

Jakub Dunin-Borkowski

Uniwersytet Warmińsko-Mazurski w Olsztynie

ORCID: <https://orcid.org/0000-0001-6676-8862>

e-mail: jakub.dunin.borkowski@uwm.edu.pl

Variation and stability of Polish Retraction in Optimality Theory

Zmienność i stabilność polskiego procesu cofania /i/ w Teorii Optymalności

Abstract

In classic Optimality Theory (OT), constraints are ranked categorically. Such an approach predicts that outputs of a given grammar do not exhibit variation. Variation is, however, present in natural languages. In order to address this issue in OT, it has been proposed that categorical ranking should be abandoned in favor of probabilistic ranking. Probabilistic ranking, however, encounters problems when a process applies variably to some forms, but categorically to other forms. In such cases, lexically indexed constraints are invoked in order to account for the observed exceptionality. While lexical indexing is a successful formal device, it does not predict the behavior of novel forms. This paper reanalyzes the process of Retraction found in Polish. Retraction applies categorically across morphological boundaries, but it may apply categorically or variably inside morphemes. A solution to this issue is found in the interaction of phonology with morphology and the lexicon. In this view, abstract Underlying Representations coexist with listed words. The proposed model employs probabilistic constraint ranking and whole-word storage, retaining the formal convention of lexical and morphological indexing. The model is compatible with the framework of Derivational OT, which is necessary to account for the opaque interaction of Retraction and palatalization.

Keywords: Derivational Optimality Theory, probabilistic ranking, variation, exceptions, whole-word storage, Polish Retraction

Abstrakt

W klasycznej Teorii Optymalności (OT) rankingi ograniczeń optymalności są kategoriyczne. Takie podejście przewiduje, że wyniki wewnątrz danej gramatyki nie wykazują zmienności. Zmienność procesów jest jednak obecna w językach naturalnych. Aby rozwiązać ten problem w OT, zaproponowano porzucenie rankingów kategoriycznych na rzecz rankingów probabilistycznych. Rankingi probabilistyczne napotykają jednak problemy, gdy dany proces zachodzi zmiennie w niektórych formach, ale kategoriycznie w innych. W takich przypadkach przywołuje się ograniczenia optymalności indeksowane

leksykalnie w celu opisania zaobserwowanych wyjątków. Chociaż indeksowanie leksykalne jest skutecznym narzędziem formalnym, nie przewiduje ono zachowania nowych form. W tym artykule omówiono proces cofania /i/ w języku polskim. Cofanie zachodzi w sposób kategoriowy na granicy morfemów, zaś w środku morfemów może zachodzić w sposób kategoriowy lub zmienny. Rozwiązanie tego problemu można znaleźć w interakcji fonologii z morfologią oraz leksykonem. Z tego punktu widzenia abstrakcyjne formy głębokie współlistnieją z całymi słowami. Proponowany model wykorzystuje probabilistyczne rankingi ograniczeń oraz przechowywanie gotowych słów, zachowując formalną konwencję indeksowania morfologicznego i leksykalnego. Model jest kompatybilny z Derywacyjną Teorią Optymalności, która jest konieczna do opisanie nietransparentnej interakcji cofania z palatalizacją.

Słowa kluczowe: Derywacyjna Teoria Optymalności, probabilistyczny ranking, zmienność, wyjątki, przechowywanie gotowych słów, cofanie /i/ w polszczyźnie

1. Introduction

Exceptional behavior of forms *vis-à-vis* (morpho)phonological processes may be divided into two main categories: stable lexical exceptionality and variable application of processes. Given a synchronic process P that applies in the context C, a stable exceptional form does not undergo P even though it contains C. Forms that meet the structural description of a given rule of grammar but in fact do not undergo this rule are found across natural languages. In rule-based generative accounts, such as *The Sound Pattern of English* (SPE; Chomsky & Halle 1968), diacritics/indexes (or other lexical specifications) are used in order to match some rules to specific morphemes. For instance, the rule of Velar Softening in English is characterized as functioning in Greek/Latinate words (Chomsky & Halle 1968; Rubach 1984b). Consequently, all forms that are not marked as ‘Greek/Latinate’ remain untouched by this rule, even if they meet the structural description of Velar Softening. In Optimality Theory (OT; Prince & Smolensky 2004; McCarthy & Prince 1995), a more recent framework within Generative Phonology, lexical exceptionality is principally accounted for by lexically indexed constraints (Fukuzawa 1999; Itô & Mester 1999; Pater 2000). The mechanism of deriving exceptionality is to some extent similar to rule-based frameworks. For instance, a lexically indexed constraint is active only for outputs that bear the same index as the constraint.

