Step 1.2: late Lechitic lengthening*

prom ⬇ (above 𝓜 V̆C̬#) *dem* ⬇ (above 𝓕 dur)

bɔp$_1$ bɔb$_5$ bɔːp$_2$ bɔːb$_6$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$	𝓕 voi	𝓜 V̆C̬#	𝓕 dur	𝓜 C̬#	𝓜 Vː
a. bɔp$_1$ bɔb$_5$ bɔːp$_2$ bɔːb$_6$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$		×!		×^{2}	×^{4}
✓ b. bɔp$_1$ bɔːp$_2$ bɔːb$_{5,6}$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$			×	×	×^{4}
c. bɔp$_{1,5}$ bɔːp$_{2,6}$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$	×^{2}!				×^{3}

The violations of each constraint remains the same as in (28). All that has changed is the relative ranking of the constraints in the hierarchy, causing (29a) to be ruled out by the newly promoted 𝓜-V̆C̬# constraint, and allowing (29b) with required lengthening of vowels before word-final voiced consonants, making early Lechitic [bɔb] be pronounced [bɔːb], with a long vowel:

(30)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.5.2 Step 2a: Old Polish word-final obstruent devoicing

As derived in the previous section, late Lechitic was characterized by a contrast in voicing and in vowel duration, except before word-final voiced consonants, where only long vowels are allowed. A language learner during early Old Polish would attempt to construct a constraint hierarchy that maps □ to the best approximation of the previous generation's outputs. As always, a reasonable subset of □ suffices for the purposes of linguistic analysis. I use the same input [bɔ(ː){p,b}(i)] used in the previous section for Lechitic:

(31) *Step 2a.1: early Old Polish (before devoicing)*

bɔp$_1$ bɔb$_5$ bɔːp$_2$ bɔːb$_6$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$	$\mathcal{F}$ voi	$\mathcal{M}$ V̆C̬#	$\mathcal{F}$ dur	$\mathcal{M}$ C̬#	$\mathcal{M}$ Vː
a. bɔp$_1$ bɔb$_5$ bɔːp$_2$ bɔːb$_6$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$		✗!		✗^{2}	✗^{4}
✓ b. bɔp$_1$ bɔːp$_2$ bɔːb$_{5,6}$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$			✗	✗	✗^{4}
c. bɔp$_{1,5}$ bɔːp$_{2,6}$ bɔpi$_3$ bɔbi$_7$ bɔːpi$_4$ bɔːbi$_8$	✗^{2}!				✗^{3}

The grammar that is derived for early Old Polish is similar to the previous generation's grammar (late Lechitic), with a few constraint pairs that are unranked. The constraint violations are the same, and the selection of (31b) as the output for early Old Polish follows similarly as the selection of (29b) for late Lechitic.

Because there is a change from the input to the output (the mapping of /bɔb/ to [bɔːb]), strong lexicon optimization will have an obvious effect that was not seen with Lechitic, with the underlying forms listed in the lexicon chosen because of their identity to the output. The crucial example is the word 'bean'. During the early Old Polish stage, the output used for comparison is that of the previous generation, late Lechitic [bɔːb]. As seen in (31), there are two inputs which map to this output, /bɔb/

and /bɔːb/, so both must be considered as possible underlying representations. By strong lexicon optimization, the most faithful input /bɔːb/ is chosen as the underlying representation and listed in the lexicon. But what about the plural form, [bɔbi]? There is only one input for it, /bɔbi/. By strong lexicon optimization, /bɔbi/ is stored in the lexicon as an underlying representation for ‘beans’ (or, /bɔb/ could be stored as a second underlying representation for ‘bean’, marked for usage in the plural; this is in fact the approach to morphology advocated in Chapter 4).

This differs from Prince and Smolensky’s (1993/2002) weak lexicon optimization with lexical minimization, in which ‘bean’ would only get one underlying representation, /bɔb/, to be used for both the singular and the plural. This is the crucial piece of the analysis, because the input for sound changes comes from the lexicon rather than the set of all words ▯, so only those forms stored by strong lexicon optimization will appear as inputs. Thus, since both /bɔb/ and /bɔːb/ map to the same output [bɔːb], only the faithful /bɔːb/ will be stored by strong lexicon optimization and used as an input word to later sound changes. The competing /bɔb/ is no longer considered as a possible input for ‘bean’. Under weak lexicon optimization, /bɔb/ is stored as the underlying representation and serves as the input to later sound changes. In particular, it undergoes word-final obstruent devoicing, resulting in transparent and ungrammatical ×[bɔp], a familiar problem when dealing with opacity.

There are two sound changes that must be accounted for in Old Polish: word-final devoicing and mid vowel raising. There is no clear evidence to determine which occurred first, or even if there was a strict chronological ordering. It is possible that some speakers acquired them in one order (devoicing first), others acquired them in the opposite order (raising first), and the rest acquired them simultaneously. For simplicity, I analyze them here one at a time, beginning with word-final devoicing, with the understanding that the analysis would work just as well if they were analyzed in the opposite order or simultaneously.

The significance of word-final devoicing to this research is the creation of opacity: vowels were required to be lengthened before word-final voiced consonants, but word-final obstruents could not be voiced. Modern □ó□, the spelling of the vowel that undergoes alternation, is derived from Old Polish *[ɔː]. Thus, the modern opacity that is apparent in the back vowel alternation has its origins in the Old Polish sound change of final devoicing. Devoicing can arise by promoting 𝓜-C̬# over 𝓕-voicing (the constraints 𝓜-V̆C̬# and 𝓕-duration are not violated by the relevant candidates, so they are not shown in this tableau):

(32) *Step 2a.2: late Old Polish devoicing*

bɔp$_1$ bɔːp$_2$ bɔːb$_5$ bɔpi$_3$ bɔbi$_6$ bɔːpi$_4$ bɔːbi$_7$	*prom* ⬇ 𝓜 C̬#	*dem* ⬇ 𝓕 voi	𝓜 Vː
a. bɔp$_1$ bɔːp$_2$ bɔːb$_5$ bɔpi$_3$ bɔbi$_6$ bɔːpi$_4$ bɔːbi$_7$	×!		×^{4}
✓ b. bɔp$_1$ bɔːp$_{2,5}$ bɔpi$_3$ bɔbi$_6$ bɔːpi$_4$ bɔːbi$_7$		×	×^{3}

The fully faithful candidate (32a) violates the newly promoted 𝓜-C̬# because it contains the word [bɔːb], with a word-final voiced obstruent. Candidate (32b) is selected as the output of this sound change because it satisfies 𝓜-C̬# (despite violating lower-ranked constraints), representing the word-final obstruent devoicing that occurred in late Old Polish, causing 'bean' to be pronounced [bɔːp]:

(33)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

Note that if weak lexicon optimization had been used instead of strong lexicon optimization, with the underlying representation of 'bean' stored as /bɔb/, then the sound change of devoicing would produce the wrong result, since devoicing of /bɔb/ causes it to map directly to [bɔp], with a short vowel. This is classic opacity: there is no reason for the [ɔ] in [bɔp] to be long since devoicing has altered the environment on the surface. Yet, both lengthening and devoicing need to apply, so that 'bean' is pronounced [bɔːp]. In FDM-OT, this is achieved by having the long vowel already encoded in the input due to strong lexicon optimization in the early grammar prior to the sound change that creates the late grammar. Thus, the relevant input for 'bean' is /bɔːb/, which maps to the correct [bɔːp] after devoicing. Weak lexicon optimization requires some sort of intermediate synchronic stage to get the proper output.

A prediction of strong lexicon optimization is that borrowings from other languages similar to /bɔb/ that entered Polish around this time (circa 1500) would be pronounced transparently without lengthening from an underlying voiced obstruent, rather than opaquely. That is, opacity cannot apply productively to new forms; it exists solely in the stored lexicon. I have not yet done a study to determine if this was the case, but given this analysis, it seems feasible that there are modern exceptions to the back vowel alternation that can trace their history to this period. The use of weak lexicon optimization does not predict the existence of such words, so if there were relevant borrowings during late Old Polish, it could weigh in as an advantage of strong lexicon optimization (and thus, FDM-OT) over weak lexicon optimization with lexical minimization.

2.5.3 Step 2b: Old Polish mid vowel raising

In addition to voicing and vowel length contrasts, late Lechitic had three levels of vowel height:

(34)

high	i		u
mid	ɛ		ɔ
low		a	

Vowel height is crucial in Old Polish because long mid vowels eventually raised to an intermediate height between high and mid, symbolized here by [eː oː]. The formal analysis here treats raising as a sound change which followed final devoicing, but as stated previously, this ordering is not crucial. Since height is relevant, it must be taken into account in the input. Vowel length is also important, but nothing else is required, since only specific values of vowel height and length triggered raising. For simplicity, I will just use inputs consisting of a single vowel, varying in height and duration.

First, a learner during early Old Polish must create a hierarchy which will produce late Lechitic outputs. I consider the following five input vowels, equally spaced in the acoustic dimension of height from low to high: [a ɒ̝ ɔ o u] (there are of course an unlimited number of possible vowels heights, but these five are sufficient to illustrate the general idea).[6] I assume that vowel length increases the quality of a height contrast, since it gives the speaker a longer time to hear the relevant cues to height. Thus, a pair of long vowels have a better contrast than their short counterparts. I further assume that a pair of vowels with one short and one long vowel contrast as well as a pair with two long vowels (that is, a single long vowel is sufficient to enhance the height contrast between two vowels; nothing crucial hinges on this assumption, which is designed only to simplify the analysis). The following

[6] I assume without analysis that the relevant back vowels must increase in rounding as they increase in height in order to satisfy perceptual demands on vowel color . Thus, the vowel just higher than [a] is [ɒ̝], rather than [ɐ] or [ɑ̝].

table lists all the relevant pairs of vowels, grouped by quality of height contrast and the $\mathcal{D}$-constraints they violate:

(35)

short pairs	*(short-)long pairs*	*ruled out by*
	[aː uː]	$\mathcal{D}_8$-height
[a u]		$\mathcal{D}_7$-height
	[aː oː] [ɒ̝ː uː]	$\mathcal{D}_6$-height
[a o] [ɒ̝ u]		$\mathcal{D}_5$-height
	[aː ɔː] [ɒ̝ː oː] [ɔː uː]	$\mathcal{D}_4$-height
[a ɔ] [ɒ̝ o] [ɔ u]		$\mathcal{D}_3$-height
	[aː ɒ̝ː] [ɒ̝ː ɔː] [ɔː oː] [oː uː]	$\mathcal{D}_2$-height
[a ɒ̝] [ɒ̝ ɔ] [ɔ o] [o u]		$\mathcal{D}_1$-height

In the FDM-OT tableau below, the ten input vowels are subjected to a constraint hierarchy that distills them down to the three heights in late Lechitic and early Old Polish. It is not crucial at this point to be concerned with which input vowel maps to which output vowel, so I have simply selected one possible way to map a five-level contrast to a three-level contrast (I assume $\mathcal{F}$-duration is undominated by these constraints, to ensure that all input vowels emerge with faithful vowel length):

(36) *Step 2b.1: early Old Polish (before raising)*

	a$_1$ ɒ̯$_2$ ɔ$_3$ o$_4$ u$_5$ aː$_6$ ɒ̯ː$_7$ ɔː$_8$ oː$_9$ uː$_{10}$	$\mathcal{D}_1$ hgt	$\mathcal{D}_2$ hgt	$\mathcal{F}$ hgt	$\mathcal{D}_3$ hgt	$\mathcal{M}$ ɔː
a.	a$_1$ ɒ̯$_2$ ɔ$_3$ o$_4$ u$_5$ aː$_6$ ɒ̯ː$_7$ ɔː$_8$ oː$_9$ uː$_{10}$	✗^{4}!	✗^{8}		✗^{11}	✗
✓ b.	a$_{1,2}$ ɔ$_3$ u$_{4,5}$ aː$_{6,7}$ ɔː$_8$ uː$_{9,10}$			✗^{4}	✗^{2}	✗
c.	a$_{1,2}$ u$_{3,4,5}$ aː$_{6,7}$ ɔː$_8$ uː$_{9,10}$			✗^{5}!		✗
d.	a$_{1,2}$ ɔ$_3$ u$_{4,5}$ aː$_{6,7}$ oː$_{8,9}$ uː$_{10}$		✗!	✗^{4}	✗^{3}	
e.	a$_{1,2}$ ɔ$_3$ u$_{4,5}$ aː$_6$ ɔː$_{7,8}$ oː$_9$ uː$_{10}$		✗^{2}!	✗^{3}	✗^{4}	✗
f.	a$_1$ ɔ$_{2,3}$ o$_4$ u$_5$ aː$_6$ ɔː$_{7,8}$ oː$_9$ uː$_{10}$	✗^{2}!	✗^{4}	✗^{2}	✗^{6}	✗

The completely faithful candidate (36a) fatally violates the highly ranked $\mathcal{D}_1$-height constraint which penalizes all of the closest vowel pairs, such as [a ɒ̯]. Since (36a) is faithful to the input, it contains all of these pairs. (36f) only contains two of the poorly contrastive pairs, but the remaining candidates (36b–e) do not contain any. Both candidates (36d) and (36e) contain the next set of poor contrasts, this time in the long vowels, exemplified by the pair [oː uː]. Neither (36b) nor (36c) contain these pairs, so they are the only candidates left in competition. The next constraint in the hierarchy is $\mathcal{F}$-height, so the most faithful of the remaining candidates will be selected as the output. Candidate (36b) only has four mergers of input vowels, while (36c) has five. Thus, candidate (36b) is selected as the output for early Old Polish.

As always in FDM-OT, strong lexicon optimization stores the output as the input. This means that the underlying representation for [bɔːp] 'bean' is /bɔːp/ and that the input to late Old Polish will be (36b), the output of early Old Polish.

In order to derive raising of *[ɔː] to *[oː], I follow claims by Donegan (1978) and Archangeli and Pulleyblank (1994) that long tense vowels are more harmonic than long lax vowels, thus motivating part of the ranking in the following articulatory markedness hierarchy (the remainder of the ranking comes from the special difficulty of low round vowels, and the affinity between lowness and duration):

(37) $\mathcal{M}$-ɔː, $\mathcal{M}$-ɒ̝ː ≫ $\mathcal{M}$-uː, $\mathcal{M}$-oː ≫ $\mathcal{M}$-aː

To achieve raising, $\mathcal{M}$-ɔː must be promoted over $\mathcal{F}$-height:

(38) *Step 2b.2: late Old Polish raising*

	$\mathcal{D}_1$ hgt	*prom* ↓ $\mathcal{M}$ ɔː	*dem* ↓ $\mathcal{D}_2$ hgt	$\mathcal{F}$ hgt	$\mathcal{D}_3$ hgt
a$_1$ ɔ$_2$ u$_3$ aː$_4$ ɔː$_5$ uː$_6$					
a. a$_1$ ɔ$_2$ u$_3$ aː$_4$ ɔː$_5$ uː$_6$		×!			×^{2}
b. a$_1$ u$_{2,3}$ aː$_4$ ɔː$_5$ uː$_6$		×!		×^{2}	
✓ c. a$_1$ ɔ$_2$ u$_3$ aː$_4$ oː$_5$ uː$_6$			×^{2}	×	×^{4}
d. a$_1$ ɔ$_2$ u$_3$ aː$_4$ uː$_{5,6}$			×^{2}	×^{2}!	×^{4}

The fully faithful candidate (38a) violates the newly high ranked constraint $\mathcal{M}$-ɔː, which bans the long lax mid vowel [ɔː]. Candidate (38b), which eliminates short [ɔ], but not the undesirable long [ɔː], also violates this constraint. Candidates (38c) and (38d) eliminate the offending vowel, but in different ways. The slightly more faithful (38c) only changes [ɔː] to [oː], a one step change in height, while (38d) changes [ɔː] to [uː], a more severe violation of $\mathcal{F}$-height. Thus, the winning candidate is (38c), which represents the late Old Polish raising of [ɔː] to [oː], as in the pronunciation of 'bean' as [boːp]:

(39)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.5.4 Step 3: Middle Polish vowel shortening

The loss of long vowels in Middle Polish has a straightforward analysis. In early Middle Polish, both long and short vowels were allowed. The following constraint ranking (much of which is the same as for late Old Polish) derives the correct vowel system. By richness of the base, the input during this early phase of acquisition is □ (here, the same subset consisting of ten vowels used for early Old Polish):

(40) *Step 3.1: early Middle Polish (before shortening)*

	$\mathcal{D}_1$ hgt	$\mathcal{M}$ ɔː	$\mathcal{F}$ hgt	$\mathcal{F}$ dur	$\mathcal{M}$ Vː	$\mathcal{D}_2$ hgt
a$_{1}$ ɒ̯$_{2}$ ɔ$_{3}$ o$_{4}$ u$_{5}$ aː$_{6}$ ɒ̯ː$_{7}$ ɔː$_{8}$ oː$_{9}$ uː$_{10}$						
a. a$_{1}$ ɒ̯$_{2}$ ɔ$_{3}$ o$_{4}$ u$_{5}$ aː$_{6}$ ɒ̯ː$_{7}$ ɔː$_{8}$ oː$_{9}$ uː$_{10}$	✗^{4}!	✗!			✗^{5}	✗^{8}
✓ b. a$_{1,2}$ ɔ$_{3}$ u$_{4,5}$ aː$_{6,7}$ oː$_{8,9}$ uː$_{10}$			✗^{4}		✗^{3}	✗
c. a$_{1,2,6,7}$ ɔ$_{3,8}$ u$_{4,5,9,10}$			✗^{4}	✗^{5}!		
d. a$_{1,2}$ ɔ$_{3}$ u$_{4,5}$ aː$_{6,7}$ uː$_{8,9,10}$			✗^{6}!		✗^{2}	
e. a$_{1,2}$ ɔ$_{3}$ u$_{4,5}$ aː$_{6,7}$ ɔː$_{8}$ uː$_{9,10}$		✗!	✗^{4}		✗^{3}	

The fully faithful candidate (40a) contains vowel pairs such as [o u] that are too close perceptually, violating highly ranked $\mathcal{D}_1$-height. Simply eliminating those pairs is not sufficient, as candidate (40e) shows, incurring a violation of $\mathcal{M}$-ɔː for having the long lax mid vowel [ɔː]. Candidates (40b–d) satisfy these two high ranked constraints, at the expense of faithfulness to the input with respect to vowel height. Candidate (40d) is the worst offender because it changes both long mid vowels /ɔː/ and /oː/ to a high vowel, while candidates (40b) and (40c) only change the height of one of the mid vowels. These two candidates differ with respect to length preservation: candidate (40b) is faithful to input vowel length, while candidate (40c) shortens all of the vowels. Since $\mathcal{F}$-duration outranks $\mathcal{M}$-Vː, candidate (40b) is the output. Strong lexicon optimization then stores this output as the underlying lexicon, allowing it to serve as the input to the late Middle Polish sound change of vowel shortening.

Vowel shortening is analyzed by promoting $\mathcal{M}$-Vː and $\mathcal{F}$-height over $\mathcal{D}_1$-height (which already outranks $\mathcal{F}$-duration by virtue of the ranking from the early grammar):

(41) *Step 3.2: late Middle Polish shortening*

	prom ↑	*prom* ↑	*dem* ↓		
a$_1$ ɔ$_2$ u$_3$ aː$_4$ oː$_5$ uː$_6$	$\mathcal{M}$ Vː	$\mathcal{F}$ hgt	$\mathcal{D}_1$ hgt	$\mathcal{F}$ dur	$\mathcal{D}_2$ hgt
a. a$_1$ ɔ$_2$ u$_3$ aː$_4$ oː$_5$ uː$_6$	×³!				×
✓ b. a$_{1,4}$ ɔ$_2$ o$_5$ u$_{3,6}$			×²	×³	×²
c. a$_{1,4}$ ɔ$_2$ u$_{3,5,6}$		×!		×³	
d. a$_{1,4}$ ɔ$_{2,5}$ u$_{3,6}$		×!		×³	

The faithful candidate (41a) cannot surface because of the newly high ranked $\mathcal{M}$-Vː, which prohibits all long vowels. Candidates (41b–d) have no long vowels, satisfying $\mathcal{M}$-Vː. Candidate (41b) is more faithful to the input vowel height than candidates (41c) and (41d), and since $\mathcal{F}$-height is ranked high, candidate (41b) is the output of late Middle Polish, reflecting the shortening sound change that occurred, changing the pronunciation of 'bean' to [bop]:

(42)

Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
Step 2b	Old Polish	1350–1500	first vowel raising	boːp
Step 3	Middle Polish	1500–1750	vowel shortening	bop
Step 4	Modern Polish	1750–present	second vowel raising	bup

2.5.5 Step 4: Modern Polish higher mid vowel raising

By richness of the base, the four-height vowel system [a ɔ o u] from late Middle Polish must be derived in early Modern Polish from □, the set of all possible words (in this case, I consider only the subset of □ consisting of the vowels [a ɒ̞ ɔ o u]). This can be achieved with the constraint ranking shown in the following tableau:

(43) *Step 4.1: early Modern Polish (before raising)*

	a$_1$ ɒ̞$_2$ ɔ$_3$ o$_4$ u$_5$	𝓜 ɒ̞	𝓕 hgt	𝓓$_1$ hgt	𝓜 a,...
a.	a$_1$ ɒ̞$_2$ ɔ$_3$ o$_4$ u$_5$	✗!		✗^{4}	✗^{4}
✓ b.	a$_{1,2}$ ɔ$_3$ o$_4$ u$_5$		✗	✗^{2}	✗^{4}
c.	a$_{1,2}$ ɔ$_3$ u$_{4,5}$		✗^{2}!		✗^{3}
d.	a$_{1,2}$ ɔ$_{3,4}$ u$_5$		✗^{2}!		✗^{3}

The fully faithful candidate (43a) violates the highly ranked constraint 𝓜-ɒ̞ because it contains the low rounded vowel [ɒ̞], which is the most marked of the five vowels shown (the low jaw height required for low vowels makes it especially difficult to sufficiently round the lips). The 𝓜-constraints which rule out the other vowels are represented by the cover constraint 𝓜-a,…, which is ranked lower than 𝓜-ɒ̞. The remaining candidates (43b–d) satisfy high ranking markedness by eliminating the marked vowel /ɒ̞/, merging it with the less marked /a/. Such merger violates 𝓕-height of course, but since it is ranked lower than 𝓜-ɒ̞, all candidates must be unfaithful to some extent. Candidate (43b) only violates 𝓕-height once, leaving the remaining four vowels unchanged, while candidates (43c) and (43d) violate it twice in an attempt to better satisfy 𝓓$_1$-height by eliminating /o/, which is perceptually close to both [ɔ] and [u]. Since 𝓕-height is ranked higher than 𝓓$_1$-height, the more faithful but less dispersed candidate (43b) is selected as the output. Strong lexicon optimization then stores this output as the underlying lexicon for use as the input to later sound changes. This affects the word for ‘bean’ by storing it as /bop/.

Because [o] is equidistant between [ɔ] and [u] with respect to perceptual height, it is not clear what would cause /o/ to raise to [u] rather than lower to [ɔ]. Since this is exactly what happened in Modern Polish, some other factor besides height must be at work. I assume here that the relevant factor is tenseness; that is, /o/ changes to [u] and not [ɔ] in order to preserve its tenseness. While the property of tenseness does intermingle with height (and thus, might appear to merely be a subset of height), it also has its own properties and is often the target of generalizations irrespective of height (e.g. some vowel harmonies, as in Tangale (Jungraithmayr 1971, Kidda 1985, van der Hulst and van de Weijer 1995), ensure that harmonizing vowels have the same tenseness). Thus, it seems reasonable to posit a constraint $\mathcal{F}$-tenseness that is independent of $\mathcal{F}$-height. This constraint determines which direction /o/ goes when $\mathcal{D}_1$-height is reranked over $\mathcal{F}$-height to pare down the vowel inventory from four to three in late Modern Polish:

(44) *Step 4.2: late Modern Polish raising*

	$\mathcal{D}_1$ hgt (*prom* ↑)	$\mathcal{F}$ hgt (*dem* ↓)	$\mathcal{F}$ tense
a_1 $ɔ_2$ o_3 u_4			
a. a_1 $ɔ_2$ o_3 u_4	✗²!		
✓ b. a_1 $ɔ_2$ $u_{3,4}$		✗	
c. a_1 $ɔ_{2,3}$ u_4		✗	✗!

The fully faithful candidate (44a) is not sufficiently dispersed, violating $\mathcal{D}_1$-height because of the vowel pairs [ɔ o] and [o u]. The remaining two unfaithful candidates (44b) and (44c) are equally unfaithful for vowel height, changing /o/ one step by raising it to [u] (44b) or lowering it to [ɔ] (44c). However, they do differ with respect to tenseness: candidate (44b) preserves the tenseness of /o/ by mapping it to [u], while candidate (44c) changes tense /o/ to lax [ɔ], fatally violating $\mathcal{F}$-tense. Thus, candidate (44a) is selected as the output, and early Modern Polish [o] maps to [u], making the final pronunciation of 'bean' [bup] through raising of the vowel:

(45)	Step 1	Lechitic	1100–1350	vowel lengthening	bɔːb
	Step 2a	Old Polish	1350–1500	word-final devoicing	bɔːp
	Step 2b	Old Polish	1350–1500	first vowel raising	boːp
	Step 3	Middle Polish	1500–1750	vowel shortening	bop
	Step 4	Modern Polish	1750–present	second vowel raising	bup

2.6 Comparison with Contrast Preservation Theory

One recent proposal to account for the back vowel alternation is Łubowicz's (2001) Contrast Preservation Theory (CPT). Like FDM-OT, CPT uses sets of words, rather than individual words, as inputs and outputs. Łubowicz analyzes the back vowel alternation synchronically, treating it as a preservation or redistribution of contrast, compensating for the surface loss of the underlying voicing contrast in word-final obstruents. Her analysis requires PRESERVE CONTRAST (PC) constraints which require that input contrasts be optimally preserved in the output:[7]

[7] Her full analysis is a bit more complicated than the version I present here, as Łubowicz distinguishes between input- and output-oriented PC constraints. The distinction is not relevant here, and I collapse the two types into one.

(46) PRESERVE CONTRAST-*P*

For every pair of input words *x* and *y* that contrast for property *P*, count a violation if *x* and *y* map to the same word in the output. Further, for every word in the output that corresponds to a contrastive pair in the input, also count a violation (in other words, minimize the number of words which are the result of merger by 'piling up' the mergers onto a single output).

The effects of PC constraints can be seen in the following tableau, which represents a synchronic analysis of the back vowel alternation in Polish, adapted from Łubowicz (2001):

(47) *Contrast Preservation Theory: Modern Polish back vowel alternation*

bɔp₁ bup₃ bɔb₂ bub₄	𝓜 Ç̬#	PC voi	PC high
a. bɔp₁ bup₃ bɔb₂ bub₄	✗²!		
✓ b. bɔp₁ bup₂,₃,₄		✗³	✗²
c. bɔp₁,₂ bup₃,₄		✗⁴!	

Candidate (47a) violates high ranking 𝓜-Ç̬# by allowing word-final voiced obstruents in the words [bɔb] and [bub]. Candidates (47b) and (47c) bypass this constraint by merging the relevant inputs to words with voiceless obstruents, thereby losing the underlying voicing contrast. This loss of contrast incurs violations of PC-voice. Both candidate (47b) and candidate (47c) incur two violations of PC-voice due to the lost contrasts (one each for the pairs /bɔp bɔb/ and /bup bub/). In addition, violations must be counted for the number of words in the output used as mergers for the voicing contrast. In candidate (47b), the merging inputs only merge to one word, [bup], incurring one extra violation of PC-voice, for a total of three. But in candidate

(47c), both [bɔp] and [bup] are used for merging inputs, adding two violations, for a fatal total of four. Thus, candidate (47b) defeats the more balanced candidate (47c) because of the nature of PC constraints: neutralized contrasts prefer to merge to the same output.

There are at least four problems facing the CPT analysis that need to be addressed. First, there is no satisfactory way to account for the back vowel alternation before oral sonorants, since there is no loss of voicing contrast to trigger raising of /ɔ/ to [u]:

(48) *Contrast Preservation Theory: the problem with sonorants*

bɔp₁ bup₃ bɔb₂ bub₄ bɔr₅ bur₆	$\mathcal{M}$ C̬#	PC voi	PC high
☠ a. bɔp₁ bup₂,₃,₄ bɔr₅ bur₆		✗³	✗²
☹ b. bɔp₁ bup₂,₃,₄ bur₅,₆		✗³	✗⁴!

The losing candidate (48b) represents raising of /ɔ/ to [u] before word-final oral sonorants and voiced obstruents, while (48a) only has raising before voiced obstruents. If the back vowel alternation is synchronically productive, then (48b) should be the output. However, it incurs extra violations of PC-high due to the unmotivated loss of the height contrast between /bɔr/ and /bur/ (one for the lost input contrast, and one for the output word which hosts the merge). As Łubowicz (personal communication) points out, this problem could be solved by introducing a constraint that bans surface [ɔ] (but not other vowels!) before word-final oral sonorants. Of course, such a constraint is purely stipulative, as no other language besides Polish seems to make use of it. What is the motivation for such a constraint? Why should it only apply to [ɔ] and not other vowels?

The second problem with CPT is that it has no inherent locality: it cannot explain why other instances of [ɔ] in the word do not raise in order to preserve the underlying voicing contrast that is lost word-finally. That is, why does /bɔbɔb/ shift to [bɔbup] and not to [bubɔp]? The voicing contrast is just as well preserved, and the constraint violations are the same:

(49) *Contrast Preservation Theory: the problem with locality*

	bɔbɔp$_1$ bɔbɔb$_2$	bɔbup$_3$ bɔbub$_4$	bubɔp$_5$ bubɔb$_6$	bubup$_7$ bubub$_8$	$\mathcal{M}$ C̬#	PC voi	PC high
✓ a.	bɔbɔp$_1$	bɔbup$_{2,3,4}$	bubɔp$_5$	bubup$_{6,7,8}$		✘^{6}	✘^{4}
✓ b.	bɔbɔp$_1$	bɔbup$_3$	bubɔp$_{2,5,6}$	bubup$_{4,7,8}$		✘^{6}	✘^{4}

The two candidates are identical, except that in candidate (49a), which represents actual Modern Polish, the merger occurs in the second syllable, while in candidate (49b), the merger occurs in the first syllable. While this could be solved in this particular case by increasing PC violations gradiently by distance from the point of lost contrast, it does not solve the problem generally if the lost contrast occurs word-medially, and the merger could occur on either side.

A third problem is determining the direction of the shift. Consider the following tableau:

(50) *Contrast Preservation Theory: the problem with direction*

bɔp₁ bup₃ bɔb₂ bub₄	𝓜 C̬#	PC voi	PC high
✓ a. bɔp$_1$ bup$_{2,3,4}$		✗^{3}	✗^{2}
✓ b. bɔp$_{1,2,3}$ bup$_4$		✗^{3}	✗^{2}

Both candidates fare equally with respect to every relevant constraint, since they both lose the same number of contrasts and merge them to the same number of outputs, with the same relative distribution. There is nothing inherent to CPT to explain why /bɔb/ should raise instead of /bub/ lowering.

The fourth problem with CPT is that it fails to explain why there are lexical exceptions to the back vowel alternation (§2.2) and why native speakers do not apply it productively to nonce forms (§2.3). In FDM-OT, this falls out straightforwardly from direct mapping: when the alternation was rendered opaque in Old Polish, it ceased being productive because direct mapping by nature cannot allow productive opacity. Any new words borrowed after that time should fail to alternate, adding non-alternating forms to the lexicon. In addition, once the alternation ceased to be productive, it could not be applied to nonce forms. These predictions of FDM-OT are supported by the actual facts discussed in §2.2–2.3, and any analysis such as CPT which seeks to analyze the back vowel alternation as synchronically productive must explain these facts.

See also §5.5 of Chapter 5 for discussion of other frameworks which have been used to analyze opacity.

2.7 Conclusion

In this chapter, I have supplied data from both lexical exceptions and experiments on the phonology of nonsense words which suggest that the back vowel alternation in Modern Polish is not synchronically productive. Thus, its opaque interaction with word-final obstruent devoicing is not a problem for monostratal OT, which cannot account for synchronic opacity.

Since the back vowel alternation is still pervasive in the extant Polish lexicon, I have constructed an analysis of the alternation based on its historical origins within the framework of FDM-OT. One novel piece of my analysis is strong lexicon optimization, which selects underlying representations that are phonologically identical to their outputs. By having strong lexicon optimization interspersed between serially ordered diachronic sound changes, the analysis maintains the serialism required to account for opacity without sacrificing monostratality in the synchronic grammar. The trade-off is an increased burden on lexical storage. It remains to be seen whether this burden is less desirable than a multistratal synchronic grammar, but it does achieve the result of preserving opacity in the lexicon while keeping it from being synchronically productive (cf. Buckley 2001, which utilizes a lexical solution based on underspecification).

There are a few outstanding issues that need further consideration. The back vowel alternation also occurs in feminine and neuter nouns, as in [pagɔda] 'pagoda (NOM SG)' with [ɔ], but [pagut] '(GEN PL)', with the same trigger for the alternation: an underlying word-final voiced oral consonant. This alternation is supposedly productive (Anna Łubowicz and Jerzy Rubach, personal communication). However, I conducted a short informal experiment in which a native speaker of Polish was asked to produce the genitive plural of a few nonce nominative singular feminine and neuter nouns which should exhibit the back vowel alternation. The results were consistent with the results of the experiment in §2.3: no alternation was observed, with [ɔ] appearing in all of the subject's responses. In addition, there are a few lexical exceptions (though notably, not nearly as many as there are for masculine nouns). Clearly, a full experiment should be conducted to confirm whether or not the back vowel alternation is synchronically productive for feminine and neuter nouns (Baranowski and Buckley (2003) conducted an experiment and found that the alternation occurs to some extent for nonce words that resemble actual words that do undergo the alternation. However, they did not test nonce forms that are similar to no extant words, so it is difficult to separate true productivity from some form of real-time analogy based on similarity). Confounding the issue is that some nouns show the alternation without the conditioning environment, so an experiment to determine the productivity of the back vowel alternation must also be sensitive to possibility that it might not be triggered (only) by underlying word-final voiced oral consonants.

A final problem arises from the analysis of raising in late Modern Polish. Why did the front vowel [e] lower to [ɛ] rather than raise to [i] (though note that in some dialects, [ɛ] did raise to [i])? One solution could rely on the differences in height between front vowels and back vowels. The front vowels are farther apart in height than the back vowels are, so [e] would have farther to travel to raise all the way to [i], while [o] has a slightly shorter distance. This solution would require that the initial raising that took place in Old Polish kept [e] closer to [ɛ], rather than making it equidistant between [ɛ] and [i]. Further study is required to adequately solve this problem.

Chapter 3: The Evolution of the Nasal Vowels

By the end of the period of disintegration (Chapter 1, §1.5.3), there were two nasal vowels, *[ɛ̃] and *[ɔ̃]. In most Slavic languages, the nasal vowels were eventually denasalized, leaving no trace of their historical nasal character. However, in Polish, the nasal vowels have continued to maintain some degree of nasality. In Lechitic, the two nasal vowels merged into a single central colorless nasal vowel, low *[ã] (or as I argue in §3.3.2, mid central *[ɜ̃]). As with all vowels during this time period, the single nasal vowel was subject to contextual lengthening, resulting in the same type of quantity-based alternation discussed in Chapter 2. Because this alternation was originally based on vowel length, it was triggered by (among other things) word-final voiced consonants. Thus, as word-final obstruents devoiced in Old Polish, the nasal vowel alternation was rendered opaque, just like the back vowel alternation. In addition, the nasal vowels underwent yet another color shift. This time, the long nasal vowel backed and rounded to its modern color [ɔ̃], while the short nasal vowel fronted to its modern color [ɛ̃], so that the opaque nasal vowel alternation became one of color. This alternation in nasal vowel color can be seen today in Modern Polish.

In this chapter, I begin by describing one instance of the nasal vowel alternation in Modern Polish (§3.1), followed by a discussion of numerous lexical exceptions in the modern vocabulary (§3.2). In §3.3, I provide background on the specific historical sound changes which led to the modern nasal vowel alternation, and then I provide an analysis of the alternation within the FDM-OT framework (§3.4). In particular, I show how FDM-OT is especially suited to an analysis of the Polish nasal vowels in two ways. First, the opaque nature of the alternation is analyzed by the interaction of strong lexicon optimization with sound change, which accounts both for the existence of the alternation in the many lexical items and for lack of synchronic productivity. Second, the merger and subsequent split in nasal vowel color is easily and naturally explained with $\mathcal{D}$-constraints and the treatment of inputs and candidates as sets of words. As I show in §3.5, a framework like standard OT, which treats candidates as individual words, has difficulty explaining both the merger and the split of the nasal vowels. Finally, I summarize and discuss the major points and implications of this analysis in §3.6.

