Strong Lexicon Optimization*

Nathan Sanders, nsanders@williams.edu
Williams College and University of Massachusetts, Amherst

Overview

As has been previously noted in numerous sources (Prince and Smolensky 2004, McCarthy and Prince 1993, Roca 1997, etc.), opacity[1] poses a problem for some versions of Optimality Theory (OT), prompting a proliferation of modifications to OT designed to solve this apparent problem. In this talk, I argue that, contrary to popular opinion, the basic architecture of OT does not need to be modified to account for opacity.

I begin by showing that given certain reasonable assumptions about generative phonology in general and about OT specifically, opacity is highly restricted in the productive phonological grammar (§1). I demonstrate how a minor modification—a strong interpretation of lexicon optimization—allows opacity to arise anyway and be encoded in the grammar (§2). I then explore further opaque phenomena from Polish present some problems that lead to minor changes in the initial conception of strong lexicon optimization (§3).

1 Should Opacity Exist?

Opacity typically arises due to certain types of interactions between two or more unrelated phonological processes. In a serial theory with ordered rules (Chomsky and Halle 1968, and their intellectual descendants), opacity is not only expected but welcomed, as it confirms the necessity of serial derivations. Cases of opacity were often embraced without sufficient questioning.

With the advent of OT, this mentality persisted, despite the conflict between opacity and OT's basic structure. In light of the radically different nature of OT from its predecessors, it is important to critically re-examine the role of opacity in phonology.

*This talk is an update of the core analysis of Sanders (2003, particularly Chapter 4). However, where Sanders (2003) uses a version of Dispersion Theory (Flemming 1995), this talk uses a more classic form of OT in order to focus on the issue of strong lexicon optimization.

[1]I use 'opacity' to refer to the class of phonological generalization that are derivationally opaque, i.e., counterfed or counterbled by another generalization (Kiparsky 1968, 1971).

1.1 Assumptions

One of the most fundamental assumptions in modern linguistics is that many aspects of language are **universal**.[2] In OT, the universal components are things like EVAL, CON, and GEN, with languages differing only in their lexicons and constraint hierarchies $\mathcal{H}$.

If grammars are ultimately built from the same atomic units organized in the same possible ways, then it follows that many grammatical patterns should be **typologically robust**, arising in multiple languages.

A general drive in theory-building is to avoid unnecessary complication. In linguistics, this can be expressed as an assumption about the general **concreteness** of grammar, adding abstraction only when necessary. Obviously, much of linguistic theory is filled with abstraction, some of which seems to be necessary, but where possible, the grammar should be concretely grounded. This may also be seen as intrinsically related to universality, especially if innateness is rejected.

A consequence of concreteness and a fundamental assumption of basic OT is **strict parallelism**, direct mapping between the input and output, with no abstract intermediate representations.

For phonology specifically, concreteness can also be interpreted as **(phonetic) naturalness**, with the grammar being based as much as possible on acoustics, articulation, aerodynamics, etc. In OT, this means that markedness constraints should be phonetically grounded.

The final two assumptions are related to language acquisition. For L1 acquisition, it is assumed that there are no language-specific constraints on possible inputs. This is known as **richness of the base (ROTB)**.

For L2 acquisition, it has generally been assumed since Lado (1957) that L2 acquisition shows **transfer effects** from the L1 grammar. In OT, this is often taken to mean that the L1 $\mathcal{H}$ is active during L2 acquisition, so that phonological processes in L1 will appear during L2 acquisition.

[2]For many linguists, especially those of a Chomskyan persuasion, an even more fundamental assumption is *innateness*, of which universality is simply a direct consequence. However, innateness isn't necessary here, so I proceed from the weaker and less controversial assumption of universality.

1.2 Predictions

If any $\mathcal{H}$ generates a phonological process $\mathscr{P}$, then:

(1) $\mathscr{P}$ is likely to be typologically robust;

(2) $\mathscr{P}$ cannot require intermediate representations;

(3) $\mathscr{P}$ must be natural;

(4) $\mathscr{P}$ must not be dependent on choice of input (that is, there should be no lexical exceptions to $\mathscr{P}$, and $\mathscr{P}$ should apply to nonce forms); and

(5) $\mathscr{P}$ is likely to transfer to L2 acquisition.

For a given $\mathscr{P}$, if we find that (1)–(5) do not hold, then we should conclude that $\mathscr{P}$ is not generated by any $\mathcal{H}$.

1.3 Polish ɔ-Raising

In Polish, underlying /ɔ/ raises to [u] before word-final voiced oral sonorants (6a) and obstruents (6b), but voiced obstruents are also devoiced word-finally, which opaquely masks raising of /ɔ/:

(6)

	stem UR	*NOM SG*	*NOM PL*	*gloss*
a.	/dvɔr/	dvur	dvɔrɪ	'mansion'
	/bɔl/	bul	bɔlɛ	'ache'
	/pɔkɔj/	pɔkujɛ	pɔkɔjɛ	'room'
	/stɔw/	stuw	stɔwɪ	'table'
b.	/bɔb/	bup	bɔbɪ	'bean'
	/rɔv/	ruf	rɔvɪ	'ditch'
	/lɔd/	lut	lɔdɪ	'ice'
	/dɔvɔz/	dɔvus	dɔvɔzɪ	'supply'
	/nɔž/	nuš	nɔžɛ	'knife'
	/rɔg/	ruk	rɔǧi	'horn'

In a serial analysis:

(7)

	'bean'	'peasant'	'club'	'purchase'
	/bɔb/	/xwɔp/	/klub/	/skup/
ɔ-Raising	bub	—	—	—
Devoicing	bup	—	klup	—
	[bup]	[xwɔp]	[klup]	[skup]

ɔ-Raising is an opaque process, while Devoicing is transparent. Does this correlate in any way to the predictions in (1)–(5)?