The second type of exceptionality, variable application, includes variation in the surface representation of a single input regarding some phonological mapping. In other words, the process P applies to the form V, which contains the context C, with a certain probability. As a result, P may or may not apply to V, even though the required context is met. An example of a variably

applying process is Ilokano metathesis: the input /taʔo+en/ ‘to repopulate’ may surface as [taʔwen] or [taʔen] (Boersma & Hayes 2001: 58). In OT, synchronic variation may be accounted for by, for instance, probabilistic constraint ranking (Boersma 1997).

Using the example of Polish Retraction, this paper argues that, first, individual strategies of generating lexical exceptionality are unable to predict the behavior of novel forms, whose status *vis-à-vis* a given index/specification is unknown. Second, it is shown that probabilistic constraint ranking in OT is not sufficient to account for the stability of the observed mappings. In other words, a probabilistic ranking that derives variable application of P predicts that P never applies categorically. This, however, is not always the case, as shown by the example of Retraction. In order to salvage the probabilistic approach to the OT grammar, it is necessary to posit a mechanism that would assure the simultaneous stability and variation of a given process across different forms (e.g., Zuraw 2000, 2010). This goal is achieved by coupling the probabilistic grammar with dual-route model of lexical access (e.g., McQueen & Cutler 1998; Hay 2003), where entire words/phrases may be listed in the lexicon. Third, the aforementioned components are combined in a unified framework that makes reference to morphology in order to account for the complexity of Retraction. The analysis employs lexically and morphologically indexed constraints in a probabilistic derivational OT framework, a combination of probabilistic OT (Boersma 1997) and Derivational OT (Rubach 1997), to account for the data regarding Retraction, including its asymmetric application across different morphological domains and opaque interaction with palatalization (e.g., Rubach 2017).

2. Retraction in rule-based phonology and classic OT

In Polish, the high front vowel /i/ becomes [+back] ([±bk], henceforth) after hard coronals (excluding /l/) (Rubach 1984a). After Rubach (1984a, 2003a), I assume that the so called *hard* consonants in Polish bear the feature [+bk], while palatal/palatalized consonants are characterized by the feature [-bk].¹ Consider the phonological rule of Retraction, given schematically in (1).

- (1) Retraction (Rubach 1984a)
 $i \rightarrow i / [+coronal +back -lateral] _$

¹ [-bk] consonants involve /ɛ z tɕ ɲ j/ as well as all palatalized consonants marked with a superscript *j*.

As reported by Rubach (1984a), the effects of Retraction can be found inside words regardless of the morphological boundaries. In the framework of Lexical Phonology (Kiparsky 1982; Booij & Rubach 1987), Retraction must be a postcyclic rule, that is, a rule that operates inside words, but has no effect across word boundaries (Rubach 1984a: Section 7.3). The rule in (1) is considered a synchronic process for two reasons. First, alternations are found between morphemes sharing the Underlying Representation (UR). Second, Retraction functions in borrowings.

(2) Retraction – examples (Rubach 1984a: 202, 204, 205)²

a. ALTERNATIONS

/xmu_z+i+te/ ‘become cloudy’

/xɔləndz+i/ ‘Duchman’ nom.pl.

vs.

/ko_e+i+te/ ‘mow’

/frantsuz+i/ ‘Frenchman’ nom.pl.

b. BORROWINGS

[i] > [ɨ]

limousine (Fr.) > limu[zi]na ‘limousine’

chifre (Fr.) > [ʃi]fr ‘code’

cretin (Fr.) > kre[ti]n ‘idiot’

The data in (2a) show that the verbalizing and the virile plural suffix, which otherwise surface as /i/, undergo Retraction when preceded by /z/. Additionally, in borrowings, as documented in (2b), the high front vowel retracts to /i/ when preceded by hard coronals.

Regardless of the assumed phonological framework, any account of Retraction must somehow relate to the words that do not conform to its requirements. In other words, there are a number of borrowings/proper names that are exceptions to Retraction, with some forms exhibiting variation.³ The data come from my observation of Polish native speakers in everyday life situations as well as from publicly available audio materials.

(3) Exceptions to Retraction

a. [ripɔsta]	‘retort’	b. [gril] ~ [gril]	‘grill’
[sinus]	‘sine’	[britnej] ~ [britnej]	‘Britney’
[tir]	‘truck’	[rivjera] ~ [rivjera]	‘proper name’
[dinɔzawɾ]	‘dinosaur’	[sarkɔzi] ~ [sarkɔzi]	‘Sarkozy’
		[awdit] ~ [awdit]	‘audit’

In classic rule-based theories, such as SPE or Lexical Phonology, arbitrary notation conventions are used in order to distinguish between structurally similar forms that behave differently *vis-à-vis* the same phonological rule.