3.1 Opacity in Modern Polish

The Modern Polish nasal vowels, spelled ⟨ę⟩ and ⟨ą⟩, are usually pronounced as nasal diphthongs consisting of a nasalized mid vowel, [ɛ̃] and [ɔ̃] respectively, followed by a nasal stop or nasal glide (both symbolized here generally as ‘N’).[1] As the following data show, the underlying front nasal vowel is pronounced as back and round [ɔ̃N] when it is followed by an underlyingly voiced consonant at the end of the word. When the following voiced consonant is not word-final, the underlying front nasal vowel is pronounced faithfully as [ɛ̃N]:

[1] Nasalization on the vowel is optional, and before [l] and [w], the orthographic nasal vowels are pronounced identically to their oral counterparts [ɛ] and [ɔ], with no nasal diphthongization and no assimilatory nasalization. Despite this discrepancy, I refer to the sounds spelled with ⟨ę⟩ and ⟨ą⟩ as ‘nasal vowels’ in every environment they occur, whether they are produced with nasality or not, opting for the more traditional (and concise) terminology at the expense of clumsier phonetic accuracy.

(1)	*stem UR*	*[ɔ̃N]*	*[ɛ̃N]*	*gloss*
	/z**ɛ̃m**b/	z**ɔ̃m**p	z**ɛ̃m**bɪ	‘tooth/teeth’
	/g**ɛ̃m**b/	g**ɔ̃m**p	g**ɛ̃m**ba	‘snout (GEN PL/NOM SG)’
	/uz̢**ɛ̃n**d/	uz̢**ɔ̃n**t	uz̢**ɛ̃n**dɪ	‘office(s)’
	/z̢**ɛ̃n**d/	z̢**ɔ̃n**t	z̢**ɛ̃n**dɪ	‘row(s)’
	/v**ɛ̃w̃**z̢/	v**ɔ̃w̃**s̢	v**ɛ̃w̃**z̢ɛ	‘snake(s)’
	/m**ɛ̃w̃**z̢/	m**ɔ̃w̃**s̢	m**ɛ̃w̃**z̢ɛ	‘husband(s)’
	/kɕ**ɛ̃ŋ**g/	kɕ**ɔ̃ŋ**k	kɕ**ɛ̃ŋ**ga	‘book (GEN PL/NOM SG)’
	/kr**ɛ̃ŋ**g/	kr**ɔ̃ŋ**k	kr**ɛ̃ŋ̟**g̟i	‘circle(s)’

It is clear that the underlyingly voiced obstruent is the trigger, since the alternation does not occur with voiceless obstruents:[2]

(2)	*stem UR*	*[ɛ̃N]*	*[ɛ̃N]*	*gloss*
	/k**ɛ̃m**p/	k**ɛ̃m**p	k**ɛ̃m**pa	‘cluster (GEN PL/NOM SG)’
	/skr**ɛ̃n**t/	skr**ɛ̃n**t	skr**ɛ̃n**tɪ	‘turn(s)’
	/k**ɛ̃w̃**s/	k**ɛ̃w̃**s	k**ɛ̃w̃**sɪ	‘morsel(s)’
	/t**ɛ̃n**ts̢/	t**ɛ̃n**ts̢	t**ɛ̃n**ts̢a	‘rainbow (GEN PL/NOM SG)’
	/g**ɛ̃j̃**ɕ/	g**ɛ̃j̃**ɕ	g**ɛ̃j̃**ɕi	‘goose (SG/PL)’
	/s**ɛ̃ŋ**k/	s**ɛ̃ŋ**k	s**ɛ̃ŋ̟**k̟i	‘knot (in wood)’
	/v**ɛ̃w̃**x/	v**ɛ̃w̃**x	v**ɛ̃w̃**xɪ	‘smell(s)’

(3)	*stem UR*	*[ɔ̃N]*	*[ɔ̃N]*	*gloss*
	/z̢**ɔ̃m**p/	z̢**ɔ̃m**p	z̢**ɔ̃m**pɪ	‘sump(s)’
	/xɔm**ɔ̃n**t/	xɔm**ɔ̃n**t	xɔm**ɔ̃n**tɔ	‘stirrup (GEN PL/NOM SG)’
	/v**ɔ̃w̃**s/	v**ɔ̃w̃**s	v**ɔ̃w̃**sɪ	‘moustache(s)’
	/b**ɔ̃ŋ**k/	b**ɔ̃ŋ**k	b**ɔ̃ŋ̟**k̟i	‘horsefly (SG/PL)’

[2] Actually, the alternation does occur in some words with final voiceless obstruents, like [rɔ̃ŋk]/[rɛ̃ŋka] ‘hand (GEN PL/NOM SG)’. Such alternations seem to have come about due to other historical triggers of vowel length, like the neo-acute accent, which I am not analyzing here.

For historical reasons, the modern nasal vowels did not develop before [r], [j], or nasal stops. Thus, the only sonorants which can follow a nasal vowel are [l] and [w] (though before [l] and [w], the orthographic nasal vowels are pronounced without nasality or diphthongization, leaving the orthography as the only clue to their nasal history). There seem to be no words in which a nasal vowel is followed by word-final [l], and only masculine third-person past tense verbs allow a nasal vowel to be followed by word-final [w]. Thus, there is insufficient data to show a transparent alternation in nasal vowels before most sonorants.[3] Because word-final obstruents must be voiceless (Chapter 2, §2.1.1), the triggering environment for the nasal vowel alternation (word-final underlyingly voiced obstruents) is obscured, resulting in the same type of opacity seen with the [ɔ]~[u] alternation.

The data below clarify that this alternation involves a process of backing and rounding of the front nasal vowel rather than fronting and derounding of the back nasal vowel, since the back nasal vowel can be found in both relevant environments:

(4)	*stem UR*	*[ɔ̃N]*	*[ɔ̃N]*	*gloss*
	/tr**ɔ̃m**b/	tr**ɔ̃m**p	tr**ɔ̃m**ba	'trumpet (GEN PL/NOM SG)'
	/z̨**ɔ̃n**d/	s**ɔ̃n**t	s**ɔ̃n**dɪ	'judgment(s)'
	/z̨**ɔ̃n**dz/	z̨**ɔ̃n**ts	z̨**ɔ̃n**dza	'lust (GEN PL/NOM SG)'
	/v^{j}**ɔ̃w̃**z/	v^{j}**ɔ̃w̃**s	v^{j}**ɔ̃w̃**zɪ	'elm(s)'
	/tɕ**ɔ̃w̃**ʐ/	tɕ**ɔ̃w̃**ʂ	tɕ**ɔ̃w̃**ʐa	'pregnancy (GEN PL/NOM SG)'
	/s**ɔ̃ŋ**g/	s**ɔ̃ŋ**k	s**ɔ̃ŋ̟**ɡ̟i	'log(s)'

[3] The alternation before [w] in verbal pairs such as [zatʐɔw]/[zatʐɛwa] 'he/she began' is productive, but it seems possible to analyze this as morphological alternation rather than phonological. Since this exhaustively represents the possible instances of a transparent nasal vowel alternation before a word-final sonorant, I ignore it and focus on the opaque alternation before word-final voiced obstruents.

In a rule-based theory, this instance of opacity can be modeled with the following rules (the Devoicing rule is identical to the one in Chapter 2):

(5) ɛ̃-Backing: ɛ̃ → ɔ̃ / ___ [+nasal][C,–nasal,+voice] #

(6) Devoicing: [–sonorant] → [–voice] / ___ #

The ɛ̃-Backing rule must be ordered before Devoicing in order to produce the correct output with a back nasal vowel, as seen in the derivation for [zɔ̃mb] 'tooth' in (7). If the rules are ordered incorrectly, as in (8), the wrong output ×[zɛ̃mp], with a front nasal vowel, is obtained instead:

(7)	UR	/zɛ̃mb/	(8)	UR	/zɛ̃mb/
	ɛ̃-Backing	zɔ̃mb		Devoicing	zɛ̃mp
	Devoicing	zɔ̃mp		ɛ̃-Backing	—
	output	[zɔ̃mp]		output	×[zɛ̃mp]

This is the same type of opacity seen in Chapter 2 which is problematic for monostratal versions of OT because of the need for intermediate representations.

3.2 Lexical exceptions

As with the [ɔ]~[u] alternation analyzed in Chapter 2, the nasal vowel alternation is not productive (cf. Westfal 1956), as evidenced by two sets of lexical exceptions which show no alternation: orthographic ⟨VN⟩ sequences and some cases of non-alternating true historical nasal vowels.

3.2.1 Non-alternating orthographic ⟨VN⟩ sequences

Tautosyllabic orthographic ⟨VN⟩ sequences are generally pronounced the same as an orthographic nasal vowel ⟨ę⟩ or ⟨ą⟩. But ⟨VN⟩ sequences have a different history than the true nasal vowels, having been ignored by the relevant sound changes which led to the nasal vowel alternation. Thus, words spelled with ⟨ęN⟩ before a stem-final voiced consonant do not alternate, as seen in the following data:

(9)	*stem UR*	*[ɛ̃N]*	*[ɛ̃N]*	*gloss*
	/lɛg**ɛ̃n**d/	lɛg**ɛ̃n**t	lɛg**ɛ̃n**da	'legend (GEN PL/NOM SG)'
	/vik**ɛ̃n**d/	vik**ɛ̃n**t	wik**ɛ̃n**dɪ	'weekend(s)'
	/kɔm**ɛ̃n**d/	kɔm**ɛ̃n**t	kɔm**ɛ̃n**da	'command (GEN PL/NOM SG)'
	/pr**ɛ̃ŋ**g/	pr**ɛ̃ŋ**k	pr**ɛ̃ŋ**ga	'stripe (GEN PL/NOM SG)'

There are also non-crucial words spelled with ⟨oN⟩ that also do not alternate, such as [bɔmp]/[bɔmba] 'bomb (GEN PL/NOM SG)'. Of course, this is to be expected regardless of the orthography, since underlying [ɔ̃N] does not alternate anyway (4).

3.2.2 Non-alternating ⟨ę⟩ and ⟨ą⟩

There are also numerous words spelled with ⟨ę⟩ or ⟨ą⟩ which do not alternate in the expected environment, with front [ɛ̃N] appearing instead of back [ɔ̃N] before a word-final consonant that is underlyingly voiced (forms marked with (S78) are from Stanisławski 1978):

(10)	*stem UR*	*[ɛ̃N]*	*[ɛ̃N]*	*gloss*
	/ʑ**ɛ̃m**b/	ʑ**ɛ̃m**p	ʑ**ɛ̃m**ba	'finch (GEN PL/NOM SG)' (S78)
	/ɔbr**ɛ̃m**b/	ɔbr**ɛ̃m**p	ɔbr**ɛ̃m**bɪ	'extent(s)'
	/sp**ɛ̃n**d/	sp**ɛ̃n**t	sp**ɛ̃n**dɪ	'round-up(s)' (S78)
	/kɔl**ɛ̃n**d/	kɔl**ɛ̃n**t	kɔl**ɛ̃n**da	'carol (GEN PL/NOM SG)'
	/vʲ**ɛ̃ɲ**ʑ/	vʲ**ɛ̃ɲ**ɕ	vʲ**ɛ̃ɲ**ʑɛ	'bond(s)'
	/krav**ɛ̃ɲ**dʑ/	krav**ɛ̃ɲ**tɕ	krav**ɛ̃ɲ**dʑi	'handful(s)'
	/pɔt**ɛ̃ŋ**g/	pɔt**ɛ̃ŋ**k	pɔt**ɛ̃ŋ**ga	'power (GEN PL/NOM SG)'
	/pr**ɛ̃ŋ**g/	pr**ɛ̃ŋ**k	pr**ɛ̃ŋ**ga	'stripe (GEN PL/NOM SG)' (S78)

Though I have not conducted the same type of experimental work as in Chapter 2, it seems reasonable to conclude that the nasal vowel alternation is not productive. New borrowings with potential nasal vowels are invariably spelled with non-alternating ⟨VN⟩ sequences rather than with ⟨ę⟩ or ⟨ą⟩, and to my knowledge, there have been no disputes of Westfal's (1956) claim that the nasal vowel alternation is not productive in Modern Polish. Under the analysis developed in this dissertation, a non-productive alternation such as the one seen with the nasal vowels is relegated to the underlying forms stored in the lexicon and is not derived by the phonology (pace Gussmann (1980), who argues that those forms which alternate should still be derived rather than listed lexically, though he admits to the general lack of productivity of the alternation).

3.3 The relevance of historical change

In this section, I give an overview of the sound changes affecting the nasal vowels that emerged during the period of disintegration, summarized in the following table. Each sound change is listed with its progressive effects on the early Lechitic word [zɛ̃b] 'tooth'. This table is repeated within the discussion of each sound change, with the relevant row highlighted:

(11)	Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
	Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
	Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
	Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

3.3.1 Step 1a: Lechitic vowel lengthening

As discussed in §2.4.1 of Chapter 2, Lechitic vowels could generally contrast for length in most positions, but vowels in final syllables with a voiced coda could only be long (Stieber 1973:28, Carlton 1991:216–217). The following data show how this lengthening affected Lechitic nasal vowels (*[ʒɛ̃ːd] ‘row’ and *[dɔ̃ːb] ‘oak’) and oral vowels (*[bɔːb] ‘bean’), triggered by a word-final voiced coda. The remaining forms have no lengthening either because the final coda is not voiced (*[nɔs] ‘nose’) or the relevant vowel is not in the final syllable (the plural forms):

(12)

	‘row’		‘oak’	
	NOM SG	*NOM PL*	*NOM SG*	*NOM PL*
Proto-Slavic	*rændu	*rænduː	*dɑmbu	*dɑmbuː
⋮ (PWS)	*ʒɛ̃dʊ	*ʒɛ̃dɨ	*dɔ̃bʊ	*dɔ̃bɨ
Lechitic	*ʒɛ̃ːd	*ʒɛ̃dɨ	*dɔ̃ːb	*dɔ̃bɨ

	‘bean’		‘nose’	
	NOM SG	*NOM PL*	*NOM SG*	*NOM PL*
Proto-Slavic	*bɑbu	*bɑbuː	*nɑsu	*nɑsuː
⋮ (PWS)	*bɔbʊ	*bɔbɨ	*nɔsʊ	*nɔsɨ
Lechitic	*bɔːb	*bɔbɨ	*nɔs	*nɔsɨ

For the model word ‘tooth’ used throughout this section, the pronunciation as a result of this sound change is [zɛ̃ːb], with a long vowel triggered by the voiced coda:

(13)

Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

The long vowels created by this sound change (as well as long vowels from other sources) remained long until at least the mid-15th century, when double letters were used to represent long vowels, as in Jakub Parkoszowic’s *Trakut o ortografii polskiej [Treatise on Polish orthography]* (written circa 1440) (de Bray 1969:231).

3.3.2 Step 1b: Lechitic loss of nasal vowel color

By the end of the Lechitic era, the spelling of the two nasal vowels *[ɛ̃] and *[ɔ̃] had changed. Originally they were distinguished with ⟨en⟩ or ⟨em⟩ for the front nasal vowel and ⟨an⟩ or ⟨am⟩ for the back one. But later spellings blurred the distinction between the two. The two vowels were more and more frequently spelled the same way, typically as ⟨an⟩ or ⟨am⟩, indicating that a difference in pronunciation was no longer being maintained (at least not consistently). For example, as early as 1303, the bailiff of Oświęcim spelled his city's name as ⟨Oswancime⟩, with the digraph ⟨an⟩ used for what should have been the front nasal vowel (Stieber 1968:13). In addition, a new grapheme ⟨□⟩ was adopted during the 14th century as a possible representation for both nasal vowels. Use of ⟨en⟩ and ⟨em⟩ for the front nasal continued to decrease; court documents from the end of the century use only ⟨□⟩ and the digraphs ⟨an⟩ and ⟨am⟩ for both formerly front and back nasal vowels (Stieber 1968:12–13, de Bray 1980:230–231). Since spelling must necessarily lag behind sound change, it is safe to assume that the loss of color in the nasal vowels had begun by the early 1300s, gaining completion by the end of the 14th century, marking it as a relatively late Lechitic or very early Old Polish sound change.

The standard analysis of this merger is that the front and back nasal vowels of PWS both became central low *[ã], matching the quality of the oral vowel *[a], as suggested by the spellings ⟨an⟩ and ⟨am⟩. However, I argue that the nasal vowels merged to mid central *[ɜ̃] rather than low central *[ã] (see §3.4.2 for discussion and analysis), making the word 'tooth' be pronounced [zɜ̃ːb]:

(14)

Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

The following charts show how the overall system of vowel quality changed during Lechitic, with the front and back nasal vowels merging to a central nasal vowel:

(15) *early Lechitic*

i		u
ɛ ɛ̃		ɔ̃ ɔ
	a	

late Lechitic

i		u
ɛ	ɜ̃	ɔ
	a	

In addition, all of the Lechitic vowels existed in both long and short variants, both phonemically and allophonically.

3.3.3 Step 2: Old Polish word-final obstruent devoicing

As discussed in §2.4.2 from Chapter 2, evidence from misspellings such as ⟨bok⟩ for ⟨Bóg⟩ 'God' in the *Kazania Gnieźnieńskie [Gniezno Sermons]* of the late 14th century point to an early Old Polish sound change in which word-final obstruents became voiceless (Stieber 1968:77). Thus, 'tooth' came to be pronounced as *[zɜ̃ːp]:

(16)

Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

3.3.4 Step 3: Middle Polish colorization of nasal vowels

During Middle Polish (and even perhaps very late in Old Polish, at least in some dialects), the colorless nasal vowels split back into colored vowels, with short [ɜ̃] fronting and long [ɜ̃ː] backing and rounding, resulting in a new color contrast in place of an old length contrast. The beginnings of this split is attested in the *Psałterz Puławski [Pulavian Psalter]* (circa 1450), in which the short nasal vowel is spelled with the new symbol ⟨ę⟩, while the long nasal vowel was spelled the same as it had been before, with ⟨ą⟩. This new symbol suggests a fronter pronunciation for the short nasal vowel, though Entwistle and Morison (1949:300) state '[t]he distinction implied was mainly one of length'. This is consistent with the fact that ⟨ę⟩ is unattested in some contemporary documents, such as Parkoszowic's treatise, suggesting that the color split had not yet become widespread. Other new spellings of the nasal vowels began to emerge in a few parts of Poland during the 15th century, with ⟨e⟩, ⟨en⟩, and ⟨em⟩ used to represent the short nasal vowel, and ⟨u⟩, ⟨un⟩, and ⟨um⟩ to represent the long nasal vowel. By the mid-16th century, after the advent of printing and resulting attempts at consistent spelling, ⟨ę⟩ and ⟨ą⟩ became fairly ubiquitous across all of Poland, used to distinguish the color of the nasal vowels as in Modern Polish (Stieber 1968:23–25, de Bray 1980:230–231).

As best as I can determine, no attempt has been made in the literature to explain *why* this sound change occurred; it is merely accepted as a fact of the history of Polish. In §3.4, I give an FDM-OT analysis of this sound change, concluding that the late Middle Polish back nasal vowel was a nasal diphthong *[ɔ̃w̃] (matching the modern word-final and pre-fricative pronunciation), contrary to the standard analysis which posits an essentially 'pure' nasal vowel *[ɔ̃]. With this sound change, the word 'tooth' was pronounced [zɔ̃w̃p] in late Middle Polish:

(17)	Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
	Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
	Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
	Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

This analysis is based on articulatory and acoustic factors which favor a back-round nasal diphthong over a long central nasal vowel. Through enhancement of the color contrast between the two nasal vowels, the short nasal vowel is forced to be front. This novel explanation for the colorization of the nasal vowels relies crucially on the use of $\mathcal{D}$-constraints, highlighting the need for such a mechanism in OT.

The following charts show how the vowel system evolved in Middle Polish in the manner I claim here, with the nasal vowels being first distinguished by length in early Middle Polish, and then by color and diphthongization in late Middle Polish:

(18) *early Middle Polish*

i iː		u uː
ɛ ɛː	**ɜ̃ ɜ̃ː**	ɔ ɔː
	a aː	

late Middle Polish

i iː		u uː
eː		oː
ɛ **ɛ̃**		**ɔ̃w̃** ɔ
	a aː	

Also recall from the previous chapter that the long mid vowels [ɛː] and [ɔː] independently raised to [eː] and [oː] due to tensing of long mid vowels.

3.3.5 Later developments of the nasal vowels

While the back nasal vowel underwent obligatory diphthongization, the front nasal vowel became a diphthong only in some dialects and with little regularity. The difference between the two nasal vowels can be seen in Middle Polish poetry of the Little Poland region, in which [ɛ̃] was frequently rhymed with [ɛ], but [ɔ̃] was rarely rhymed with either [ɔ] or [a] (Stieber 1968:35–36). Variations in spelling in Great

Poland court records also support this hypothesis. Scribes used the digraph ⟨am⟩ about 50% of the time to represent word-final [ɔ̃] (the traditional ⟨ą⟩ appears only slightly less often). This digraph is suggestive of the proposed diphthongal nature of the back nasal vowel, indicating both the nasality and the bilabial place of articulation of the nasal glide [w̃] to the best of the ability of the orthography of the time. In comparison, word-final [ɛ̃] could never be spelled with ⟨em⟩ or ⟨en⟩ in Middle Polish, only with ⟨ę⟩ or ⟨e⟩ (each used with an approximate frequency of 50%), suggesting that [ɛ̃] did not diphthongize like [ɔ̃] did, and in fact, was in the process of losing its nasality instead (Stieber 1968:33–34).

However, the digraphs were common variants for both nasal vowels when they occurred before stops, with ⟨am⟩ and ⟨em⟩ used before labial stops and ⟨en⟩ and ⟨an⟩ used before coronal and velar stops. For the back nasal vowel, this seems quite reasonable, with the nasal glide assimilating in place of articulation to the following stop, as in Modern Polish. It is not clear why the short front nasal vowel also underwent obligatory diphthongization before consonants, but only optionally at the end of a word or before fricatives. Whatever the cause, it is still undeniable that the front and back nasal vowels evolved differently with respect to diphthongization in a way consistent with the analysis proposed here, since the modern back nasal vowel has long been known to be pronounced as a nasal diphthong, while a diphthongal pronunciation for the front nasal vowel is erratic, dialectally inconsistent, and considered obsolete in modern standard Polish. An explanation for the optional diphthongization of the front nasal is not germane to the discussion at hand, so I leave this problem as an open question for further research.

3.4 FDM-OT analysis of the nasal vowel alternation

3.4.1 Step 1a: Lechitic vowel lengthening

This sound change was analyzed in §2.5.1 of Chapter 2, so I only provide a summary of the essential portions of the analysis. In early Lechitic, there was a general contrast between short and long vowels. Such a contrast must exist in the input (by richness of the base), so its survival in the output means that $\mathcal{F}$-duration must outrank $\mathcal{D}_1$-duration (which punishes vowel length contrasts), as well as any $\mathcal{M}$-constraints that penalize particular specific vowel lengths (including $\mathcal{M}$-V̌C̬#, which bans short vowels before word-final voiced consonants, and $\mathcal{M}$-Vː, which bans all long vowels):

(19) *Step 1a.1: early Lechitic (prior to lengthening)*

zɛ̃p$_1$ zɛ̃b$_5$ zɛ̃ːp$_2$ zɛ̃ːb$_6$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$	$\mathcal{F}$ voi	$\mathcal{F}$ dur	$\mathcal{D}_1$ dur	$\mathcal{M}$ V̌C̬#	$\mathcal{M}$ Vː
✓ a. zɛ̃p$_1$ zɛ̃b$_5$ zɛ̃ːp$_2$ zɛ̃ːb$_6$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$			✗^{4}	✗	✗^{4}
b. zɛ̃p$_1$ zɛ̃ːp$_2$ zɛ̃ːb$_{5,6}$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$		✗!	✗^{3}		✗^{4}
c. zɛ̃p$_{1,5}$ zɛ̃ːp$_2$ zɛ̃ːb$_6$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$	✗!		✗^{3}		✗^{4}

In order to model the late Lechitic sound change of lengthening before word-final voiced codas, as represented by candidate (19b) above, the markedness constraint 𝓜-V̆C̬# must be ranked over 𝓕-duration. This ranking overcomes the ill-formedness of a short vowel before a word-final voiced consonant by violating the lower ranked 𝓕-duration, which means that the vowel is lengthened:

(20) *Step 1a.2: late Lechitic lengthening*

		prom ⬇	*dem* ⬇		
zɛ̃p$_1$ zɛ̃b$_5$ zɛ̃ːp$_2$ zɛ̃ːb$_6$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$	𝓕 voi	𝓜 V̆C̬#	𝓕 dur	$\mathcal{D}_1$ dur	𝓜 Vː
a. zɛ̃p$_1$ zɛ̃b$_5$ zɛ̃ːp$_2$ zɛ̃ːb$_6$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$		×!		×^{4}	×^{4}
✓ b. zɛ̃p$_1$ zɛ̃ːp$_2$ zɛ̃ːb$_{5,6}$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$			×	×^{3}	×^{4}
c. zɛ̃p$_{1,5}$ zɛ̃ːp$_2$ zɛ̃ːb$_6$ zɛ̃pɨ$_3$ zɛ̃bɨ$_7$ zɛ̃ːpɨ$_4$ zɛ̃ːbɨ$_8$	×!			×^{3}	×^{4}

The fully faithful candidate (20a) violates the newly promoted constraint banning short vowels before word-final voiced codas. As with early Lechitic, devoicing (20c) is not a viable option to avoid violating 𝓜-V̆C̬#, due to high ranking 𝓕-voicing.

This leaves candidate (20b) as the grammatical output, with vowel lengthening before word-final voiced codas and 'tooth' pronounced as [zɛ̃ːb]:

(21)

Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

3.4.2 Step 1b: Lechitic loss of nasal vowel color

It is well established that color contrasts in general decrease in quality as height decreases, as mirrored by the traditional trapezoidal vowel space. Thus, high vowels have better color contrasts than their mid counterparts, which in turn have better color contrasts than their low counterparts. This fact can be formalized with a scale of ordered $\mathcal{D}$-color constraints that each rule out particular vowel pairs, as schematized below (this is an expansion of the use of $\mathcal{D}$-color put forth in §1.3.2 of Chapter 1, with new subscripts required in order to account for the different sizes of color divisions for high, mid, and low vowels):

(22)

	[y ɨ],[ɨ ɯ]	<	[i y],[y ɯ],[ɯ u]	<	**[i ɨ]**,[ɨ u]	<	[i ɯ],[y u]	<	[i u]
ruled out by:	$\mathcal{D}_0$		$\mathcal{D}_2$		$\mathcal{D}_4$		$\mathcal{D}_6$		$\mathcal{D}_8$

	[œ ɜ],[ɜ ʌ];[ɛ œ],[œ ʌ],[ʌ ɔ]	<	**[ɛ ɜ]**,[ɜ ɔ]	<	[ɛ ʌ],[œ ɔ]	<	[ɛ ɔ]
ruled out by:	$\mathcal{D}_0$		$\mathcal{D}_2$		$\mathcal{D}_4$		$\mathcal{D}_6$

	[ɶ a],[a ɑ];[æ ɶ],[ɶ ɑ],[ɑ ɒ];**[æ a]**,[a ɒ]	<	[æ ɑ],[ɶ ɒ]	<	[æ ɒ]
ruled out by:	$\mathcal{D}_0$		$\mathcal{D}_2$		$\mathcal{D}_4$

According to this formalization, a vowel system with a front and central high vowel pair [i ɨ] (in bold on the first line) will incur a violation of $\mathcal{D}_4$-color. The same type of color contrast (front/central) for the mid vowels [ɛ ɜ] violates $\mathcal{D}_2$-color (which is ranked higher than $\mathcal{D}_4$-color), while the comparable low vowel pair [æ a] violates the highest ranked $\mathcal{D}_0$-color.

The apparent skipping of odd-numbered $\mathcal{D}$-constraints is crucial to the analysis presented in the section and relates to the effect of nasality on the perception of vowel quality. Beddor (1993) notes that nasal vowels generally have poorer quality contrasts with each other than oral vowels do, citing studies (Bond 1975, Butcher 1976, Mohr and Wang 1968, Wright 1986, among others) which found that the quality of true nasal vowels and vowels extracted from nasal contexts tend to be more easily confused or judged more similar than their oral counterparts. The effect is mostly seen for vowel height, but color contrasts are also impacted. Within this analysis, I formalize these results by having pairs of particular nasal vowels of the same height be ruled out by $\mathcal{D}_{(n-1)}$-color, where $\mathcal{D}_n$-color is the (even-numbered) constraint which rules out the oral version of that pair (cf. Padgett 1997, in which a similar proposal is made). Thus, the odd-numbered $\mathcal{D}$-constraints fill in the perceptual gaps in the $\mathcal{D}$-color scale that correspond to the nasal vowels. For example, for the pair [ɛ̃ ɔ̃], the oral counterpart is [ɛ ɔ], which is ruled out by $\mathcal{D}_6$-color. Therefore, the nasal pair will be ruled out by $\mathcal{D}_{(6-1)}$-color = $\mathcal{D}_5$-color. This can be seen explicitly for all nasal vowel pairs below:

(23) [ỹ ɨ̃],[ɨ̃ ɯ̃] < [ĩ ỹ],[ỹ ɯ̃],[ɯ̃ ũ] < [ĩ ɨ̃],[ɨ̃ ũ] < [ĩ ɯ̃],[ỹ ũ] < [ĩ ũ]

ruled out by: $\mathcal{D}_0$ $\mathcal{D}_1$ $\mathcal{D}_3$ $\mathcal{D}_5$ $\mathcal{D}_7$

[œ̃ ɜ̃],[ɜ̃ ʌ̃];[ɛ̃ œ̃],[œ̃ ʌ̃],[ʌ̃ ɔ̃] < [ɛ̃ ɜ̃],[ɜ̃ ɔ̃] < [ɛ̃ ʌ̃],[œ̃ ɔ̃] < [ɛ̃ ɔ̃]

ruled out by: $\mathcal{D}_0$ $\mathcal{D}_1$ $\mathcal{D}_3$ $\mathcal{D}_5$

[ɶ̃ ã],[ã ɑ̃];[æ̃ ɶ̃],[ɶ̃ ɑ̃],[ɑ̃ ɒ̃];[æ̃ ã],[ã ɒ̃] < [æ̃ ɑ̃],[ɶ̃ ɒ̃] < [æ̃ ɒ̃]

ruled out by: $\mathcal{D}_0$ $\mathcal{D}_1$ $\mathcal{D}_3$

It is also important to consider how the oral vowels and nasal vowels contrast with each other. The simplest solution, which I will adopt here, is that they contrast for nasality only and do not contrast at all for any other perceptual dimension (in particular, they do not contrast for color). A plausible alternative is to assume that nasality itself is a type of color (or perhaps enhances color contrast; cf. Wright 1986), so that, for example, the mixed oral/nasal pair [ɛ ɜ̃] has a significantly better color contrast than either the purely nasal [ɛ̃ ɜ̃] or oral [ɛ ɜ] pairs (and thus, it violates fewer $\mathcal{D}$-color constraints). This alternative is reasonable and worth pursuing, but it is beyond the scope of this dissertation and not necessary for the present analysis.

In early Lechitic, a two-way color contrast was allowed for mid (both oral and nasal) and high vowels, but not for low vowels. Thus, by richness of the base, the constraint hierarchy of the early Lechitic grammar must be able to map the universal input $\mathcal{I}$ to a language that only allows the vowel system [i u ɛ ɔ ɛ̃ ɔ̃ a]. Since I am only concerned with vowels, I only consider inputs and outputs that are subsets of the twelve-vowel set {i,ɨ,u,ɛ,ɜ,ɔ,æ,a,ɒ,ɛ̃,ɜ̃,ɔ̃}, which has three heights for oral vowels, but only one for nasal vowels (I assume high ranking $\mathcal{D}$-height constraints allow a three-level height contrast for oral vowels, but only a one-level height contrast in the nasal vowels; cf. Padgett 1997), and three colors for both oral and nasal vowels (I ignore vowel length for now, assuming that the analysis presented here holds for both short and long vowels). The following tableau shows how early Lechitic is derived:

(24) *Step 1b.1: early Lechitic (before nasal vowel color loss)*

i ɨ u ɛ ɜ ɔ ɛ̃ ɜ̃ ɔ̃ æ a ɒ	$\mathcal{D}_4$ color	$\mathcal{F}$ color	$\mathcal{D}_5$ color	$\mathcal{D}_6$ color
a. i ɨ u ɛ ɜ ɔ ɛ̃ ɜ̃ ɔ̃ æ a ɒ	✗^{9}!		✗^{10}	✗^{11}
b. i ɨ u ɛ ɜ ɔ ɛ̃ ɜ̃ ɔ̃ a	✗^{6}!	✗^{2}	✗^{7}	✗^{8}
c. i ɨ u ɛ ɔ ɛ̃ ɔ̃ a	✗^{2}!	✗^{4}	✗^{3}	✗^{4}
✓ d. i u ɛ ɔ ɛ̃ ɔ̃ a		✗^{5}	✗	✗^{2}
e. i u ɛ ɔ ɜ̃ a		✗^{6}!		✗

The fully faithful candidate (24a) has too many poor color contrasts. The three-level color contrast in the oral low vowels is particularly bad, since even the maximal two-level color contrast [æ ɑ] violates $\mathcal{D}_4$-color. Thus, the low vowels contribute three violations of $\mathcal{D}_4$-color. The central oral vowels also present problems for the mid and high vowel contrasts, since the mid pairs [ɛ ɜ] and [ɜ ɔ] and the high pairs [i ɨ] and [ɨ u] also violate $\mathcal{D}_4$-color, adding four more violations for candidate (24a) (in fact, the mid vowel pairs also violate the higher constraints $\mathcal{D}_2$-color and $\mathcal{D}_3$-color). Finally, the nasal vowel pairs [ɛ̃ ɜ̃] and [ɜ̃ ɔ̃] violate $\mathcal{D}_1$-color, which means they also violate all lower ranked $\mathcal{D}$-color constraints, including $\mathcal{D}_4$-color, giving candidate (24a) a grand total of nine violations of $\mathcal{D}_4$-color.

Candidate (24b) has only one low vowel, so it has three fewer violations for the low vowels than (24a) does, but it maintains all of the other poor contrasts, so it still incurs six violations of $\mathcal{D}_4$-color. Candidate (24c) improves the mid vowel contrasts for both oral and nasal vowels, losing the four mid vowel violations that (24a) and (b) have. However, it still has two violations for the high vowels because of the central vowel [ɨ] which contrasts poorly enough with [i] and [u] to violate $\mathcal{D}_4$-color. Candidates (24d) and (24e) satisfy $\mathcal{D}_4$-color by having no central high or mid vowels, and no attempt at a color contrast in the low vowels. Candidate (24d) defeats candidate (24e) (despite having a worse nasal vowel contrast per the low ranked $\mathcal{D}_5$-color), due to candidate (24e)'s unmotivated (but prescient) merger of the nasal vowels, which gives it more violations of $\mathcal{F}$-color than candidate (24d) incurs. Candidate (24d) is the winning candidate and represents the output of early Lechitic speakers, before any sound changes occurred. By strong lexicon optimization, this output language will be stored in the lexicon, allowing it to serve as the input for later sound changes, such as the merger of the nasal vowels.

In late Lechitic, the two colors of nasal vowels merged to a single vowel color. It is not completely clear what the exact quality of the merged nasal vowel was. The standard assumption by Slavicists is that the late Lechitic nasal vowel was *[ã], with the earlier nasal vowels not only losing their color, but also lowering from mid to low. However, the evidence for such lowering is inconsistent, and it is not easily captured formally (especially taking into account the later split and possible raising of this merged vowel during Middle Polish; see §3.4.4). I argue that the merged nasal vowel of late Lechitic was in fact mid *[ɜ̃] rather than low *[ã], with the merger causing color loss only, preserving the original height of the early Lechitic nasal vowels.