1.3.1 *Typological Robustness*

No language other than Polish seems to have the specific process of ɔ-Raising; a small number of languages exhibit some connection between vowel height and voicing, but the actual process is often radically different from Polish ɔ-Raising, affecting different types of vowels, tensing rather than raising, or in the case of Canadian Raising, having the trigger be *voiceless* consonants rather than voiced: /raɪt/ → [rʌɪt], but /raɪd/ → [raɪd], *[rʌɪd]).

In comparison, Devoicing is found in numerous languages, having risen independently in most other Slavic languages, as well as in German, Catalan, Turkish, Wolof, etc., with only minor variation (syllable-final instead of word-final, devoicing of stops specifically rather than obstruents, etc.).

1.3.2 *Intermediate Representations*

As seen in (7), ɔ-Raising takes the input and produces a form that is not the ultimate output. This intermediate form is necessary, so that Devoicing does not block application of ɔ-Raising. There is no way to revise ɔ-Raising to avoid production of an intermediate form, without increasing abstraction somewhere else (for example, allowing outputs to covertly code information with unpronounced features that nonetheless are abstractly present).

In comparison, Devoicing produces no intermediate representation: its output is the actual final output.

1.3.3 *Naturalness*

There is no direct phonetic motivation for ɔ-Raising as is. Vocal cord vibration, a raised velum, and tongue height are physically unrelated, and standard feature geometry has [voice], [nasal], [high] about as far from each other as features can be.

It's possible there is some perceptual effect of voicing that could cause vowel raising. However, the relevant acoustic properties are asynchronous, which may prove problematic for a perceptual analysis. Further, why should only /ɔ/ raise, when other vowels have the approximately the same F1 /ɛ ɛ̃ ɔ̃/? Why is raising not triggered by a following nasal, which is also voiced?

Another possible naturalness argument is functional: ɔ-Raising shifts the contrast from the voicing (lost through Devoicing) to vowel height, allowing the underlying voicing to be recoverable (cf. Łubowicz 2003). Such an analysis raises numerous questions: Why doesn't /u/ lower instead? Why don't any other vowels shift to preserve the voicing contrast?

Worse, note that while surface [ɔp] will be unambiguous, voicing is still not recoverable from surface [up], and in fact, now vowel height isn't either. However, with ɔ-Raising taken out of the picture, both [ɔp] and [up] have no ambiguity in their underlying vowels, while the voicing contrast is lost, just like it is elsewhere in the language. In other words, ɔ-Raising *increases* contrast loss, rather than mitigating it, so this functional explanation seems incorrect.

But even if all these problems could somehow be solved, ɔ-Raising also occurs before oral sonorants, where no voicing contrast is lost at all, so a completely different explanation is still required.

In comparison, Devoicing has a completely natural explanation. Vocal cord vibration requires airflow through the glottis. Obstruents have relatively narrow constrictions that impede this airflow, and airflow naturally decreases near the end of a word, so aerodynamically, it is expected that word-final obstruents will devoice (Ohala 1983).

1.3.4 Lexical Exceptions and Nonce Forms

There are many lexical exceptions to ɔ-Raising:

(8)	*stem UR*	*NOM SG*		*gloss*
	pɔr	pɔr	*pur	'leek'
	xɔl	xɔl	*xul	'lobby'
	kɔvbɔj	kɔvbɔj	*kɔvbuj	'cowboy'
	ɔɕɔw	ɔɕɔw	*ɔɕuw	'donkey'
	glɔb	glɔp	*glup	'globe'
	kɔd	kɔt	*kut	'code'
	nɛkrɔlɔg	nɛkrɔlɔk	*nɛkrɔluk	'obituary'
	xɔwd	xɔwt	*xuwt	'homage'

In comparison, there are no lexical exceptions to Devoicing in Polish, not even words which are exceptions to ɔ-Raising.

In nonce word experiments (Sanders 2003), when native Polish speakers were given sentences containing morphologically plural made-up words such as *ʑnabɔdɪ* and asked to produce new sentences that required the corresponding singular, they invariably produced *ʑnabɔt*, never **ʑnabut*, **ʑnabud*, or **ʑnabɔd*, demonstrating that ɔ-Raising is not productive, while Devoicing is.[3]

1.3.5 L2 Acquisition

I have no formal data yet. However, my informal impressions of the English spoken by native Polish speakers reveals pervasive Devoicing but no instances of ɔ-Raising.

1.3.6 Another Difference: Diachronic Complexity

Modern Polish ɔ-Raising is the result of an intricate set of unrelated natural sound changes. A very simplified chronology is given below:

(9)			
	Pre-12th c.		dvɔr
	12th–14th c.	V > Vː / __ $[+\text{voi}]_{\text{coda}}$	dvɔːr
	14th–16th c.	Vː > [+tense]	dvoːr
	16th–18th c.	Vː > V̆	dvor
	18th–20th c.	[+tense] > [+high]	dvur

In comparison, Devoicing occurred as a single change in the 14th–16th centuries, perhaps as a gradual process of progressive devoicing: d > d̥) > d(̥) > t.