² I ignore the effects of palatalization (e.g., r → ʒ) and return to this issue in Section 4.

³ Both inter- and intraspeaker variation.

In other words, even though the forms in (3) meet the structural description of Retraction, they are/may be unaffected by this rule and hence must be representationally distinct from the forms in (2). A notation convention may involve assigning a diacritic to phonological rules as well as to the forms/segments that are (not) targeted by these rules (Chomsky & Halle 1968: 172–177). The derivation in (4) uses the diacritic $[\pm R]$ to indicate whether the rule of Retraction applies.

- (4) Rule-based derivation
 UR: $//\text{sinus}_{[\pm R]}//$
 --- Retraction (*applies to forms bearing $[\pm R]$*)
 Surface: $[\text{sinus}]$

As shown in (4), the attested output is derived with the use of diacritic marking. However, such a derivation does not account for variation. In order to arrive at the two alternative outputs for ‘grill’, given in (3), we must assume that Retraction applies optionally to $[\pm R]$ and that ‘grill’ is represented as $//\text{gril}_{[\pm R]}//$. One of the problems with such an approach is that Retraction applies categorically to many morphemes, including the verbalizing and the virile plural suffix (see (2)). Therefore, treating Retraction as optional would also necessitate positing an additional copy of this rule, ‘Retraction’, that would categorically target such morphemes.

The problem of lexical exceptionality and variation is not entirely resolved in OT (Prince & Smolensky 2004), a later development within generative phonology. In spite of many advantages of parallel evaluation of output candidates over rule-based derivations, classic OT still fails to predict the outcome of some processes when any inconsistency of application is involved. For instance, Rubach (2017) analyzes Retraction as an effect of the interaction of the constraints listed in (5).

- (5) PAL-*i* (Rubach 2000)
 A consonant and a following unrounded high vowel must agree in $[\pm \text{bk}]$.
 IDENTVOWEL (IDV)
 The features on an input vowel must be preserved in its output correspondent.

PAL-*i* (Rubach 2000) is a markedness constraint requiring the agreement in backness between a consonant and the following high front vowel /i/ or the high back vowel /i/. The competing faithfulness constraint is IDV, which bars any featural changes on vowels between the input and the output. Consider the evaluation of candidates for /sinus/ → [sinus] ‘sine’ and /xmuz_t+i+te/ → [xmuz_tite] ‘become cloudy’, given in (6).

(6) Retraction – failed evaluation⁴

<i>In</i>	<i>Out</i>	PAL- <i>i</i>	IDV
sinus	a. ☹️sinus	*!	
	b. ☹️sinus		*
xmuz _i +i+tɛ	c. xmuz _i tɛ	*!	
	d. ☹️xmuz _i tɛ		*

The ranking PAL-*i* >> IDV incorrectly predicts Retraction both in ‘sine’ and in ‘become cloudy’. A solution to this problem within classic OT is provided by lexically indexed constraints (Fukuzawa 1999, Itô & Mester 1999, Pater 2000).⁵ For instance, *sinus* [sinus] ‘sine’, which does not undergo Retraction, can be assigned a diacritic F(aithful). Morphemes bearing such a diacritic are targeted by a lexically indexed faithfulness constraint, IDV-F. If we posit a ranking IDV-F >> PAL-*i* >> IDV, Retraction will apply to forms without diacritics, while forms bearing the diacritics will remain unchanged, as demonstrated in (7).

(7) Retraction – lexically indexed constraints

<i>In</i>	<i>Out</i>	IDV-F	PAL- <i>i</i>	IDV
sinus _F	a. ☹️sinus _F		*	
	b. sinus _F	*!		*
xmuz _i +i+tɛ	c. xmuz _i tɛ		*!	
	d. ☹️xmuz _i tɛ			*

While the ranking argument in (7) selects the attested outputs, it does not resolve the issue of variation. For instance, in order to arrive at the two alternative outputs for ‘grill’ ([gril] ~ [gril]) without abandoning categorical ranking, we would need to posit two alternative inputs, one indexed with the diacritic F, /gril_F/, and one unindexed, /gril/. The predictions of such a model are unclear. It seems that, by default, novel forms entering the lexicon should be unindexed, as indexing refers to lexical exceptions. Under these assumptions, however, novel forms would always undergo Retraction, which is counterfactual. If, on the other hand, novel forms were to be marked as exceptional by default, they would never undergo Retraction. Such an outcome is also unwelcome, as shown by the example of ‘grill’.