There are several reasons for taking this unconventional position. Within the FDM-OT framework, some $\mathcal{D}$- or $\mathcal{M}$-constraint(s) must be promoted over some $\mathcal{F}$-constraint in order to trigger any sound change (by definition, a sound change cannot be caused by a promoted $\mathcal{F}$-constraint, since a sound change requires a *difference* between early and late phases of the same stage of a language, not increased similarity). No $\mathcal{D}$-constraint can force the Lechitic mid nasal vowels to lower, because there are no other nasal vowels for them to be more perceptually distinct from (and I am assuming that nasal vowels do not get involved with any contrasts with oral vowels except nasality). If there were also high nasal vowels in Lechitic, it would be expected that $\mathcal{D}$-height could force the mid nasal vowels to lower, but no such high nasal vowels existed. This leaves $\mathcal{M}$-constraints as the only possible impetus for lowering of the mid nasal vowels. Since the oral vowels did not lower, any $\mathcal{M}$-constraints used to cause lowering of the mid nasal vowels must crucially refer to the interaction of nasality and height. Specifically, lowering of the

Lechitic mid nasal vowels requires $\mathcal{M}$-ɛ̃,ɔ̃ to outrank $\mathcal{M}$-ã (this ranking must of course be universal in FDM-OT, since $\mathcal{M}$-constraints represent physiological markedness). It is not clear that this is a reasonable universal ranking, however. Low vowels require more extreme movement of the jaw than mid vowels (which are closer to rest position), so we would expect low vowels to be more articulatorily marked than mid vowels, giving us the opposite ranking: $\mathcal{M}$-ã over $\mathcal{M}$-ɛ̃,ɔ̃. Thus, a single nasal vowel should prefer to be mid and central, since there are no neighboring vowels to push it away from the physiological neutral mid-central position.

In addition, the 14th-century orthographic innovation of □□□used to represent the late Lechitic merged nasal vowel suggests that the vowel quality was different from any extant vowel of the time, rather than being similar to [a]. Otherwise, scribes might have relied solely on some variation of □a□(as with the 14th and 15th century digraphs, or the 16th century □ą□, which has been used ever since) instead of inventing a completely unrelated symbol. But if 14th-century scribes were in fact trying to capture the mid-central quality of this merged nasal vowel, a quality completely unlike any other vowel in their inventory, their choice of a novel symbol makes sense.

Finally, if the former mid nasal vowels did indeed merge and lower to *[ã], some explanation has to be given for why both the lowering and decolorization processes eventually reversed themselves later in Middle Polish. The simpler analysis I adopt claims that vowel height remains constant through both the Lechitic merger and subsequent split in Middle Polish, requiring only an account of the loss of vowel color (which must be accounted for independently of vowel height, since it is uncontroversial that the vowels merged with respect to color). Thus, I assume that the single nasal vowel quality that emerged at the end of the Lechitic period is best represented by the mid central vowel *[ɜ̃] rather than the more traditional *[ã], allowing for a simpler analysis of the nasal vowels that meshes with the formal framework of FDM-OT, notions of articulatory difficulty, and philological evidence.

The late Lechitic merger of the two early Lechitic nasal vowels can be achieved by changing the relative ranking of $\mathcal{F}$-color and $\mathcal{D}_5$-color from early Lechitic so that $\mathcal{D}_5$-color outranks $\mathcal{F}$-color, allowing the language with a single nasal (25b) to win over the faithful candidate (25a) (recall that by strong lexicon optimization, the output of early Lechitic is used as the input to the late Lechitic sound change, so the faithful candidate (25a) represents a situation with no sound change):

(25) *Step 1b.2: late Lechitic nasal vowel color loss*

prom ↓ ($\mathcal{D}_5$ color) *dem* ↓ ($\mathcal{F}$ color)

i u ɛ ɔ ɛ̃ ɔ̃ a	$\mathcal{D}_4$ color	$\mathcal{D}_5$ color	$\mathcal{F}$ color	$\mathcal{D}_6$ color
a. i u ɛ ɔ ɛ̃ ɔ̃ a		×!		×^{2}
✓ b. i u ɛ ɔ ɜ̃ a			×^{2}	×
c. i u ɜ ɜ̃ a			×^{4}!	

This is the desired output for late Lechitic, with five oral vowels and one nasal vowel (in addition, all six vowels could occur long or short, which is easily accounted for by ranking $\mathcal{F}$-duration higher than any constraints which would prevent a duration contrast). The faithful candidate (25a) violates high ranking $\mathcal{D}_5$-color because of the two mid nasal vowels, while candidates (25b) and (25c) satisfy this constraint by merging the nasal vowels to *[ɜ̃]. Candidate (25b) ultimately defeats candidate (25c) because of faithfulness; candidate (25c) merges the oral mid vowels as well, satisfying the low ranked $\mathcal{D}_6$-color constraint, but at the expense of the higher ranked $\mathcal{F}$-color, which candidate (25b) satisfies better. Thus, 'tooth' was pronounced as [zɜ̃ːb] in late Lechitic, with a central nasal vowel:

(26)	Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
	Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
	Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
	Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

3.4.3 Step 2: Old Polish word-final obstruent devoicing

This sound change was analyzed in §2.5.2 of Chapter 2, so I provide only a summary of the analysis presented there. The early Old Polish grammar must be able to derive the late Lechitic contrasts in obstruent voicing and in vowel duration (except before word-final voiced consonants, where only long vowels are allowed). The following hierarchy derives the correct output:

(27) *Step 2.1: early Old Polish (before devoicing)*

	$\mathcal{F}$ voi	$\mathcal{M}$ V̆Ç#	$\mathcal{F}$ dur	$\mathcal{M}$ Ç#	$\mathcal{M}$ V:
zɔ̃p1 zɔ̃b5 zɔ̃:p2 zɔ̃:b6 zɔ̃pɨ3 zɔ̃bɨ7 zɔ̃:pɨ4 zɔ̃:bɨ8					
a. zɔ̃p1 zɔ̃b5 zɔ̃:p2 zɔ̃:b6 zɔ̃pɨ3 zɔ̃bɨ7 zɔ̃:pɨ4 zɔ̃:bɨ8		×!		×2	×4
✓ b. zɔ̃p1 zɔ̃:p2 zɔ̃:b5,6 zɔ̃pɨ3 zɔ̃bɨ7 zɔ̃:pɨ4 zɔ̃:bɨ8			×	×	×4
c. zɔ̃p1,5 zɔ̃:p2 zɔ̃:b6 zɔ̃pɨ3 zɔ̃bɨ7 zɔ̃:pɨ4 zɔ̃:bɨ8	×!				×3

The winning candidate (27b) is stored in the lexicon via strong lexicon optimization, serving as the input for any late Old Polish sound changes, in this case, word-final devoicing of obstruents.

As with the back vowel alternation, word-final devoicing has an important impact on the nasal vowel alternation (which at this point relies on length). Since word-final voiced consonants are one of the triggers for vowel length, word-final devoicing renders the alternation opaque. Again, the modern opacity that is apparent in the nasal vowel alternation has its origins in the Old Polish sound change of final devoicing. Devoicing can arise by promoting $\mathcal{M}$-Ç# over $\mathcal{F}$-voicing (the constraints

𝓜-V̆C̬# and 𝓕-duration are not violated by the relevant candidates, so they are not shown in this tableau):

(28) *Step 2.2: late Old Polish devoicing*

	prom ⬇ 𝓜 C̬#	*dem* ⬇ 𝓕 voi	𝓜 V:
zɜ̃p1 zɜ̃:p2 zɜ̃:b5,6 zɜ̃pɨ3 zɜ̃bɨ7 zɜ̃:pɨ4 zɜ̃:bɨ8			
a. zɜ̃p1 zɜ̃:p2 zɜ̃:b5,6 zɜ̃pɨ3 zɜ̃bɨ7 zɜ̃:pɨ4 zɜ̃:bɨ8	×!		×4
✓ b. zɜ̃p1 zɜ̃:p2,5,6 zɜ̃pɨ3 zɜ̃bɨ7 zɜ̃:pɨ4 zɜ̃:bɨ8		×	×3

The fully faithful candidate (28a) violates the newly promoted 𝓜-C̬# because it contains the word [zɜ̃:b], with a word-final voiced obstruent. Candidate (28b) is selected as the output of this sound change because it satisfies 𝓜-C̬# (despite violating lower-ranked constraints), representing the word-final obstruent devoicing that occurred in late Old Polish, causing 'tooth' to be pronounced [zɜ̃:p]:

(29)

Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃:b
Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃:b
Step 2	**Old Polish**	**1350–1500**	**word-final devoicing**	**zɜ̃:p**
Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

As with the back vowel alternation analyzed in Chapter 2, this is the crucial stage at which opacity enters the system. If not for strong lexicon optimization preserving older vowel length, 'tooth' would be pronounced transparently as ✗[zɜ̃p], since there is no voicing in the surface coda to trigger vowel length. With an ordering of strong lexicon optimization before sound change, opacity can be achieved in the lexicon.

3.4.4 Step 3: Middle Polish colorization of nasal vowels

In order to satisfy the hypothesis of richness of the base, a constraint hierarchy must be constructed in early Middle Polish that can derive the late Old Polish sound system from the universal input □. I use the same subset of □ from §3.3.2, consisting of nine oral vowels (three heights and three colors) and three nasal vowels (one height and three colors). The target language has five oral vowels (the standard five vowel system) and only one nasal vowel:

(30) *Step 3.1: early Middle Polish (before nasal vowel colorization)*

i ɨ u ɛ ɜ ɔ ɛ̃ ɜ̃ ɔ̃ æ a ɒ	$\mathcal{D}_4$ color	$\mathcal{D}_5$ color	$\mathcal{F}$ color	$\mathcal{D}_6$ color
a. i ɨ u ɛ ɜ ɔ ɛ̃ ɜ̃ ɔ̃ æ a ɒ	✗^{9}!	✗^{10}		✗^{11}
b. i u ɛ ɔ ɛ̃ ɔ̃ a		✗!	✗^{5}	✗^{2}
✓ c. i u ɛ ɔ ɜ̃ a			✗^{6}	✗

The fully faithful candidate (30a) has too many poor color contrasts (the same ones as calculated for candidate (24a)). Candidate (30b) is better, but not good enough, still incurring a violation of $\mathcal{D}_5$-color because of the nasal vowels, leaving candidate (30c) as the selected output of early Middle Polish.

In addition, the early Middle Polish hierarchy must also derive contrastive vowel length. Every vowel could in principle be short or long in any position, so $\mathcal{F}$-duration must be ranked higher than all of the $\mathcal{D}$-duration constraints to allow short and long vowels in the input to emerge faithfully with respect to duration. The input set I consider is the same as in (30) plus the long counterparts of each vowel:

(31) *Step 3.1 (continued): early Middle Polish (before nasal vowel colorization)*

i ɨ u ɛ ɜ ɔ ɛ̃ ɜ̃ ɔ̃ æ a ɒ iː ɨː uː ɛː ɜː ɔː ɛ̃ː ɜ̃ː ɔ̃ː æː aː ɒː	$\mathcal{F}$ dur	$\mathcal{F}$ color	$\mathcal{D}_1$ dur	$\mathcal{D}_2$ dur
✓ a. i u ɛ ɔ ɜ̃ a iː uː ɛː ɔː ɜ̃ː aː		✗^{12}	✗	✗^{11}
b. i u ɛ ɔ ɛ̃ a iː uː ɛː ɔː ɔ̃ː aː		✗^{14}!	✗	✗^{10}
c. i u ɛ ɔ ɜ̃ a iː uː ɛː ɔː aː	✗^{3}!	✗^{12}	✗	✗^{10}
d. i u ɛ ɔ ɜ̃ a	✗^{6}!	✗^{12}		

All of the candidates involve some mergers of vowel color, as expected from (30). Candidate (31b) has two extra violations from mapping /ɔ̃/ to [ɛ̃] and /ɛ̃ː/ to [ɔ̃ː] in order to preserve vowel duration. Candidates (31c–d) alleviate potentially bad contrasts in vowel length by merging some or all of the short/long vowel pairs. But with $\mathcal{F}$-duration and $\mathcal{F}$-color ranked high, only the fully faithful candidate (31a) can win. Thus, the hierarchies in (30) and (31) derive the correct output for early Middle Polish, with ten oral vowels and two nasal vowels of the same central color

The traditional analysis of the nasal vowel split in Middle Polish is that the long nasal vowel backed and rounded to [ɔ̃ː] while the short nasal vowel fronted to [ɛ̃]. However, I could find no published analysis which provides an explanation for this sound change. I argue that the long nasal vowel diphthongized during Middle Polish in order to be more distinct from the short nasal vowel with respect to duration, under the assumption that length contrasts in nasal vowels are not as good as they are for oral vowels. This assumption is based on the tendency of nasal vowels to be longer than oral vowels because of the extra time required for velic lowering to occur, so a distinction in the inherently longer nasal vowel duration would be harder to perceive than for the shorter oral vowels.[4] Thus, in order to enhance the relatively poor duration distinction between [ɜ̃] and [ɜ̃ː], diphthongization of the long vowel was required in Middle Polish (under the reasonable assumption that a diphthong is more distinctive from a short vowel than a steady long vowel is).

[4] This is readily understood with a simple example. It is much easier to tell that one second is shorter than two seconds than it is to tell that one minute is shorter than sixty-one seconds. As the base quantity gets larger (from one second to one minute), a constant difference (of one second) has less of an effect on the perceptual distinctiveness of the base quantity and the base plus the difference. This is why perception is logarithmic rather than linear (cf. the Richter scale, astronomical magnitude, decibels, etc., all of which are logarithmic scales of perception). Thus, for nasal vowels, which are slightly longer than oral vowels, length differences will be harder to perceive than for oral vowels.

In addition, I argue that the best nasal off-glides are back [w̃] and [ɰ̃], rather than front [j̃] or central [ɰ̶̃].[5] Ohala and Ohala (1993) cite evidence from Acatlan Mixtec (Pike and Wistrand 1974), Mbay (Caprile 1968), and Vietnamese that show that nasal vowels have a tendency to be followed by a velar closure, supporting their contention that back nasal consonants are less consonantal than front nasal consonants due to diminished perceptual cues to consonantality of back nasals; that is, they are more vowel-like. Articulatory concerns seem also to play a role: nasal sounds are produced with a lowered velum, and back glides target the velum. With the velum lowered, it is easier to achieve the target, making back nasal glides better than front nasal glides.

With a requirement that there be no color contour in the nasal diphthong (for reasons of articulatory difficulty, represented by undominated $\mathcal{M}$-contour(color), which prevents changing color from one segment to the next, a maneuver that requires more effort than maintaining a constant color), the vocalic element is forced to become back as well to match the off-glide, providing motivation for half of this sound change. Finally, the resulting diphthong, by virtue of no longer being central, must contrast for color with its short counterpart. Since a central versus back contrast is not allowed in Polish for the oral mid vowels (they must be front and back-round), it is not a surprise it is not allowed for nasal vowels either. This forces the short nasal vowel to become front [ɛ̃] and the back nasal diphthong to be round [ɔ̃w̃] in order to further enhance the color contrast between them, motivating the second half of this

[5] The IPA does not have a symbol for a central glide, so I adopt the symbol [ɰ̶] on analogy with the use of the crossbar in the central vowels [ɨ] and [ʉ]. I use [ɰ] instead of [j] as the base symbol for an unrounded glide because the resulting character [ɰ̶] is more distinct from unrelated symbols than [ɟ̶] is (the latter is too similar to the IPA symbol for the palatal stop [ɟ], whereas [ɰ̶] is not likely to be confused with any unrelated symbol).

sound change. A framework such as FDM-OT is ideally suited to this analysis, since competing frameworks with no mechanism for enforcing and enhancing contrast cannot provide an account for this sound change.

By strong lexicon optimization, the input to late Middle Polish is the sound system of early Middle Polish. I only consider inputs and outputs consisting of mid vowels, since they alone underwent the crucial change. This sound change requires $\mathcal{D}_1$-duration to be promoted over $\mathcal{F}$-color:

(32) *Step 3.2: late Middle Polish nasal vowel colorization*

ε ɔ ɜ̃ εː ɔː ɜ̃ː	$\mathcal{F}$ dur	$\mathcal{D}_1$ dur (*prom* ←)	$\mathcal{F}$ color (*dem* →)	$\mathcal{D}_2$ dur
a. ε ɔ ɜ̃ εː ɔː ɜ̃ː		×!		×³
✓ b. ε ɔ ɛ̃ εː ɔː ɔ̃w̃			×²	×³
c. ε ɔ ɜ̃ εː ɔː	×!			

The fully faithful candidate (32a) has two vowels [ɜ̃] and [ɜ̃ː] which are too close together in the perceptual dimension for vowel duration, violating $\mathcal{D}_1$-duration (the duration contrast is stronger for the oral vowels, which violate low-ranked $\mathcal{D}_2$-duration, but not $\mathcal{D}_1$-duration). Candidate (32c) represents an attempt to resolve this poor contrast by eliminating it entirely. However, $\mathcal{F}$-duration is still highly ranked (it was ranked higher than $\mathcal{D}_1$-duration during early Middle Polish, so without motivation to rerank, it continues to outrank $\mathcal{D}_1$-duration in late Middle Polish), so candidate (32c) is ruled out. This leaves candidate (32b), which satisfies the top constraints at the expense of faithfulness to vowel color, which is fine, since $\mathcal{F}$-color was demoted.

The next tableau shows why the colorization of the nasal vowel happened the way it did, rather than in the opposite direction (with the long nasal vowel fronting and the short nasal vowel backing and rounding). The primary reason is articulatory markedness: the best nasal glides are back, not front or central (perceptual concerns might also play a role). Thus, the markedness constraint $\mathcal{M}$-j̃,ɯ̶̃ universally outranks the markedness constraint $\mathcal{M}$-ɯ̯̃,w̃ (these constraints are already highly ranked, since there are no nasal glides in early Middle Polish). $\mathcal{D}_1$-duration must be further promoted over $\mathcal{M}$-ɯ̯̃,w̃ to get the right results:

(33) *Step 3.2 (continued): late Middle Polish nasal vowel colorization*

			prom ⬋	*dem* ⬊	*dem* ⬊
ɛ ɔ ɜ̃ ɛː ɔː ɜ̃ː	$\mathcal{F}$ dur	$\mathcal{M}$ j̃,ɯ̶̃	$\mathcal{D}_1$ dur	$\mathcal{M}$ ɯ̯̃,w̃	$\mathcal{F}$ color
a. ɛ ɔ ɜ̃ ɛː ɔː ɜ̃ː			×!		×[2]
✓ b. ɛ ɔ ɛ̃ ɛː ɔː ɔ̃w̃				×	×[2]
c. ɛ ɔ ɜ̃ ɛː ɔː ɜ̃ɯ̶̃		×!			
d. ɛ ɔ ɔ̃ ɛː ɔː ɛ̃j̃		×!			×[2]

Candidate (33a) is still bad because of the poor duration contrast in the nasal vowels. Candidates (33c) and (33d) are ruled out because they have worse nasal glides than candidate (33b) does, which has a back nasal glide. But [ɰ̃] is as good of a back nasal glide as [w̃], so more must be said to explain why [w̃] appears instead.

Because $\mathcal{D}_2$-color outranks $\mathcal{F}$-color (a ranking already needed to get the correct colors for oral vowels), the nasal vowels are required to be more distinct from each other with respect to vowel color:

(34) *Step 3.2 (continued): late Middle Polish nasal vowel colorization*

ɛ ɔ ɜ̃ ɛː ɔː ɜ̃ː	$\mathcal{F}$ dur	$\mathcal{M}$ j̃,ɰ̶̃	$\mathcal{M}$ ɰ̃,w̃	$\mathcal{D}_2$ color	$\mathcal{F}$ color
✓ a. ɛ ɔ ɛ̃ ɛː ɔː ɔ̃w̃			×		×^{2}
b. ɛ ɔ ɛ̃ ɛː ɔː ʌ̃ɰ̃			×	×!	×^{2}
c. ɛ ɔ ɜ̃ ɛː ɔː ɔ̃w̃			×	×!	×

The vowel color contrasts in candidates (34b) and (34c) are too small, whereas those in candidate (34a) are large enough to satisfy the $\mathcal{D}_2$-color constraint, allowing it to violate $\mathcal{F}$-color. (Presumably, *[ʌ̃ɰ̃] incurs slightly fewer violations of $\mathcal{F}$-color than *[ɔ̃w̃] does, but I have not been counting faithfulness violations to a small enough detail to distinguish the two; regardless, *[ɔ̃w̃] surfaces as the correct output due to satisfaction of $\mathcal{D}_2$-color), giving us the late Middle Polish pronunciation [zɔ̃w̃p] for 'tooth':

(35)	Step 1a	Lechitic	1100–1350	vowel lengthening	zɛ̃ːb
	Step 1b	Lechitic	1100–1350	nasal decolorization	zɜ̃ːb
	Step 2	Old Polish	1350–1500	word-final devoicing	zɜ̃ːp
	Step 3	Middle Polish	1500–1750	nasal colorization	zɔ̃w̃p

3.5 Comparison with standard OT

The Lechitic merger and Middle Polish split of the nasal vowels provides an interesting problem for standard OT. The problem hinges on OT's use of individual words as inputs and candidates, which do not require any sort of $\mathcal{D}$-constraints to regulate the contrasts between unrelated words. Instead, OT is limited to just faithfulness and markedness constraints. To see why this is a problem, consider a possible OT analysis of the late Lechitic merger of the nasal vowels. Recall that early Lechitic *[ɔ̃] and *[ɛ̃] merged to *[ɜ̃]. This requires a change in vowel color, so FAITH-[color] (or alternatively, FAITH-[back] and FAITH-[round]) must be outranked by some constraint which prefers central [ɜ̃] to [ɔ̃] and [ɛ̃]. Clearly, this higher constraint cannot be a faithfulness constraint, since [ɜ̃], [ɔ̃], and [ɛ̃] are identical with respect to every property except color, and FAITH-[color] is already accounted for. Thus, we must rely on markedness constraints for these vowels, ranked so that the vowels with color are dispreferred to the colorless vowel:

(36) *OT analysis of late Lechitic loss of vowel color*

/ɛ̃ ɔ̃/	✗ɛ̃	✗ɔ̃	FAITH [color]	✗ɜ̃
a. ɛ̃ ɔ̃	✗!	✗!		
b. ɛ̃ ɜ̃	✗!		✗	✗
c. ɜ̃ ɔ̃		✗!	✗	✗
✓ d. ɜ̃ ɜ̃			✗^{2}	✗^{2}

This analysis works fine for late Lechitic, mapping both input vowels to the same colorless central vowel. However, it already has a conceptual problem. OT markedness constraints are used to derive typologies. The ranking ✗ɛ̃,✗ɔ̃ ≫ ✗ɜ̃ suggests that the nasal vowels [ɛ̃] and [ɔ̃] are more marked than [ɜ̃], and should therefore appear in fewer inventories. This is exactly opposite to what we find cross-linguistically, yet this ranking is required in order to get decolorization of the nasal vowels. If this OT analysis of late Lechitic is correct, then the typological foundation for markedness constraints must be abandoned (not a terrible conclusion, since this is exactly the stance I take with $\mathcal{M}$-constraints in FDM-OT). But there are other problems for an OT analysis of the nasal vowels.

Consider the analysis needed for late Middle Polish, when the short nasal vowel fronted to *[ɛ̃]. The constraints must be ordered as shown below in order to get the correct output:

(37) *OT analysis of late Middle Polish colorization of the short nasal vowel*

/ɜ̃/	×ɔ̃	×ɜ̃	×ɛ̃	FAITH [color]
a. ɜ̃		×!		
✓ b. ɛ̃			×	×
c. ɔ̃	×!			×

The ranking ×ɜ̃ ≫ ×ɛ̃ is exactly the opposite ranking required for late Lechitic. The implication of this reranking is that [ɛ̃] is *less* marked than [ɜ̃]. Clearly, these OT markedness constraints cannot represent anything universal, since the relative markedness of [ɛ̃] and [ɜ̃] depends on which stage of Polish we are looking at. This means that the fact that the short nasal vowel fronted is unpredictable and arbitrary; it could just as easily have stayed central [ɜ̃], or even shifted to some other vowel completely.

In the FDM-OT analysis, markedness rankings are consistent between stages of a language since they represent articulatory considerations grounded in the physical world. Two sounds cannot arbitrarily switch their relative markedness under FDM-OT because $\mathcal{M}$-constraints cannot be reranked. In addition, there is a predictable, functional motivation for [ɜ̃] to front in FDM-OT: its color needs to be more perceptually distinct from that of [ɔ̃w̃]. This is a direct consequence of treating candidates as sets of words which can influence each other's phonology. The short nasal vowel fronts precisely because the long nasal vowel backs; the two changes are intimately related in a way that standard OT cannot capture. Thus, what appears to be an arbitrary sound change in OT is actually a predictable sound change in FDM-OT.

3.6 Conclusion

In this chapter, I have constructed an analysis of the opaque nasal vowel alternation based on its historical origins within the framework of FDM-OT. A novel piece of my analysis is strong lexicon optimization, which selects underlying representations that are phonologically identical to their outputs. By having strong lexicon optimization interspersed between serially ordered diachronic sound changes, the analysis maintains the serialism required to account for opacity without sacrificing monostratality in the synchronic grammar. In addition, the set-based nature of the inputs and candidates in FDM-OT provide a principled explanation for why the back vowels evolved the way they did, especially in Middle Polish, when the nasal vowels split into two colors in order to maintain a sufficient perceptual contrast with each other. This type of analysis is unavailable to frameworks such as standard OT which treat inputs and candidates as individual words.

Chapter 4: Palatal Mutation as Allomorph Selection

Most consonants in Polish alternate in certain phonological and morphological environments as part of a general process often referred to as 'palatalization'. Though only some of the alternations truly involve articulations in the palatal region, they all originated historically through assimilation to a following front vowel and, for a period of time, actually involved truly palatalized segments with a high, front articulation symbolized in the IPA by superscript [^{j}]. Various diachronic sound changes mutated many of these newly palatalized consonants to other places and manners of articulation, masking their palatal origins. I will thus call this group of alternations 'palatal mutations' to distinguish them from uniformly true palatalization. For some consonants, the effects of the palatal mutations were so drastic that in modern Polish, they are opaque. As expected from the FDM-OT analyses of opacity in Chapters 2 and 3, the palatal mutations are no longer completely productive: front vowels do not generally require palatal mutation to occur. This implies that palatal mutation should not be encoded in the constraint hierarchy, which governs synchronically productive phonology.

However, while the palatal mutations are not technically synchronically productive, there are various morphemes, such as the locative singular [-ɛ], which do trigger the palatal mutations for all stems they attach to, including recent loanwords and nonce forms. Other morphemes, such as the instrumental singular [-ɛm] never trigger the palatal mutations. This points to a solution which the palatal mutations are encoded in the morphology. The palatal mutations could be explained by allowing the phonology and morphology to refer to each other (e.g. by having a phonological process of palatal mutation that is sensitive to particular morphological boundaries, or allowing morphemes to subcategorize for particular phonological features). In this chapter, I advocate a less powerful analysis which keeps the phonology and the morphology distinct, neither component able to directly refer to each other. The morphology is responsible for combining morphemes based solely on morphological features, and the phonology is responsible for mapping inputs to well-formed outputs based solely on phonological features. I show that even under this restricted view of morphology and phonology, the very complex phenomenon of palatal mutation can be analyzed, and thus, that there is no need to allow morphemes to refer to phonological features, or for phonology to refer to morphemes.

In §4.1, I provide the relevant data showing the modern palatal mutations. In §4.2, I explore how FDM-OT must be structured in order to account for morphological phenomena. In particular, the lexicon must be composed of morphemes (rather than fully formed words, as simplistically assumed in Chapters 2 and 3). This has implications for strong lexicon optimization, predicting that every extant morpheme has all of its surface allomorphs listed in its underlying representation. Using this richer version of FDM-OT in §4.3–6, I provide an analysis in which palatal mutation is derived by selection of allomorphs by general phonological considerations. This analysis allows for palatal mutation to be fully productive at certain morphological boundaries but unproductive in the general case. A key component of this analysis is that productive morphemes can overwrite portions of novel morphemes through coalescence, forcing them to undergo palatal mutation. Finally in §4.7, I summarize the major results of this chapter.

4.1 Palatal mutations in modern Polish

There are actually multiple palatal mutations that occurred throughout the history of Polish and the other Slavic languages, due to a variety of palatalizing sound changes at different points in time, each with their own effects. Because of the complexity of the various palatal mutations, I focus only on the later palatal mutations during the period of disintegration, or pre-West Slavic (PWS). These include (i) the true palatalization of the plain labials *[p b f v m] and plain coronals *[t d s z n l r], which was triggered by old and new front vowels,[1] and (ii) the so-called 'second velar palatalization' of *[k g x], which was triggered only by new front vowels such as PWS *[æː] < PSl *[ɑj] (since old front vowels had already mutated the velars under the 'first velar palatalization'). No other consonants during PWS could undergo the palatal mutations, either because they were already palatalized from previous sound changes or could not appear in the relevant environment, so their modern reflexes are ignored here. The resulting palatal mutations often caused a shift in both the place and manner of articulation of the affected consonant, masking its original nature.

The effects of the palatal mutations can be seen in older native stems such as [ɔ**ʐ**ɛx] 'nut' (from PSl *[ɑ**r**æːxu]), which shows the palatal mutation PSl *[r] > [ʐ]; cf. the Bulgarian, Macedonian, Russian, and Slovak cognate [ɔ**r**ɛx], with unmutated [r] in the same environment (Carlton 1991:340–341). However, the palatal mutations within stems are no longer productive in modern Polish, as evidenced by numerous lexical exceptions allowing [r] to appear unmutated before front vowels: [**r**ɛtsɛpta] 'prescription', [**r**ɛsɔr] 'spring', and even [**rʲ**iŋk] 'sporting ring', with [r] before [i]

[1] But not triggered by [j], since this had already caused the palatalizing sound change known as 'jotation' at an earlier stage.

showing true palatalization, but not palatal mutation to [ʐ]. Despite this lack of productivity within stems, certain morphemes, such as the locative singular suffix [-ɛ], trigger palatal mutation without exception, as in [mu**ʐ**ɛ] 'wall (LOC SG)' (cf. [mu**r**] 'wall (NOM SG)', a borrowing from French [myʁ] 'wall').[2] This difference in productivity is very clear in [**r**ɛsɔ**ʐ**ɛ] 'spring (LOC SG)', in which tautomorphemic [ɛ] does not trigger the palatal mutation of stem-initial [r] to [ʐ], but the [-ɛ] of the locative singular suffix does trigger palatal mutation (cf. [rɛsɔ**r**] 'spring (NOM SG)'). In this section, I cite data which exemplify the palatal mutations from the locative singular forms of noun stems.

4.1.1 Palatal mutation of labials

The palatal mutation for the modern labial consonants is similar to its historical roots; [p b f v m] palatalize to [pʲ bʲ fʲ vʲ mʲ], with epenthesis of a palatal glide [j] between the palatalized labial and a following non-high vowel (though labial-velar [w] mutates to [l]; see the discussion of velars in §4.1.3):

(1)

stem UR	*GEN SG*	*LOC SG*	*gloss*
/xwɔ**p**/	xwɔ**p**a	xwɔ**pʲj**ɛ	'peasant'
/xlɛ**b**/	xlɛ**b**a	xlɛ**bʲj**ɛ	'bread'
/ʂɛ**f**/	ʂɛ**f**a	ʂɛ**fʲj**ɛ	'boss'
/krakɔ**v**/	krakɔ**v**a	krakɔ**vʲj**ɛ	'Kraków'
/twu**m**/	twu**m**u	twu**mʲj**ɛ	'crowd'

Similar true palatalization of the labials occurs before [i] generally, as in non-native [**pʲ**ilɔt] 'pilot'. But [ɛ] does not generally trigger palatal mutation in the labials,

[2] This is claimed in numerous definitive sources on Polish morphophonology, such as Gussmann 1980, Rubach 1984a, Czaykowska-Higgins 1988, etc. Having conducted informal experiments with native speakers confirming this claim, I adopt it here without dispute.

either within stems like [**p**ɛx] ‘bad luck’ (not ×[**p^{j}j**ɛx]), or across other morpheme boundaries, such as the instrumental singular: [xwɔ**p**ɛm] ‘peasant (INS SG)’, not ×[xwɔ**p^{j}j**ɛm]. Thus, even for what seems at first to be a clear case of transparent articulatory assimilation, there is a deeper story lurking within the morphology.

4.1.2 Palatal mutation of coronals

For many of the coronals, palatal mutation is nearly as transparently connected to true palatalization as it is for the labials, with only slight mutations in place and/or manner. The dental fricatives and nasal [s z n] shift to alveolo-palatals [ɕ ʑ ɲ̟] (2a), a minor difference in place of articulation from palatalized dentals [s^j z^j n^j]. The dental stops [t d] also become alveolo-palatals rather than palatalized dentals, but undergo an additional change in manner, affricating to [tɕ dʑ] (2b):

(2)		*stem UR*	*GEN SG*	*LOC SG*	*gloss*
	a.	/tʂa**s**/	tʂa**s**u	tʂa**ɕ**ɛ	‘time’
		/vɔ**z**/	vɔ**z**u	vɔ**ʑ**ɛ	‘cart’
		/pa**n**/	pa**n**a	pa**ɲ̟**ɛ	‘gentleman’
	b.	/bra**t**/	bra**t**a	bra**tɕ**ɛ	‘brother’
		/sɔ̃j̃ɕa**d**/	sɔ̃j̃ɕa**d**a	sɔ̃j̃ɕa**dʑ**ɛ	‘neighbor’

These minor deviations from truly palatalized segments can be attributed to acoustic and articulatory factors that favor alveolo-palatals over palatalized dentals, and in the general palatal region, favor affricates over stops. Thus, these palatal mutations are not at all surprising or opaque (though in the context of the palatal mutations of the velars, the coronals are in fact opaque; see §4.1.3).

As was shown earlier in this section, the alveolar trill [r] undergoes a somewhat more serious shift in both place and manner, becoming the post-alveolar fricative [ʐ]:

(3)	*stem UR*	*GEN SG*	*LOC SG*	*gloss*
	/mu**r**/	mu**r**u	mu**ʐ**ɛ	‘wall’
	/dɔktɔ**r**/	dɔktɔ**r**a	dɔktɔ**ʐ**ɛ	‘doctor’
	/sɛ**r**/	sɛ**r**a	sɛ**ʐ**ɛ	‘cheese’

While coronal [l] does not undergo palatal mutation, it is the result of the palatal mutation of labial-velar [w], discussed in §4.1.3.

4.1.3 Palatal mutation of velars

The velars undergo more severe place changes than the coronals do, switching from a dorsal to a coronal articulation in all instances: dental for [k g], post-alveolar for [x], and alveolar for [w]. In addition, [k g] shift from stops to affricates, and [w] shifts from a central to a lateral approximant (which can be seen as a change in manner):

(4)	*stem UR*	*GEN SG*	*LOC SG*	*gloss*
	/ʐɛ**k**/	ʐɛ**k**a	ʐɛ**ts**ɛ	‘river’
	/drɔ**g**/	drɔ**g**a	drɔ**dz**ɛ	‘road’
	/mu**x**/	mu**x**a	mu**ʂ**ɛ	‘fly’
	/ʂkɔ**w**/	ʂkɔ**w**a	ʂkɔ**l**ɛ	‘school’

These data show how problematic the palatal mutations are. If dental [ts] is an allowed output of palatal mutation, why is dental [t] required to change its place (and manner) of articulation to become alveolo-palatal [tɕ], instead of remaining dental [ts]? This would involve the same change in manner, but no change in place. And if alveolo-palatal [tɕ] is a valid palatal mutation, why must velar [k] change its place of articulation as far forward as dental [ts], rather than stopping at the closer (i.e. more faithful) alveolo-palatal [tɕ]? Likewise, if [w] can mutate to [l], why is [r] unable to do so as well, since a change from [r] to [l] would involve less of a change (manner only) than from [r] to [ʐ] (manner and place)? The palatal mutations are clearly governed by principles beyond surface phonotactics and input faithfulness to the stem-final consonant.