1.4 What does this all mean?

CLAIM #1: These facts are not coincidence. Known cases of opacity tend to be typologically limited, abstract, unnatural, subject to lexical exceptions, nonproductive with nonce forms, not transferred to L2, and diachronically complex. This suggests a single analysis to unify these facts.

CLAIM #2: All of these facts can be accounted for in a basic version of OT with only a slight modification to lexicon optimization. There is no need, as in other OT solutions to opacity, to sacrifice reasonable assumptions such as universality, concreteness, strict parallelism, and naturalness.

[3] Baranowski and Buckley (2003) found that ɔ-Raising can apply to nonce words if they are similar to existing words that also undergo ɔ-Raising. However, they also found that nasal-final nonce words can undergo ɔ-Raising, so it's not quite clear what their results actually show.

2 Strong Lexicon Optimization

2.1 Lexicon Optimization and Lexical Economy

The nature of OT, especially because of ROTB, allows multiple possible inputs to be mapped onto the same phonetic output. For example, English has a process of glottalization that affects voiceless stops in codas, as in *knit* [nɪʔt]. Whether the input is /nɪt/ or /nɪʔt/, the output will still be the same, with the first input undergoing glottalization, while the second input needs no changes. From the perspective of the language user, it is simply irrelevant which input is used, since both end up with the same output.

With **lexicon optimization**, Prince and Smolensky (2004) state that in fact, the speaker will ultimately chose the most faithful input to be the UR stored in the lexicon, because it is the most harmonic with respect to the language's $\mathcal{H}$:

(10) *inputs* *output* *lexicon*

/nɪt/ —glottalization→ [nɪʔt]

/nɪʔt/ —faithful→ [nɪʔt] ···lexicon optimization···→ /nɪʔt/

This conception of lexicon optimization runs into complications when faced with morphological alternations, as in *knitted* [nɪɾəd], which shows the effects of flapping. With no other considerations taken into account, lexicon optimization would create two competing URs: /nɪʔt and /nɪɾ/.

Prince and Smolensky argue that this complication can be resolved by appeal to some principle of **lexical economy** that forces the UR to be a single phonological string that can account for all surface alternations. This yields the usual expected analysis, with /nɪt/ as the selected UR since it is properly affected by both glottalization and flapping, when the others are not.

But it's clear that the URs created by lexicon optimization with lexical economy are abstract: the UR /nɪt/ doesn't actually surface faithfully in any of the forms that help derive it! In the interest of preserving concreteness, I propose an alternative analysis called **strong lexicon optimization (SLO)**, in which lexicon optimization determines the UR without being weakened by adherence to lexical economy. This proposal is essentially a more radical version of similar "stored allomorphy" proposals in OT (Mester 1994, Burzio 1996, Kager 1996, Tranel 1998, etc.).

2.2 Consequences

SLO is **more concrete** than lexicon optimization plus lexical economy. Since URs selected by SLO are faithful to the outputs, there is a direct, concrete connection between input and output. The speaker does not need to go through multiple layers of computation over all surface forms to posit an abstract UR. Instead, each UR is selected directly from an individual output.

A further implication is that there will be **no underspecification** in the lexicon. This is again a desirable result from the perspective of concreteness, since archiphonemes, minimal feature combinations like [DOR, −voi], etc., are rather abstract.

Application of SLO will necessarily result in a **large lexicon**, since every known surface alternation will be encoded directly into the lexicon. So, for example, if every lexical item has an average of five surface allomorphs, then the lexicon's size will be five times the size of a lexicon formed through adherence to lexical economy.

Since there are multiple possible URs for a given word, most of which will typically not produce the correct output in particular contexts, the grammar must be modified in such a way to ensure correct **allomorph selection**.

SLO creates a lexicon that looks a lot like a list of groups of **suppletive allomorphs**, blurring the distinction between predictable allomorphy and true suppletion (as in *go*∼*went*, *be*∼*are*, etc.). If such a distinction is indeed important, then some way of accounting for it must also be added to the grammar.

2.3 Evidence

There is evidence from a variety of psycholinguistic phenomena that suggest that URs are more concrete and specified than allowed by lexical economy.

In **tip of the tongue effects**, when speakers are unable to recall a word, they are often aware of certain properties of the pronunciation of that word that would be predictable from the UR, such as the number of syllables and stress pattern (Brown and McNeill 1966). But since they are temporarily unable to access the UR, they have no sufficient form to submit to the grammar to derive those properties, which suggests that the properties are stored despite their predictability, allowing the speaker direct access to them.

Recognition speed of regularly inflected forms has been found to be dependent on the frequency of the inflected form, rather than the frequency of the uninflected stem (Sereno and Jongman 1997), suggesting that frequent use of an inflected form is remembered.

Children will often use particular regularly inflected words correctly without being able to apply the regular inflection to other words, including novel words (Berko 1958, MacWhinney 1978, Peters 1983), suggesting that **children memorize regularly inflected forms** before the inflectional rule has been solidified.

2.4 An Implementation

I begin with a first-pass definition of SLO that ignores morphological complexity:

(11) **Strong Lexicon Optimization (1st attempt):** Given an OT constraint hierarchy $\mathcal{H}$, an output o occurring in linguistic context $\mathbb{C}$, and the maximal set of inputs $I = \{i_k \mid \mathcal{H}\text{-}\textsc{Eval}(i_k) = o\}$, then the UR for o contains $\langle i_o, \mathbb{C} \rangle$, where $i_o \in I$ is the most faithful input for o with respect to $\mathcal{H}$.