⁴ Actually, PAL-*i* is always satisfied on the surface. The two strategies of satisfying PAL-*i* involve retraction and palatalization, the latter yielding [-bk] consonants, e.g., [sⁱinus]. For the clarity of presentation, I ignore palatalization until Section 4.

⁵ A lexically indexed constraint targets only forms that bear the same diacritic as the constraint; e.g., the constraint IDENT-X targets the outputs that bear the diacritic X.

As demonstrated in this section, neither a serial rule-based framework nor a parallel OT framework enriched with lexical indexing is able to successfully predict the behavior of novel forms *vis-à-vis* Retraction without creating major analytical problems. Moreover, the presented approaches fail to systematically generate variation. In effect, the modelling fully depends on a stipulative diacritic assignment to individual lexical entries. Additionally, some forms must involve allomorph listing, where the process applies to one allomorph, but it is blocked when the other allomorph is used.

3. Variation and probabilistic ranking in OT

A classic example of variation in phonology is nasal assimilation (Boersma 1997). For instance, the hypothetical input /an+pa/ may surface as [ampa] or [anpa]. The former involves place assimilation and the latter is a faithful parse of the input. Let us assume that assimilation is driven by AGREE(Place) which requires the agreement of the place features on adjacent consonants (cf. Lombardi 1999). The faithfulness constraint militating against changes in place features between the input and the output is ID(Place). Assuming the principle of strict domination familiar from classic OT, the ranking AGREE(Place) >> ID(Place) predicts that [ampa] is the only possible output form. The reverse ranking, on the other hand, always selects the competing [anpa]. Categorical ranking does not, however, predict variation: [ampa] ~ [anpa].

In order to account for variation, it has been proposed that categorical ranking be replaced with probabilistic ranking. For instance, Boersma (1997) represents a line of research whereby OT constraint ranking is treated as the probability of one constraint outranking another. Using probabilistic ranking (represented by a double line) of AGREE(Place) *vis-à-vis* ID(Place), we arrive at a certain probability of assimilation in the output of /an+pa/.

(8) Probabilistic ranking in OT

<i>In</i>	<i>Out</i>	AGREE(Place)	ID(Place)
anpa	a. anpa	*	
	b. ampa		*

The model presented in tableau (8) introduces a built-in mechanism that is able to successfully predict the rate in which a given process applies. Accordingly, the probability P of place assimilation in /an+pa/ equals the probability with which AGREE(Place) outranks ID(Place). Therefore, P refers to the constraint ranking rather than to the particular lexical entry. Assuming

a language where outputs (8a) and (8b) surface with the same probability of 50%, the ranking argument in (8) predicts that a similar input, e.g., /kan+pa/, will receive similar rates of assimilation: 50% for [kanpa] and 50% for [kampa]. This is due to the fact that the constraints in (8) equally affect the nasal-stop clusters in all evaluations using the same probability of one outranking the other.

A problem with such an approach arises when similar inputs do not behave similarly on the surface. Using our hypothetical examples, if variation is observed in /an+pa/, but /kan+pa/ remains stable and always surfaces as [kampa], the predictions of a probabilistic ranking in (8) become inadequate.

An example of the issue of variation and stability is found in Tagalog nasal substitution (Zuraw 2000). In Tagalog, there are established forms undergoing the process as well as established forms not undergoing the process, e.g., /maŋ+RED+bigkas/ → [mambibigkas] ‘reciter’, /maŋ+bigaj/ → [mamigaj] ‘to distribute’ (Zuraw 2010: 421). Moreover, there are forms exhibiting variation, e.g., /maŋ+blog/ → [mamlog] ~ [mamblog] ‘to blog’ (Zuraw 2010: 445). Finally, the distribution of nasal substitution in nonce forms is also probabilistic (Zuraw 2000, 2010).

In her analysis, Zuraw (2000, 2010) proposes a combination of probabilistic OT rankings and lexicalization (storage) of entire words. The latter is responsible for the stability of forms in which nasal substitution either always applies or never applies. The former, on the other hand, is responsible for on-line computation of forms not established in the lexicon. These may include novel borrowings, low frequency words or nonce forms.

The approach to grammatical processing advocated by Zuraw (2000, 2010) is promising for the Polish data regarding Retraction, discussed in Section 2. Consider the evaluation of candidates for /gril/ ‘grill’ in (9). Recall that the attested outputs involve both [gril] and [gril].

(9) gril → gril ~ gril ‘grill’

<i>In</i>	<i>Out</i>	PAL- <i>i</i>	IDV
gril	a. gril	*	
	b. gril		*

In the tableau in (9), both candidates have a chance to win. It is the job of the probabilistic ranking to establish with what probability each output emerges. A probabilistic ranking of PAL-*i* and IDV predicts that all forms that are targeted by these constraints should exhibit variation. Such a ranking, therefore, is incorrect for forms where categorical results are found.