4.1.4 Summary

The following is a summary of the set of palatal mutations discussed above (recall that the remaining consonants of Polish do not undergo this palatal mutation and are not discussed here):

(5)	*plain*	*mutated*	*place shift*	*manner shift*
	p b f v m	p^j b^j f^j v^j m^j	labial → palatalized labial	
	s z n	ɕ ʑ ɲ̟	dental → alveolo-palatal	
	t d	tɕ dʑ	dental → alveolo-palatal	stop → affricate
	r	ʐ	alveolar → post-alveolar	trill → fricative
	k g	ts dz	velar → dental	stop → affricate
	x	ʂ	velar → post-alveolar	
	w	l	labial-velar → alveolar	central → lateral

In the following chart, I give the specifications for the plain segments and their palatal mutations for the phonological properties used in the analysis in this chapter. The plus symbol '+' is used to indicate that segments in a given row have the property in the given column (or more appropriately, have the same value for that property as the other segments marked with '+', since I assume that properties can be multi-valued), while a blank cell indicates that the segments in question lack (or do not have a '+' value for) the relevant property. Shading indicates that specification for the property does not change between the plain segment and its palatal mutation, while white cells indicate that the property does change. Note that the following chart shows that the palatal mutation never changes the properties stop, narrow, or nasal, and that there is no listed property in the chart which always changes in every palatal mutation (however, not all possible properties are shown here, only those which play an interesting role in the analysis; for example, I ignore properties like voicing, which never changes under palatal mutation, and anterior, which arguably always changes, because they do not interact in a significant way with the other properties, and thus, do not warrant an explicit ranking of their corresponding $\mathcal{F}$-constraints):

(6)

		labial	coronal	palatal	dorsal	retracted	stop	frication	sonorant	narrow	nasal	lateral
p	b	+					+			+		
p^j	b^j	+	+	+	+		+			+		
f	v	+						+		+		
f^j	v^j	+	+	+	+			+		+		
	m	+					+		+	+	+	
	m^j	+	+	+	+		+		+	+	+	
t	d		+				+			+		
tɕ	dʑ		+	+	+		+	+		+		
s	z		+					+		+		
ɕ	ʑ		+	+	+			+		+		
	n		+				+		+	+	+	
	ɲ̟		+	+	+		+		+	+	+	
	r		+		+	+			+	+		
	ʐ		+		+	+		+		+		
k	g				+	+	+			+		
ts	dz		+				+	+		+		
x					+	+		+		+		
ʂ			+		+	+		+		+		
	w	+			+	+			+			
	l		+						+			+

The properties listed above generally correspond to standard distinctive features of the same name from modern generative phonology (as exemplified by Keating 1987), with the notable exceptions of five properties: *palatal* (which indicates segments articulated in the palatal region, such as true palatals, palatalized segments, and alveolo-palatals; all palatal sounds are also redundantly both coronal and dorsal), *retracted* (which can mostly be identified with the traditional feature

[+back], used to mark any sound made with a tongue dorsum retracted toward the velum; see §4.4.3 for discussion), *stop* (which can be identified with the traditional feature [–continuant]), *frication* (the property of aperiodic noise, distinguishing affricates from stops, and fricatives from sonorants), and *narrow* (a property adopted from Padgett 2002b, in which narrowness distinguishes the somewhat close [ʋ̞] from other approximants in Russian; used here in a similar fashion as a property of the trill [r]; see §4.6.3 for discussion).

Within rule-based frameworks, the alternations indicative of the palatal mutations have been analyzed with ordered rules, requiring numerous abstract intermediate representations (e.g. Gussmann 1980, Rubach 1984a, and Czaykowska-Higgins 1988). As an instance of opacity, the palatal mutations are problematic for standard OT with direct mapping between the input and output, requiring some sort of extra machinery (see for example Łubowicz 1998 and Nowak 2000, in which constraint conjunction is used). Something extra is needed for FDM-OT as well: a view of morphology in which underlying representations are composed of multiple allomorphs, which can have the palatal mutations encoded in them.

4.2 Morphology in FDM-OT

The analysis of palatal mutations in the extant Polish vocabulary follows directly from the same type of analyses in the previous chapters. The historical alternations are stored in the lexicon for every word the speaker has heard, including all the inflected forms of the same stem. For example, 'wall' will have in its lexical entry, at minimum, an unmutated form to map to the nominative singular output [mur] and a mutated form to map to the locative singular output [muz̡ɛ]. This will correctly predict the modern outputs for all of the palatal mutations for existing lexical items.

However, in the previous chapters, the lexicon was assumed to consist of fully formed words. Obviously, this is too simplistic to be the true nature of the lexicon, since there is nothing inherent to an arbitrary list that would extend to morphological derivations of new borrowings or nonce forms. Since they are productive for some morphemes (the locative singular, the vocative singular, etc.), the palatal mutations must be encoded somewhere. The morpheme used to exemplify the palatal mutations in this chapter is the locative singular, which has a characteristic [-ɛ] that historically triggered the changes that led to the modern palatal mutations. That fact that this morpheme has a front vowel on the surface is not sufficient to trigger the palatal mutations, as there are other [ɛ]-initial suffixes, such as the instrumental singular [-ɛm], which invariably do not trigger palatal mutation. Thus, we have [mu**r**ɛm] 'wall (INS SG)', not ×[mu**ʐ**ɛm] in contrast to [mu**ʐ**ɛ] 'wall (LOC SG)'. Since the palatal mutations are not productive within stems or across morpheme boundaries generally, only at specific morpheme boundaries such as the locative singular, the palatal mutations are idiosyncratic quirks of particular *morphemes* and not a general part of the phonology.

In order to allow some morphemes to make productive changes in nonce words and new borrowings, the lexicon must be composed of morphemes (the reader should note that this is not a novel proposition). While significantly different from the word-based lexicon used for expository simplicity in Chapters 2 and 3, the morpheme-based lexicon advocated here does not impact those analyses, since they do not rely on generalizations triggered by specific morphological environments. To see how the morpheme-based lexicon operates in FDM-OT, consider the following mini-lexicon of English: {/tæk/$_{\text{TACK}}$, /tæks/$_{\text{TAX}}$, /-Ø/$_{\text{SG}}$, /-z/$_{\text{PL}}$}. Subscripts mark the meanings of the morphemes (the morpheme whose meaning is ‘tack’ has the underlying representation /tæk/, while the singular morpheme is null, etc.). The output then consists of the set {[$tæk_{TACK}$-$Ø_{SG}$], [$tæk_{TACK}$-s_{PL}], [$tæks_{TAX}$-$Ø_{SG}$], [$tækz_{TAX}$-$ɨz_{PL}$]} (that is, the set consisting of the singular and plural forms of the words ‘tack’ and ‘tax’), where the mapping between underlying representations and output substrings is marked by identical subscripts.

As used in previous chapters, the principle of strong lexicon optimization (SLO) requires a word’s lexical listing to match its output form. Since the lexicon is a set of morphemes, the effect of SLO on the lexicon is more complex than assumed in Chapters 2 and 3. By denying the principle of lexical minimization, FDM-OT allows a single morpheme to have more than one underlying allomorph. Some work in OT, such as Mester 1994, Burzio 1996, Kager 1996, 1999b, and Rubach 2001, allow some amount of stored allomorphy in the lexicon. Taking this allowance one step further, I claim that the underlying representation of a given morpheme in FDM-OT consists of *all* of its surface allomorphs. Thus, the English plural would not be stored simply as /-z/ as given above, but rather as the set consisting of all of the ways it can be realized in actual words: {/-z/,/-s/,/-ɨz/,/-Ø/,/-ən/,...}.

It comes as no surprise that the plural allomorphs /-Ø/ and /-ən/ need to be stored in the lexicon, since they only occur with limited sets of nouns: {*deer*, *fish*, *sheep*, ...} and {*ox*, *child(r)*}, respectively. Few researchers would dare posit that these realizations of the plural morpheme are derived productively from a single underlying phonological string that also derives surface [-z], [-s], and [-ɨz]. As has been shown in the previous chapters, there is no problem with also storing productive information in the lexicon. By the hypothesis of richness of the base, the language learner must still construct a constraint hierarchy that derives productive alternations. It is only later, once forms come into use, that the speaker no longer needs to derive a word from scratch every time it is uttered and can rely on past pronunciations to bypass the constraint hierarchy. As discussed in §1.1.5 of Chapter 1, such behavior is supported by psycholinguistic experimentation, leading credence to SLO. The lexicon in FDM-OT, as a set of morphemes shaped by the effects of SLO, is a set consisting of exhaustive sets of allomorphs. Allomorphs used to build extant words are selected by morphological feature-matching or subcategorization. For example, for the English plural, each allomorph listed in the underlying representation is marked for when it is used: /-z/ is marked for use with 'tag', while /-s/ is marked for use with 'tack'. This use of SLO does not preclude the need for phonology, since there still must be an explanation for how allomorphs are selected for nonce forms, such as [wʌg]. The grammar of English productively requires obstruents clusters to agree in voicing, so the constraint hierarchy still forces [wʌg] to take /-z/ in the plural in order to satisfy the phonology of English.

In summary, each morpheme in the lexicon in FDM-OT has an underlying representation that consists of all of its observed surface allomorphs. The correct allomorphs are selected either *morphologically* by means of morphological markings from prior usage (in the case of extant words) or *phonologically* by means of the constraint hierarchy (in the case of nonce words).

4.3 FDM-OT analysis of the palatal mutations of labials

Recall from §4.1.1 that, for stems ending in a labial consonant, the palatal mutation involves both a secondary palatal articulation and epenthesis of the palatal glide [j]:

(7)	*stem UR*	*GEN SG*	*LOC SG*	*gloss*
	/xwɔ**p**/	xwɔ**p**a	xwɔ**p**^{j}**j**ɛ	‘peasant’
	/xlɛ**b**/	xlɛ**b**a	xlɛ**b**^{j}**j**ɛ	‘bread’
	/ʂɛ**f**/	ʂɛ**f**a	ʂɛ**f**^{j}**j**ɛ	‘boss’
	/krakɔ**v**/	krakɔ**v**a	krakɔ**v**^{j}**j**ɛ	‘Kraków’
	/twu**m**/	twu**m**u	twu**m**^{j}**j**ɛ	‘crowd’

Also recall that in general, [ɛ] does not trigger the palatal mutation, so outputs can contain labial+[ɛ] sequences, as in the stem [pɛx] ‘bad luck’ and in instrumental singular forms of labial-final stems, such as [xwɔpɛm] ‘peasant (INS SG)’. Thus, [pɛ], [bɛ], [fɛ], etc., are allowed on the surface with neither palatalization nor the palatal glide that are indicative of palatal mutation.

4.3.1 The general grammar

Because the palatal mutations are not fully productive, but do occur in some cases, the grammar must correctly predict that both [pɛ] and [p^jjɛ] are valid sequences in the output, while [pjɛ] and [p^jɛ] are not. Labials are always palatalized when followed by the front high vocoids [j] or [i], and never when followed by any other segment, including other palatalized sounds (Nowak 2001). Both [p^j] and [j] are palatal, and [p] and [ɛ] are not, where palatal sounds are characterized by a high, front articulation of the tongue.[3] The undominated constraint $\mathcal{M}$-Cj rules out a sequence of a non-palatal consonant and [j], preventing [pjɛ] from surfacing (this constraint is likely more general, ruling out non-palatal consonants before [i] as well, since only palatal consonants are allowed before [j] and [i] in Polish). Other relevant constraints to the

[3] Nowak (2001) argues that palatal sounds are also distributed, in order to also capture the difference between high and front [i] (distributed, therefore palatal) and [ɪ], spelled ⟨y⟩ (not distributed, so not palatal, despite being high and front).

analysis include: $\mathcal{F}$-segment (which has the combined effect of McCarthy and Prince's (1995) MAX and DEP constraints from Correspondence Theory, punishing both deletion and epenthesis), $\mathcal{F}$-palatal (which punishes changes in the specification of the palatal property between corresponding segments in the input and output), and the $\mathcal{M}$-constraints $\mathcal{M}$-p^j and $\mathcal{M}$-p (which respectively ban occurrences of palatalized [p^j] and plain [p] in the output), with $\mathcal{M}$-p^j universally outranking M-p because [p^j] involves more complicated—and therefore more difficult—articulation than [p] does.

The following tableau shows the required ranking, with $\mathcal{M}$-Cj undominated and $\mathcal{F}$-segment ranked very high. The sets of words used are defined by the expression [p(^{j})(j)ε]:

(8)

pε$_1$ p^jε$_3$ / pjε$_2$ p^jjε$_4$	$\mathcal{M}$ Cj	$\mathcal{F}$ seg	$\mathcal{M}$ p^j	$\mathcal{F}$ pal	$\mathcal{M}$ p
a. pε$_1$ p^jε$_3$ / pjε$_2$ p^jjε$_4$	×!		×^{2}		×^{2}
✓ b. pε$_{1,3}$ / p^jjε$_{2,4}$			×	×^{2}	×
c. pε$_1$ p^jε$_3$ / p^jjε$_{2,4}$			×^{2}!	×	×
d. pε$_{1,2}$ / p^jjε$_{3,4}$		×^{2}!	×		×
e. pε$_{1,2,3,4}$		×^{2}!		×^{2}	×

The fully faithful candidate (8a) violates $\mathcal{M}$-Cj because word 2 is an instance of a non-palatal consonant followed by [j]. The remaining candidates do not have this offending word, satisfying $\mathcal{M}$-Cj by merging various inputs in the output to avoid an ill-formed sequence. Candidates (8d) and (8e) involve mergers which require the insertion or deletion of [j] from the input to the output, violating highly ranked $\mathcal{F}$-segment. This leaves candidates (8b) and (8c) in contention as the output. Because $\mathcal{M}$-p^j is ranked over $\mathcal{F}$-palatal, spurious instances of [p^j] decide between these two candidates. Candidate (8b) only contains one instance of [p^j], required in order to satisfy the higher ranked constraints. Candidate (8c) contains this same instance of [p^j], but it also contains an extra [p^j] in the word [p^jɛ]. Thus, candidate (8b) is selected at the winner, having palatalized [p^j] only before [j]. Note that this while this analysis is just for [p], it holds equally well for all of the labials, since they undergo the same type of palatal mutation, modulo their specifications for voicing, nasalization, continuancy, etc.

4.3.2 The lexicon

There are many possible analyses of the morphological composition of outputs like [xwɔp^jjɛ] 'peasant (LOC SG)'. For clarity of exposition, I will assume that the surface string [xwɔp^j] is the realization of 'peasant' and [jɛ] is the realization of the locative singular (in §4.6, I revise this analysis; I simplify here in order to focus on other facets of the analysis). For the instrumental singular [xwɔpɛm], the string [xwɔp] is the realization of 'peasant', and [ɛm] is the realization of the instrumental singular. By SLO, the lexicon should look something like (9), with the surface allomorphs stored in the underlying representations:

(9)

PEASANT	=	/xwɔp/	_____ + {INS SG,…}
		/xwɔp^j/	_____ + {LOC SG,…}
LOC SG	=	/-jɛ/	{PEASANT,…} + _____
INS SG	=	/-ɛm/	{PEASANT,…} + _____

The morphological contexts that dictate which environments the various morphemes are used come directly from SLO. For example, the speaker consistently hears and uses [xwɔp] with the instrumental singular (and other morphemes), so when SLO takes effect, /xwɔp/ is stored in the lexicon marked for use with the instrumental singular (and whatever other morphemes it is used with).[4] The relevant outputs are formed by concatenation of the necessary morphemes: /xwɔp^j/ + /-jɛ/ → [xwɔp^{j}jɛ] for the locative singular of 'peasant', etc.

4.3.3 Palatal mutation in the locative singular

The above morphological analysis is based only on extant noun stems, however, and does not cover how a speaker will pronounce the locative and instrumental singulars of new borrowings or nonce forms. This must be a part of the productive synchronic phonology. As stated previously, the palatal mutations are productive for the locative singular, so the synchronic grammar must be able to derive the same results for an

[4] In a fully connectionist model of the lexicon in which each use of a particular word strengthens the neural connections between the morphemes comprising that word, morphemes would likely be specified for which other morphemes they are used with (or more precisely, which ones they have predominately been used with in previous utterances) in a more gradient fashion dependent on frequency of use. Thus, two morphemes which are often used together to form a frequent word would have stronger specifications than uncommon morphemes that are not used together very often. Such gradient morphological specification is well beyond the scope of this dissertation, and I simplistically assume only a categorical specification provided by a single instance of SLO that applies equally to the entire lexicon.

arbitrary labial-final input as it does for stored stems like 'peasant', which has two lexical allomorphs that are pre-specified for which contexts they are used in.

For the purposes of this discussion, I assume that possible outputs are defined by the expression [x(w)ɔp(^{j})(j)ɛ(m)], with the [m] only appearing in the instrumental singular, and the optional [w] distinguishing between extant stems (that have [w]) and novel stems (without [w]). The input is a bit larger, allowing for different possible underlying representations for the relevant stems and affixes. For visual ease, I list stems graphically above affixes, separated by a dotted line. Stems are subscripted with numbers (allomorphs of the same stem are subscripted with the same number), while the affixes are mnemonically subscripted with IS and LS, for instrumental singular and locative singular, respectively.

In the following tableau, the subscript 1 is used for the extant stem /xwɔp(^{j})/ 'peasant', while 2 and 3 are used for the possible nonce stems /xɔp/ and /xɔp^{j}/ (which are expected to exist according to richness of the base). To help visually distinguish extant stems from nonce stems, nonce stems are italicized. In all of the output candidates, the only forms of [xwɔp(^{j})] 'peasant' allowed to surface are the allomorphs contextually chosen by the morphological marking in the lexicon (9). This marking and the resulting allomorph choice is assumed to hold without exception for extant words. However, since new stems do not have the same type of lexical entries (because they do not have the benefit of prior usage solidifying particular allomorph shapes in various environments), they cannot be marked for use with any particular morphemes, so all possible combinations must be considered. The constraint hierarchy, i.e. the phonology, must rule out the ungrammatical combinations (such as palatal mutation in the instrumental singular and lack of palatal mutation in the locative singular).

I also introduce a new constraint, $\mathcal{F}$-uniqueness, which preserves the underlying phonological uniqueness of stems into all of their output forms. For example, if underlying stem *x* is phonologically distinct from underling stem *y*, then the locative singular form of *x* must be also be phonologically distinct from the locative singular form of *y*. $\mathcal{F}$-uniqueness does not play a crucial role in the following tableau, but in later stages of the analysis, it is required in order to ensure that the proper, distinct, allomorphs are selected, rather than creating homophonous forms.

Because the palatal mutations never change specifications for the properties of voicing, stop, narrowness, or nasality, I assume for simplicity that the $\mathcal{F}$-constraints for these properties are undominated by other constraints in the analysis, so no ranking arguments will be given for these four constraints.

(10)	𝓜 Cj	𝓕 seg	𝓜 p^j	𝓕 uniq	𝓕 pal	𝓜 p
xwɔp$_1$ xwɔp$^j{}_1$ *xɔp*$_2$ *xɔp*$^j{}_3$ -ɛm$_{IS}$ -jɛ$_{LS}$						
a. xwɔp$_1$-ɛm$_{IS}$ xwɔp$^j{}_1$-jɛ$_{LS}$ *xɔp*$_2$-ɛm$_{IS}$ *xɔp*$_2$-jɛ$_{LS}$ *xɔp*$^j{}_3$-ɛm$_{IS}$ *xɔp*$^j{}_3$-jɛ$_{LS}$	✗!		✗^{3}			✗^{3}
✓ b. xwɔp$_1$-ɛm$_{IS}$ xwɔp$^j{}_1$-jɛ$_{LS}$ *xɔp*$_2$-ɛm$_{IS}$ *xɔp*$^j{}_2$-jɛ$_{LS}$ *xɔp*$_3$-ɛm$_{IS}$ *xɔp*$^j{}_3$-jɛ$_{LS}$			✗^{2}	✗^{4}	✗^{2}	✗^{2}
c. xwɔp$_1$-ɛm$_{IS}$ xwɔp$^j{}_1$-jɛ$_{LS}$ *xɔp*$_2$-ɛm$_{IS}$ *xɔp*$^j{}_2$-jɛ$_{LS}$ *xɔp*$^j{}_3$-ɛm$_{IS}$ *xɔp*$^j{}_3$-jɛ$_{LS}$			✗^{3}!	✗^{2}	✗	✗^{3}
d. xwɔp$_1$-ɛm$_{IS}$ xwɔp$^j{}_1$-jɛ$_{LS}$ *xɔp*$_2$-ɛm$_{IS}$ *xɔp*$_2$-ɛ$_{LS}$ *xɔp*$_3$-ɛm$_{IS}$ *xɔp*$_3$-ɛ$_{LS}$		✗^{2}!	✗	✗^{4}	✗^{2}	✗^{3}

The fully faithful candidate (10a) has palatal mutation only from historical sources, preserved in the lexicon and forced to surface due to morphological markings on extant vocabulary such as ‘peasant’. For novel input stems like /xɔp/ and /xɔp^j/, with no prior disposition towards palatal mutation, the stems (and the affixes) surface faithfully to their inputs, with no deletion or insertion of [j] (which would violate 𝓕-segment), with no changes in palatalization (which would violate 𝓕-palatal), and with no homophonous forms of distinctive stems (which would violate 𝓕-uniqueness). However, this candidate violates the highly ranked constraint 𝓜-Cj, which punishes [xɔpjɛ] (the locative singular of stem 2), with plain [p] next to [j], so it cannot emerge as the winner.

In the remaining candidates, $\mathcal{M}$-Cj is satisfied, but at the expense of lower ranked constraints. In candidate (10b), palatalization in the novel stems is neutralized, so that the input stems 2 and 3 have the same instrumental singular forms and the same locative singular forms. This incurs four violations of $\mathcal{F}$-uniqueness: one each for the instrumental singulars and locative singulars of stems 2 and 3, none of which is unique on the surface. This candidate also violates $\mathcal{F}$-palatal: one violation for the palatalization of stem 2 in the locative singular and one for the depalatalization of stem 3 in the instrumental singular. In comparison, candidate (10c) only changes the palatalization of stem 2 in the locative singular; stem 3's underlying palatalization is allowed to surface, giving (10c) only a single violation of $\mathcal{F}$-palatal. $\mathcal{F}$-uniqueness is also violated less, since only the locative singular forms are homophonous. However, the extra instance of [p^j] in (10c) is enough for (10b) to defeat it, since $\mathcal{M}$-p^j outranks $\mathcal{F}$-uniqueness and $\mathcal{F}$-palatal. Candidate (10b) only has two violations of $\mathcal{M}$-p^j because, crucially, $\mathcal{M}$-constraints only evaluate phonetic strings, not separate, but homophonous, words. Since the locative singulars of stems 2 and 3 are pronounced exactly the same in (10b), as [xɔp^jjɛ], only one instance of [p^j] is counted for this phonetic string.

Finally, in candidate (10d), the stems are fully neutralized to have only plain [p] in all forms, to avoid violating $\mathcal{M}$-p^j. As a consequence, the initial /j/ of the locative singular must be deleted, since it would obligatorily trigger palatalization of [p]. This deletion incurs two violations of $\mathcal{F}$-segment, one for each separate word in which underlying /j/ does not surface. Candidate (10b) involves no deletion of /j/, so it does not violate $\mathcal{F}$-segment, which outranks $\mathcal{M}$-p^j. Therefore, (10b) is chosen by the grammar as the output. In this winning candidate, the instrumental singular /-ɛm/

does not trigger palatal mutation, while the locative singular /-ɛ/ does. Thus, FDM-OT is able to derive the palatal mutation of the labials in the locative singular by selecting the correct allomorphs of the stems and affixes based on morphological specification (for extant stems) and the constraint hierarchy (for nonce stems).

4.4 FDM-OT analysis of the palatal mutations of coronals

In §4.1.2, it was shown that coronal consonants follow three patterns of palatal mutation: the fricatives and nasals change only in place of articulation, from dental to alveolo-palatal (11a); the oral stops undergo the same change in place of articulation but also change in manner of articulation, from stops to affricates (11b); and the alveolar trill [r] mutates to the post-alveolar fricative [ʐ], changing both place and manner of articulation (11c):

(11)

	stem UR	*GEN SG*	*LOC SG*	*gloss*
a.	/tʂa**s**/	tʂa**s**u	tʂa**ɕ**ɛ	‘time’
	/vɔ**z**/	vɔ**z**u	vɔ**ʑ**ɛ	‘cart’
	/pa**n**/	pa**n**a	pa**ɲ**ɛ	‘gentleman’
b.	/bra**t**/	bra**t**a	bra**tɕ**ɛ	‘brother’
	/sɔ̃j̃ɕa**d**/	sɔ̃j̃ɕa**d**a	sɔ̃j̃ɕa**dʑ**ɛ	‘neighbor’
c.	/mu**r**/	mu**r**u	mu**ʐ**ɛ	‘wall’

As with the labials, I can ignore the voiced obstruents since they pattern analogously to their voiceless counterparts. However, other distinctions (continuancy, nasality, trill) cannot be ignored and are fully represented in the following discussion.

4.4.1 The general grammar

As evidenced by words such as [sɛr] ‘cheese’, [nɛrpa] ‘seal’, [tɛmu] ‘ago’, and [rɛsɔr] ‘spring’, as well as the instrumental singular forms of coronal-final stems, such as [tʂasɛm] 'time', [panɛm] 'gentleman', [bratɛm] 'brother', and [murɛm] 'wall', [ɛ] is not generally required to trigger the palatal mutation of a preceding coronal; palatal mutation only occurs with certain morphemes, such as the locative singular (11). This result for the coronals is identical to that for labials. Thus, the basic grammar of modern Polish should derive [sɛ], [nɛ], etc. as allowable sequences, not required to change to [ɕɛ], [ɲ̟ɛ], etc. (which must also exist in the output, since they are allowed to appear, for example, in the locative singular).

I begin by showing that before [ɛ], the coronal obstruents and nasals can have either dental or alveolo-palatal places of articulation, except the stops, which can only be dental. This is derived from the following assumptions, which seem relatively uncontroversial. First, I consider alveolo-palatals to have the palatal property (for the purposes of constraints such as $\mathcal{F}$-palatal), while dentals are not palatal. Second, though affricates are generally more difficult than stops formed at the same place of articulation (since affricates involve more muscle movement by the formation of a stop closure followed by delayed opening for the fricated release), I claim that the alveolo-palatal stop is universally harder than alveolo-palatal fricatives and affricates, because a palatal closure has a large area of contact between the tongue and the roof of the mouth. In order to release a stop with such a large contact area, more effort is required to move the tongue quickly enough to prevent audible frication in the release (Ladefoged 2001:144). Thus, I assume the universal ranking $\mathcal{M}$-c̟ ≫ $\mathcal{M}$-ɕ,$\mathcal{M}$-tɕ. By ranking a constraint such as $\mathcal{F}$-frication below $\mathcal{M}$-c̟, alveolo-palatal stops are forced to become affricates:[5]

[5] The astute reader will note that /c̟/ could just as easily become a fricative under this ranking. The exact transformation of /c̟/ depends on the ranking of $\mathcal{F}$-stop. Since it does not matter for this analysis whether /c̟/ maps to a fricative or an affricate, I do not analyze this detail.

(12) tε$_{1}$ ç̟ε$_{4}$ sε$_{2}$ ɕε$_{5}$ tsε$_{3}$ tɕε$_{6}$	ℱ pal	ℳ ç̟	ℱ fric	ℳ ɕ	ℳ tɕ
a. tε$_{1}$ ç̟ε$_{5}$ sε$_{2}$ ɕε$_{5}$ tsε$_{3}$ tɕε$_{6}$		×!		×	×
✓ b. tε$_{1}$ sε$_{2}$ ɕε$_{5}$ tsε$_{3}$ tɕε$_{4,6}$			×	×	×
c. tε$_{1,4}$ sε$_{2,5}$ tsε$_{3,6}$	×^{3}!				

The fully faithful candidate (12a) contains the highly marked alveolo-palatal stop and is ruled out by ℳ-ç̟. Candidates (12b) and (12c) avoid [ç̟] by merging it with some other segment. In (12b), /ç̟/ surfaces as an affricate, keeping its place of articulation, while in (12c), /ç̟/ and all other alveolo-palatals surface as dentals, keeping their manners of articulation. Because ℱ-palatal outranks ℱ-frication, (12b) wins, and Polish is prevented from having alveolo-palatal stops.

4.4.2 The lexicon

Coronal-final stems do not have clearly defined morphemic boundaries in the locative singular, which means that there are many possible underlying representations for the noun stems and the locative singular morpheme. For example, for the form [ts̪aɕε] 'time (LOC SG)', there are at least three possible partial lexicons that can result from SLO. As shown in (13), the [ɕ] that appears in the locative singular (bold-faced in the following examples) could be stored as part of an allomorph for 'time', while the locative singular is stored as [-ɛ] and [-jɛ] (the allomorph as determined in the previous section). Also shown is the UR for the instrumental singular, which remains the same as for labial-final stems:

(13)

TIME =	/tʂas/	_____ + {INS SG,…}
	/tʂa**ɕ**/	_____ + {LOC SG,…}
LOC SG =	/-ɛ/	{TIME,…} + _____
	/-jɛ/	{PEASANT,…} + _____
INS SG =	/-ɛm/	{TIME,PEASANT,…} + _____

This lexicon requires only a straightforward linear concatenation of the morphemes to produce the correct output for the locative singular [tʂaɕ-ɛ]. However, this analysis means that both the locative singular and the instrumental singular begin with [ɛ], and there would be no explanation for why the /ɛ/ in the locative singular triggers palatal mutation in nonce forms while the /ɛ/ in the instrumental singular does not. Thus, the lexicon in (13) must be rejected.

The next logical possibility to consider is that the [ɕ] that surfaces is actually part of the underlying representation of the locative singular and not a part of the stem. This means that the stem has an allomorph with no final consonant for use in the locative singular, while the locative singular has an allomorph beginning with /ɕ/ (shown in bold):

(14)

TIME	=	/tʂ̻as/	_____ + {INS SG,…}
		/tʂ̻a/	_____ + {LOC SG,…}
LOC SG	=	/-**ç**ɛ/	{TIME,…} + _____
		/-jɛ/	{PEASANT,…} + _____
INS SG	=	/-ɛm/	{TIME,PEASANT,…} + _____

The lexicon in (14) makes better predictions than the one in (13), since the [ç] that is expected to surface in the palatal mutation of nonce stems ending in /s/ is part of the productive locative singular morpheme, showing a clear phonological distinction between the mutating locative singular and the non-mutating instrumental singular, a distinction that is not available from the lexicon in (13). However, what this lexicon cannot predict is why the locative singular would cause the stem-final consonant for a nonce stem like /dras/ to be missing from the surface, either through deletion or merger (note that /tʂ̻as/ comes with a pre-truncated allomorph /tʂ̻a/, a benefit of extant stems that nonce stems do not share). The grammar cannot force a consonant to be deleted or merged when it is followed by its own palatal mutation, because there are words in which a consonant plus its palatal mutation are allowed to surface without consonantal loss, e.g. [bɛ**çç**ɛ] 'price drop (DAT)' (cf. [bɛ**ss**a] 'price drop (NOM)'), [stara**ɟ̊ɲɲ**ɛ] 'carefully' (cf. [stara**nn**ɪ] 'careful'), and [lɛ**kts**ɛvaz̻ɔ̃ntsɪ] 'disrespectful'. Other repairs might occur (such as the assimilation in [bɛççɛ] and [staraɟ̊ɲɲɛ]), but crucially, the grammar does not induce deletion or merger, which would be required in order for this to be the correct lexicon. Therefore, this lexicon must also be rejected.

The final option worth consideration is that the surface [ç] is part of the realizations of *both* the stem and the locative singular, meaning that it comes from two instances of underlying /ɕ/, one in the UR for locative singular and one in the UR for the stem, as shown in (15):

(15)

TIME	=	/tʂ̥as/	_____ + {INS SG,…}
		/tʂ̥a**ɕ**/	_____ + {LOC SG,…}
LOC SG	=	/**ɕ**ɛ/	{TIME,…} + _____
		/jɛ/	{PEASANT,…} + _____
INS SG	=	/ɛm/	{TIME,PEASANT,…} + _____

Unlike with labial-final stems, simple concatenation of the two morphemes does not obtain the correct output for the coronals, yielding ✗[tʂ̥aɕɕɛ] with a geminate [ɕɕ] instead of the correct output [tʂ̥aɕɛ] with a singleton [ɕ]. What is needed is for both instances of /ɕ/ in the input to merge in the output, coalescing to a singleton [ɕ] which is part of the realizations of both the stem and the locative singular. However, geminates are generally allowed in Polish, as in [ɔddatɕ] 'give back', [lɛkkɔ] 'lightly', and the examples cited in the preceding paragraph.

In order to distinguish gemination from coalescence of identical underlying segments, I propose that they have slightly different representations. In the case of geminates, where two underlying sounds emerge as two sounds (or a single sound with two timing slots), I assign full segmental status to the two underlying sounds. So for an output like [ɔddaw], the URs for the two morphemes involved are /ɔd-/ 'give' and /datɕ/ 'give', where each underlying /d/ is a full segment with a root node, timing slot, etc. For coalescence in the locative singular, however, the relevant underlying sounds are not both full segments. Following a core idea in the predominant generative analysis of the palatal mutations in Polish (as exemplified by Gussmann 1980, Rubach 1984a, and Czaykowska-Higgins 1988, and adopted in OT by Nowak 2001), I assume that the initial sound of the relevant underlying allomorph of the locative singular is actually not a full segment. Rather, it is a floating segment with no timing slot, which I represent here in a superscript grey font to iconically suggest the 'floating' nature of this abstract sound.[6] Thus, one of the allomorphs listed in the UR for the locative singular is /$^{\text{ɕ}}$ε/:

(16)

TIME	=	/tʂ̻as/	_____ + {INS SG,…}
		/tʂ̻aɕ/	_____ + {LOC SG,…}
LOC SG	=	/-$^{\text{ɕ}}$ε/	{TIME,…} + _____
		/-jε/	{PEASANT,…} + _____
INS SG	=	/-εm/	{TIME,PEASANT,…} + _____

In order for this floating /$^{\text{ɕ}}$/ to surface as a full segment, it must be anchored to some timing slot in the output. Rather than creating a new one from scratch, the grammar prefers for the floating segment to dock onto a pre-existing timing slot; in this case, it docks onto the final consonant of the stem. The palatal mutations of the locative singular forms of the other coronal-final stems will also similarly be stored as allomorphs in the lexicon, with /-$^{\text{ɲ}}$ε/ used for [pan] 'gentleman', /-$^{\text{tɕ}}$ε/ used for [brat] 'brother', etc. The following partial lexicon shows the full range of palatal mutations for coronal-final stems (except voiced obstruents, because their behavior can be analogized from the voiceless obstruents, assuming a highly ranked $\mathcal{F}$-voicing constraint), the coronal and labial allomorphs of the locative singular discussed so far, and the sole instrumental singular UR /-εm/ for comparison to the locative singular:

[6] A similar alternative which I do not pursue here, is that this floating segment is actually just a specification for the formant transitions into the vowel. That is, /-$^{\text{ɕ}}$ε/ could really mean an /ε/ with alveolo-palatal format transitions (on the left). This would avoid the need for coalescence and would allow the morphemes to simply be concatenated together.

(17)

TIME	=	/tʂas/	_____ + {INS SG,…}
		/tʂaɕ/	_____ + {LOC SG,…}
GENTLEMAN	=	/pan/	_____ + {INS SG,…}
		/paɲ̟/	_____ + {LOC SG,…}
BROTHER	=	/brat/	_____ + {INS SG,…}
		/bratɕ/	_____ + {LOC SG,…}
WALL	=	/mur/	_____ + {INS SG,…}
		/muʐ/	_____ + {LOC SG,…}
LOC SG	=	/-ɕɛ/	{TIME,…} + _____
		/-ɲ̟ɛ/	{GENTLEMAN,…} + _____
		/-tɕɛ/	{BROTHER,…} + _____
		/-ʐɛ/	{WALL,…} + _____
		/-jɛ/	{PEASANT,…} + _____
INS SG	=	/-ɛm/	{TIME,GENTLEMAN,…} + _____

In the remainder of this section, I show how various $\mathcal{F}$-constraints are responsible for the proper selection of stem and affix allomorph combinations in order to generate productive palatal mutations, based on an optimal merger of similar, but distinct, underlying sounds: the final consonants of the stems and the initial floating consonants of the matching allomorph of the locative singular.

4.4.3 Palatal mutation in the locative singular

As with the analysis of the labial-final stems, the palatal mutation of the extant coronal-final stems comes from selection of the appropriate allomorphs of both the stem and the locative singular as determined by morphological marking in the lexicon. From here on out, I ignore extant stems since their analysis does not change. Since palatal mutation in the locative singular is productive with nonce stems, the palatal mutation of the nonce stems must come from selection of the appropriate allomorph of the locative singular only, since nonce stems by design cannot be guaranteed to have lexical allomorphs marked for use with particular morphemes.