For Polish specifically, the word [bup] 'bean (NOM SG)', which seems to undergo ɔ-Raising, would have a UR that looks something like ⟨bup, NOM SG⟩, while [bɔbɪ] 'bean (NOM PL)' with no ɔ-Raising would have the UR ⟨bɔbɪ, NOM PL⟩. Note also that the fully predictable effects of Devoicing are also encoded in the URs.

$\mathcal{H}$ is structured to generate the effects of Devoicing but not ɔ-Raising. This accounts for the fact that Devoicing is exceptionless and fully productive, while ɔ-Raising is not. If the speaker attempts to posit an input like /bob/ with a final voiced obstruent, $\mathcal{H}$ would always enforce Devoicing but never enforce ɔ-Raising, resulting in the output [bop]. Thus, for both possible lexical exceptions and nonce forms, SLO gets the correct results.

When a sound change occurs, the resulting outputs will ultimately be stored as URs. As more and more sound changes are encoded into the lexicon, opaque words are expected to arise. Crucially, however, the opacity is never part of the productive grammar at any moment in time; it is always just an epiphenomenon of the interaction between SLO and historical sound change across multiple generations:

(12)

	12th–14th c.	*14th–16th c.*	*16th–18th c.*	*18th–20th c.*
i_E	/bɔb/	/bɔːb/	/boːb/	/bop/
$\mathcal{H}$-EVAL$_E$	↓	↓	↓	↓
o_E	[bɔb]	[bɔːb]	[boːb]	[bop]
SLO$_E$	⇣	⇣	⇣	⇣
UR$_E$	/bɔb/	/bɔːb/	/boːb/	/bop/
$\mathcal{H}$-EVAL$_L$	↓	↓	↓	↓
o_L	[bɔːb]	[boːb]	[bop]	[bup]
SLO$_L$	⇣	⇣	⇣	⇣
UR$_L$	/bɔːb/	/boːb/	/bop/	/bup/

In this diagram, a subscript E indicates the *early* stage of the language, prior to the sound change of that time period, while L indicates the *late* stage, after sound change. The ordering of devoicing and vowel shortening in the 16th–18th centuries is irrelevant, so I have simply collapsed them.

As we can see, the opaque pronunciation [bup] 'bean' is carried through as each intervening sound change updates the lexicon. In a way, the interaction between SLO and sound changes mirrors a serial derivation. The key difference is that here, the seriality is not productive; speakers do not have access to older grammars or older inputs, so they cannot productively apply the history of their language to new words. But as long as children faithfully adopt their parent's outputs, the lexical opacity will persist.

3 Opacity in Polish Nominal Inflections

The previous analysis works just fine if all words are stored as monolithic units in the lexicon, regardless of their internal morphological structure. However, while this may be partly true (especially for high-frequency words), there is plenty of reason to believe that individual morphemes are also stored separately. Indeed, further evidence from Polish shows that this is required.

3.1 The Feminine and Neuter Vocative Singular

Polish has a complex system of productive noun inflectional endings depending on case, gender (including a distinction in the masculines between inanimates, animates, and humans), number, and the final consonant of the stem. Consider first the VOC SG endings for FEM and NEUT nouns ending in the consonants listed in the header row:

(13)

	t͡ɕ d͡ʑ ɲ *ɕ ʑ j*	*l*	*t͡s d͡z* *š ž t͡š d͡ž*	*p b m f v* *t d n s z w*	*r*	*x*	*k g*
FEM VOC SG	-ɔ	-ɔ	-ɔ	-ɔ	-ɔ	-ɔ	-ɔ
NEUT VOC SG	-ɛ	-ɛ	-ɛ	-ɔ	-ɔ	-ɔ	-ɔ

SLO as formulated can't deal with morphemes, so let's informally assume an intuitive version of SLO that can identify morphemes and morpheme realizations, but otherwise functions normally. In that case, we get the following partial UR:

(14) $\textsc{voc sg} = \langle \text{-ɔ}, \textsc{fem} \rangle, \left\langle \begin{Bmatrix} \text{-ɔ} \\ \text{-ɛ} \end{Bmatrix}, \textsc{neut} \right\rangle$

This gets the result that the FEM VOC SG always ends in [-ɔ], but we need some way to ensure that the correct allomorph is chosen for the NEUT. Since the trigger is the stem-final consonant of the noun, it makes sense for the choice to be made by the phonology, i.e., by $\mathcal{H}$. At minimum, $\mathcal{H}$ must contain subrankings like *t^{s}ɔ ≫ *t^{s}ɛ and *sɔ ≫ *sɛ, to force the correct choice of allomorph:

(15)

$\ldots t^{s}+\begin{Bmatrix} \text{-ɔ} \\ \text{-ɛ} \end{Bmatrix}$	*t^{s}ɔ	*t^{s}ɛ
a. … t^{s}ɔ	*!	
☞ b. … t^{s}ɛ		*

(16)

$\ldots s+\begin{Bmatrix} \text{-ɔ} \\ \text{-ɛ} \end{Bmatrix}$	*sɛ	*sɔ
☞ a. … sɔ		*
b. … sɛ	*!	

It's not clear what natural explanation could motivate these rankings. What connection is there between consonant friction and vowel quality such that dental affricates prefer to be followed by a front vowel, while dental fricatives prefer to be followed by a back vowel? However, assuming a natural explanation can be found, there is no serious problem here—SLO stores both allomorphs, and the phonology (somehow) picks the correct one.