Similarly to the case of Tagalog nasal substitution outlined above, a solution to this issue can be found in the theory of memory representation. In classic generative accounts, it is assumed that lexical representations are minimal and redundancy-free (Chomsky & Halle 1968; Kenstowicz 1994). Albeit OT introduces some changes to the view on the lexicon, such as Lexicon Optimization (Prince & Smolensky 2004), it still assumes that, in principle, morphemes possess a single UR. However, numerous works indicate that lexical storage need not be limited to morphemes, but it can also include derived words (Bybee 2001; Pierrehumbert 2001; Booij 2002). Moreover, even classic generative accounts admit storage of allomorphs or even entire words/phrases if no other option is available (Mester 1994; Kager 1996; Rubach & Booij 2001). This suggests that the hypothesis of entire words/phrases being stored in the lexicon and invoked as such does not necessarily stand in opposition to the philosophy of the generative grammar. Importantly, phonological processing must be part of the mental grammar regardless of the theory of memory representation. It is necessary for deriving, for instance, novel forms. However, treating processing and rich storage as mutually excluding may be a fallacy (Langacker 1987; Zuraw 2000, 2010; Sanders 2003; Czaplicki 2014).

The approach outlined above can be applied to Polish Retraction. Throughout the ensuing analysis, two principal assumptions are maintained: (i) phonological computation is necessary in order to derive new words and (ii) whole-word storage is necessary in order to account for the established forms. Accordingly, only new (or forgotten/infrequent) forms undergo full processing. The established forms are invoked from the lexicon and effectively skip part of the phonological computation. Finally, the proposed analysis assumes that the listed words must also possess URs. Otherwise, it would be difficult to draw generalizations over patterns in the grammar.

In the evaluations in (10), the indexed constraint IDENT-LISTED (ID-L) assures that the stored input is preserved in the output (Zuraw 2000). Importantly, the diacritic (L)isted is assigned to the whole words stored in the lexicon and not to individual morphemes. Additionally, I assume that stored words have priority over individual morphemes as inputs to evaluations. This assumption is in accordance with the dual-route models of lexical access (e.g., McQueen & Cutler 1998; Hay 2003). In this view, phonological computation of abstract URs is more costly than invoking an entire word from the lexicon. Tableau (10i) uses the listed word /sinus_L/ ‘sine’ as input; in contrast, tableau (10ii), where a recent borrowing /gril/ is evaluated, must use the UR as input since no stored form is available.

Although in a probabilistic grammar all constraints are by definition ranked probabilistically, the crucial probabilistic ranking that may play a role in the evaluations is marked by a double line.

(10) Polish Retraction – lexical storage and probabilistic ranking

<i>In</i>	<i>Out</i>	ID-L	PAL- <i>i</i>	IDV
(i) sinus _L	a. sinus_L		*	
	b. sinus_L	*!		*
(ii) gril	c. gril		*	
	d. gril			*

Words listed in the lexicon are stable and effectively yield categorical results. As shown in (10i), the undominated constraint ID-L ensures the victory of the faithful candidate, (10a). In (10ii), on the other hand, ID-L is mute and the probabilistic ranking comes into play. Effectively, both candidates, (10c) and (10d), have a certain probability of winning. Such a model predicts that novel forms exhibit variation as they are not listed in the lexicon. Since it falls outside the scope of the current paper, I do not elaborate on the specific probabilities in the ranking. These are most likely established on the basis of type frequency of the alternation (e.g., Bybee 2001).

It is the coexistence of whole-word storage and probabilistic constraint ranking that not only generates the stable attested outputs, but also predicts the outcome of processes exhibiting variation. The cost of such an analysis is similar to the cost of the analyses based on lexically indexed constraints and/or allomorph listing (see Section 2), both widely admitted in the framework of OT. The upshot of the current approach, however, is that both lexical exceptionality and phonological variation are accounted for in a uniform manner.

4. Morphological conditioning and palatalization

So far, two important facts about Retraction have been largely ignored. First, Retraction applies variably to strings inside words, but it applies categorically across morpheme boundaries. Second, Retraction counterbleeds palatalization that occurs across morpheme boundaries.⁶ These facts are not readily expressed using the model presented in the preceding section. A solution proposed in this section employs constraints

⁶ Rule B counterbleeds rule A if rule A applies first and rule B destroys the environment/structural change of rule B (see Baković 2011).

referencing morphological affiliations of segments. It is shown that the proposed approach is compatible with the existing analyses couched in the framework of Derivational OT (Rubach 1997). The discussion sets off with the analyses proposed by Rubach (2003a, 2017), who employs only categorical rankings that generate Retraction. The final model presented in this section combines Derivational OT with the probabilistic approach to the constraint ranking outlined in the preceding section.