To simplify the tableaux, I analyze the coronal-final stems one pair at a time, since that generates only four candidates (a tableau with all of the stems would be quite large!). I begin with the nonce stems /dras/ and /drat/ (nonce stems are given in italics for consistency with previous tableaux). Subscripts separated by commas indicate coalescence of two underlying segments. Recall from §4.1.4 that $\mathcal{F}$-stop is undominated because the palatal mutations never change the property stop; the ranking of $\mathcal{F}$-palatal over $\mathcal{F}$-frication is from (12):

(18)

dras$_1$ *drat*$_2$ -εm$_{IS}$ -ɕε$_{LS}$ -tɕε$_{LS}$	$\mathcal{F}$ stop	$\mathcal{F}$ uniq	$\mathcal{F}$ pal	$\mathcal{F}$ fric
✓ a. *dras*$_1$-εm$_{IS}$ *dra*$_1$-*ɕ*$_{1,LS}$-ε$_{LS}$ *drat*$_2$-εm$_{IS}$ *dra*$_2$-*tɕ*$_{2,LS}$-ε$_{LS}$			×²	×
b. *dras*$_1$-εm$_{IS}$ *dra*$_1$-*tɕ*$_{1,LS}$-ε$_{LS}$ *drat*$_2$-εm$_{IS}$ *dra*$_2$-*ɕ*$_{2,LS}$-ε$_{LS}$	×²!		×²	×
c. *dras*$_1$-εm$_{IS}$ *dra*$_1$-*ɕ*$_{1,LS}$-ε$_{LS}$ *drat*$_2$-εm$_{IS}$ *dra*$_2$-*ɕ*$_{2,LS}$-ε$_{LS}$	×!	×²	×²	×
d. *dras*$_1$-εm$_{IS}$ *dra*$_1$-*tɕ*$_{1,LS}$-ε$_{LS}$ *drat*$_2$-εm$_{IS}$ *dra*$_2$-*tɕ*$_{2,LS}$-ε$_{LS}$	×!	×²	×²	×

In candidate (18a), the final /s/ of stem 1 /dras/ remains a fricative in the output, surfacing as faithfully as possible given the possible locative singular allomorphs. The final /t/ of stem 2 /drat/ surfaces as [tɕ], the result of a merger with the /-tɕɛ/ allomorph of the locative singular. This requires a change from a stop to an affricate (which I treat as having both the frication and stop properties). Thus candidate (18a) incurs a violation of $\mathcal{F}$-frication for stem 2 because the underlying stop /t/, with no frication, surfaces as an affricate [tɕ], with frication. At the same time, $\mathcal{F}$-stop is satisfied, since the surface affricate preserves the stop property of the input stop. In comparison, candidate (18b) changes the fricative /s/ in stem 1 /dras/ to an affricate, triggering a violation of $\mathcal{F}$-stop. The stop /t/ in stem 2 /drat/ surfaces as the fricative [ɕ], incurring a second violation of $\mathcal{F}$-stop as well as a violation of $\mathcal{F}$-frication. Candidate (18c) violates $\mathcal{F}$-uniqueness by mapping both stems to the same locative singular form, with the change from final /t/ in stem 2 to [ɕ] incurring violations of $\mathcal{F}$-stop and $\mathcal{F}$-frication as well. Candidate (18d) has the same violations as (18c): one for $\mathcal{F}$-stop for the change from fricative /s/ to the affricate [tɕ] in stem 1, two for $\mathcal{F}$-uniqueness for mapping both stems to the same phonetic string in the locative singular, and one for $\mathcal{F}$-frication for the change from the stop /t/ to the affricate [tɕ] in stem 2. Because candidates (18b–d) all violate undominated $\mathcal{F}$-stop, candidate (18a) is selected as the output by the grammar. This analysis can be analogized to the nonce pair /draz/ and /drad/ since this pair differs from the analyzed pair only in voicing of the final consonant.

The next pair of coronal-final nonce stems I consider is the pair /dran/ and /drar/. Recall from §4.1.4 that $\mathcal{F}$-nasal is undominated:

(19)

dran$_1$ *drar*$_2$ -εm$_{\text{IS}}$ -ɲ̟ε$_{\text{LS}}$ -ʐε$_{\text{LS}}$	$\mathcal{F}$ nas	$\mathcal{F}$ son	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric
✓ a. *dran*$_1$-εm$_{\text{IS}}$ *dra*$_1$-*ɲ̟*$_{1,\text{LS}}$-ε$_{\text{LS}}$ *drar*$_2$-εm$_{\text{IS}}$ *dra*$_2$-*ʐ*$_{2,\text{LS}}$-ε$_{\text{LS}}$		×			×	×
b. *dran*$_1$-εm$_{\text{IS}}$ *dra*$_1$-*ʐ*$_{1,\text{LS}}$-ε$_{\text{LS}}$ *drar*$_2$-εm$_{\text{IS}}$ *dra*$_2$-*ɲ̟*$_{2,\text{LS}}$-ε$_{\text{LS}}$	×^{2}!	×		×^{2}	×	×
c. *dran*$_1$-εm$_{\text{IS}}$ *dra*$_1$-*ɲ̟*$_{1,\text{LS}}$-ε$_{\text{LS}}$ *drar*$_2$-εm$_{\text{IS}}$ *dra*$_2$-*ɲ̟*$_{2,\text{LS}}$-ε$_{\text{LS}}$	×!		×^{2}	×	×^{2}	
d. *dran*$_1$-εm$_{\text{IS}}$ *dra*$_1$-*ʐ*$_{1,\text{LS}}$-ε$_{\text{LS}}$ *drar*$_2$-εm$_{\text{IS}}$ *dra*$_2$-*ʐ*$_{2,\text{LS}}$-ε$_{\text{LS}}$	×!	×^{2}	×^{2}	×		×^{2}

Candidate (19a) violates $\mathcal{F}$-sonorant and $\mathcal{F}$-frication for the change of the unfricated sonorant /r/ to the fricated obstruent [ʐ] for stem 2 /drar/, and violates of $\mathcal{F}$-palatal for the change of dental /n/ to alveolo-palatal [ɲ̟] for stem 1 /dran/. Candidate (19b) has violations of the same constraints, plus it violates $\mathcal{F}$-nasal and $\mathcal{F}$-retraction twice each by changing the non-retracted nasal /n/ to retracted oral [ʐ] in stem 1 and by changing retracted oral /r/ to the non-retracted nasal [ɲ̟] in stem 2. Candidates (19c) and (19d) (which violate $\mathcal{F}$-uniqueness because of their identical locative singulars for stems 1 and 2) have only one violation each of $\mathcal{F}$-nasal and $\mathcal{F}$-retraction. Because $\mathcal{F}$-nasal is undominated, candidate (19a) defeats candidates (19b–d). This reasoning means there is no need to do further comparison between coronal-final stems when one ends in /n/ since there is only one locative singular allomorph that can merge with /n/ to satisfy $\mathcal{F}$-nasal.

Next, I give an analysis of the nonce stems /drar/ and /draz/ because their palatal mutations involve very similar segments, [ʐ] and [ʑ]:

(20)

drar$_1$ *draz*$_2$ / -εm$_{IS}$ -ʐε$_{LS}$ -ʑε$_{LS}$	$\mathcal{F}$ son	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric
✓ a. *drar*$_1$-εm$_{IS}$ *dra*$_1$-ʐ$_{1,LS}$-ε$_{LS}$ / *draz*$_2$-εm$_{IS}$ *dra*$_2$-ʑ$_{2,LS}$-ε$_{LS}$	×			×	×
b. *drar*$_1$-εm$_{IS}$ *dra*$_1$-ʑ$_{1,LS}$-ε$_{LS}$ / *draz*$_2$-εm$_{IS}$ *dra*$_2$-ʐ$_{2,LS}$-ε$_{LS}$	×		×^{2}!	×	×
c. *drar*$_1$-εm$_{IS}$ *dra*$_1$-ʐ$_{1,LS}$-ε$_{LS}$ / *draz*$_2$-εm$_{IS}$ *dra*$_2$-ʐ$_{2,LS}$-ε$_{LS}$	×	×^{2}!	×!		×
d. *drar*$_1$-εm$_{IS}$ *dra*$_1$-ʑ$_{1,LS}$-ε$_{LS}$ / *draz*$_2$-εm$_{IS}$ *dra*$_2$-ʑ$_{2,LS}$-ε$_{LS}$	×	×^{2}!	×!	×^{2}	×

In candidate (20a), the change from the trill /r/ to the fricative [ʐ] in stem 1 incurs violations of $\mathcal{F}$-sonorant and $\mathcal{F}$-frication, and the change of dental /z/ to alveolo-palatal [ʑ] in stem 2 incurs a violation of $\mathcal{F}$-palatal. Candidate (20b) has violations of the same three constraints for the change from the alveolar sonorant /r/ to the alveolo-palatal fricative [ʑ] in stem 1. The tie-breaking violations come from $\mathcal{F}$-retracted. Trills, like [r], and retroflex fricatives, like the Polish post-alveolar [ʐ], involve a retracted articulation of the tongue dorsum toward the velum or pharynx (on [r], see Delattre 1971 and Rochoń 2001; on [ʐ], see Bhat 1974, Catford 1977, Hamilton 1980, Spencer 1984, Hamann 2002a,b). Because [r] and [ʐ] are both retracted, while [z] and [ʑ] are not, candidate (20b) incurs two violations of $\mathcal{F}$-retracted for the change from /r/ to [ʑ] in stem 1 and from /z/ to [ʐ] in stem 2. In comparison, candidate (20a) does not involve any changes in retraction (retracted /r/ changes to retracted [ʐ] while non-retracted /z/ changes to non-retracted [ʑ]), so it satisfies ℱ-retraction.

In candidate (20c), the two stems have the same palatal mutation [ʐ] in the locative singular. This incurs a violation of ℱ-retraction due to the change from /z/ to [ʐ], as well as two violations of ℱ-uniqueness for the two homophonous locative singular forms that derive from distinct stems. Since candidate (20c) satisfies ℱ-palatal while the desired output (20a) violates it, at least one of ℱ-retraction and ℱ-uniqueness must be ranked higher than ℱ-palatal in order for (20a) to be selected at the winner. Candidate (20d) also violates ℱ-uniqueness twice for mapping both stems to the same locative singular form with the palatal mutation [ʑ], but it also violates enough other constraints that it is harmonically bounded by the winner (20a).

The final coronal-final pairs to consider are not given a full analysis here since their analysis follows from previous discussion. For example, the tableau in (20) can be extended quite transparently to the nonce pair /drar/ and /drad/. If /r/ will not mutate to [ʑ] because of a fatal violation of ℱ-retraction (20b), then it certainly will not mutate to [dʑ], because such a change involves all of the same violations, plus an additional violation of undominated ℱ-stop for changing a trill to a stop.

In summary, here is a chart listing the possible coronal pairs. Pairs that are compared in a tableau are indicated by the number of the tableau; pairs which are not compared have one or more letters indicating an explanation for why no tableau comparison is necessary (the legend is listed to the right of the table):

(21)

	s/ɕ	z/ʑ	n/ɲ	t/tɕ	d/dʑ
r/ʐ	*v*	(20)	(19)	*vs*	*s*
d/dʑ	*vs*	a_{18}	*ñ*	*v*	
t/tɕ	(18)	*vs*	*ñ*		
n/ɲ	*vsñ*	*vñ*			
z/ʑ	*v*				

v = $\mathcal{F}$-voi is undominated
s = $\mathcal{F}$-stop is undominated
ñ = $\mathcal{F}$-nas is undominated
a_{18} = analogy to (18)

4.4.4 Asymmetry in coalescence

In the analysis of the coronals, the locative singular morpheme does not change from the input to the output (except in gaining a timing slot), but it triggers a change in the final consonant of nonce stems. This seems to require a split in the $\mathcal{F}$-constraints, with some constraints governing faithfulness within affixes ($\mathcal{F}_{affix}$), while others govern faithfulness within stems ($\mathcal{F}_{stem}$). This is not a new insight (see for example, Beckman 1995, 1997, 1998, McCarthy and Prince 1995, Selkirk 1995, Urbanczyk 1996, Alderete 1997b). However, since the locative singular suffix does not change, but the stems do, Polish seems to require a ranking of $\mathcal{F}_{affix} \gg \mathcal{F}_{stem}$, which runs counter to the usual assumption that $\mathcal{F}_{stem}$ is ranked higher.

The more accepted ranking $\mathcal{F}_{stem} \gg \mathcal{F}_{affix}$ can be maintained if the split behavior is not attributed to a stem/affix distinction, but rather to the relative age of the morphemes. Note that extant stems do not technically change in the locative singular; they have underlying allomorphs which already have the palatal mutations encoded in them, allowing them to surface faithfully. Nonce stems are different, since they do not have the benefit of underlying allomorphs and must change in the locative singular. Thus, the split could be between extant versus nonce, with $\mathcal{F}_{extant} \gg \mathcal{F}_{nonce}$. This type of analysis would get the right results, with the asymmetry in coalescence favoring faithfulness to the extant locative singular suffix, with nonce stems changing to suit the coalescence. However, this potentially contradicts Ito and Mester's (1995, 1999) core-periphery analysis of lexical strata, in which newer words are more faithful than older words because older words are generally more restricted by phonological generalizations. Other possible solutions also exist, but a full analysis of this asymmetry in coalescence is beyond the scope of this dissertation.

4.5 FDM-OT analysis of the palatal mutation of velars

The palatal mutations of the velars were discussed in §4.1.3, and the relevant data are repeated below:

(22)	*stem UR*	*GEN SG*	*LOC SG*	*gloss*
	/z̻ɛ**k**/	z̻ɛ**k**a	z̻ɛ**ts**ɛ	'river'
	/drɔ**g**/	drɔ**g**a	drɔ**dz**ɛ	'road'
	/mu**x**/	mu**x**a	mu**ʂ̻**ɛ	'fly'
	/ʂ̻kɔ**w**/	ʂ̻kɔ**w**a	ʂ̻kɔ**l**ɛ	'school'

Recall that in general, [ɛ] does not trigger the palatal mutation, so [kɛ], [gɛ], etc. are allowed as surface sequences, as in [kɛks] 'tea cake', [gɛst] 'motion', [xɛm^{j}ja] and 'chemistry', and [wɛp] 'head'.

The convenient comparison between the locative singular and the instrumental singular that was useful for analyzing the palatal mutations of labials and coronals breaks down for stems ending in velar obstruents (it still holds for stems ending in the velar glide /w/). The instrumental singular marker for feminine stems is [-ɔ̃w̃], not [-ɛm], and the locative singular form of masculine and neuter stems ending in velar obstruents is [-u], not [-ɛ].[7] Thus, there are in fact no stems ending in velar

[7] There is a small class of nouns that have a mixed declension, with feminine endings in the singular and masculine endings in the plural. Since I am only concerned here with the singular, these stems can be treated as feminine.

obstruents which take [-ɛ] in the locative singular and [-ɛm] in the instrumental singular. This is not a serious problem, however, since the only purpose of looking at the instrumental singular forms is to show that palatal mutation is not triggered by every suffix beginning with [ɛ]. Thus, in the remaining tableau, I ignore the instrumental singular forms, focusing solely on the palatal mutation in the locative singular, with the understanding that the instrumental singular does not trigger palatal mutation.

4.5.1 The general grammar

Since the velars and their palatal mutations are all allowed before [ɛ] in Polish, the grammar must create sequences such as [kɛ], [gɛ], etc., as well as [tsɛ], [dzɛ], etc. The palatal mutations of the labials and coronals are very similar to their unmutated counterparts and significantly compete with each other perceptually: [p] is similar to and thus competes with [p^jj], [t] competes with [tɕ], etc. In contrast, the velars and their palatal mutations are very different from each other and are not in significant perceptual competition: [k] does not compete with [ts], [w] does not compete with [l], etc. Because of this lack of competition, I do not provide a constraint ranking to derive the sequences [kɛ], [gɛ], etc. and [tsɛ], [dzɛ], etc.. Instead, I simply stipulate that they are allowed to exist in Polish.

4.5.2 The lexicon

In the following partial lexicon, I give two allomorphs for the velar-final stems (plain and palatal mutation; note that even though the instrumental singular marker for these stems is not [-ɛm], the allomorph used is still the one without palatal mutation), as well as for the coronal- and labial-final stems and the allomorphs of the locative singular discussed so far. I leave out the instrumental singular, since it does not play a role in the remaining analysis:

(23)

RIVER	=	/ʐɛk/	______ + {INS SG,…}
		/ʐɛts/	______ + {LOC SG,…}
FLY	=	/mux/	______ + {INS SG,…}
		/muʂ/	______ + {LOC SG,…}
SCHOOL	=	/ʂkɔw/	______ + {INS SG,…}
		/ʂkɔl/	______ + {LOC SG,…}
TIME	=	/tʂas/	______ + {INS SG,…}
		/tʂaɕ/	______ + {LOC SG,…}
GENTLEMAN	=	/pan/	______ + {INS SG,…}
		/paɲ/	______ + {LOC SG,…}
BROTHER	=	/brat/	______ + {INS SG,…}
		/bratɕ/	______ + {LOC SG,…}
WALL	=	/mur/	______ + {INS SG,…}
		/muʐ/	______ + {LOC SG,…}
PEASANT	=	/xwɔp/	______ + {INS SG,…}
		/xwɔp^j/	______ + {LOC SG,…}
LOC SG	=	/-tsɛ/	{RIVER,…} + ______
		/-ʂɛ/	{FLY,…} + ______
		/-lɛ/	{SCHOOL,…} + ______
		/-ɕɛ/	{TIME,…} + ______
		/-ɲɛ/	{GENTLEMAN,…} + ______
		/-tɕɛ/	{BROTHER,…} + ______
		/-ʐɛ/	{WALL,…} + ______
		/-jɛ/	{PEASANT,…} + ______

As with the labials and coronals, the choice of which locative singular allomorph for nonce stems ending in a velar is determined by various ℱ-constraints within the constraint hierarchy.

4.5.3 Palatal mutation in the locative singular

The first set of velar-final stems to consider is the nonce pair /drak/ and /drax/. Recall that at least one of ℱ-uniqueness and ℱ-retraction must outrank ℱ-palatal (20), which itself is required to outrank ℱ-frication (12):

(24)

$drak_1$ $drax_2$ -tsε_{LS} -ʂε_{LS}	ℱ stop	ℱ uniq	ℱ retr	ℱ pal	ℱ fric	ℱ cor
✓ a. $dra_1\text{-}ts_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}ʂ_{2,LS}\text{-}\varepsilon_{LS}$			×		×	×^{2}
b. $dra_1\text{-}ʂ_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}ts_{2,LS}\text{-}\varepsilon_{LS}$	×^{2}!		×		×	×^{2}
c. $dra_1\text{-}ʂ_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}ʂ_{2,LS}\text{-}\varepsilon_{LS}$	×!	×^{2}			×	×^{2}
d. $dra_1\text{-}ts_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}ts_{2,LS}\text{-}\varepsilon_{LS}$	×!	×^{2}	×^{2}		×	×^{2}

Both of the palatal mutations in candidate (24a) involve a change from a velar to a coronal, which violates ℱ-coronal, and the change from the velar stop /k/ to the non-retracted coronal affricate [ts] in stem 1 involves violations of ℱ-retraction and ℱ-frication. In candidate (24b), the same constraints are violated for similar reasons: both palatal mutations are from velar to coronal (two violations of ℱ-coronal), the change from the stop /k/ to the fricative [ʂ] involves a change in $\mathcal{F}$-frication, and the change from velar /x/ to dental [ts] in stem 2 violates $\mathcal{F}$-retraction. In addition, both of the palatal mutations in (24b) incur a violation of $\mathcal{F}$-stop, since the stop /k/ changes to a fricative [ʂ] in stem 1 and the fricative /x/ changes to an affricate [ts] in stem 1. Thus, candidate (24a) harmonically bounds candidate (24b). In candidate (24c), both stems have the same palatal mutation [ʂ], yielding two violations of $\mathcal{F}$-uniqueness. In addition, by mapping the stop /k/ to the fricative [ʂ] in stem 1, candidate (24c) incurs a violation of $\mathcal{F}$-stop and $\mathcal{F}$-frication. Candidate (24d) incurs two violations of $\mathcal{F}$-uniqueness by mapping both input stems to homophonous locative singular forms, with the palatal mutation [ts]. The change from the stop /k/ to the affricate [ts] in stem 1 violates $\mathcal{F}$-fricative, the change from the fricative /x/ to the affricate [ts] in stem 2 violates $\mathcal{F}$-frication, and the change from the two velars to the non-retracted coronal [ts] violates $\mathcal{F}$-retracted twice. Because $\mathcal{F}$-stop is undominated, candidate (24a) is selected as the output over candidates (24b–d).

The next nonce stems to be considered are /drag/ and /draw/:

(25)

$drag_1$ $draw_2$ $\text{-}\mathbf{dz}\varepsilon_{LS}$ $\text{-}\mathbf{l}\varepsilon_{LS}$	$\mathcal{F}$ stop	$\mathcal{F}$ stop	$\mathcal{F}$ son	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric	$\mathcal{F}$ cor
✓ a. $dra_1\text{-}dz_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}l_{2,LS}\text{-}\varepsilon_{LS}$					×^{2}		×	×^{2}
b. $dra_1\text{-}l_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}dz_{2,LS}\text{-}\varepsilon_{LS}$	×^{2}!	×^{2}!	×^{2}		×^{2}		×	×^{2}
c. $dra_1\text{-}l_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}l_{2,LS}\text{-}\varepsilon_{LS}$	×!	×!	×	×^{2}	×^{2}			×^{2}
d. $dra_1\text{-}dz_{1,LS}\text{-}\varepsilon_{LS}$ $dra_2\text{-}dz_{2,LS}\text{-}\varepsilon_{LS}$	×!	×!	×	×^{2}	×^{2}		×^{2}	×^{2}

In candidate (25a), there is a change from a stop /g/ to an affricate [dz] in stem 1 (a violation of $\mathcal{F}$-frication) and two changes from velars /g w/ to non-retracted coronals [dz l] (two violations each of $\mathcal{F}$-retracted and $\mathcal{F}$-coronal). Candidate (25b) violates the same constraints, for changing the glide /w/ to the affricate [dz] in stem 2 ($\mathcal{F}$-frication) and for the two velars /g w/ changing to two non-retracted coronals [l dz] ($\mathcal{F}$-retracted and $\mathcal{F}$-coronal). In addition, candidate (25b) has more constraint violations: two violations of each of $\mathcal{F}$-stop, $\mathcal{F}$-narrow, and $\mathcal{F}$-sonorant (for changing the narrow stop /g/ into the wide sonorant [l] in stem 1 and the wide sonorant /w/ into the narrow affricate [dz] in stem 2). Candidate (25c) changes both of the velars to [l] in the locative singular, incurring two violations of $\mathcal{F}$-uniqueness. In addition, the change from the narrow stop /g/ to the wide sonorant [l] in stem 1 also violates $\mathcal{F}$-stop, $\mathcal{F}$-narrow, and $\mathcal{F}$-sonorant. In candidate (25d), the velars both mutate to [dz], violating $\mathcal{F}$-uniqueness twice, and violating $\mathcal{F}$-stop, $\mathcal{F}$-narrow, and $\mathcal{F}$-sonorant once each (for changing the wide sonorant /w/ to the narrow affricate [dz]), as well as other lower-ranked constraints. Candidate (25a) satisfies both of the undominated constraints 𝓕-stop and 𝓕-narrow, while the other three candidates (25b–d) violate at least one each, so candidate (25a) and is selected as the winner.

In fact, all of the velar comparisons can be decided by undominated constraints. The following table summarizes the possible velar pairs, with numbers indicating tableaux showing actual comparisons made in this section, and letters indicating no comparison is needed due to some undominated constraint:

(26)

	k/ts	g/dz	x/ʂ
w/l	*vsn*	(25)	*vn*
x/ʂ	(24)	*vs*	
g/dz	*v*		

v = 𝓕-voicing is undominated
s = 𝓕-stop is undominated
n = 𝓕-narrow is undominated

4.6 Comparison across the three groups

The analysis is incomplete if the only comparisons made are within groups based on places of articulation. It is necessary to compare the palatal mutations *across* these groups as well (labial versus coronal, labial versus velar, and coronal versus velar).

4.6.1 Comparison of labials and coronals

Recall that the locative singular allomorph used in the analysis of labial-final stems is /-jɛ/, whereas for coronal-final stems, the available allomorphs are /-tɕɛ/, /-dʑɛ/, /-ɕɛ/, /-ʑɛ/, /-ɲɛ/, and /-ʐɛ/. In comparing the locative singular form of labial- and coronal-final stems, all of these allomorphs must be considered. An important question immediately arises: why would a stem like /brat/ ‘brother’ select the locative singular allomorph /-tɕɛ/, which forces a change in place and manner of articulation, when /jɛ/ is available (the sequence [tj] is allowable in Polish, as in [tjara] ‘tiara’)?

The clearest solution to this problem in the context of the previous analysis is that for each of the labials [p b f v m], there is another allomorph of the locative singular with a floating segment, just as for the coronals and velars, rather than the simple /-jɛ/ used previously. Thus, a partial lexical entry for the locative singular (consisting of allomorphs used for labial- and coronal-final stems) looks like this:

(27)

		/-pʲjɛ/	{PEASANT,…} + ______
		/-bʲjɛ/	{BREAD,…} + ______
		/-fʲjɛ/	{BOSS,…} + ______
		/-vʲjɛ/	{KRAKÓW,…} + ______
		/-mʲjɛ/	{CROWD,…} + ______
LOC SG	=	/-ɕɛ/	{TIME,…} + ______
		/-ʑɛ/	{CART,…} + ______
		/-ɲ̟ɛ/	{GENTLEMAN,…} + ______
		/-tɕɛ/	{BROTHER,…} + ______
		/-dʑɛ/	{NEIGHBOR,…} + ______
		/-ʐɛ/	{WALL,…} + ______

With these allomorphs, the correct palatal mutations fall out quite easily. I do not show this for every labial-coronal pair of stems since the basic analysis is the same; a single example suffices to show how the palatal mutations of the labials must be different from that of the coronals. In the following tableau, I give the derivation of the locative singular forms of the nonce stems /drap/ and /drat/:

(28) $drap_1$ $drat_2$ -p^jjε_{LS} -tɕε_{LS}	$\mathcal{F}$ stop	$\mathcal{F}$ son	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric	$\mathcal{F}$ cor	$\mathcal{F}$ lab
✓ a. dra_1-$p^j_{1,LS}$-jε_{LS} dra_2-$tɕ_{2,LS}$-ε_{LS}					×^{2}	×	×	
b. dra_1-$tɕ_{1,LS}$-ε_{LS} dra_2-$p^j_{2,LS}$-jε_{LS}					×^{2}	×	×	×^{2}!
c. dra_1-$p^j_{1,LS}$-jε_{LS} dra_2-$p^j_{2,LS}$-jε_{LS}			×^{2}!		×^{2}		×	×
d. dra_1-$tɕ_{1,LS}$-ε_{LS} dra_2-$tɕ_{2,LS}$-ε_{LS}			×^{2}!		×^{2}	×^{2}	×	×

Candidate (28a) harmonically bounds candidate (28b) because candidate (28b) changes labial /p/ to coronal [tɕ] and coronal /t/ to labial [p^j], incurring two violations of $\mathcal{F}$-labial. In comparison, candidate (28a) satisfies $\mathcal{F}$-labial (note that secondary palatalization in [p^j] counts as coronal, so changing /p/ to either [p^j] or [tɕ] violates $\mathcal{F}$-coronal). All of the other violations between these two candidates are the same. Candidates (28c–d) violate highly ranked $\mathcal{F}$-uniqueness by mapping both input stems to the same locative singular form. This leaves candidate (28a) as the winner. This same analysis can be generalized to any labial-coronal pair, since any swap of their palatal mutations would require extraneous violations of $\mathcal{F}$-labial. Nothing is to be gained by changing labiality since allomorphs of the locative singular are available for both labials and coronals which allow them to surface faithfully with respect to labiality.

The following table summarizes the possible comparisons between labial-final and coronal-final stems, with a number indicating the tableau in which an explicit comparison is made, and with letters indicating the reason why no tableau is given:

(29)

	p/pʲ	b/bʲ	f/fʲ	v/vʲ	m/mʲ
t/tɕ	(28)	*v*	*s*	*vs*	*vñ*
d/dʑ	*v*	a_{28}	*vs*	*s*	*ñ*
s/ɕ	*s*	*vs*	a_{28}	*v*	*vsñ*
z/ʑ	*vs*	*s*	*v*	a_{28}	*sñ*
n/ɲ	*vñ*	*ñ*	*vsñ*	*sñ*	a_{28}
r/ʐ	*vs*	*s*	*v*	a_{28}	*sñ*

v = $\mathcal{F}$-voi is undominated
s = $\mathcal{F}$-stop is undominated
n = $\mathcal{F}$-narr is undominated
ñ = $\mathcal{F}$-nas is undominated
a_{28} = analogy to (28)

4.6.2 Comparison of labials and velars

For a comparison of the palatal mutations of labials and velars, the allomorphs of the locative singular that must be considered are as follows:

(30)

LOC SG =

/-pʲjɛ/	{PEASANT,…} + ______
/-bʲjɛ/	{BREAD,…} + ______
/-fʲjɛ/	{BOSS,…} + ______
/-vʲjɛ/	{KRAKÓW,…} + ______
/-mʲjɛ/	{CROWD,…} + ______
/-tsɛ/	{RIVER,…} + ______
/-dzɛ/	{ROAD,…} + ______
/-ʂɛ/	{FLY,…} + ______
/-lɛ/	{SCHOOL,…} + ______

The first comparison I make is with the nonce stems /drap/ and /drak/ (this analysis can also be analogized to their voiced counterparts /drab/ and /drag/):

(31) *drap*$_1$ *drak*$_2$ -p^jjε_{LS} -tsε_{LS}	$\mathcal{F}$ stop	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric	$\mathcal{F}$ cor	$\mathcal{F}$ lab	$\mathcal{F}$ dor
✓ a. dra_1-$p^j_{1,LS}$-jε_{LS} dra_2-$ts_{2,LS}$-ε_{LS}			×	×	×	×²		×²
b. dra_1-$ts_{1,LS}$-ε_{LS} dra_2-$p^j_{2,LS}$-jε_{LS}			×	×	×	×²	×²!	
c. dra_1-$p^j_{1,LS}$-jε_{LS} dra_2-$p^j_{2,LS}$-jε_{LS}		×²!	×	×²		×²	×	×
d. dra_1-$ts_{1,LS}$-ε_{LS} dra_2-$ts_{2,LS}$-ε_{LS}		×²!	×		×²	×²	×²	×

Candidate (31a) changes the dorsal specification in the palatal mutation for both /p/ and /k/: labial /p/ becomes palatalized [p^j] (which involves a dorsal articulation), and the velar /k/ becomes the coronal [ts]. In comparison, losing candidate (31b) does not violate $\mathcal{F}$-dorsal at all (labial /p/ maps to dental [ts], and velar /k/ maps to palatalized [p^j]). All other constraint violations are the same for these two candidates, except for $\mathcal{F}$-labial, which candidate (31b) violates (for changing labial /p/ to coronal [ts] and velar /k/ to labial [p^j]) and candidate (31a) does not. Thus, $\mathcal{F}$-labial must outrank $\mathcal{F}$-dorsal in order for candidate (31a) to win over candidate (31b). The two candidates (31c–d) violate $\mathcal{F}$-uniqueness while (31a) does not. Recall from (20) that either $\mathcal{F}$-uniqueness or $\mathcal{F}$-retraction must outrank $\mathcal{F}$-palatal. Since candidate (31d) satisfies $\mathcal{F}$-palatal but candidate (31a) does not, $\mathcal{F}$-palatal must be ranked below $\mathcal{F}$-uniqueness in order for candidate (31a) to be selected as the output.

Next I compare the palatal mutations of the labial and velar fricatives using the nonce stems /draf/ and /drax/:

(32)

draf$_1$ *drax*$_2$ — -fʲjε$_{LS}$ -ʂε$_{LS}$	$\mathcal{F}$ stop	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric	$\mathcal{F}$ cor	$\mathcal{F}$ lab	$\mathcal{F}$ dor
✓ a. *dra*$_1$-*f*$^j_{1,LS}$-jε$_{LS}$ *dra*$_2$-*ʂ*$_{2,LS}$-ε$_{LS}$				✗		✗^{2}		✗
b. *dra*$_1$-*ʂ*$_{1,LS}$-ε$_{LS}$ *dra*$_2$-*f*$^j_{2,LS}$-jε$_{LS}$			✗^{2}!	✗		✗^{2}	✗^{2}	✗
c. *dra*$_1$-*ʂ*$_{1,LS}$-ε$_{LS}$ *dra*$_2$-*ʂ*$_{2,LS}$-ε$_{LS}$		✗^{2}!	✗!			✗^{2}	✗^{2}	✗
d. *dra*$_1$-*f*$^j_{1,LS}$-jε$_{LS}$ *dra*$_2$-*f*$^j_{2,LS}$-jε$_{LS}$		✗^{2}!	✗!	✗^{2}		✗^{2}	✗	✗

Candidate (32a) changes plain /f/ to palatalized [fʲ] in stem 1 (violating $\mathcal{F}$-palatal, $\mathcal{F}$-coronal, and $\mathcal{F}$-dorsal) and velar /x/ to coronal [ʂ] in stem 2 (a second violation of $\mathcal{F}$-coronal). Candidate (32b) violates the same constraints as candidate (32a), plus $\mathcal{F}$-retraction and $\mathcal{F}$-labial for the changes from the non-retracted labial /f/ to the retracted coronal [ʂ] and from velar /x/ to the palatalized labial [fʲ], Thus, candidate (32b) is harmonically bounded by candidate (32a). High ranking of $\mathcal{F}$-uniqueness prevents candidates (32c–d) from being selected as the output, since they violate this constraint by merging the two stems into the same locative singular form. This leaves candidate (32a) as the selected output.

The following table summarizes the comparisons of the palatal mutations of the possible labial-velar pairs, with the usual numbers and letters indicating an explicit analysis from a tableau and the reasons why an explicit comparison is not given:

(33)

	k/ts	g/dz	x/ʂ	w/l
p/pʲ	(31)	*v*	*s*	*vsn*
b/bʲ	*v*	a_{31}	*vs*	*sn*
f/fʲ	*s*	*v*	(32)	*vn*
v/vʲ	*vs*	*s*	*v*	*n*
m/mʲ	*vñ*	*ñ*	*vsñ*	*snñ*

v = 𝓕-voicing is undominated
s = 𝓕-stop is undominated
n = 𝓕-narrow is undominated
ñ = 𝓕-nasal is undominated
a_{31} = analogy to (31)

4.6.3 Comparison of coronals and velars

The final comparisons of palatal mutations that need to made are between coronal-final and velar-final stems. The first coronal-velar pair of nonce stems I analyze is the pair /drar/ and /draw/:

(34)

*drar*₁ *draw*₂ / -ʐε$_{LS}$ -lε$_{LS}$	𝓕 narr	𝓕 son	𝓕 uniq	𝓕 retr	𝓕 pal	𝓕 fric	𝓕 cor	𝓕 dor
✓ a. *dra*₁-*ʐ*$_{1,LS}$-ε$_{LS}$ *dra*₂-*l*$_{2,LS}$-ε$_{LS}$		×		×		×	×	×
b. *dra*₁-*l*$_{1,LS}$-ε$_{LS}$ *dra*₂-*ʐ*$_{2,LS}$-ε$_{LS}$	×²!	×		×		×	×	×
c. *dra*₁-*l*$_{1,LS}$-ε$_{LS}$ *dra*₂-*l*$_{2,LS}$-ε$_{LS}$	×!		×²	×²		×²	×	×²
d. *dra*₁-*ʐ*$_{1,LS}$-ε$_{LS}$ *dra*₂-*ʐ*$_{2,LS}$-ε$_{LS}$	×!	×²	×²	×²		×²	×	

In candidate (34a), the narrow trill /r/ is mapped to the narrow fricative [ʐ], while the wide approximant /w/ is mapped to the wide approximant [l], completely satisfying 𝓕-narrow. Its violations of lower constraints (𝓕-sonorant and 𝓕-frication for changing the sonorant /r/ to the fricative [ʐ]; 𝓕-retracted, 𝓕-coronal, and 𝓕-dorsal for changing velar /w/ to the non-retracted coronal [l]) do not matter, since 𝓕-narrow is undominated and the other candidates violate 𝓕-narrow: candidate (34b) changes the narrowness of both underlying stems, incurring two violations of 𝓕-narrow, while candidates (34c–d) each change one value for narrowness. Thus, candidate (34a) is selected as the output.