3.2 The Neuter Dative and Locative Singular

In the following data for NEUT nouns, the symbol ⊕ indicates that the final consonant of the stem undergoes a change, as listed below. Traditionally, the ⊕ changes are called 'palatalization', but to avoid confusion with true palatalization (a secondary high, front articulation), I call them *palatal mutations*:

(17)

	t͡ɕ d͡ʑ ɲ *ɕ ʑ j*	*l*	*t͡s d͡z* *š ž t͡š d͡ž*	*p b m f v* *t d n s z w*	*r*	*x*	*k g*
NEUT DAT SG	-u	-u	-u	-u	-u	-u	-u
NEUT LOC SG	-u	-u	-u	⊕-ɛ	⊕-ɛ	-u	-u

(18)

plain hard	*⊕ form*
p b f v m	pʲ bʲ fʲ vʲ mʲ
t d	t͡ɕ d͡ʑ
s z n	ɕ ʑ ɲ
r	ž
w	l
k g	t^{s} d^{z}
x	š / ɕ (depends on morpheme!)

This seems to just be analogous to the VOC SG, with URs like the following:

(19) $\textsc{dat sg} = \langle \text{-u}, \textsc{neut} \rangle \qquad \textsc{loc sg} = \left\langle \begin{Bmatrix} \text{-u} \\ \text{-ɛ} \end{Bmatrix}, \textsc{neut} \right\rangle$

The DAT SG will be chosen very straightforwardly, with [-u] always surfacing.

For the LOC SG ending, however, it's not so simple. First, consider the behavior of stem-final /ɕ/ (the ranking *ɕɔ ≫ *ɕɛ follows from the rankings needed for the NEUT VOC SG):

(20)

$\ldots ɕ+\begin{Bmatrix} \text{-u} \\ \text{-ɛ} \end{Bmatrix}$	*ɕɔ	*ɕɛ	*ɕu
☞ a. … ɕu			*
b. … ɕɛ		*!	

For stem-final /s/, however, it is not simply a matter of choosing the correct ending. In addition, the stem consonant must mutate to [ɕ]. A ranking such as *sɛ ≫ IDENT-[back] will trigger the necessary palatal mutation:

(21)

…s+ɛ	*sɛ	ID-[bk]
a. …sɛ	*!	
☞ b. …ɕɛ		*

But this assumes that correct allomorph is chosen to begin with. When the allomorphy is factored in, we need rank *su high enough to prevent /-u/ from being selected:

(22)

…s+{-u, -ɛ}	*su	*sɛ	ID-[bk]
a. …sɛ		*!	
☞ b. …ɕɛ			*
c. …su	*!		

But recall from (20) that *ɕɛ ≫ *ɕu, which yields the wrong result:

(23)

…s+{-u, -ɛ}	*su	*sɛ	ID-[bk]	*ɕɛ	*ɕu
a. …sɛ		*!			
b. …ɕɛ			*	*!	
c. …su	*!				
✘ d. …ɕu			*		*

A purely phonological solution with strict parallelism is simply not going to work here: the underlying consonant needs to select an inflectional allomorph that just happens to trigger palatal mutation of the consonant to one that needs to select the other inflectional allomorph. Since I am assuming strict parallelism, this means that the correct inflectional ending cannot be selected phonologically. That is, $\mathcal{H}$ **is *not* responsible for allomorph selection**.

But if allomorphs are not determined phonologically, how are they determined? As we have seen before with the VOC SG, an allomorph can be selected by satisfying the linguistic context $\mathbb{C}$ of the UR. So far, we have only considered morphological contexts, but it seems clear that **we must allow $\mathbb{C}$ to contain *phonological* contexts** as well.

This yields the following new URs:

(24) $$\text{VOC SG} = \langle \text{-ɔ}, \text{FEM} \rangle, \left\langle \text{-ɔ}, \text{NEUT}\& \ldots \left\{ \begin{matrix} \text{p} \\ \text{b} \\ \vdots \end{matrix} \right\} \right\rangle, \left\langle \text{-ɛ}, \text{NEUT}\& \ldots \left\{ \begin{matrix} \text{ʨ} \\ \text{ʥ} \\ \vdots \end{matrix} \right\} \right\rangle$$

(25) $$\text{LOC SG} = \left\langle \text{-ɛ}, \text{NEUT}\& \ldots \left\{ \begin{matrix} \text{p} \\ \text{b} \\ \vdots \end{matrix} \right\} \right\rangle, \left\langle \text{-u}, \text{NEUT}\& \ldots \left\{ \begin{matrix} \text{ʨ} \\ \text{ʥ} \\ \vdots \end{matrix} \right\} \right\rangle$$

With particular allomorphs selected by the morphology prior to being subjected to the phonology, the rankings derived before are irrelevant. $\mathcal{H}$ doesn't need to be structured in such a way to select the correct allomorph, which is a good thing: the rankings derived above don't seem very natural.

However, the palatal mutations in the stem-final consonant must still be accounted for. Unlike the inflectional endings, palatal mutations must still be encoded in $\mathcal{H}$ somehow, because it applies fully productively, including to recent borrowings and nonce forms.

That is, while extant nouns may have complex URs created by SLO, listing both the plain and mutated allomorphs (each of which are selected by the morphology to appear in the correct environments), this cannot be the case for nonce forms, because they do not have the benefit of prior use for SLO to encode the necessary complexity into the URs.