As discussed in Section 2, the motivation for the process of Retraction in Polish comes from the morphophonological alternation on the one hand and the incorporation of borrowings into the language, on the other hand. While variation is observed in some borrowings, e.g., [gril] ~ [gril] ‘grill’, the morphophonological process is obligatory and variation is not attested (see Section 2). Whole-word storage does not resolve this issue since the suffixes undergoing Retraction are fully productive and they can readily attach to any novel word, producing categorical results. In other words, while Retraction inside morphemes in novel forms is predicted to exhibit variation, Retraction affecting suffixes applies categorically whenever its context is met. This fact creates problems for the constraint set proposed in Section 3, which treats Retraction as a homogeneous process affecting all strings regardless of the morphological boundaries.

The second problem regarding the fully parallel analysis of Retraction is its interaction with palatalization. The actual mappings found in the data familiar from Section 2 are given in (11).

(11) Retraction and palatalization (Rubach 1984a: 202)

- | | | |
|----|----------------------|---------------------------|
| a. | xmur+a ‘cloud’ | xmuz+i+te ‘become cloudy’ |
| | xələnder ‘Dutchman’ | xələndz+i nom.pl. |
| b. | kəs+a ‘scythe’ | koe+i+te ‘mow’ |
| | frantsuz ‘Frenchman’ | frantsuz+i nom.pl. |

As visible in (11), apart from Retraction, the data exhibit an additional stem alternation affecting stem-final segments. According to classic analyses, this is an effect of palatalization triggered by a high front vowel across morpheme boundaries (e.g., Steele 1973; Gussmann 1978; Rubach 1984a, 2003a, 2003b). In the examples in (11b), palatalization is overt; that is, the triggering /i/ is present on the surface and the output of palatalization involves [-bk] consonants. In the retracting examples in (11a), on the other hand, palatalization is opaque, including a double counterbleeding effect. First, spelled-out palatalization yields non-palatalized [+bk] segments, so the focus of palatalization is opaque. Second, the trigger of palatalization, the

high front vowel /i/, is retracted to /i/, so the environment of palatalization is opaque. Such a mapping is impossible to model in a strictly parallel OT framework (Rubach 2003a, 2003b).

Let us first consider the asymmetric application of Retraction inside morphemes and across morpheme boundaries, ignoring palatalization effects. The evaluation in (12) uses the familiar examples /gril/ ‘grill’ and /xmur+i/ ‘become cloudy’ (3rd sg.) to illustrate this issue in a strictly parallel framework that employs probabilistic constraint ranking, outlined in Section 3. The morpheme-internal string /ri/ may variably undergo Retraction; however, variation is not permitted across the morphological boundary after adding the verbalizing suffix and hence the only possibility should be palatalization and Retraction, [...z̥iɛ]. The tableau in (12) is highly simplified.

(12) Non-homogeneity of Retraction – failed evaluation

<i>In</i>	<i>Out</i>	PAL- <i>i</i>	IDV
(i) gril	a. gril	*	
	b. gril		*
(ii) xmur+i+tɛ	c. $\dots z̥iɛ$	*	
	d. $\dots z̥iɛ$		*

The model in (12), with a crucially probabilistic ranking of PAL-*i* and IDV, incorrectly predicts variation in the suffix. A solution involves a morphologically conditioned identity constraint, IDV- $\sqrt{\quad}$, targeting vowels affiliated with morphological roots (McCarthy & Prince 1995; Kager 2000). A probabilistic ranking of PAL-*i* and IDV- $\sqrt{\quad}$, with a low ranked generic IDV, assures that root-internal vowels variably undergo Retraction, while suffix vowels are retracted categorically.

(13) Non-homogeneity of Retraction

<i>In</i>	<i>Out</i>	PAL- <i>i</i>	IDV- $\sqrt{\quad}$	IDV
(i) gril	a. gril	*		
	b. gril		*	*
(ii) xmur+i+tɛ	c. $\dots z̥iɛ$	*!		
	d. $\dots z̥iɛ$			*

As demonstrated in (13i), both candidates can win the evaluation due to a possibility of reranking PAL-*i* and IDV- $\sqrt{\quad}$. The constraints PAL-*i* and IDV, on the other hand, are ranked apart with a zero probability of reranking, which assures that Retraction always applies across morpheme boundaries, as shown in (13ii).