The next palatal mutations I analyze are for the coronal and velar fricatives, using the pair of nonce stems /dras/ and /drax/:

(35)

dras$_1$ *drax*$_2$ -ɕ$ε_{LS}$ -ʂ$ε_{LS}$	𝓕 stop	𝓕 uniq	𝓕 retr	𝓕 pal	𝓕 fric	𝓕 cor	𝓕 dor
✓ a. *dra*$_1$-*ɕ*$_{1,LS}$-$ε_{LS}$ *dra*$_2$-*ʂ*$_{2,LS}$-$ε_{LS}$				×		×	×
b. *dra*$_1$-*ʂ*$_{1,LS}$-$ε_{LS}$ *dra*$_2$-*ɕ*$_{2,LS}$-$ε_{LS}$			×^{2}!	×		×	×
c. *dra*$_1$-*ʂ*$_{1,LS}$-$ε_{LS}$ *dra*$_2$-*ʂ*$_{2,LS}$-$ε_{LS}$		×^{2}!	×!			×	×
d. *dra*$_1$-*ɕ*$_{1,LS}$-$ε_{LS}$ *dra*$_2$-*ɕ*$_{2,LS}$-$ε_{LS}$		×^{2}!	×!	×^{2}		×	×

Candidate (35a) incurs one violation each of 𝓕-palatal and 𝓕-dorsal for the change from coronal /s/ to alveolo-palatal [ɕ] in stem 1 and one violation of 𝓕-coronal for the change from velar /x/ to the retracted coronal [ʂ] in stem 2. Candidate (35b) has the same three violations, plus two violations of $\mathcal{F}$-retraction for the change from the non-retracted coronal /s/ to the retracted coronal [ʂ] in stem 1 and for the change from velar /x/ to the non-retracted coronal [ɕ] in stem 2. This means that candidate (35b) is harmonically bounded by candidate (35a) and cannot be selected as the output. The two candidates (35b–d) in which both stems have the same palatal mutation in the locative singular violate high ranking $\mathcal{F}$-uniqueness and are also prevented from winning. Thus, candidate (35a) is the winner.

Finally, I analyze the pair of nonce stems /drat/ and /drak/:

(36)

drat$_1$ *drak*$_2$ -tɕɛ$_{LS}$ -tsɛ$_{LS}$	$\mathcal{F}$ stop	$\mathcal{F}$ uniq	$\mathcal{F}$ retr	$\mathcal{F}$ pal	$\mathcal{F}$ fric	$\mathcal{F}$ cor	$\mathcal{F}$ dor
☹ a. *dra*$_1$-*tɕ*$_{1,LS}$-ɛ$_{LS}$ *dra*$_2$-*ts*$_{2,LS}$-ɛ$_{LS}$			×	×	×^{2}	×	×^{2}!
☠ b. *dra*$_1$-*ts*$_{1,LS}$-ɛ$_{LS}$ *dra*$_2$-*tɕ*$_{2,LS}$-ɛ$_{LS}$			×	×	×^{2}	×	
c. *dra*$_1$-*ts*$_{1,LS}$-ɛ$_{LS}$ *dra*$_2$-*ts*$_{2,LS}$-ɛ$_{LS}$		×^{2}!	×		×^{2}	×	×
d. *dra*$_1$-*tɕ*$_{1,LS}$-ɛ$_{LS}$ *dra*$_2$-*tɕ*$_{2,LS}$-ɛ$_{LS}$		×^{2}!	×	×^{2}	×^{2}	×	×

In this tableau, the actual output of the language, candidate (36a), is predicted to lose to candidate (36b) because of extra violations of $\mathcal{F}$-dorsal. In candidate (36a), the changes from dental /t/ to alveolo-palatal [tɕ] in stem 1 and from velar /k/ to dental [ts] in stem 2 require the specification for dorsal for both stems to change, whereas candidate (36b) maps each stem to a locative singular form that has a matching specification for the dorsal property (dental /t/ is mapped to dental [ts] in stem 1 and velar /k/ is mapped to alveolo-palatal [tɕ] in stem 2). This is a problem for the analysis developed so far, as there is no ranking of these constraints in which candidate (36a) can defeat candidate (36b) (candidates (36c–d) are not a factor, since they violate highly ranked $\mathcal{F}$-uniqueness).

What is needed to further distinguish these two candidates is some property that can group a dental with an alveolo-palatal, while at the same time grouping a velar with a dental. Obviously, this must be some property that is based on place of articulation; all other properties behave identically in both candidates, so they are not relevant. One of the primary phonetic correlates of place of articulation is transitions in the formants of neighboring vowels. The following table (based on discussion in Ladefoged 2001:179–193) summarizes the loci of the transitions of the second and third formants for consonants having velar, dental, and alveolo-palatal places of articulation. A 'low' locus indicates that the formant is lower near the consonant, a 'high' locus indicates that the formant is raised near the consonant, and a 'mid' locus is intermediate between the two, often resulting in formant transitions that are heavily dependent on the neighboring vowel (for example, a mid F2 locus has a lower transition for front vowels, because front vowels have high values for F2, but it has a raised transition for back vowels, which have a lower F2).

(37)

formant	velar	dental	alv-pal
F2	high	mid	high
F3	low	mid	high

Only F3 fits the needed requirement of grouping dentals with alveolo-palatals and velars with dentals, while keeping alveolo-palatals and velars ungrouped; F2 treats alveolo-palatals and velars as a natural class, which is contrary to the desired grouping. Thus, the crucial property is the value of the F3 locus, with velars having the lowest locus, dentals the next, and alveolo-palatals the highest. For simplicity, I assume a three-way division in F3 loci, so that a change from velar to dental or from dental to alveolo-palatal is one violation of $\mathcal{F}$-locus$_{F3}$, while a change from velar to alveolo-palatal is two violations:

(38)

	1 violation of $\mathcal{F}$-loc$_{F3}$		1 violation of $\mathcal{F}$-loc$_{F3}$	
k		t/ts		tɕ
low loc$_{F3}$		mid loc$_{F3}$		high loc$_{F3}$
		2 violations of $\mathcal{F}$-loc$_{F3}$		

However, adding $\mathcal{F}$-locus$_{F3}$ is not quite sufficient to get the correct output:

(39)

drat$_1$ *drak*$_2$ -tɕε_{LS} -tsε_{LS}	$\mathcal{F}$ loc$_{F3}$	$\mathcal{F}$ dor
☹ a. *dra*$_1$-*tɕ*$_{1,LS}$-ε_{LS} *dra*$_2$-*ts*$_{2,LS}$-ε_{LS}	✖^{2}	✖^{2}!
☠ b. *dra*$_1$-*ts*$_{1,LS}$-ε_{LS} *dra*$_2$-*tɕ*$_{2,LS}$-ε_{LS}	✖^{2}	

Both candidates incur two violations of $\mathcal{F}$-locus$_{F3}$, so this constraint has no effect on the grammar's selection. Candidate (39a)'s two violations of $\mathcal{F}$-dorsal still prevent it from defeating candidate (39b), so the correct output is not obtained. What is important to note is that while these two candidates have the same number of total violations of $\mathcal{F}$-locus$_{F3}$, their specific violations are not the same. Candidate (39a) has one violation for one one-step change in F3 locus from low to mid, plus another violation for a one-step change from mid to high. In comparison, candidate (39b) has two violations for one instance of a two-step change from low to high. This is a case of a *chain shift*, in which *A* is compelled to change to *B*, and *B* is compelled to change to *C*, but *A* does not skip directly to *C*. Within OT, chain shifts have received many analyses, most notably the use of constraint conjunction of faithfulness constraints (Kirchner 1996) and faithfulness to scalar adjacency (Gnanadesikan 1997). I adopt an analysis informed by both of these theories, with a constraint $\mathcal{F}^2$-locus$_{F3}$ that is violated by two-step changes in F3 locus. Compare this constraint to the conjunction $\mathcal{F}$-locus$_{F3}$&$\mathcal{F}$-locus$_{F3}$ (following basic ideas in Kirchner 1996) or to the constraint IDENT-ADJ [locus$_{F3}$] (in the spirit of Gnanadesikan 1997).[8]

With $\mathcal{F}^2$-locus$_{F3}$ outranking $\mathcal{F}$-dorsal, the correct output for the stems /drak/ and /drat/ can be obtained:

[8] See §5.5.1 of Chapter 5 for arguments against the need for general constraint conjunction. Those arguments do not apply here since $\mathcal{F}^2$-locus$_{F3}$ need not be viewed as conjunction.

(40)

$drat_1$ $drak_2$ -tɕ$ɛ_{LS}$ -ts$ɛ_{LS}$	$\mathcal{F}^2$ loc_{F3}	$\mathcal{F}$ loc_{F3}	$\mathcal{F}$ dor
✓ a. dra_1-$tɕ_{1,LS}$-$ɛ_{LS}$ dra_2-$ts_{2,LS}$-$ɛ_{LS}$		×^{2}	×^{2}
b. dra_1-$ts_{1,LS}$-$ɛ_{LS}$ dra_2-$tɕ_{2,LS}$-$ɛ_{LS}$	×!	×^{2}	

Candidate (40a) does not violate $\mathcal{F}^2$-locus$_{F3}$ because its changes in F3 locus (from mid (dental) /t/ to high (alveolo-palatal) [tɕ] in stem 1 and from low (velar) /k/ to mid (dental) [ts] in stem 2) are all one-step changes. Candidate (40b) involves a change from low (velar) /k/ to F3 (alveolo-palatal) [tɕ] in stem 2, which is a two-step change in F3 locus, so it does violate $\mathcal{F}^2$-locus$_{F3}$. Since $\mathcal{F}^2$-locus$_{F3}$ is ranked over $\mathcal{F}$-dorsal, candidate (40a)'s violations of $\mathcal{F}$-dorsal are not enough to prevent to it from being selected over candidate (40b) as the correct output.[9]

The following table summarizes the comparisons for the palatal mutations of coronals and velars, with the familiar notation referring to explicit analyses in tableaux or to reasons why no explicit analysis is necessary:

[9] This is not the only possible way in this line of reasoning to get the right results. Rather than a separate constraint $\mathcal{F}^2$-locus$_{F3}$ which punishes two-step violations, it might be the case that two-step violations are universally worse than one-step violations. A different notation could be used for multiple instances of violations, perhaps 2× to represent candidate (40a)'s violations of locus$_{F3}$, leaving ×^{2} to represent two-step violations like that for candidate (40b). Such a universal system of violation counting would obviate the need for an additional constraint. I leave this matter open to future research.

(41)

	k/ts	g/dz	x/ʂ	w/l
t/tɕ	(40)	*v*	*s*	*vsn*
d/dʑ	*v*	a_{40}	*vs*	*sn*
s/ɕ	*s*	*vs*	(35)	*vn*
z/ʑ	*vs*	*s*	*v*	*n*
n/ɲ̟	*vñ*	*ñ*	*vsñ*	*snñ*
r/ʐ	*vs*	*s*	*v*	(34)

v = $\mathcal{F}$-voicing is undominated
s = $\mathcal{F}$-stop is undominated
n = $\mathcal{F}$-narrow is undominated
ñ = $\mathcal{F}$-nasal is undominated
a_{40} = analogy to (40)

4.7 Conclusion

In this chapter, I have put forth an analysis of Polish palatal mutations based on selection of underlying allomorphs. These allomorphs are stored via strong lexicon optimization and can be selected in two ways. For a word composed entirely of extant morphemes, the allomorphs are marked via strong lexicon optimization with which morphological environments they occur in. This morphological subcategorization ensures that extant stems will emerge with the proper palatal mutations in the locative singular. Since nonce stems by definition have no prior usage, strong lexicon optimization cannot have created allomorphs for them, so the productive palatal mutations must come solely from the locative singular morpheme. Since this requires changing the nonce stem, various $\mathcal{F}$-constraints are violated. The specific ranking of $\mathcal{F}$-constraints determines which locative singular allomorphs are selected. In general, this analysis should be applicable to any case of productive morphological alternations, and it allows for such alternations to be opaque, since the opacity is encoded in the morphemes and not required to be a part of the constraint hierarchy.

A key component of this analysis is that a single instance of a two-step violation must be fundamentally different than two instances of a one-step violation. Since any chain shift seems to require that this distinction be made, this is not a drawback for FDM-OT. Every theory of phonology must come to terms with chain shifts, and the proposed analysis given here (which combines elements from Kirchner 1996 and Gnanadesikan 1997) is reasonable and potentially useful in other areas. For example, ALIGNMENT constraints (McCarthy and Prince 1995) are notorious for the difficulties they present in counting constraint violations. Is a single instance of misalignment by 2 units equally bad as two instances of misalignment by 1 unit? The analysis proposed here suggests that there is a way to distinguish these types of ALIGNMENT violations, with no amount of 1-unit violations able to 'gang up' on a single 2-unit violation. It remains to be seen if such conception of violation counting would be useful outside the realm of chain shifts.

A significant advantage of this analysis is that morphology and phonology are separate modules in the grammar, unable to directly refer to each other. The constraint hierarchy does not need to have different constraints or rankings depending on which morphemes are being evaluated, and the allomorphs listed in the lexicon are not subcategorized for phonological properties, only morphological ones. Why is such a separation a virtue for this framework? Because if this more restrictive separatist approach can derive the correct results, then there is no need for a more permissive theory with the superfluous power of allowing morphology and phonology to refer to each other.

Chapter 5: A Typology of Opacity

In the previous chapters, I have shown how FDM-OT can account for cases of opacity in Polish using a framework with direct mapping between the input and output. Because a framework with direct mapping lacks the abstract intermediate forms required to facilitate opacity, FDM-OT predicts that opacity cannot arise synchronically, and thus, that any apparent case of synchronically productive opacity is either not fully productive or is not really opaque.

In this chapter, I put forth a typology of the cases of opacity which can be analyzed in FDM-OT despite direct mapping and give a brief overview of prominent cases of opacity which fit into this typology (§5.1), exemplified by analyses of specific cases of opacity in languages other than Polish: refined Low German (§5.2), Turkish (§5.3), and Tuyuca (§5.4). In §5.5, I compare FDM-OT to other frameworks which have been used to analyze opacity. Finally, I conclude in §5.6 with a summary of the major results of this dissertation and directions for future research.

5.1 The typology

Because FDM-OT makes use of components such as strong lexicon optimization and $\mathcal{D}$-constraints, it can account for at least three types of opacity which cannot be analyzed in classic OT.

5.1.1 Unproductive opacity

The simplest case of opacity to consider is that which is in fact not synchronically productive, deriving from ordered diachronic sound changes. Because the results of these sound changes are stored directly in the lexicon via strong lexicon optimization, any opacity derived this way permeates the lexicon, and upon cursory examination of the language, might appear to be productive. However, in FDM-OT, direct mapping between the input and the output predicts that, even with a nearly ubiquitous presence in the lexicon, true opacity cannot be synchronically productive for arbitrary forms (i.e. for any word regardless of its morphological composition).

This type of opacity is seen with the [ɔ]~[u] alternation analyzed in Chapter 2, and the nasal vowel alternation analyzed in Chapter 3. In §5.2, I provide an analysis of a similar case of opacity from refined Low German, in which [ɡ] opaquely alternates with [x/ç]. I believe that a large number of cases of opacity fit into this category, and for most of these, it should be relatively straightforward to determine the absence or presence of synchronic productivity. For opacity in dead languages, however, it would be exceedingly difficult, if not entirely impossible, to test for productivity. Some well-known cases from the literature which plausibly seem to be of this type include:

- shortening and lowering of long high vowels in the Yawelmani dialect of Yokuts and the resulting interaction with vowel harmony (Newman 1944, Kuroda 1967, Kenstowicz 1994);
- Rendaku (voicing of the initial segment of the second member of a compound) in Japanese, which is blocked by [ŋ] when it derives from underlying /ɡ/ in the standard dialect of Tokyo (Ito and Mester 1997a); and
- lowering of Maltese vowels before underlying /ʕ/, which never surfaces (Brame 1972, Borg 1997).

5.1.2 Morphological opacity

It is also possible for an otherwise unproductive instance of opacity to be fully productive at certain morphological boundaries. As described above, truly opaque alternations cannot be synchronically productive tautomorphemically in a framework with direct mapping. However, because strong lexicon optimization stores allomorphic alternations in the lexicon, opacity can be encoded as part of productive affixes. When these opaquely allomorphic affixes are combined with novel stems (which do not have the benefit of history to encode the alternation in their underlying representations), the opaque alternations may still be allowed to surface.

This type of opacity is seen in the palatal mutations analyzed in Chapter 4, as well as in the opaque interaction of epenthesis and velar deletion in Turkish, which I analyze in §5.3. Other cases from the literature which seem to fit into this class of opacity include:

- palatalization and spirantization before certain morphemes in English (Rubach 1984b, Borowsky 1986, Kang 2002);
- in various dialects of Spanish, the plural marker /-s/ closes open syllables, which triggers laxing of mid vowels but /s/ cannot surface in codas (Saporta 1965, Hooper 1976);
- the [a]~[e] alternation (triggered by high vowels) in the definite marker in the Baztan dialect of Basque and its interaction with raising of mid vowels (Hualde 1991, Kenstowicz 1994); and
- umlaut triggered by [ʏ] in Icelandic is not triggered by an epenthesized [ʏ] (Orešnik 1972, Karvonen and Sherman 1997).

5.1.3 Transparent 'opacity'

A third type of opacity is not really opacity at all, given the right analysis. There are many synchronically productive alternations which are seemingly opaque, but when analyzed in a different way (by utilizing new theoretical tools such as $\mathcal{D}$-constraints, or by paying attention to other surface properties such as prosodic structure, precise phonetic implementation, etc.), they can be treated transparently. I present such a reanalysis in §5.3 using $\mathcal{D}$-constraints for the neutral behavior of obstruents within Tuyuca nasal harmony. There are other examples of transparent reanalyses of cases of opacity from the literature, such as:

- Bruening's (2000) argument that epenthesis in Tiberian Hebrew (Prince 1975) is triggered transparently by morphological templates which demand certain surface prosody, rather than by avoidance of complex codas;
- Padgett's (2002b) proposal that the Russian fricative [v], which behaves opaquely with respect to voicing assimilation, should actually be considered to be [ʋ̞] (a 'narrow' approximant), with transparent voicing behavior;
- Lee-Schoenfeld's (2002) analysis of German hypocoristics, which have been previously analyzed as augmented truncation that seems to depend on an unaugmented truncated form that never appears on the surface (Ito and Mester 1997b); and
- Various analyses within versions of Correspondence Theory (McCarthy and Prince 1995) that account for patterns of reduplication in many languages, many of which can be thought of instances of opacity (see Spaelti 1997, Struijke 1998, 2000, Inkelas and Zoll 2000, and many others).

Similar reanalyses might also be available for such cases as:

- neutral segments in any harmony system, such as [i] and [e] in Finnish vowel harmony (Ringen 1975, Kiparsky 1981, van der Hulst and van de Weijer 1995), accounted for by generalizing the analysis of Tuyuca nasal harmony in §5.4 to arbitrary types of harmony;
- [a] → [ɛ] → [e] → [i] in Nzɛbi (Guthrie 1968) and similar chain shifts, accounted for with multi-valued properties and $\mathcal{F}^2$-constraints (cf. Kirchner 1996, Gnanadesikan 1997, and §4.6.3 of Chapter 4).

5.1.4 Summary

The following chart summarizes the typology of opacity presented in this section. Italicized references contain a transparent analysis of traditionally opaque data:

(1) *unproductive opacity*

Polish	[ɔ]~[u] alternation	*Chapter 2*
	nasal vowel alternation	*Chapter 3*
refined Low German	[g]-spirantization	*§5.2*
Yawelmani Yokuts	vowel lowering and shortening	Newman 1944, Kuroda 1967, Kenstowicz 1994
Tokyo Japanese	[ŋ]-blocked Rendaku	Ito and Mester 1997a
Maltese	vowel lowering before /ʕ/	Brame 1972, Borg 1997

morphological opacity

Polish	palatal mutations	*Chapter 4*
Turkish	velar deletion	*§5.3*
English	palatalization and spirantization	Rubach 1984b, Borowsky 1986, *Kang 2002*
Spanish dialects	laxing and /s/-deletion	Saporta 1965, *Hooper 1976*
Baztan Basque	[a]~[e] alternation	Hualde 1991, Kenstowicz 1994
Icelandic	[ʏ]-umlaut	Orešnik 1972, Karvonen and Sherman 1997

transparent 'opacity'

Tuyuca	neutrality in nasal harmony	*§5.4*
Tiberian Hebrew	epenthesis	Prince 1975, *Bruening 2000*
Russian	[ʋ̥] voicing patterns	*Padgett 2002b*
German	hypocoristic truncation	Ito and Mester 1997b, *Lee-Schoenfeld 2002*
various languages	reduplication patterns	*McCarthy and Prince 1995, et alia*
Finnish, etc.	neutrality of [i/e] in vowel harmony	Ringen 1975, Kiparsky 1981, van der Hulst and van de Weijer 1995
Nzɛbi, etc.	[a/ɛ/e/i] chain shift	Guthrie 1968, *Kirchner 1996*

5.2 Refined Low German

Cross-linguistically, there are numerous examples of opacity with lexical exceptions. Generally, the lexical exceptions are numerous enough that the opaque generalization is deemed not to be productive. However, it is often the case that few enough lexical exceptions exist that a case of opacity is reported in the literature without noting how productive it may or may not be (this is especially true of analyses in serial frameworks, in which the existence of opacity provides supporting evidence for the framework itself). As shown in the analyses of Polish given in Chapters 2 and 3, instances of opacity which are not synchronically productive can still exist robustly in the lexicon due to strong lexicon optimization. But as long as only a small number of new words have entered the lexicon since the opaque generalization was rendered unproductive, it would be very difficult to determine productivity from a lexical search alone.

In this section, I provide an analysis of the alternation between velar stops and fricatives in refined Low German (the 'colloquial' Northern German of Weise 1996 and Ito and Mester 2003). Low German is the general name for modern dialects of

German spoken in the northern lowlands of Germany.[1] 'Refined' or *vornehm* Low German is the language that resulted from speakers of pure Low German dialects (for example, those around Hannover, Bielefeld, and Detmold) who integrated aspects of the southerly, more prestigious High German dialects for higher register speech (Pilch 1966:253, Robinson 2001:6, Armin Mester, personal communication). The resulting language has an alternation between [g] and [x/ç] that interacts opaquely with obstruent devoicing in codas. This case of opacity has recently been given a synchronically productive analysis within OT by Ito and Mester 2003. However, I show that there are lexical exceptions to the velar alternation, and thus, despite the overwhelming occurrence of opacity in the lexicon, it does not need a synchronic analysis. I demonstrate how FDM-OT can derive both the numerous opaque forms and the rarer transparent forms, and how it predicts that the alternation is not synchronically productive, a fact supported by the lexical exceptions and native speaker intuitions for nonce words.

5.2.1 Opacity in refined Low German

Like standard German, refined Low German does not allow voiced obstruents in codas; they must be voiceless (Weise 1996). This can be seen in the following nouns, in which the stem-final obstruent is in a coda in the nominative singular, where it must be voiceless, and is in an onset in the plural, where it can be voiced:

[1] 'Low German' is often also used to refer to any Germanic language that does not derive from High German, such as Dutch or English (Wells 1985:42). That usage is not intended here.

(2)	*stem UR*	*NOM SG*	*NOM PL*	*gloss*	
	/tʀiː**b**/	tʀiː**p**	tʀiː**b**ə	'sprout'	
	/ʀiː**d**/	ʀiː**t**	ʀiː**d**ə	'reed'	
	/ɡʀaɪ**z**/	ɡʀaɪ**s**	ɡʀaɪ**z**ə	'old man'	(Klatt et al. 1983)

As an obstruent, /ɡ/ is required to devoice in a coda. However, there is a further complication for /ɡ/ which prevents it from just simply devoicing to [k] in the Low German dialects (refined or otherwise).

In standard Low German dialects, postvocalic /ɡ/ appears as a fricative, spirantizing to either [ɣ] or [ʝ], which must devoice to [x] and [ç] respectively in codas (Viëtor 1904, Zhirmunskii 1962, Pilch 1966, Robinson 2001; the exact place of the resulting fricative is determined by the previous segment[2]):

(3)	*stem UR*	*NOM SG*	*NOM PL*	*gloss*
	/taː**ɡ**/	taː**x**	taː**ɣ**ə	'day'
	/veː**ɡ**/	veː**ç**	veː**ʝ**ə	'way'
	/bɛʀ**ɡ**/	bɛʀ**ç**	bɛʀ**ʝ**ə	'mountain'

Because no [ɡ] appears in any form, these forms are probably more appropriately analyzed with an underlying fricative as the stem-final consonant, instead of /ɡ/ (that is, as /taːɣ/, /vɛːʝ/, and /bɛʀʝ/). However, a more interesting situation is found in the refined Low German dialects. Under social pressure to conform to High German, any /ɡ/ in an onset is pronounced as a stop, whether preceded by a vowel or not, (as in High German), while /ɡ/ in a coda is spirantized (as in standard Low German), creating an alternation which seems to require underlying /ɡ/:

---

[2] The trigger behind this allophony is generally attributed to place assimilation to the previous vowel, with velar (or sometimes uvular) [ɣ/x] appearing after back vowels and palatal [ʝ/ç] appearing after front vowels. After consonants (of which, only /l/ and /ʀ/ may precede /ɡ/), [ʝ/ç] appears (Weise 1996, Robinson 2001). This allophony is not relevant here and will be assumed without analysis.

(4)

stem UR	*NOM SG*	*NOM PL*	*gloss*	
/taː**g**/	taː**x**	taː**g**ə	‘day’	
/veː**g**/	veː**ç**	veː**g**ə	‘way’	
/køːnɪ**g**/	køːnɪ**ç**	køːnɪ**g**ə	‘king’	(Wiese 1996:206)

Since surface [k] deriving from underlying /k/ is not subject to spirantization in the same position (for example, [plastɪ**k**] ‘plastic’, not ×[plastɪ**ç**]), this is an instance of opacity: a phonological generalization that cannot be described purely in terms of the surface form. In a serial analysis, such as in Weise 1996, this is easily captured through rule ordering, with Spirantization preceding Devoicing:

(5)

UR	/taːg/	/plastɪk/
Spirantization	taːɣ	—
Devoicing	taːx	—
output	[taːx]	[plastɪk]

This ordering is crucial. If /g/ devoices too early in the derivation, it merges with /k/, and there is no way to ensure that former /g/ eventually spirantizes while former /k/ does not. Under the incorrect ordering, either both /g/ and /k/ spirantize (in which case /plastɪk/ surfaces incorrectly as ×[plastɪç]) or neither do (and /taːg/ surfaces incorrectly as ×[taːk], as it does in standard German), depending on the precise formalization of Spirantization (i.e. whether it affects velar stops generally, or just voiced /g/):

(6)

UR	/taːg/	/plastɪk/	*or*	/taːg/	/plastɪk/
Devoicing	taːk	—		taːk	—
Spirantization	taːx	plastɪç		—	—
output	[taːx]	×[plastɪç]		×[taːk]	[plastɪk]

This is a case of opacity requiring an abstract intermediate representation (in this example, /taːɣ/), a requirement that is problematic for classic OT. Consider the following tableau, with ✗C̬]σ and ✗g]σ as cover constraints for the markedness conditions which trigger, respectively, devoicing of obstruents in a coda and spirantization of /g/ in a coda:

(7)

/taːg/	✗C̬]σ	✗g]σ	IDENT [voi]	IDENT [cont]
a. taːg	✗!	✗!		
b. taːɣ	✗!			✗
☠ c. taːk			✗	
☹ d. taːx			✗	✗!

The fully faithful candidate (7a) and the candidate with spirantization (7b) both contain voiced obstruents in the coda, and thus are disallowed (specifically, because of the ranking of ✗C̬]σ over IDENT-[voice], which is required to trigger obstruent devoicing in codas). This leaves the candidate with devoicing (7c) and the candidate with both devoicing and spirantization (7d) in contention. Both are unfaithful to the input with respect to voicing, but (7d) is also unfaithful with respect to continuancy. By having a subset of (7d)'s violations, transparent (7c) defeats opaque (7d) under any ranking of the relevant constraints, and the output of /taːg/ is predicted to be ✗[taːk], with no evidence of opacity.

Ito and Mester 2003 solve this problem by means of constraint conjunction (Smolensky 1993) of markedness and faithfulness constraints, along the lines of Łubowicz 1998. However, by attempting to preserve certain types of opacity in the synchronic grammar, this analysis weakens direct mapping in OT, and as I show below, makes incorrect empirical predictions concerning the (lack of) synchronic productivity of velar spirantization in refined Low German.

5.2.2 Lexical exceptions and nonce forms

While difficult to find, there do appear to be some lexical exceptions to velar spirantization. I collected the following data from native speakers of German familiar with refined Low German.[3] The singular forms are pronounced without spirantization to [ç] or [x], despite having a coda consonant which is assumed to be an underlying /g/ (as evidenced by a combination of the orthography, the original pronunciation of the loanword, and/or any related forms in which [g] surfaces in an onset):

(8)	*stem UR*	*NOM SG*	*related form*	*gloss*
	/tʀɪ**g**/	tʀɪ**k**	tʀɪ**g**onometʀiə	‘trig(onometry)’
	/monoloː**g**/	monoloː**k**	monoloː**g**ə	‘monologue(s)’
	/bʀɪ**g**/	bʀɪ**k**	—[4]	‘brig (type of jail)’
	/gɪ**g**/	gɪ**k**	—	‘gig (musician’s job)’
	/mɪ**g**/	mɪ**k**	—	‘MiG (type of fighter jet)’
	/eʀbɛ**g**/	eʀbɛ**k**	—	‘airbag’

In addition, while I have not conducted a formal experiment as in Chapter 2 for Polish, I asked my consultants about their judgments for nonce stems ending in /g/ by giving them plurals such as [blɪgə] and asking for singulars. Monosyllabic stems are

---

[3] Native speakers are hard to come by, as Low German is mostly relegated to older generations. My consultants were children and grandchildren of native speakers.

[4] There are very few related forms which put the final consonant of the stem into the onset of a syllable, where [g] may surface. This is due to the fact that the affixes in the declensions of loanwords and proper names are often restricted to -Ø (null) and [-s].

invariably pronounced with final [k], rather than with fricatives ×[ç] or ×[x]. Polysyllabic forms also show no spirantization, except when /g/ is preceded by /ɪ/. This is likely due to reanalysis of final /ɪg/ as the adjectival and adverbial suffix [-ɪç]. Aside from this exception, final /g/ in nonce stems does not undergo spirantization.

More research using native speakers of refined Low German is called for, but these results suggest that spirantization is not synchronically productive, obviating the need for a synchronic analysis. Of course, the fact that spirantization still exists so prominently in the lexicon still needs to be accounted for, and strong lexicon optimization interspersed between diachronic sound changes can accomplish this.

5.2.3 Strong lexicon optimization

As with any analysis of lexical opacity based on the interaction of multiple sound changes, it is necessary to look at the relative chronology and the specific effects of the sound changes in question. The first of these is an alternation in Proto-Germanic between voiced fricatives *[v ð ɣ] and voiced stops *[b d g]: fricatives appeared after vowels if not geminated, while stops appeared elsewhere (Waterman 1966:24–27, 56, Russ 1978:31–34):[5]

(9)		'day'				'back'	
	Proto-Germanic	**d**a**ɣ**	×da**g**	×**ð**a**ɣ**	×**ð**a**g**	ru**gg**i	×ru**ɣɣ**i

By richness of the base, this alternation must accept any possible UR (e.g. /dag/, /daɣ/, /ðag/, and /ðaɣ/ for 'day') and change voiced fricatives to stops (occlusion) and voiced stops to fricatives (spirantization) in the correct positions (e.g. /ðag/ → [daɣ]).

[5] This alternation was lost in Old High German as the voiced fricatives became stops in all positions (Russ 1978:50).

In the following FDM-OT tableau, $\mathcal{F}$-continuancy punishes both occlusion and spirantization (changes in the feature [continuant]), and $\mathcal{M}$-#ð and $\mathcal{M}$-Vg are cover constraints which ban, respectively, any word-initial voiced fricative, and a vowel followed by any voiced stop (the specific constraints are not relevant here; a more precise analysis is certainly available):

(10) *Proto-Germanic occlusion/spirantization*

dag$_1$ day$_3$ ðag$_2$ ðay$_4$	$\mathcal{M}$ #ð	$\mathcal{M}$ Vg	$\mathcal{F}$ cont
a. dag$_1$ day$_3$ ðag$_2$ ðay$_4$	×^{2}!	×^{2}!	
✓ b. day$_{1234}$			×^{3}
c. dag$_{12}$ day$_{34}$		×!	×^{2}
d. day$_{13}$ ðay$_{24}$	×!		×^{2}

The fully faithful candidate (10a) violates the highly ranked $\mathcal{M}$-constraints by having the voiced fricative [ð] at the beginning of words and the voiced stop [g] after vowels. Candidate (10b) represents the actual output of Proto-Germanic, with both occlusion of initial voiced fricatives and spirantization of post-vocalic voiced stops, satisfying both of the $\mathcal{M}$-constraints at the expense of lowly ranked $\mathcal{F}$-continuancy. Candidate (10c) has only occlusion, and candidate (10d) has only spirantization; both violate one of the $\mathcal{M}$-constraints and thus, cannot emerge as the output under this ranking, despite being more faithful than the actual output (10b).

Language learners during early Old Low German (OLG) would hear Proto-Germanic outputs like [daɣ] and construct a grammar similar to (10). Eventually, OLG speakers would optimize the lexicon by storing underlying representations that are as faithful as possible to the output, replacing any less faithful potential underlying representations created by richness of the base. Thus, [daɣ] 'day' would be stored as the most faithful input /daɣ/, rather than the other contenders such as /dag/, /ðag/, and /ðaɣ/. This is the crucial step in the creation of opacity, as the newly optimized lexical entries act as intermediate representations, which are a necessary component of opacity.

Later in OLG, a sound change emerged in which obstruents in coda position devoiced, mirroring the same change in contemporaneous versions of High German (Lockwood 1976:61, 80; Russ 1978:65):

(11)			'day'
	early OLG		da**ɣ**
	late OLG	devoicing	da**x**

The FDM-OT analysis of word-final devoicing for Polish in Chapter 2 can easily be adapted to coda devoicing in OLG. The following tableau shows what happens to early OLG forms with voiced codas, voiceless codas, and no codas (note the lack of input forms with post-vocalic [g], such as [dag]; the input is the lexicon that resulted from strong lexicon optimization of early OLG, which eliminated such forms):

(12) *late Old Low German devoicing*

			$\mathcal{M}$ C̬]σ	$\mathcal{F}$ voi
	dak$_1$ dakə$_2$	daɣ$_3$ dax$_4$ daɣə$_5$ daxə$_6$		
a.	dak$_1$ dakə$_2$	daɣ$_3$ dax$_4$ daɣə$_5$ daxə$_6$	×!	
✓ b.	dak$_1$ dakə$_2$	dax$_{3,4}$ daɣə$_5$ daxə$_6$		×

The fully faithful candidate (12a) violates high ranking $\mathcal{M}$- C̬]σ because of the form [daɣ] with a voiced obstruent in the coda, leaving candidate (12b) as the output of late OLG, with coda devoicing.

Coda devoicing interacts opaquely with Proto-Germanic spirantization, as in the form [dax] 'day'. In FDM-OT (and classical OT), this kind of opaque interaction cannot exist in the synchronic grammar. A proposed input such as /dag/ would map to [dak] rather than [dax], since it does not have the benefit of being stored via strong lexicon optimization in early OLG. Both [dak] and [dax] are well-formed on the surface (both are in (12b)), so the choice between them comes down to their faithfulness to the input. Of the two, [dak] is clearly the most faithful to /dag/, changing only the voicing specification of the final segment. In comparison, [dax] involves *two* changes, one in voicing and one in manner. Thus, transparent [dak], rather than opaque [dax], will be the output of /dag/ in late OLG. In general, new words entering late OLG or later (including the modern dialects) would only undergo devoicing, without spirantization.

Later loanwords such as *Mig* and nonce words are then predicted to be pronounced transparently in all Low German dialects, for example, as [mɪk], rather than opaquely, like ✗[mɪç], since spirantization can no longer be synchronically productive in FDM-OT. While opaque spirantization is not productive, it is still pervasive in the lexicon. This is achieved by strong lexicon optimization in early OLG, which stored the effects of spirantization. This storage persisted in the lexicon into modern Low German, leaving evidence of historical opacity embedded in the lexicon, as with opaque [taːx] 'day'. Nonetheless, spirantization is no longer synchronically productive, and the analysis sketched in this section accounts for both the presence of opaque forms in the lexicon and the lack of productivity of the older process of spirantization.