Some further data for consideration:

(26)

	ʨ ʥ ɲ *ɕ ʑ j*	*l*	*ʦ ʣ* *š ž t͡š d͡ž*	*p b m f v* *t d n s z w*	*r*	*x*	*k g*
MAS LOC SG	-u	-u	-u	⊕-ɛ	⊕-ɛ	-u	-u
FEM GEN SG	-i	-i	-ɪ	-ɪ	-ɪ	-ɪ	-i
FEM LOC SG	-i	-i	-ɪ	⊕-ɛ	⊕-ɛ	⊕-ɛ	⊕-ɛ
NONVIR NOM PL	-ɛ	-ɛ	-ɛ	-ɪ	-ɪ	-ɪ	-i
VIR NOM PL	-ɛ	-ɛ	-ɛ	⊕-i	⊕-ɪ	⊕-i	⊕-ɪ

3.3 Palatal Mutation and Phonological Context

Giving a concrete, strictly parallel phonological account of palatal mutation seems to be impossible. Consider the results and triggers of palatal mutations ('—' indicates that the consonant does not mutate):

(27)

	NEUT LOC SG	*FEM LOC SG*	*VIR NOM PL*	*NONVIR NOM PL*
	-ε	*-ε*	*-ɪ/i*	*-ɪ/i*
p b f v m	p̡ b̡ f̡ v̡ m̡	p̡ b̡ f̡ v̡ m̡	p̡ b̡ f̡ v̡ m̡	—
t d	ȶ^{ɕ} ȡ^{ʑ}	ȶ^{ɕ} ȡ^{ʑ}	ȶ^{ɕ} ȡ^{ʑ}	—
s z n	ɕ ʑ ȵ	ɕ ʑ ȵ	ɕ ʑ ȵ	—
r	ž	ž	ž	—
w	l	l	l	—
k g	—	t^{s} d^{z}	t^{s} d^{z}	—
x	š	š	ɕ	—

As is clear from the LOC SG endings for nouns ending in velar stops, an analysis based purely on surface segments will not explain why the NEUT doesn't mutate while the FEM does, even though the inflectional ending is phonologically /-ε/ for both.

This is even more obvious in the MASC PL endings for all nouns: the VIR ('virile', male human) nouns mutate, whereas the NONVIR ('non-virile', non-human MASC) nouns do not mutate, despite both taking /-ɪ or -i/ (actual choice depends on the stem-final consonant; VIR and NONVIR stems ending in /r/ both take /-ɪ/). Simple linear combinations of phonological segments just cannot account for these asymmetries.

One solution would be to posit some abstract marker like $\oplus$ on the mutating inflections, and make $\mathcal{H}$ sensitive to that. However, continuing along with the assumption of concreteness, no abstraction is allowed unless absolutely necessary.

So what is concrete about the palatal mutations and the inflections that trigger them? The actual outputs are concrete. That is, [p̡] is a concrete indication that palatal mutation has applied to a stem ending in /p/.

And how can morphemes differ from each other? With $\mathcal{H}$ being purely phonological, that leaves only the URs as a place where morphemes can be distinguished. Thus, the logical conclusion is that **the concrete outputs of the mutating inflectional endings must be encoded in their lexical entries**!

A first naive attempt would be to encode the UR as follows:

(28) LOC SG = ⟨-p̡ε, NEUT&...p⟩, ⟨-b̡ε, NEUT&...b⟩, ...

This would cause the morphology to build inputs like /...p+p̡ε/, and if the phonology forced coalescence, everything would be fine. Of course, the story isn't quite that simple. In the general case, Polish allows a consonant to be followed by its own palatal mutation (with obligatory regressive assimilation of true palatalization for coronals):

(29) bεɕɕε 'price drop (DAT)'
staraȵȵε 'carefully'
lekt^{s}εvažɔnt^{s}ɪ 'disrespectful'

3.4 Floating Segments

Taking a cue from the predominant generative analysis of palatal mutations in Polish (Gussmann 1980, Rubach 1984, etc.), I propose that **the palatal mutations are encoded in the URs as floating segments** (segments with no timing slot, notated with a superscript slanted grey font: p̡), which merge with the stem-final consonant and overwrite its features.

3.4.1 Palatal Mutation of the Labials

All of the labial palatal mutations undergo the same change: the addition of a high front secondary articulation (IPA [^{j}]). I only give an analysis for /p/, with palatal mutation of the other labials /b f v m/ following straightforwardly.

I assume the following undominated constraints:

(30) *FLOAT: Floating segments are banned from outputs.

(31) DEP-X: All timing slots in the output must have an input correspondent.

(32) INTEGRITY-X: Distinct timing slots in the output cannot correspond to the same input timing slot.

(33) MAX-seg: All segments in the input must have an output correspondent.

With these constraints undominated, there will be no need to consider any candidates in which a floating segment appears, in which a floating segment is given its own timing slot from out of nowhere, in which an input timing slot is split into two output timing slots, or in which any input segments are deleted. In short, these constraints force any floating segments in the input to coalesce with some other (non-floating) segment.

Further constraints that are needed include:

(34) UNIFORMITY: Distinct segments in the input cannot correspond to the same output segment.

(35) IDENT-[palatal]: Corresponding segments in the input and output cannnot differ in palatal articulation. This is a cover constraint for IDENT-[high] and IDENT-[back].