The model presented so far is compatible with previous analyses of Polish palatalization and Retraction in OT (Rubach 2003a, 2003b, 2017), which show that parallel one-step OT is unable to account for the complexity of the Polish data. This is due to a double counterbleeding effect, where both the focus and the environment of palatalization are rendered opaque. A solution is offered by Derivational OT (Rubach 1997), which casts the OT evaluation in three (or four) levels of computation that roughly correspond to lexical and postlexical strata, familiar from Lexical Phonology (Kiparsky 1982; Booij & Rubach 1987). In Derivational OT, the output of the previous level becomes the input to the next level. Constraints may be reranked between levels.

As argued by Rubach (2003a, 2003b, 2017), Retraction is part of a conspiracy to satisfy the requirements of PAL-*i*. The two main strategies of satisfying PAL-*i* involve Retraction and palatalization. Palatalization applies across morpheme boundaries as well as in the postlexical phonology (across the board to all strings of consonants followed by /i/, yielding palatalized consonants).⁷ Retraction, on the other hand, applies in the domain of the phonological word.

In the case of velars as well as /r/, the outcome of palatalization is an unexpected [+bk] consonant. This is due to an interplay of two additional markedness constraints at the lexical level (e.g., Rubach 2003a).

- (14) POSTERIOR (Rubach 2003b: 219)
 Do not be [Coronal, +anterior, -bk].
 HARD (Rubach 2003a: 609)
 Posterior stridents must be [+bk].

POSTERIOR penalizes palatalized anterior coronals, such as /r^j/. HARD, on the other hand, militates against palatalized posterior fricatives and affricates, that is, /s^j z^j t^s d^z/. The change /ri/ → /z^ji/ satisfies PAL-*i* at the stem level, where HARD is low ranked. However, this effect is counterbled by a further change taking place at the word level, where palatalized posterior stridents, such as /z^j/, are rejected by a high ranked HARD. This reranking between Level 1 and Level 2 also contributes to Retraction, which proves the most optimal strategy at this point, effectively yielding a mapping /z^ji/ → /zⁱ/. Finally, phrase level phonology again favors palatalization over retraction, changing the specification of all consonants followed by /i/ to [-bk]. This is expressed in Derivational OT by a demotion of POSTERIOR (Rubach 2003a, 2003b, 2017).

⁷ Polish palatalization can be descriptively divided into phonemic and allophonic. However, in OT, both are triggered by a single constraint, PAL-*i* (Rubach 2003b: 204).

gril	g. gril							*	
	h. $\text{gr}^{\text{j}}\text{il}$		*!		*				
	i. gril					*			*
	j. $\text{gz}^{\text{j}}\text{il}$			*!	*	*			
	k. gzil					*!	*		*
sinus _L	l. sinus_L							*	
	m. sinus_L	*!					*		*

At the stem level, the evaluation proceeds exactly as predicted in Rubach (2003a, 2017). At the word level, however, a crucially probabilistic ranking is established between IdV[-bk]- $\sqrt{\quad}$ and PAL-*i*. Such a ranking is able to generate variable Retraction inside words without endangering the categorical nature of Retraction across morpheme boundaries. The latter is assured by the low ranking of the generic IdV[-bk]. At the phrase level (not included for the reasons of space), Surface Palatalization (Rubach 1984a) applies across the board to all consonants followed by /i/, e.g., deriving [gr^jil] from /gril/ ‘grill’ as well as [s^jinus] from /sinus/ ‘sine’ (cf. Rubach 2003a). This is achieved by the promotion of PAL-*i* and the vowel faithfulness constraints to an undominated position.

The upshot of the proposed analysis is that it draws a systematic distinction between the morphologically conditioned Retraction and the phonological Retraction, without losing the generalization that the driver of both phenomena is PAL-*i*. The current model predicts that the stem-level Retraction will remain a categorical process. This is corroborated by the formation of verbs based on novel nouns. For instance, the Polish slang expression *hardcore*, which is used as a masculine noun [xartkɔr], may serve as a base to form a verb. After adding the verbalizing suffix, the only possible outcome is [xartkɔz^jitɕ], with the palatalization of the stem-final consonant and the retraction of the suffix /i/.⁸

Word-level Retraction, on the other hand, may be subject to synchronic variation as well as substantial diachronic change, the latter being expressed by the decreasing probability of PAL-*i* outranking IDV[-bk]- $\sqrt{\quad}$. For instance, Rubach (1984a: 203–206) discusses the incorporation of borrowings from the perspective of Retraction, pointing out that older borrowings tend to comply with Retraction, while recent borrowings tend to exhibit variation.