The results of this case of opacity are present in both standard and refined Low German, since both pronounce 'day' as [taːx]. In a typical generative analysis of standard Low German, this is actually a transparent form because the UR is /taːɣ/, as evidenced by the plural form [taːɣə]. However, in refined Low German, opacity is apparent in such an analysis due to a sound change of velar occlusion which has resulted in the plural being pronounced [taːgə], requiring the single underlying stem to be /taːg/. Velar occlusion arose some time after High German dialects had come to be seen as dialects of prestige. The Low German dialects evolved a refined version in which standard Low German [ɣ] in an onset was occluded [g], mimicking the High German pronunciation (which derives from one of the major splits between OLG and OHG: the loss of the Proto-Germanic voiced fricatives in OHG). Thus, standard Low

German [taːɣə] 'days' is pronounced [taːgə] in refined Low German. As with any sound change, lexicon optimization stores the results (a now very obviously opaque alternation between [g] and [x/ç]), directly into the lexicon, but only for extant words. Any new words with /g/ in the relevant environment will be pronounced transparently with an alternation between [g] and [k], with no spirantization.

5.3 Turkish

5.3.1 Productive opacity in Turkish

In Turkish, there is an opaque interaction between epenthesis within word-final consonant clusters and deletion of intervocalic /k/ (Lewis 1967, Zimmer and Abbott 1978, Kenstowicz and Kisseberth 1979, Sezer 1981, Inkelas and Orgun 1995).[6] Both processes are fully productive, even marked in the spelling (deleted /k/ is spelled ⟨ğ⟩ in the Turkish orthography, representing an older stage of the language in which was a voiced fricative [ɣ], a pronunciation still heard in some dialects):

(13) a. *epenthesis within word-final clusters*

/baʃ-m/	[baʃɯm]	'head (1SG POSS)'
/ad-n/	[adɯn]	'name (2SG POSS)'

b. *intervocalic /k/-deletion*

/ajak-ɯ/	[ajaɯ]	'foot (3SG POSS)'
/kabak-ɯ/	[kabaɯ]	'pumpkin (ACC)'

c. *opaque interaction between epenthesis and deletion*

/ajak-m/	[ajaɯm]	'foot (1SG POSS)'
/kabak-n/	[kabaɯn]	'pumpkin (2SG POSS)'

[6] I assume here that intervocalic /k/ is in fact phonetically null. However, the result of this lenition process can vary. For example, between front vowels, a weak [j] can be heard (Zimmer and Orgun 1999).

In a multistratal model, the analysis of this data is quite simple: Epenthesis applies first, breaking up the consonant cluster, and the resulting vowel then acts as part of the trigger for /k/-deletion:

(14)		
	UR	/ɑjɑk-m/
	Epenthesis	ɑjɑkɯm
	/k/-Deletion	ɑjɑɯm
	output	[ɑjɑɯm]

However, in frameworks with direct mapping between the input and output, like standard OT and FDM-OT, an analysis of these data is not so straightforward. Consider the following tableau, in which ×VkV and ×CC# are cover constraints punishing intervocalic /k/ and word-final consonant clusters respectively (whether these constraints are covers for markedness or dispersion constraints, or some combination, is not important and does not change the analysis or the difficulty OT has with this case of opacity), and MAX and DEP are the faithfulness constraints which ban deletion and epenthesis, respectively:

(15)

ɑjɑk-m	×VkV	×CC#	MAX	DEP
a. ɑjɑkm		×!		
b. ɑjɑkɯm	×!			
☹ c. ɑjɑɯm			×	×!
☠ d. ɑjɑɯm			×	

The fully faithful candidate (15a) contains a word-final consonant cluster, while the candidate with epenthesis (15b) contains an intervocalic [k]. Both of these structures are banned by the highly ranked constraints ×VkV and ×CC#, so these candidates are ruled out. Candidates (15c) and (15d) both have deletion of underlying /k/, so they tie with respect to MAX, at one violation each. This leaves DEP to determine the output. Since candidate (15c) has the epenthesized vowel [ɯ], it violates DEP. Candidate (15d) has no epenthesis and satisfies DEP. Thus, this tableau shows that the transparent, but incorrect, candidate (15d) is selected instead of the opaque, but grammatical, candidate (15c). Indeed, because (15c) is harmonically bounded by (15d), no ranking of these constraints can get the desired output (15c) to emerge as the winner.

5.3.2 The role of morphology

Since the loss of intervocalic /k/ is triggered by the suffixes /-m/ '1SG POSS' and /-n/ '2SG POSS', which can be productively attached to any noun stem (barring pragmatic restrictions), deletion of /k/ must be productive on some level. However, it is not fully productive in the language as a whole, as the following representative data show:

(16)	/birt**ɑkɯ**m/	birt**ɑkɯ**m	✗birt**ɑɯ**m	'several'
	/j**ɑkɯ**-m/	j**ɑkɯ**m	✗j**ɑɯ**m	'blister (1SG POSS)'
	/bɯr**ɑk-ɯ**jor/	bɯr**ɑkɯ**jor	✗bɯr**ɑɯ**jor	'leave (3SG PROG)'
	/ger**ek-i**r/	ger**eki**r	✗ger**ei**r	'deservedly'
	/bɑnkɑ-d**ɑ-ki**/	bɑnkɑd**ɑki**	✗bɑnkɑd**ɑi**	'which is in the bank'

These data suggest that deletion of /k/ is not fully productive for all surface forms, only at certain morphological boundaries. In this respect, Turkish /k/-deletion is similar to Polish palatal mutation, which leads to an analysis within FDM-OT in which the morphological alternations are stored via strong lexicon optimization and allomorphs are selected, rather than created, by the grammar, as shown in Chapter 4.

5.3.3 Turkish epenthesis as allomorphic selection

The 1SG POSS suffix has two surface allomorphs: [-m], which attaches to vowel-final stems, as in [kedim] 'cat (1SG POSS)' (cf. [kedi] 'cat'); and [-ɯm], which attaches to consonant-final stems, as in [ɑtɯm] 'horse (1SG POSS)' and opaque [ɑjɑɯm] 'foot (1SG POSS)' (cf. unpossessed [ɑt] 'horse' and [ɑjɑk] 'foot').[7] At first glance, one might posit that the lexical entry for '1SG POSS' would contain /-m/ and /-ɯm/, identical to the two surface allomorphs. However, this is not sufficient to derive Turkish. Since intervocalic deletion of /k/ is not productive, it is not reflected in the constraint hierarchy. Thus, there would be nothing to ensure that a nonce stem such as /pɑjɑk/ would lose its /k/ when it takes /-ɯm/ (the allomorph it would be forced to take, since /-m/ would result in an ill-formed consonant cluster). Because velar deletion is not productive except with certain morphemes, those morphemes must have deletion encoded in them somehow. Following the analysis of palatal mutation in Polish given in Chapter 4, I assume that changes to the stem can be achieved via floating segments, indicated in a superscripted grey font. These floating segments coalesce with the appropriate stem segments, creating the necessary changes.

For stems ending in non-velar consonants, the corresponding allomorphs of the 1SG POSS morpheme have a floating segment identical to the stem-final consonant. Thus, the stem /ɑt/ 'horse' takes the allomorph /-^{t}ɯm/, while /kɯz/ 'girl' takes /-^{z}im/. The merger of the stem-final consonant with the floating segment of the affix is the same as in Polish: the floating segment has no timing slot of its own, and the constraint hierarchy prefers coalescence to creation of a new timing slot, so the

[7] I am ignoring the mechanics of vowel harmony in this analysis, since it is not relevant to the case of opacity under consideration. Technically, there are more allomorphs for 1SG POSS, since there are four possible vowels [i y ɯ u], which surface in order to harmonize with the vowels in the stem.

floating segments latches onto the final consonant of the stem. Since the two segments are identical, there are no violations of $\mathcal{F}$-constraints (except whichever $\mathcal{F}$-constraint prevents coalescence, which is ranked low enough not to matter here).

For velar-final stems, the situation is a bit more complex. Since there is no 'null segment' that could be used as the floating segment for velar-final stems to induce deletion, the stem-final /k/ must merge with something else (and it must be merger, since intervocalic deletion of /k/ is not productive). I propose that the final /k/ of the stem merges with a floating vowel in the affix. In other words, the relevant allomorph for '1SG POSS' is /-ɯm/. The stem-final /k/ and the floating affixal /ɯ/ coalesce to [ɯ] in the output, resulting in violations of all of the properties that distinguish [k] from [ɯ]: consonantality, voicing, stop, sonorant, etc. Crucially, the property that does not change is dorsal, or more generally, place of articulation. This explains why other consonant-final stems do not merge with /-ɯm/: if $\mathcal{F}$-place is highly ranked, only consonants which match the floating vowel's place of articulation (dorsal) can merge with it.

For vowel-final stems, /-m/ is sufficient to capture the right generalization. Since most stems do not undergo any alternation, they have only one allomorph listed in the lexicon. However, extant velar-final stems will have two allomorphs, one with final /k/ and one without. This means that a partial lexicon for Turkish is as follows:

(17)

HORSE	=	/ɑt/	_____ + {1SG POSS,PL,…}
HEAD	=	/inek/ /ine/	_____ + {PL,…} _____ + {1SG POSS,…}
CAT	=	/kedi/	_____ + {1SG POSS,PL,…}
1SG POSS	=	/-$^{\text{t}}$ɯm/ /-$^{\text{ɯ}}$m/ /-m/	{HORSE,…} + _____ {HEAD,…} + _____ {CAT,…} + _____

If we only needed to consider extant forms, there would be no problem: the morphological specification in the lexicon (given by strong lexicon optimization) would tell the speaker which allomorphs to select. However, deletion of intervocalic /k/ does not apply just to stored morpheme; it is productive with any stem combined with the 1SG POSS, so the analysis must ensure that even if the stem is not specified for deletion of /k/ when combined with the 1SG POSS (as is the case for nonce forms and new loanwords), the correct allomorph of '1SG POSS' is selected and the stem-final /k/ is deleted. In addition, the analysis must prevent deletion of non-velar consonants and ensure that vowel-final stems take the correct allomorph. I consider the sub-language of Turkish consisting of the nonce stems /pɑjɑt/, /pɑjɑk/, and /pɑjɑ/ (all italicized in the following tableau; extant stems are not considered since their output is determined by the morphology and not the phonology), and the 1SG POSS allomorphs /-$^{\text{t}}$ɯm/, /-$^{\text{ɯ}}$m/, and /-m/ (all subscripted with 1SP):

(18) *pajat*$_1$- *pajak*$_2$- *paja*$_3$- -ᵗɯm$_{1SP}$ -ᵚm$_{1SP}$ -m$_{1SP}$	𝓜 CC#	𝓕 seg	𝓕 place	𝓕 cons (etc.)	𝓜 VkV
✓ a. *paja*$_1$-*t*$_{1,1SP}$-ɯm$_{1SP}$ *paja*$_2$-*ɯ*$_{2,1SP}$-m$_{1SP}$ *paja*$_3$-m$_{1SP}$				✗^{N}	
b. *paja*$_1$-*ɯ*$_{1,1SP}$-ɯm$_{1SP}$ *paja*$_2$-*t*$_{2,1SP}$-ɯm$_{1SP}$ *paja*$_3$-m$_{1SP}$			✗^{2}!	✗^{N}	
c. *paja*$_1$-*t*$_{1,1SP}$-ɯm$_{1SP}$ *paja*$_2$-*t*$_{2,1SP}$-ɯm$_{1SP}$ *paja*$_3$-m$_{1SP}$			✗!		
d. *paja*$_1$-*t*$_{1,1SP}$-ɯm$_{1SP}$ *paja*$_2$-m$_{1SP}$ *paja*$_3$-m$_{1SP}$		✗!			
e. *paja*$_1$-*t*$_{1,1SP}$-ɯm$_{1SP}$ *pajak*$_2$-ɯ-m$_{1SP}$ *paja*$_3$-m$_{1SP}$		✗!			✗
f. *paja*$_1$-*t*$_{1,1SP}$-ɯm$_{1SP}$ *pajak*$_2$-m$_{1SP}$ *paja*$_3$-m$_{1SP}$	✗!				

The grammatical candidate (18a) that represents true Turkish has the correct allomorphs selected for each nonce stem. Stem 1 /pɑjɑt/ takes /-ᵗɯm/, with /t/ and /ᵗ/ merging completely faithfully (modulo 𝓕-uniformity or whatever low-ranking constraint bans merger). Stem 2 /pɑjɑk/ takes /-ᵚm/, with /k/ and /ᵚ/ merging, violating all the 𝓕-constraints that distinguish [k] from [ɯ] (𝓕-consonant, 𝓕-stop, 𝓕-voicing, etc.). Stem 3 /pɑjɑ/ takes /-m/, with simple concatenation and no violated 𝓕-constraints.

Candidate (18b) shows what happens if stem 1 /pɑjɑt/ tries to take /-ɯm/, while stem 2 /pɑjɑk/ takes /-tɯm/ (there is no /-kɯm/ for it to take, since that is not a possible allomorph of the 1SG POSS morpheme). This results in many constraint violations. The merger of /t/ with /ɯ/ violates not only the low-ranked 𝓕-constraints, but also high-ranking 𝓕-place. Additionally, the merger of /k/ with /t/ increases the number of 𝓕-place violations. Candidate (18c) is slightly better, with stem 1 /pɑjɑt/ correctly taking /-tɯm/. However, it also has stem 2 /pɑjɑk/ taking /-tɯm/ in an attempt to circumvent the massive change that would result from /k/ merging with /ɯ/. Since 𝓕-place is so highly ranked, merging /k/ with /t/ is still worse than merging /k/ with /ɯ/, since the /k-t/ merger requires a switch from dorsal to coronal that the /k-ɯ/ merger does not. Thus, neither (18b) nor (18c) can defeat (18a).

In candidates (18d–f), stems 1 /pɑjɑt/ and 3 /pɑjɑ/ take the correct allomorphs as in (18a), while stem 2 /pɑjɑk/ takes /-m/. To avoid the ill-formed final consonant cluster that would result if both /k/ and /m/ were allowed to surface, candidate (18d) involves deletion of the stem-final /k/ (a violation of 𝓕-segment), and candidate (18e) breaks up the potential [km] with an epenthetic [ɯ] (also a violation of 𝓕-segment).[8] Because 𝓕-segment is ranked over the other 𝓕-constraints, neither (18d) nor (18e) cannot surface as the grammatical output. Candidate (18f) makes no attempt to repair the final [km] cluster, letting it surface faithfully. This cluster violates 𝓜-CC#, which is ranked high enough to prevent (18f) from defeating (18a), leaving candidate

[8] Note that candidate (18e) also has an intervocalic [k], which violates the constraint 𝓜-VkV. However, since intervocalic [k] is generally allowed in Turkish, this constraint is low-ranked and has no effect.

(18a) as the winner. This represents the correct output for Turkish, with velar-final stems losing their final /k/ in the 1SG POSS form, while vowel-final and other consonant-final stems emerge unchanged. Other affixes which trigger deletion of /k/ would be analyzed the same way, leaving intervocalic /k/ as a possible surface structure (cf. (16)).

5.4 Tuyuca

A neutral segment is one that neither undergoes nor blocks harmony, allowing segments on both sides of the intervening neutral segment to harmonize with each other unhindered.[9] The behavior of neutral segments is often analyzed as an instance of opacity, since neutral segments should undergo harmony in order to allow harmony to pass through them, but they do not harmonize on the surface. If this is correct, then they represent one unquestionably productive instance of opacity. In this section, I show that neutral segments can be derived through the use of $\mathcal{D}$-constraints by providing an analysis of Tuyuca, a Tucanoan language spoken in Columbia and Brazil.

[9] This usage of 'neutral segment' is non-standard, since it typically refers to any segment which does not undergo harmony, whether it allows the harmony to pass through or not (cf. van der Hulst and van de Weijer 1995, Walker 1998, etc.). However, the more usual terms 'transparent segment' and 'opaque segment' are confusing in the context of this dissertation, since 'transparent segments' involve phonological opacity, while 'opaque segments' behave transparently. To avoid confusion, I use 'neutral segment' in place of the more traditional 'transparent segment'.

5.4.1 Tuyuca nasal harmony

Words in Tuyuca harmonize for nasality; generally, they are either all nasal or all oral (all data in this section are from Walker 1998; also see Barnes and Takagi de Silzer 1976 and Barnes 1996):[10]

(19)

all oral		*all nasal*	
waa	'to go'	w̃ãã	'to illuminate'
hoo	'banana'	h̃ɨ̃ɨ̃r̃ĩ	'watch out or you'll get burned!'
wati	'dandruff'	j̃õr̃ẽ	'little chicken'

However, the voiceless obstruents in Tuyuca, [p t k s], do not undergo nasal harmony, due to the relatively high markedness of nasalized obstruents (including the impossibility of nasalized obstruent stops; see Walker 1998 and references therein). Instead, they remain oral (underlying voiced obstruents do harmonize by becoming nasal sonorants). In addition, voiceless obstruents are neutral segments, allowing the nasal harmony to spread through them:

(20)

mĩpĩ	'badger'
w̃ãtĩ	'demon'
ãkã	'choke on a bone'
j̃õsõ	'bird'

Neutral segments act like 'holes' in an otherwise continuous domain of harmony. In order to avoid problems of skipping (see Walker 1998 and Padgett and Ní Chiosáin 2001 for discussion of this problem), most recent analyses of neutral segments require some type of opacity to account for their behavior. A typical

[10] This description breaks down for some morphemes, which carry their own nasality or orality. I ignore that complication here.

opaque analysis requires the neutral segments to harmonize in some representation other than the surface form, such as an intermediate level (as in a serial framework) or a sympathetic candidate (as in Walker's OT-based analysis of Tuyuca), so that the harmony spreads throughout the entire word, including the neutral segments. The ill-formed segments which result from the neutral segments harmonizing are then 'repaired' at the surface, forced to change back to their original, disharmonic, form, leaving the rest of the harmonized segments alone. The following derivation of the Tuyuca word [ãkã] 'choke on a bone' illustrates this type of analysis:[11]

(21)	UR	/ãka/	final /a/ begins as an oral vowel
	Harmony	ãk̃ã	/k/ harmonizes to marked (impossible) /k̃/
	Repair	ãkã	[k̃] is not allowed on the surface and is repaired to [k]
	output	[ãkã]	

The end result is a word in which the segments before and after [k] harmonize for nasality, while [k] itself stands out as disharmonic with respect to the rest of the word. This type of opacity is the kind that is problematic for a framework with direct mapping between input and output, such as classic OT and FDM-OT.

5.4.2 Motivating harmony

The skeletal analysis of Tuyuca nasal harmony given in the previous section assumes that the harmony in question must spread to intervening segments in order to reach segments farther along in the word. This interpretation necessarily requires a post-harmonic repair in order to achieve the effects of neutral segments, and thus, will always require some abstract representation that is neither pronounced nor stored in

[11] For simplicity of exposition, I follow Walker in assuming here that nasality spreads outwards from the initial vowel, though this assumption is not crucial.

the lexicon. This is classic opacity. This case of opacity relies on the crucial assumption that harmonic spreading must be continuous. This assumption is a natural extension of a framework based on markedness, since it is hypothesized that adjacent segments tend to share features due to a decrease in articulatory difficulty. It seems reasonable in such a framework to assume that this is the primary motivation for harmony.

As Jaye Padgett (personal communication) points out, simply requiring adjacent segments to agree is not sufficient to force harmony. Some domain of agreement (syllable, foot, etc.) must be specified. For Tuyuca, the relevant domain is the word. I use $\mathcal{M}$-$\square_{\square}$nasal to be the constraint that requires every continuous string of nasal segments to span the entire word, with violations counted segmentally for each nasal string. Thus, a word like [ãka] would have two violations of $\mathcal{M}$-$\square_{\square}$nasal (for the two segments that are not part of the continuous nasal string consisting only of the first [ã]), while the word [ãkã] would have four violations (the same two violations for the first [ã], plus the mirror violations for the final [ã]; note that despite being nasal themselves, the two nasal vowels do not count as part of each other's *continuous* string of nasal segments).

However, in FDM-OT, markedness is not the only way to force segments to change (or in constraint terms, to force violations of $\mathcal{F}$-constraints). I argue that $\mathcal{D}$-constraints can be used to cause harmony, by requiring that entire words, not just individual segments, must be sufficiently distinct along some perceptual dimension. Thus, there are two potential sources for harmony: $\mathcal{M}$-constraints like $\mathcal{M}$-$\square_{\square}$nasal, which result in continuous domains of harmony containing no neutral segments; and

𝒟-constraints, which result in perceptual harmony, which can have neutral segments since it is not bound by physical limitations on coarticulation.

5.4.3 Tuyuca nasal harmony as perceptual harmony

As used so far, 𝒟-constraints have been used to measure perceptual distinctiveness at the segmental level, but in fact, the definition as given in Chapter 1 (repeated below) actually requires *words* to be sufficiently distinct:

(22) $\mathcal{D}_n$-P 'Every pair of words *x* and *y* in the output which contrast for property *P* must be at least as far apart as the *n*th from smallest allowable perceptual distance for *P*.'

Thus, while [tapa] and [tapã] contrast for the property of nasality, so do [tapa] and [tãpã]. For the purposes of this analyses, we need to consider the set of eight words shown in (23), consisting of the segmental string [aka], with each segment potentially being nasal or oral. This set represents the universal input □, the set of all possible words, and by richness of the base, serves as the input to the grammar prior to strong lexicon optimization (which is not crucial here).

(23)

aka	akã
ãka	ãkã
ak̃a	ak̃ã
ãk̃a	ãk̃ã

Without the benefit of 𝒟-constraints, ℱ- and ℳ-constraints can be used to derive articulatory harmony with no neutral segments:

(24) *transparent articulatory nasal harmony without neutral segments*

aka$_1$ akã$_5$ ãka$_2$ ãkã$_6$ ak̃a$_3$ ak̃ã$_7$ ãk̃a$_4$ ãk̃ã$_8$	$\mathcal{M}$ k̃	$\mathcal{M}$ $\square_{\square}$nas	$\mathcal{F}$ nas
a. aka$_1$ akã$_5$ ãka$_2$ ãkã$_6$ ak̃a$_3$ ak̃ã$_7$ ãk̃a$_4$ ãk̃ã$_8$	×^{4}!	×^{12}	
✓ b. aka$_{1,3,5,7}$ ãka$_{2,4,6,8}$		×^{2}	×^{8}
c. aka$_{1,3,5,7}$ ãkã$_{2,4,6,8}$		×^{4}!	×^{8}
d. aka$_{1,3,5,7}$ ãk̃ã$_{2,4,6,8}$	×!		×^{8}

The fully faithful candidate (24a) and the candidate with full harmony (24d) cannot surface because they contain the marked segment [k̃], which is banned by high-ranking $\mathcal{M}$-k̃. The two remaining candidates both have forms which are not fully harmonic, [ãka] in (24b) and [ãkã] in (24c). These forms violate $\mathcal{M}$-$\square_{\square}$nasal, which forces nasality to continuously span the entire word. In (24b), the segment [k] blocks harmony; nasality cannot spread past it, so the final [a] is oral. But in candidate (24c), [k] is a neutral segment, allowing the vowels on both sides to harmonize with each other for nasality. This candidate represents Tuyuca. Because it allows discontinuous harmony, candidate (24c) incurs more violations of $\mathcal{M}$-$\square_{\square}$nasal, and thus loses to candidate (24b) which contains no neutral segments.

Luckily for Tuyuca, FDM-OT is not limited to just $\mathcal{F}$- and $\mathcal{M}$-constraints. Comparing candidates (24b) and (24c), it seems intuitive how a $\mathcal{D}$-constraint might help differentiate them, assuming the graphical layout of the words has some correlation to their relative perceptual distinctiveness: the words in candidate (24b) are closer together than the words in (24c). More specifically, the words in candidate (24b) differ from each other by only one instance of nasality, while the words in candidate (24c) differ by two instances. Thus, if we assume that there is some sort of $\mathcal{D}$-nasal constraint which measures such differences between words, then Tuyuca with its neutral obstruents can be derived in FDM-OT. Here, I simplify the analysis by assuming a single $\mathcal{D}$-nasal constraint which punishes words that differ by only one instance of nasality, but allows words which differ by two or more:

(25) *opaque perceptual nasal harmony with neutral segments*

aka$_1$ akã$_5$ ãka$_2$ ãkã$_6$ ak̃a$_3$ ak̃ã$_7$ ãk̃a$_4$ ãk̃ã$_8$	$\mathcal{M}$ k̃	$\mathcal{D}$ nas	$\mathcal{M}$ □nas	$\mathcal{F}$ nas
a. aka$_1$ akã$_5$ ãka$_2$ ãkã$_6$ ak̃a$_3$ ak̃ã$_7$ ãk̃a$_4$ ãk̃ã$_8$	×^{4}!	×^{12}!	×^{12}	
b. aka$_{1,3,5,7}$ ãka$_{2,4,6,8}$		×!	×^{2}	×^{8}
✓ c. aka$_{1,3,5,7}$ ãkã$_{2,4,6,8}$			×^{4}	×^{8}
d. aka$_{1,3,5,7}$ ãk̃ã$_{2,4,6,8}$	×!			×^{8}

The highly marked segment [k̃] still prevents the fully faithful (25a) and the fully harmonic (25d) candidates from winning. The new constraint 𝒟-nasal enforces a minimum perceptual distance within the candidates, causing (25b) to lose due to the very similar pair of words, [aka] and [ãka], which differ only in a single instance of nasality. Candidate (25c), which represents Tuyuca, satisfies the 𝒟-nasal constraint, because its words, [aka] and [ãkã], are sufficiently distinct with respect to nasality. Thus, it is possible to derive the seemingly opaque behavior of neutral segments in a harmony system in a framework with direct mapping between the input and output, using the independently motivated family of 𝒟-constraints.

5.4.4 Excursus into harmony systems in FDM-OT

Clearly, this is not all that needs to be said about the analysis of harmony in FDM-OT. However, the basic idea, distinguishing transparent articulatory harmony from (potentially) opaque perceptual harmony, has been set forth and shown to work for a highly idealized case. Harmonies with no neutral segments can be motivated by articulation, using ℳ-constraints such as ℳ-□□nasal to force continuous harmony. Harmonies with neutral segments, such as in Tuyuca, can be motivated by perception, using 𝒟-constraints to force pairs of words to be perceptually distinct with respect to the harmonizing property.

There are many questions that need to be answered to fully realize this kind of analysis of harmony. How do the 𝒟-constraints needed for harmony (in which differences are measured by the number of segments) fit into the families of 𝒟-constraints that were developed in the previous chapters (in which differences are measured between segments along a scale of similarity)? How can these two types of 𝒟-constraints be reconciled with each other, whether they are different or the same?

Other than harmony systems, can we find other uses of these 'new' $\mathcal{D}$-constraints? Is there evidence (psycholinguistic, morphological, diachronic, etc.), beyond the mere existence of neutral segments, that there is a difference between articulatory harmony and perceptual harmony? I believe this line of research is very promising and deserves significant work in the future to answer the questions it raises.

5.5 Other proposals for analyzing opacity

Numerous proposals have been offered that modify basic OT in such a way that opacity can be accounted for synchronically. With that goal in mind, many of these approaches have been successful, creating OT-based frameworks which are capable of generating synchronic opaque interactions. However, as I have argued in this dissertation, there is no need to account for synchronic opacity because any generalization which appears to be synchronically opaque is (i) not synchronically productive, (ii) only productive at certain morphemic boundaries, or (iii) can be reanalyzed transparently (perhaps with $\mathcal{D}$-constraints). Consequently, any theory which allows general synchronic opacity would appear to be too powerful, and thus, needlessly complex and abstract. In comparison, the level of abstractness that FDM-OT does have (most obviously, by having inputs and candidates be sets of words instead of individual words as in standard OT) is necessary, independent of an analysis of opacity (for example, see §3.5 of Chapter 3, in which I show that it is impossible to analyze both the Old Polish merger and the Middle Polish split of the nasal vowels without sets for inputs and candidates).

In the remainder of this section, I discuss specific notable frameworks from the literature which have been used to incorporate synchronic opacity into OT. In the discussion, I note how each analysis increases abstractness (and how such an increase compares to that for FDM-OT), and what benefits (beyond a synchronic account of opacity) each analysis brings with it.

5.5.1 Constraint conjunction

One theoretical tool that has been used to provide analyses of opacity (among other phenomena) is *(local) constraint conjunction*, in which the special constraint $C_1\&C_2$ is violated exactly when the constraints C_1 and C_2 are both violated (in the same domain). Constraint conjunction was proposed by Smolensky (1993) and has been applied by numerous researchers, including Kirchner (1996) for chain shifts, Łubowicz (1998) for derived environment effects, Alderete (1997a) and Ito and Mester (1998) for dissimilation, Sanders (1999) for restricting application of truncation, Ito and Mester (2003) for opacity in German codas, and many others.

Constraint conjunction functions by subverting the principle of strict dominance within classic OT (Prince and Smolensky 1993/2002). Normally, if some constraint A dominates both C_1 and C_2, then a candidate that satisfies A (26a) will defeat a candidate that violates A (26b–d), regardless of how many violations of C_1 and C_2 it incurs, all else being equal. That is, constraint ranking is more important than number of violations; multiple violations of low-ranked constraints cannot overpower a high-ranked constraint:

(26)

input	A	C_1	C_2
✓ a. *cand*$_1$		×	×
b. *cand*$_2$	×!	×	
c. *cand*$_3$	×!		×
d. *cand*$_4$	×!		

However, if the conjoined constraint $C_1\&C_2$ outranks A, then violations of the lower constraints *do* matter and can lead to the elimination of a candidate that would not ordinarily lose under strict dominance. In this example, $cand_1$ (27a) violates $C_1\&C_2$ because it violates both C_1 and C_2. None of the other candidates (27b–c) violate both, which means they all satisfy $C_1\&C_2$, so (27a) loses to all of them:

(27)

input	$C_1\&C_2$	A	C_1	C_2
a. $cand_1$	×!		×	×
b. $cand_2$		×	×!	
c. $cand_3$		×		×!
✓ d. $cand_4$		×		

One of the fundamental problems with constraint conjunction is specifying exactly what types of conjunction are allowed: what constraints can and cannot be conjoined, which domains they can/must be confined to, etc. These many problems lead Padgett (2002a) to argue that conjunction itself may not be a necessary component of the grammar at all. Rather, he attributes the epiphenomenon of conjoined constraints to a projection of a universal scale of difficulty (for example, a coronal place of articulation is easier than dorsal or labial) to a grounded constraint subhierarchy (in this case, ×Dorsal, ×Labial ≫ ×Coronal; cf. universal rankings of $\mathcal{M}$-constraints in FDM-OT or universal subhierarchies in Prince and Smolensky 1993/2002). Padgett explores many examples of constraint conjunction and shows how they can be reduced similar grounded constraint subhierarchy. However, he explicitly does not address the use of constraint conjunction in analyzing opacity.

If Padgett is correct, there is no motivation for true constraint conjunction, except perhaps to analyze opacity. But as I have argued in this dissertation, opacity need not be given a synchronic analysis and that any analytic tool which is motivated solely by a drive to create synchronic opacity is superfluous. Thus, it would seem that constraint conjunction is an unnecessary modification to OT (which favors an interpretation of $\mathcal{F}^2$-constraints (Chapter 4, §4.6.3) based on multi-valued properties, along the lines of Gnanadesikan 1997).

5.5.2 Multistratality and Sympathy

Because opacity requires intermediate representations, a logical way to modify OT to allow opacity is to incorporate intermediate representations directly into the theory. Inkelas and Orgun (1995) argue for a general multistratal OT framework, in which each serially ordered stratum is a constraint hierarchy which takes the output of the previous hierarchy as its input:

(28) *multistratal evaluation in OT*

$\text{Eval}_1(I_1) = O_1$	first stratum; I_1 is the underlying representation
$\text{Eval}_2(O_1) = O_2$	second stratum; previous output is the current input
$\vdots$	
$\text{Eval}_n(O_{n-1}) = O_n$	nth stratum

In principle, each stratum could have a different constraint hierarchy, so that $\text{Eval}_i \neq \text{Eval}_j$ for some i and j. This general model is far too powerful and quite likely impossible for a human child to learn (Kager 199a:385). However, it serves as a good starting point for a discussion of multistratal OT frameworks.

A more restricted multistratal framework is Kiparsky's (1998) Lexical Phonology and Morphology in OT (LPM-OT), which limits the number of possible strata to principled morphological constructs but still allows for potentially different constraint hierarchies in each stratum, with the understanding that a grammar with differences in constraint rankings between strata is more marked than a grammar with identical constraint rankings in every strata. LPM-OT has some attractive properties which make it worth consideration. Most notably for the present discussion, LPM-OT distinguishes between the lexical processes that affect words and the postlexical processes that affect phrases. Here we encounter a situation in which there is clear motivation for added complexity to the theory: words and phrases are different entities, and it seems reasonable that they could be treated differently by the grammar (though by no means is the grammar required to do so). Notable work supporting a lexical/postlexical bifurcation in OT include Bermúdez-Otero 1999, Ito and Mester 2001, 2002, Herrick 2001. I have not yet determined whether a postlexical stratum is actually required in FDM-OT or if there is any type of synchronically productive opacity which would arise from the existence of such a stratum that could not be accounted for with direct mapping. Postlexical opacity in FDM-OT is clearly an area which needs further study but is beyond the scope of this dissertation.

In addition to a distinction between lexical and postlexical strata, LPM-OT also includes a division within the lexical phonology, with various strata allowed to apply within words (at minimum, a stratum for uninflected stems and a stratum for fully inflected words). This type of framework creates a situation in which opacity within words easily arises. As I have shown in this dissertation, allowing opacity does not seem to be necessary, so any framework which allows opacity should have independent justification. However, it is not clear that there is any such justification for within-word strata. If the justification is based on phonological processes, either they are transparent and can be analyzed by direct mapping, or they are opaque and the 'independent' justification is not really independent at all.

Another extension to OT is Sympathy (first proposed in McCarthy 1997, with modifications in Ito and Mester 1997a, Karvonen and Sherman 1997, Walker 1998, and McCarthy 1999, among others; Sanders 1997 is a formal overview of the types of Sympathy that are allowed). Sympathy is arguably a multistratal framework, which is so restricted that at first it does not even appear to be multistratal. While proponents of Sympathy generally deny its multistratality, I argue that Sympathy is clearly multistratal, because it is equivalent to a framework with two strata.[12]

In Sympathy, one of the candidates created by Gen is selected to be the *sympathy candidate* (indicated by the symbol ❀), which is allowed to shape the ultimate output via special sympathetic faithfulness constraints between the sympathy candidate and the output (called ❀O-FAITH here). The sympathy candidate is purely abstract: it need not be identical to the input, output, or any other actual word in the language (which means that Sympathy is not easily compatible with FDM-OT); it is merely the candidate which best satisfies a specified *selector constraint* (marked by a superscript ❀), and if necessary, the rest of the hierarchy. The following tableau gives a simplified sample Sympathy analysis for the [ɔ]~[u] alternation in Polish:

[12] Sympathy could potentially be modified so that candidates are candidate pairs, with one member of the pair evaluated as a possible sympathy candidate and the other member evaluated as a possible output. To my knowledge, Sympathy has yet to be formalized this way, so it is not clear that such a monostratal version of Sympathy would produce the same results as the current multistratal version.

(29)

/lɔd/	×ɔd#	×d#	❀O-FAITH [high]	FAITH [high]	FAITH [voice]❀
a. lɔd	×!	×!	×!		
❀ b. lud		×!		×	
c. lɔt			×!		×
✓ d. lut				×	×

The fully faithful candidate (29a) and the sympathy candidate (29b) both best satisfy the selector constraint FAITH-[voice]. Of the two, (29a) violates the highly ranked markedness constraints, disqualifying it as the sympathy candidate. Thus, (29b) is selected as the sympathy candidate and is the form that the output must be faithful to in order to satisfy ❀O-FAITH. High-ranking markedness prevents (29a) and (29b) from being the actual output, so we see that the sympathy candidate is not 'real', in the sense that it is not an input and is not an output. Of the remaining two candidates, transparent (29c) (which would be the selected output if Sympathy were not involved) is not faithful to the vowel height of the sympathy candidate, and is ruled out by ❀O-FAITH-[high]. This leaves opaque (29d) as the selected winner.