As it stands, these constraints are not sufficient to account the palatal mutation:

(36)

. . . p+p̡ε	UNIF	ID-[pal]
✘ a. . . . pε	*	*
☞ b. . . . p̡ε	*	*

The apparent difference here is between changing the features of the stem versus changing the features of the affix, which can be accomplished by relativizing faithfulness constraints to stems and affixes (Beckman 1995, 1997, 1998, etc.). However, the required ranking for Polish palatal mutations is ID$_{\text{affix}}$-[pal] $\gg$ ID$_{\text{stem}}$-[pal], which runs counter to the usual assumption that FAITH$_{\text{stem}}$ outranks FAITH$_{\text{affix}}$ (McCarthy and Prince 1995).

The solution lies in realizing that we are dealing only with nonce stems here; any extant stems will already have the palatal mutations stored in their own lexical entries, and the phonology is never needed. But for nonce stems, there is no lexical entry yet, so $\mathcal{H}$ is required to generate the palatal mutations. This is then not a faithfulness asymmetry between stems and affixes, but rather, **a universal faithfulness asymmetry between nonce stems and extant affixes, with FAITH$_{\text{extant}}$ ranked higher than FAITH$_{\text{nonce}}$**:

(37)

. . . p+p̡ε	UNIF	ID$_{\text{ext}}$-[pal]	ID$_{\text{nonce}}$-[pal]
a. . . . pε	*	*!	
☞ b. . . . p̡ε	*		*

This not only gets the right results, but a universal ranking of FAITH$_{\text{extant}}$ $\gg$ FAITH$_{\text{nonce}}$ seems to make sense: when new words are borrowed or created, they do not trigger new alternations in existing affixes; rather, the nonce forms themselves are subject to the restrictions of the language's morphology and phonology.[4]

[4]There is a intuitive conflict here with Ito and Mester's (1995, 1999) core-periphery model of the lexicon, in which FAITH$_{\text{foreign}}$ $\gg$ FAITH$_{\text{native}}$. A resolution to this potential contradiction is beyond the scope of this talk.

3.4.2 *Palatal Mutation of the Coronals*

The ranking already established for labials yields the correct results for coronal fricatives and nasals, since they too simply palatalize (I no longer show the ranking of UNIFORMITY, since it is not crucially ranked because every candidate involves coalescence):

(38)

. . . s+ɕε	ID$_{\text{ext}}$-[pal]	ID$_{\text{nonce}}$-[pal]
a. . . . sε	*!	
☞ b. . . . ɕε		*

For coronal stops, another constraint is required:

(39) IDENT-[affrication]: Corresponding segments in the input and output cannnot differ in affrication. This is a cover constraint for whatever constraint punishes changes between stops and affricates.

(40)

. . . t+ȶɕε	ID$_{\text{ext}}$ [pal]	ID$_{\text{ext}}$ [affr]	ID$_{\text{nonce}}$ [pal]	ID$_{\text{nonce}}$ [affr]
a. . . . tε	*!	*!		
b. . . . tˢε	*!			*
c. . . . ȶε		*!	*	
☞ d. . . . ȶɕε			*	*

For the palatal mutation of /r/ to [ž], the place of articulation changes from alveolar to post-alveolar, violating IDENT-[anterior], and the manner of articulation changes from trill to fricative, violating IDENT-[sonorant]:

(41)

. . . r+žε	ID$_{\text{ext}}$ [ant]	ID$_{\text{ext}}$ [son]	ID$_{\text{nonce}}$ [ant]	ID$_{\text{nonce}}$ [son]
a. . . . rε	*!	*!		
b. . . . zε	*!			*
c. . . . ɻε		*!	*	
☞ d. . . . žε			*	*

Even this drastic change falls out fairly simply from the general ranking of FAITH$_{\text{extant}}$ over FAITH$_{\text{nonce}}$. However, as these coronals show, it seems like *every* FAITH$_{\text{extant}}$ must be ranked over every FAITH$_{\text{nonce}}$, not just individual matching constraints.

3.4.3 Palatal Mutation of the Velars

The velar stops are analyzed in much the same way as the labials and coronals:

(42)

k+$^{t^s}$ɛ	IDext [COR]	IDext [DOR]	IDext [affr]	IDnonce [COR]	IDnonce [DOR]	IDnonce [affr]
a. kɛ	*!	*!	*!			
b. k^{x}ɛ	*!	*!				*
c. ʔɛ	*!		*!		*	
d. ƫɛ		*!	*!	*		
e. ʔ^{h}ɛ	*!				*	*
f. ƫ^{ɕ}ɛ		*!		*		*
g. tɛ			*!	*	*	
☞ h. t^{s}ɛ				*	*	*

The velar fricative /x/ has two possible palatal mutations, depending on which inflectional ending is being used. For the LOC SG, the input will contain /-^{š}ɛ/:

(43)

x+^{š}ɛ	IDext [COR]	IDext [DOR]	IDnonce [COR]	IDnonce [DOR]
a. xɛ	*!	*!		
b. çɛ	*!			*
c. hɛ		*!	*	
☞ d. šɛ			*	*

While the VIR NOM PL input will contain /-^{ɕ}i/:

(44)

x+^{ɕ}i	IDext [COR]	IDext [pal]	IDnonce [COR]	IDnonce [pal]
a. xi	*!	*!		
b. x̟i	*!			*
c. si		*!	*	
☞ d. ɕi			*	*

Finally, the bizarre mutation of /w/ to [l] is also easily accounted for:

(45)

w+^{l}ɛ	IDext [COR]	IDext [DOR]	IDext [lat]	IDnonce [COR]	IDnonce [DOR]	IDnonce [lat]
a. wɛ	*!	*!	*!			
b. ʟɛ	*!	*!				*
c. ɦɛ	*!		*!		*	
d. ɟɛ		*!	*!	*		
e. ʔɛ	*!				*	*
f. ɬɛ		*!		*		*
g. ɹɛ			*!	*	*	
☞ h. lɛ				*	*	*

4 Summary

☆ I have assumed that the phonological component of grammar obeys universality, typological robustness, concreteness, strict parallelism, and naturalness.