The choice of sympathy candidate is dependent on the input, since it is selected from the set of candidates built from the input by Gen. Every input has its own sympathy candidate which is selected by means of a constraint hierarchy that differs from the 'normal' hierarchy in some specific way. In terms of strata, the sympathy candidate can be considered the output of the first stratum and the input to the second stratum, and thus is an intermediate representation. This fits nicely with the asymmetric nature of the sympathy candidate: it shapes the output, but is itself not affected by the output (if Sympathy were truly parallel, we would expect winning outputs to influence the choice of the sympathy candidate, but this is not possible in Sympathy). This is precisely the kind of effect that a multistratal model is designed to produce, and so it seems natural to reanalyze Sympathy in a multistratal way, as follows:

(30) *multistratal characterization of Sympathy*

$\text{Eval}_1(I) = S$	first stratum; *S* is the sympathy candidate
$\text{Eval}_2(I,S) = O$	second stratum

The first stratum involves a constraint hierarchy with the selector constraint undominated. For the example in (29), this means the first stratum would look like the following, with the selector constraint FAITH-[voice] ranked at the top of the hierarchy (cf. Walker's (1998) Harmonic Sympathy):

(31)

/lɔd/	FAITH [voice]	×ɔd#	×d#	FAITH [high]
a. lɔd		×!	×	
✓ b. lud			×	×
c. lɔt	×!			
d. lut	×!			×

The winner (30b) is the sympathy candidate and becomes the input to the next stratum, where FAITH-[voice] is demoted below FAITH-[high]:

(32)

lud	×ɔd#	×d#	FAITH [high]	FAITH [voice]
a. lɔd	×!	×!	×	
b. lud		×!		
c. lɔt			×!	×
✓ d. lut				×

In principle, the ranking of the selector constraint is the only difference between $Eval_1$, where it is undominated, and $Eval_2$, where it is ranked normally. However, the unrestricted potential for multiple different selector constraints essentially eliminates any requirement on ranking conservation between strata. Regardless of how it is viewed (as monostratal or multistratal), Sympathy is designed solely to account for opacity. There is no external evidence (e.g. from language acquisition or speech processing) to independently motivate Sympathy.[13]

Besides lack of independent motivation, a problem facing multistratal OT frameworks is over-generation of unattested language types, the same problem encountered by pre-OT serial frameworks which OT was designed in part to solve. As Kager (1999a:385) notes, it would theoretically be possible for each stratum to have completely different rankings, but generally, multistratal analyses need only minor changes in constraint ranking between strata to account for even the most complex phenomena. There are no clear reasons why this should be the case,

[13] In addition, Sympathy traditionally requires faithfulness between each stratum and the original input. This transstratal faithfulness adds another layer of unnecessary power beyond ordinary multistratal frameworks. Sympathy is so powerful, in fact, that it cannot easily be described within model theory in a way consistent with the rest of OT (Potts and Pullum 2002).

certainly not within the theories themselves, and there seems to be no means of ensuring minimal reranking without ad hoc stipulation.

A key insight of multistratal frameworks like LPM-OT is that the shape of a morphologically complex word can depend on the shape of its components. This insight, discussed in the next section, has been incorporated into OT as faithfulness between morphologically related outputs and requires no abstract multistratality.

5.5.3 Output-output faithfulness and paradigm uniformity

An important research direction for OT, exemplified by influential works such as Benua 1995, 1997, Burzio 1996, Kenstowicz 1996, and Steriade 1996, expanded the concept of faithfulness to govern identity between morphologically related outputs. This line of research is based on the observation that related outputs seem to influence each other. This is formalized by means of output-output faithfulness constraints, symbolized here as $\mathcal{F}_{OO}$-constraints (the terminology for this concept varies from proposal to proposal, but the core idea is the same). These constraints behave like standard (input-output) $\mathcal{F}$-constraints, but instead of enforcing identity between the input and output, $\mathcal{F}_{OO}$-constraints enforce identity between two outputs which are morphologically related to each other.

While not explicitly designed to account for opacity, $\mathcal{F}_{OO}$-constraints can be used to create the serialism required for opacity to arise under the right circumstances, by using one output as the intermediate form. In Palestinian Arabic (Brame 1974), there is a general process of syncope which deletes [i] in non-final unstressed open syllables, triggered by a constraint ranking such as $\mathcal{M}$-i]$_{\square}$ $\gg$ $\mathcal{F}$-segment. However, syncope is opaquely blocked in forms like [f**i**ˈhimna] ‘he understood us’. The boldfaced [i] is in the correct environment for syncope, so it should be deleted if syncope applies blindly to all surface forms. However, there is a related form [ˈfihim] ‘he understood’ which syncope does not apply to. This form can act as the intermediate form for opacity since [fiˈhimna] is derived from it. By ranking $\mathcal{F}_{OO}$-segment over $\mathcal{M}$-i]$_{\square}$, the relevant [i] in the base form [ˈfihim] is guaranteed to survive in the derived form [fiˈhimna], despite syncope. In contrast, there is syncope in [ˈfhimna] ‘we understood’ because it is not derived from [ˈfihim], so $\mathcal{F}_{OO}$-segment plays no role. See Kenstowicz 1996, Steriade 1996, and Kager 1999a for more detailed analyses of this data, and Kiparsky 1998 for an analysis within LPM-OT).

Since candidates in FDM-OT are sets of words containing all related forms for any particular word, $\mathcal{F}_{OO}$-constraints are compatible with FDM-OT and would not introduce any additional abstractness. The required related output forms are already available in every candidate, so $\mathcal{F}_{OO}$-constraints need not reference anything beyond what is already present in FDM-OT candidates (as Potts and Pullum (2002) note, FDM-OT’s sets-as-candidates can avoid the problems that crop up when trying to find a model-theoretic interpretation of conventional OT $\mathcal{F}_{OO}$-constraints). Despite the compatibility, it is not entirely clear to me that $\mathcal{F}_{OO}$-constraints are needed in FDM-OT because inputs in FDM-OT are subject to strong lexicon optimization. As a result, inputs generally look exactly like outputs; any information needed from any output form can be found in the lexicon. This means that standard $\mathcal{F}$-constraints could be capable subsuming the power of $\mathcal{F}_{OO}$-constraints. It remains to be seen if $\mathcal{F}_{OO}$-constraints could be used in FDM-OT in a meaningfully different way from $\mathcal{F}$-constraints. If indeed they are required, $\mathcal{F}_{OO}$-constraints can be folded easily into FDM-OT, as there would be no substantial increase to the level of abstractness in the theory, beyond the addition of an extra family of constraints.

5.5.4 Other frameworks

There are many other frameworks that can handle opacity which I do not have space to discuss in detail. Containment Theory (the original PARSE/FILL model of faithfulness presented in Prince and Smolensky 1993/2002) and turbidity (Goldrick 1998, 1999, Goldrick and Smolensky 1999) bypass the need for intermediate representations by encoding unpronounced information in the output, allowing it to be accessed by the constraint hierarchy. Wilson's (2000) targeted constraints modify constraints and constraint evaluation to limit violations to certain subsets of candidates, somewhat reminiscent of both Sympathy and constraint conjunction. Contrast Preservation Theory (Łubowicz 2001) is discussed and argued against in §2.6 of Chapter 2. The list of modifications to OT designed to account for opacity is continually growing; some are flashes in the pan, while others manage to secure a devoted following. However, I have argued in this dissertation that there is no true synchronic, monomorphemic opacity. Thus, if a framework is motivated solely by a need to account for such opacity, with no independent support, it is superfluous. Additionally, even if not solely motivated by opacity, if a framework even allows synchronically productive opacity, it is likely too powerful and should be restricted.

5.6 Conclusion

If anything should be taken away from a reading of this chapter, it is that an apparent case of phonological opacity need not be analyzed as such. Many of the examples of opacity in the literature can be translated into one of the types analyzed in this chapter. I suspect that a large number of these cases only appear to be productive simply because of the proportion of the lexicon that is affected, but in fact, are not actually synchronically productive. Treating these cases as productive ends up unnecessarily encoding the language's history in the synchronic grammar. Upon experimentation with nonce forms and a more thorough search for lexical exceptions, they will probably turn out to be unproductive, as seen for refined Low German in §5.2 and for Polish in Chapters 2 and 3, and thus can be analyzed using strong lexicon optimization (in FDM-OT or another compatible framework).

Of course, it is possible that an instance of opacity will actually be fully productive, as with opacity due to morphology such as in Turkish (§5.3) and in Polish (Chapter 4), or with neutral segments in harmony systems such as in Tuyuca (§5.4). Such cases need careful consideration. Does the phenomenon at hand only arise at certain morphological boundaries? Or might it be motivated by dispersion of contrast or some other overlooked surface property? If either of the questions are answered affirmatively, then it is likely that the types of FDM-OT analyses sketched in §5.3 and 5.4 can be extended. While I hope that this is the final word on opacity, it is possible that there are holes remaining in the FDM-OT framework developed in this dissertation. These holes consist of 'true' opacity, which:

- is fully synchronically productive, and thus applies to nonce forms and does not have (real or potential) lexical exceptions;
- applies to all forms regardless of the presence or lack of particular morpheme boundaries; and
- cannot be explained by transparently appealing to surface properties of a word or of any contrastive or morphologically related surface forms.

I have yet to find a case of opacity that unquestionably satisfies all of these requirements. The search for true opacity is an important line of research in and of itself, and if true opacity can be found, it would provide an interesting challenge to FDM-OT.

References

Alderete, John. 1997a. Dissimilation as local conjunction. In Kiyomi Kusumoto, ed. *Proceedings of the Northeast Linguistic Society 27*. Amherst, MA: GLSA. 17–32. [ROA 175]

Alderete, John. 1997b. Root-controlled accent in Cupeño. Manuscript, University of Massachusetts, Amherst. [ROA 253]

Archangeli, Diana, and Douglas Pulleyblank. 1994. *Grounded phonology*. Cambridge, MA: MIT Press.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Baranowski, Maciej, and Eugene Buckley. 2003. Lexicalization and analogy in Polish o-Raising. Paper presented at the annual meeting of the Linguistic Society of America, Atlanta, GA.

Barnes, Janet. 1996. Autosegments with three-way contrasts in Tuyuca. *International Journal of American Linguistics*. 62:31–58.

Barnes, Janet, and Sheryl Takagi de Silzer. 1976. Fonolgía del tuyuca. *Sistemas fonológicos de idiomas colombianos* 3. Bogotá, Colombia: Ministerio de Gobierno. 123–137.

Beckman, Jill N. 1995. Shona height harmony: Markedness and positional identity. In Beckman, et al. 53–75.

Beckman, Jill N. 1997. Positional faithfulness, positional neutralization, and Shona height harmony. *Phonology* 14:1–46.

Beckman, Jill N. 1998. *Positional faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA 234]

Beckman, Jill N., Laura Walsh Dickey, and Suzanne Urbanczyk, eds. 1995. *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*. Amherst, MA: GLSA.

Beddor, Patrice Speeter. 1993. Perception of nasal vowels. In Marie K. Huffman and Rena A. Krakow, eds. 1993. *Nasals, nasalization, and the velum*. Phonetics and phonology, vol. 5. San Diego: Academic Press.. 171–196.

Benua, Laura. 1995. Identity effects in morphological truncation. In Beckman, et al. 1995. 77–136. [ROA 74]

Benua, Laura. 1997. *Transderivational identity: Phonological relations between words*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA 259]

Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change [Opacity and globality in phonological change]. Doctoral dissertation, University of Manchester and Universidad de Santiago de Compostela.

Bethin, Christina Y. 1992. *Polish syllables: The role of prosody in phonology and morphology*. Columbus, OH: Slavica.

Bhat, D. N. S. 1974. Retroflexion and retraction. *Journal of Phonetics* 2:233–237.

Bidwell, Charles, E. 1963. *Slavic historical phonology in tabular form*. The Hague: Mouton.

Bielec, Dana. 1998. *Polish: An essential grammar*. London: Routledge.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul, and David Weenink. *Praat: Doing phonetics by computer*. 1992/2003. [`www.praat.org`]

Bond, Z. S. 1975. Identification of vowels excerpted from neutral and nasal contexts. *Journal of the Acoustical Society of America* 59:1229–1232.

Borg, Alexander. 1997. Maltese phonology. In Alan S. Kaye, ed. *Phonologies of Asia and Africa*. Winona Lake, IN: Eisenbrauns. 245–285.

Borowsky, Toni. 1986. *Topics in the lexical phonology of English*. Doctoral dissertation, University of Massachusetts, Amherst.

Brame, Michael. 1972. On the abstractness of phonology: Maltese 9. In Michael Brame, ed. *Contributions to generative phonology*. Austin: University of Texas Press. 22–61.

Brame, Michael. 1974. The cycle in phonology: Stress in Palestinian, Maltese, and Spanish. *Linguistic Inquiry* 5:39–60.

Brown, R., and D. McNeill. 1966. The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5:325–337.

Bruening, Benjamin. 2000. Morphological templates and phonological opacity. Handout of talk given at MIT.

Buckley, Eugene. 2001. Polish o-Raising and phonological explanation. Paper presented at the annual meeting of the Linguistic Society of America, Washington, DC.

Burzio, Luigi. 1996. Surface constraints versus underlying representations. In Durand and Laks 1996. 97–122.

Burzio, Luigi. 1998. Multiple correspondence. *Lingua* 104:79–109.

Butcher, A. 1976. The influence of the native language on the perception of vowel quality. *Arbeitsberichte* 6:1–137.

Bybee, Joan L. 1988. Morphology as lexical organization. In Michael Hammond and Michael Noonan, eds. *Theoretical morphology*. San Diego: Academic Press. 119–141.

Bybee, Joan L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10:425–455.

Caprile, Jean Pierre. 1968. Essai de la phonologie d'un parler mbay: Les emprunts arabes en Mbay. *Bulletin de la Société pour l'Étude des Langues Africaines* 8.

Carbonell, Joan F., and Joaquim Llisterri. 1999. Catalan. In IPA 1999. 61–65.

Carlson, Rolf, Björn Granström, and Gunnar Fant. 1970. Some studies concerning perception of isolated vowels. *Speech Transmission Laboratory Quarterly Progress Status Report* 2/3. Royal Institute of Technology, Stockholm. 19–35.

Carlton, Terence R. 1991. *Introduction to the phonological history of the Slavic languages*. Columbus, OH: Slavica.

Catford, John C. 1977. *Fundamental problems in phonetics*. Edinburgh: University Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Cambridge, MA: MIT Press.

Chung, Sandra. 1983. Transderivational relationships in Chamorro phonology. *Language* 59:35–65.

Čiževskij, Dmitrij. 1971. *Comparative history of Slavic literatures*. Noel Richard Porter, and Martin P. Rice, trans. Serge A. Zenkovsky, ed. Vanderbilt University Press.

Clements, George N. 1991. Vowel height assimilation in Bantu languages. *Working papers of the Cornell Phonetics Laboratory* 5:37–76.

Cutler, Ann. 1986. Phonological structure in speech recognition. *Phonology Yearbook* 3: 161–178.

Czaykowska-Higgins, Ewa. 1988. *Investigations into Polish morphology and phonology*. Doctoral dissertation, Massachusetts Institute of Technology.

de Bray, R. G. A. 1969. *Guide to the Slavonic languages*. London: J. M. Dent and Sons.

de Bray, R. G. A. 1980. *Guide to the West Slavonic languages*. Guide to the Slavonic languages, Part 2. 3rd ed. London: J. M. Dent and Sons.

de Saussure, Ferdinand. 1959. *Course in general linguistics*. New York: McGraw-Hill.

Delattre, Pierre. 1971. Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English. *Phonetica* 23:129–155.

Dogil, G. 1990. Hissing and hushing fricatives: A comment on non-anterior spirants in Polish. Manuscript, Stuttgart University.

Donegan, Patricia. 1978. *The natural phonology of vowels*. Doctoral dissertation, Ohio State University.

Durand, Jacques, and Bernand Laks, eds. 1996. *Current trends in phonology: Models and methods*. Salford, UK: European Studies Research Institute, University of Salford Publications.

Entwistle, William J., and W. A. Morison. 1949. *Russian and the Slavonic languages*. London: Faber and Faber.

Fay, D., and Ann Cutler. 1977. Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8:505–520.

Flemming, Edward. 1995. *Auditory representations in phonology*. Doctoral dissertation, University of California, Los Angeles.

Fromkin, Victoria A., ed. 1973. *Speech errors as linguistic evidence*. The Hague: Mouton.

Gess, Randall. 1996. *Optimality Theory in the diachronic phonology of French*. Doctoral dissertation, University of Washington.

Glass, Gene V., and Kenneth D. Hopkins. 1996. *Statistical methods in education and psychology*. 3rd ed. Boston: Allyn & Bacon.

Gnandesikan, Amalia E. 1997. *Phonology with ternary scales*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA 195]

Goldrick, Matthew. 1998. Optimal opacity: Covert structure in phonology. Manuscript, Johns Hopkins University.

Goldrick, Matthew. 1999. Turbid output representations and the unity of opacity. In M. Hirotani, A. Coetzee, N. Hall, and J.-Y. Kim, eds. *Proceedings of the Northeast Linguistics Society* 30. vol. 1. Amherst, MA: GLSA. 231–245. [ROA 368]

Goldrick, Matthew, and Paul Smolensky. 1999. Opacity and turbid representations in Optimality Theory. Paper presented at the 35th meeting of the Chicago Linguistics Society, University of Chicago.

Goldsmith, John. 1996. Tone in Mituku: How a floating down nailed down an intermediate level. In Durand and Laks 1999. 267–280.

Gotteri, Nigel. Polish. 1998. In Glanville Price, ed. 1998. *Encyclopedia of the languages of Europe*. Oxford: Blackwell. 360–364.

Gussmann, Edmund. 1980. *Studies in abstract phonology*. Cambridge, MA: MIT Press.

Guthrie, M. 1968. Notes on Nzɛbi. *Journal of African Linguistics* 7.2:101–129.

Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.

Hamann, Silke. 2002a. Postalveolar fricatives in Slavic languages as retroflexes. In S. Baauw, M. Huiskes, and M. Schoorlemmer, eds. *OTS Yearbook 2002*. Utrecht Institute of Linguistics. 105–127.

Hamann, Silke. 2002b. Retroflexion and retraction revised. In T. A. Hall, B. Pompino-Marschall, and M. Rochoń, eds. *ZAS Paper in Linguistics* 28:3–26.

Hamilton, William S. 1980. *Introduction to Russian phonology and word structure*. Columbus: Slavica.

Hayward, Katrina, and Richard J. Hayward. 1999. Amharic. In IPA 1999. 45–50.

Herrick, Dylan. 2001. Levels and edges: A study of Catalan place assimilation and cluster simplification. Paper presented at the Trilateral Weekend, Stanford University.

Holt, Eric. 1997. *The role of the listener in the historical phonology of Spanish and Portuguese: An Optimality-Theoretic account*. Doctoral dissertation, Georgetown University. [ROA 278]

Hooper [Bybee], Joan. 1976. *An introduction to Natural Generative Phonology*. New York: Academic Press.

Hualde, José. 1991. *Basque phonology*. London: Routledge.

Hudson, Grover. 1975. *Suppletion in the representations of alternations*. Doctoral dissertation, University of California, Los Angeles.

Idsardi, William. 1998. Tiberian Hebrew spirantization and phonological derivations. *Linguistic Inquiry* 29:37–73.

Inkelas, Sharon, and C. Orhan Orgun. 1995. Level ordering and economy in the lexical phonology of Turkish. *Language* 71:763–793.

Inkelas, Sharon, and Cheryl Zoll. 2000. Reduplication as morphological doubling. Manuscript, University of California, Berkeley, and Massachusetts Institute of Technology. [ROA 412]

IPA [International Phonetic Association], ed. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

Ito, Junko, and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Beckman, et al. 181–209.

Ito, Junko, and Armin Mester. 1997a. Featural sympathy: Feeding and counterfeeding interactions in Japanese. In Walker, et al. 1997. 29–36.

Ito, Junko, and Armin Mester. 1997b. Sympathy theory and German truncation. In Viola Miglio and Bruce Moréen, eds. *Selected phonology papers from Hopkins Optimality Theory Workshop 1997/University of Maryland Mayfest* 1997. University of Maryland working papers in linguistics 5. 117–139. [ROA 211]

Ito, Junko, and Armin Mester. 1998. Markedness and word structure: OCP effects in Japanese. Manuscript, University of California, Santa Cruz. [ROA 255]

Ito, Junko, and Armin Mester. 1999. The structure of the phonological lexicon. In Natsuko Tsujimura, ed. *The handbook of Japanese linguistics*. Malden, MA: Blackwell Publishers. 62–100.

Ito, Junko, and Armin Mester. 2001. Structure preservation and stratal opacity in German. In Linda Lombardi, ed. *Segmental phonology in Optimality Theory: constraints and representations*. Cambridge University Press.

Ito, Junko, and Armin Mester. 2002. Lexical and postlexical phonology in Optimality Theory. In Gisbert Fanselow and Caroline Féry, eds. *Linguistische Berichte Sonderheft* 11.

Ito, Junko, and Armin Mester. 2003. On the sources of opacity in OT: Coda processes in German. In Caroline Féry and Ruben van de Vijver, eds. *The syllable in Optimality Theory*. Cambridge University Press. 271–303.

Jacobs, Haike. 1995 Optimality Theory and sound change. In Jill N. Beckman, ed. *Proceedings of the Northeast Linguistic Society* 25. vol. 2. Amherst, MA: GLSA. 219–232. [ROA 129]

Janecki, Klara. 2000. *301 Polish verbs*. Hauppauge, NY: Barron's Educational Series.

Jaślan, Janina, and Jan Stanisławski. 1993. *Wiedza powszechna compact Polish and English dictionary*. Izabella Jastrzębska-okoń and Katarzyna Billip, eds. Lincolnwood, IL: NTC Publishing Group.

Jensen, John Thayer. 1977. *Yapese reference grammar*. Honolulu: University Press of Hawaii.

Jungraithmayr, Herrmann. 1971. The Tangale vowel system reconsidered. *Journal of African Linguistics* 10:28–33.

Kager, René. 1996. On affix allomorphy and syllable counting. In U. Kleinhenz, ed. *Interfaces in phonology*. Studia Grammatica 41. Berlin: Akademie Verlag. 155–171.

Kager, René. 1999a. *Optimality Theory*. Cambridge University Press.

Kager, René. 1999b. Ternary alternations and lexical allomorphy. Paper presented at the Northeast Linguistic Society 30. Rutgers University.

Kang, Seok-keun. 2002. An optimality theoretic account of phonological opacity in English. *Studies in phonetics, phonology, and morphology* 6.2:307–332.

Karvonen, Daniel, and Adam Sherman [Ussishkin]. 1997. Sympathy, opacity, and u-umlaut in Icelandic. In Walker, et al. 1997. 37–48.

Kaye, Jonathan. 1974. Opacity and recoverability in phonology. *Canadian Journal of Linguistics* 19:134–149.

Kaye, Alan S. 1990. Arabic. In Bernard Comrie, ed. *The world's major languages*. New York: Oxford University Press. 664–685.

Keating, Patricia. 1987. A survey of phonological features. *UCLA Working Papers in Phonetics* 66. Bloomington: Indiana University Linguistics Club. 124–150.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, MA: MIT press.

Kenstowicz, Michael. 1996. Base-identity and uniform exponence. In Durand and Laks 1996. 363–393. [ROA 103]

Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology: Description and theory*. Orlando, FL: Academic Press.

Kidda, Mairo. 1985. *Tangale phonology: A descriptive analysis*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In Emmon Bach and Robert T. Harms, eds. *Universals in linguistic theory*. New York: Holt. 171–202.

Kiparsky, Paul. 1971. Historical linguistics. In Osamu Fujimura, ed. *A survey of linguistic science*. College Park: University of Maryland Linguistics Program. 576–642.

Kiparsky, Paul. 1981. Vowel harmony. Manuscript, Massachusetts Institute of Technology.

Kiparsky, Paul. 1998. Paradigm effects and opacity. Manuscript, Stanford University.

Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27:341–350.

Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1:291–306.

Klatt, Edmund, Dietrich Roy, Gisela Klatt, and Heinz Messinger. 1983. *Langenscheidt's standard German dictionary*. New York: Langenscheidt.

Kuroda, Sige-Yuki. 1967. *Yawelmani phonology*. Cambridge, MA: MIT Press.

Ladefoged, Peter. 2001. *A course in phonetics*. 4th ed. Boston: Heinle and Heinle.

Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Ladefoged, Peter, and Anthony Traill. 1994. Clicks and their accompaniments. *Journal of Phonetics* 22:33–64.

Lee, Hyun Bok. 1999. Korean. In IPA 1999. 120–123.

Lee-Schoenfeld, Vera. 2002. The elimination of opacity in German truncations. Manuscript, University of California, Santa Cruz.

Levelt, W. J. M., 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lewis, Geoffrey L. 1967. *Turkish grammar*. Oxford University Press.

Lindau, Mona. 1978. Vowel features. *Language* 54:541–563.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. In John J. Ohala and Jeri J. Jaeger, eds. 1986. *Experimental phonology*. Orlando, FL: Academic Press. 13–44.

Lindblom, Björn. 1990. On the notion of 'possible speech sound'. *Journal of Phonetics* 18:135–152.

Lockwood, W. B. 1976. *An informal history of the German language*. London: André Deutsch Limited.

Lombardi, Linda. 1991. *Laryngeal features and laryngeal neutralization*. Doctoral dissertation, University of Massachusetts, Amherst.

Łubowicz, Anna. 1998. Derived environment effects in OT. Manuscript, University of Massachusetts, Amherst. [ROA 239]

Łubowicz, Anna. 2001. Opacity, recoverability, and constraints on contrast. Paper presented at Generative Linguistics in Poland 3, Uniwersytet Warszawski [University of Warsaw].

Martinet, André. 1964. *Elements of general linguistics*. University of Chicago Press.

McCarthy, John. 1997. Sympathy and phonological opacity. Paper presented at the Hopkins Optimality Theory Workshop/University of Maryland Mayfest, Johns Hopkins University.

McCarthy, John. 1999. Sympathy and phonological opacity. *Phonology*16:331–399.

McCarthy, John, and Alan Prince. 1993. Prosodic morphology I: Constraint interaction and satisfaction. Manuscript, University of Massachusetts, Amherst, and Rutgers University. [ROA 482]

McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In Beckman, et al. 1995. 249–384. [ROA 60]

Mester, Armin. 1994. The quantitative trochee in Latin. *Natural Language and Linguistic Theory* 12:1–61.

Mohanan, K. P. 1993. Fields of attraction in phonology. In Goldsmith, John, ed. *The last phonological rule*. University of Chicago Press. 61–116.

Mohr, B., and W. S.-Y. Wang. 1968. Perceptual distance and the specification of phonological features. *Phonetica* 18:31–45.

Newman, Stanley. 1944. *Yokuts language of California*. Viking Fund Publications in Anthropology 2. New York: Viking Fund.

Ní Chiosáin, Máire, and Jaye Padgett. 2001. Markedness, segment realization, and locality in spreading. In Linda Lombardi, ed. *Segmental phonology in Optimality Theory: constraints and representations*. Cambridge University Press. 118–156. [ROA 503]

Nowak, Paweł M. 2000. Palatalization in Polish: An Optimality-Theoretic approach. Manuscript, Harvard University.

Nowak, Paweł M. 2001. *Consonantal and vocalic features and the nature of agreement constraints: Evidence from Polish.* Masters thesis, Harvard University.

Ohala, John J., and Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In Marie K. Huffman and Rena A. Krakow, eds. 1993. *Nasals, nasalization, and the velum*. Phonetics and phonology, vol. 5. San Diego: Academic Press.. 225–249.

Orešnik, Janez. 1972. On the epenthesis rule in modern Icelandic. *Arkiv för Nordisk Filologi* 87. 1–32.

Padgett, Jaye. 1997. Perceptual distance of contrast: Vowel height and nasality. In Walker, et al. 1997. 63–78.

Padgett, Jaye. 2002a. Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory. Manuscript, University of California, Santa Cruz. [ROA 530]

Padgett, Jaye. 2002b. Russian voicing assimilation, final devoicing, and the problem of [v]. Manuscript, University of California, Santa Cruz. [ROA 528]

Padgett, Jaye. To appear a. Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory*.

Padgett, Jaye. To appear b. The emergence of contrastive palatalization in Russian. In Eric Holt, ed. *Optimality Theory and language change*. Kluwer Academic Press.

Pike, E. V., and K. Wistrand. 1974. Step-up terrace tone in Acatlan Mixtec. In R. M. Brend, ed. *Advances in tagmemics*. Amsterdam: North Holland. 81–104.

Pilch, Herbert. 1966. Das Lautsystem der hochdeutschen Umgangssprache. *Zeitshcrift für Mundartforschung* 33. 247–266.

Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.

Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper and Row.

Potts, Christopher, and Geoffrey K. Pullum. 2002. Model theory and the content of OT constraints. Manuscript, University of California, Santa Cruz.

Prince, Alan. 1975. *The phonology and morphology of Tiberian Hebrew*. Doctoral dissertation, Massachusetts Institute of Technology.

Prince, Alan, and Paul Smolensky. 1993/2002. Optimality Theory: Constraint interaction in generative grammar. Manuscript, Rutgers University and University of Colorado, Boulder. [ROA 537]

Ringen, Catherine. 1975. *Vowel harmony: Theoretical implications*. Doctoral dissertation, Indiana University.

ROA [Rutgers Optimality Archive]. [`http://roa.rutgers.edu`]

Robinson, Orrin W. 2001. *Whose German? The* ach/ich *alternation and related phenomena in 'standard' and 'colloquial'*. Amsterdam: John Benjamins Publishing Company.

Roca, Iggy, ed. 1997. *Derivations and constraints in phonology*. Oxford: Clarendon Press.

Rochoń, Marzena. 2001. Markedness of trills: Palatalization of r in Polish. Paper presented at Generative Linguistics in Poland 3, Uniwersytet Warszawski [University of Warsaw].

Rubach, Jerzy. 1984a. *Cyclic and lexical phonology: The structure of Polish*. Dordrecht: Foris.

Rubach, Jerzy. 1984b. Segmental rules of English and cyclic phonology. *Language* 60:21–54.

Rubach, Jerzy. 2000. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31:271–317.

Rubach, Jerzy. 2001. Polish iotation: Allomorphy in Optimality Theory. Paper presented at Generative Linguistics in Poland 3, Uniwersytet Warszawski [University of Warsaw].

Russ, Charles V. J. 1978. *Historical German Phonology and Morphology*. Oxford University Press.

Sanders, Nathan. 1997. On sympathetic correspondence. In Walker, et al. 91–102.

Sanders, Nathan. 1999. Intra-Representational Correspondence and truncation. Manuscript, University of California, Santa Cruz. [ROA 394]

Sanders, Nathan. 2001. Preserving synchronic parallelism: Diachrony and opacity in Polish. Paper presented at Generative Linguistics in Poland 3, Uniwersytet Warszawski [University of Warsaw].

Sanders, Nathan. 2002a. Dispersion in OT: Middle Polish nasal vowels. In Line Mikkelsen and Christopher Potts, eds. *WCCFL 21 proceedings*. Somerville, MA: Cascadilla Press. 415–428.

Sanders, Nathan. 2002b. Preserving synchronic parallelism: Diachrony and opacity in Polish. In Andronis, Mary, Christopher Ball, Heidi Elston, and Sylvain Neuvel, eds. *CLS: The Main Session. Papers from the 37th meeting of the Chicago Linguistics Society*, vol. 1. Chicago Linguistics Society. 501–516.

Saporta, Sol. 1965. Ordered rules, dialect differences and historical processes. *Language* 41:218–224.

Schuh, Russell G., and Lawan D. Yalwa. 1999. Hausa. In IPA 1999. 90–95.

Selkirk, Elisabeth O. 1984. On the major class features and syllable theory. In Mark Aronoff and Richard T. Oehrle, eds. *Language sound structure: Studies in phonology dedicated to Morris Halle by his teacher and students*. Cambridge, MA: MIT Press. 107–113.

Sezer. Engin. 1981. The k/Ø alternation in Turkish. In George Clements, ed. *Harvard studies in phonology* 2. Bloomington: Indiana University Linguistics Club. 354–382.

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at Rutgers Optimality Workshop I, Rutgers University. [ROA 87]

Spaelti, Philip. 1997. *Dimensions of variation in multi-pattern reduplication*. Doctoral dissertation, University of California, Santa Cruz. [ROA 311]

Spencer, Andrew. 1984. Eliminating the feature [lateral]. *Journal of Linguistics* 20:23–43.

Spencer, Andrew. 1986. A non-linear analysis of vowel-zero alternations in Polish. *Journal of Linguistics* 22:249–280.

Stanisławski, Jan. 1978. *Wielki słownik polsko-angielski [The great Polish-English dictionary]*. 2 vols. Warsaw: Wiedza Powszechna.

Steriade, Donca. 1995. Positional neutralization. Manuscript, University of California, Los Angeles.

Steriade, Donca. 1996. Paradigm uniformity and the phonetics-phonology boundary. Paper presented at the 5th Conference in Laboratory Phonology, Evanston, IL.

Stieber, Zdzisław. 1968. *The phonological development of Polish*. Michigan Slavic Materials 8. E. Schwartz, trans. Ann Arbor: University of Michigan.

Stieber, Zdzisław. 1973. *A historical phonology of the Polish language*. Historical phonology of the Slavic languages 5. George Shevelov, ed. Heidelberg: Carl Winter Universitätverlag.

Struijke, Caro. 1998. Reduplicant and output TETU in Kwakwala. In H. Fukazawa, F. Morelli, C. Struijke, and Y. Su, eds. *University of Maryland Working Papers* 7. 150–178. [ROA 261]

Struijke, Caro. 2001. Existential faithfulness: A study of reduplicative TETU, feature movement, and dissimilation. Doctoral dissertation, University of Maryland, College Park.

Šuštaršič, Rastislav, Smiljana Komar, and Bojan Petek. 1999. Slovene. In IPA 1999. 135–139.

Švarný, O., and K. Zvelebil. 1955. Some remarks on the articulation of the 'cerebral' consonants in Indian languages, especially Tamil. *Archiv Orientalni* 23:374–434.

Thelwall, Robin, and M. Akram Sa'adeddin. 1999. Arabic. In IPA 1999. 51–54.

Townsend, Charles E., and Laura A. Janda. 1996. *Common and comparative Slavic: Phonology and inflection with special attention to Russian, Polish, Czech, Serbo-Croatian, Bulgarian*. Columbus, OH: Slavica.

Traill, Anthony. 1994. *A !Xóõ dictionary*. Köln: Rüdiger Köppe.

Ullman, Michael T. 1993. *The computation of inflectional morphology*. Doctoral dissertation, Massachusetts Institute of Technology.

Ullman, Michael T. 1999. Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighborhood effects. *Language and Cognitive Processes* 14:47–67.

Urbanczyk, Suzanne. 1996. *Patterns of reduplication in Lushootsheed*. Doctoral dissertation, University of Massachusetts, Amherst,

van der Hulst, Harry, and van de Weijer. 1995. Vowel harmony. In John Goldsmith, ed. *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers, Ltd. 495–534.

Vennemann, Theo. 1974. Words and syllables in Natural Generative Grammar. In Anthony Bruck, Robert A. Fox, and Michael W. La Galy, eds. *Papers from the parasession on Natural Generative Phonology*. Chicago Linguistic Society. 346–374.

Viëtor, Wilhelm. 1904. *Elemente der Phonetik des Deutschen, Englischen und Französischen*. 5th ed. Leipzig: O. R. Reisland.

Walker, Rachel. 1998. *Nasalization, neutral segments, and opacity effects*. Doctoral dissertation, University of California, Santa Cruz. [ROA 405]

Walker, Rachel, Daniel Karvonen, and Motoko Katayama, eds. 1997. *Phonology at Santa Cruz* 5.

Waterman, John T. 1966. *A history of the German language*. Seattle: University of Washington Press.

Wells, Christopher J. 1985. *German: A linguistic history to 1945*. Oxford University Press.

Westfal, Stanisław. 1956. The *ę* to *ą* alternation in Modern Polish noun declension. *Slavonic and East European Review* 34:460–484.

Wiese, Richard. 1996. *The phonology of German*. Oxford University Press.

Williamson, Kay. 1977. Multivalued features for consonants. *Language* 53:843–871.

Wilson, Colin. 2000. *Targeted constraints: An approach to contextual neutralization in Optimality Theory*. Doctoral dissertation, Johns Hopkins University.

Wright, J. T. 1986. The behavior of nasalized vowels in the perceptual vowel space. In John J. Ohala and Jeri J. Jaeger, eds. 1986. *Experimental phonology*. Orlando, FL: Academic Press. 45–67.

Zhirmunskii, Viktor Maksimovch. 1962. *Deutsche Mundartkunde: Vergleichende Lautund Formenlehre derdeutschen Mundarten*. Berlin: Akademie-Verlag.

Zimmer, Karl, and Barbara Abbott. 1978. The k/Ø alternation in Turkish: Some experimental evidence for its productivity. *Journal of Psycholinguistic Research* 7:35–46.

Zimmer, Karl, and Orhan Orgun. 1999. Turkish. In IPA 1999. 154–156.

Zubritskaya, Katya. 1997. Mechanism of sound change in Optimality Theory. *Language Variation and Change* 9:121–148.