☆ Under these assumptions, it is predicted that many types of opacity should not be generated by any possible $\mathcal{H}$, contrary to apparent data.

☆ Using a case of opacity from Polish, I showed how opaque process often are not typologically robust, are abstract, are unnatural, are not productive, are not transferable to L2, and are diachronically complex.

☆ These various facts about opacity were shown not to be coincidence, but rather the product of encoding opacity into the lexicon via strong lexicon optimization, which allows opacity to continue to flourish in the lexicon without being generated by $\mathcal{H}$.

☆ I further showed that lexical entries need to be sensitive not only to morphological contexts but also phonological contexts.

☆ In addition, strong lexicon optimization allows radical phonological changes to productively take place, as long as they are triggered by specific morphemes, which must have the outputs encoded directly in the UR.

☆ **Is phonology obsolete?!** No! We still need a phonological $\mathcal{H}$ to explain how nonce forms conform to phonotactics. We still need $\mathcal{H}$ to explain sound change. And we may still need $\mathcal{H}$ to do more with morphology than simply allow extant (memorized) morphemes to trump nonce forms.

References

Baranowski, Maciej, and Eugene Buckley. 2003. Lexicalization and raising in Polish. Annual Meeting of the Linguistics Society of America.

Beckman, Jill. 1995. Shona height harmony: Markedness and positional identity. In Beckman et al. (1995), 53–75.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona height harmony. *Phonology* 14:1–46.

Beckman, Jill. 1998. Positional faithfulness. Doctoral Dissertation, University of Massachusetts, Amherst.

Beckman, Jill, Laura Walsh Dickey, and Suzanne Urbanczyk, ed. 1995. *University of Massachusetts occasional papers in linguistics*. Amherst, MA: GLSA.

Berko, J. 1958. Processing of English inflectional morphology. *Word* 14:150–177.

Brown, R., and D. McNeill. 1966. The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior* 325–337.

Burzio, Luigi. 1996. Surface constraints versus underlying representations. In *Current trends in phonology: Models and methods*, ed. Jacques Durand and Bernard Laks, 97–122. European Studies Research Institute, University of Salford Publications.

Chomsky, Noam, and Morris Halle, ed. 1968. *The sound pattern of English*. Cambridge, MA: MIT Press.

Flemming, Edward. 1995. Acoustic and auditory representations in phonology. Doctoral Dissertation, University of California, Los Angeles.

Gussmann, Edmund. 1980. *Studies in abstract phonology*. Cambridge, MA: MIT Press.

Ito, Junko, and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Beckman et al. (1995), 181–209.

Ito, Junko, and Armin Mester. 1999. The structure of the phonological lexicon. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 62–100. Malden, MA: Blackwell Publishers.

Kager, René. 1996. On affix allomorphy and syllable counting. In *Interfaces in phonology [Studia Grammatica 41]*, ed. U. Kleinhenz, 155–171. Berlin: Akademie Verlag.

Kiparsky, Paul. 1968. Linguistic universals and language change. In *Universals in linguistic theory*, ed. E. Bach and R. Harms, 170–202. New York: Holt, Rinehart, and Winston.

Kiparsky, Paul. 1971. Historical linguistics. In *A survey of linguistic science*, ed. W.O. Dingwall, 576–642. College Park: University of Maryland Linguistics Program.

Lado, Richard. 1957. *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Łubowicz, Anna. 2003. Contrast preservation in phonological mappings. Doctoral Dissertation, University of Massachusetts, Amherst.

MacWhinney, B. 1978. Processing a first language: The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* 43.

McCarthy, John, and Alan Prince. 1993. Prosodic morphology I: Constraint interaction and satisfaction. Manuscript, University of Massachusetts, Amherst, and Rutgers University.

McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In Beckman et al. (1995), 249–384.

Mester, Armin. 1994. The quantitative trochee in Latin. *Natural Language and Linguistic Theory* 12:1–61.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The production of speech*, ed. Peter F. MacNeilage, 189–216. New York: Springer.

Peters, A.M. 1983. *The units of language acquisition*. Cambridge University Press.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell Publishers.

Roca, Iggy, ed. 1997. *Derivation and constrains in phonology*. Oxford: Clarendon Press.

Rubach, Jerzy. 1984. *Cyclic and lexical phonology: The structure of Polish*. Dordrecht: Foris.

Sanders, Nathan. 2003. Opacity and sound change in the Polish lexicon. Doctoral Dissertation, University of California, Santa Cruz.

Sereno, Joan A., and Allard Jongman. 1997. Processing of English inflectional morphology. *Memory and Cognition* 25:425–437.

Tranel, Bernard. 1998. Suppletion and OT: On the issue of the syntax/phonology interaction. In *Proceedings of West Coast Conference on Formal Linguistics 16*, ed. E. Curtis, J. Lyle, and G. Webster, 415–429. Stanford, CA: CSLI.