⁸ Although this form is mostly used in spoken language, a corpus search confirms that <hardkorzy>/<hardcorzy>, where <rz> corresponds to /z/ and <y> to /i/, is found with a joint frequency of 0.006 per million, while non-retracted forms, e.g., <hardcori>/<hardcorzi>, are not attested (Polish Web Corpus plTenTen12 accessed via Sketch Engine; Jakubiček et al. 2013).

The current model is not only compatible with these observations, but also restricts possible future changes. Namely, established words that normally resist Retraction, such as *sinus* ‘sine’, might be subject to slips of the tongue or even restructuring. While *sinus* is generally immune to Retraction, it might still surface as *[sinus] due to a slip of the tongue. In such cases, ‘sine’ would be treated as generated from an underlying form (as opposed to the listed form), which renders ID-L mute and allows for Retraction with a certain probability. The reverse, however, is impossible; that is, an established form containing an underlying back vowel /i/, e.g., [limuzina] ‘limousine’, will never surface as *[limuz^hina], since such an output does not improve on any markedness constraint and hence is harmonically bounded by the fully faithful candidate.

5. Conclusion

The analysis shows that a comprehensive account of Polish Retraction requires a combination of whole word storage (expressed by lexically indexed constraints), probabilistic constraint ranking as well as reference to morphology in the phonological computation. These strategies, when used separately, either create major analytical problems or fail to generate the attested outputs. For instance, lexical indexing correctly accounts for the behavior of the forms present in the lexicon, but fails to predict the behavior of novel forms. Probabilistic ranking is able to generate variation, but it fails to generate stability. Each of the aforementioned components is necessary since Retraction exhibits simultaneous stability and variation, an asymmetry between the application in different morphological domains as well as an opaque interaction with palatalization. The final model is able to predict the behavior of the established forms as well as novel forms, both inside morphemes and across morpheme boundaries. Moreover, it correctly predicts that phrase level phonology applies to all forms regardless of their representation in the lexicon.

Literature

- Baković E. (2011): *Opacity and ordering*. [In:] *The Handbook of Phonological Theory*. J. Goldsmith, J. Riggle & A.C.L. Yu (eds). Cambridge, MA & Oxford, pp. 40–67.
- Boersma P. (1997): *How we learn variation, optionality, and probability*. „Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam” 21, pp. 41–58.
- Boersma P. & B. Hayes (2001): *Empirical tests of the Gradual Learning Algorithm*. „Linguistic Inquiry” 32, pp. 45–86.

- Booij G. (2002): *The balance between storage and computation in phonology*. [In:] *Storage and Computation in the Language Faculty. Studies in Theoretical Psycholinguistics*. S. Nootboom, F. Weerman & F. Wijnen (eds). Dordrecht, pp. 133–156.
- Booij G.E. & J. Rubach (1987): *Postcyclic versus postlexical rules in Lexical Phonology*. „Linguistic Inquiry” 18, pp. 1–44.
- Bybee J. (2001): *Phonology and language use*. Cambridge.
- Chomsky N. & M. Halle (1968): *The sound pattern of English*. New York.
- Czaplicki B. (2014): *Lexicon Based Phonology: Arbitrariness in Grammar*. Munich.
- Fukuzawa H. (1999): *Theoretical implications of OCP effects on features in Optimality Theory*. [PhD dissertation, University of Maryland].
- Gussmann E. (1978): *Contrastive Polish-English consonantal phonology*. Warsaw.
- Hay J. (2003): *Causes and consequences of word structure*. London.
- Itô J. & A. Mester (1999): *The phonological lexicon*. [In:] *Handbook of Japanese Linguistics*. N. Tsujimura (ed.). Oxford, pp. 62–100.
- Jakubiček M., A. Kilgarriff, V. Kovář, P. Rychlý & V. Suchomel (2013): *The TenTen corpus family*. „7th International Corpus Linguistics Conference CL”, pp. 125–127.
- Kager R. (1996): *On affix allomorphy and syllable counting*. [In:] *Interfaces in phonology. Studia Grammatica 41*. U. Kleinhenz (ed.). Berlin, pp. 155–171.
- Kager R. (2000): *Surface opacity of metrical structure in Optimality Theory*. [In:] *The derivational residue in phonology*. B. Hermans & M. van Oostendorp (eds). Amsterdam, pp. 207–247.
- Kenstowicz M. (1994): *Phonology in Generative Grammar*. Cambridge, MA.
- Kiparsky P. (1982): *From Cyclic to Lexical Phonology*. [In:] *The Structure of Phonological Representations*. Vol. I. H. van der Hulst & N. Smith (eds). Dordrecht, pp. 131–75.
- Langacker R. (1987): *Foundations of cognitive grammar*. Vol. 1: *Theoretical prerequisites*. Stanford, CA.
- Lombardi L. (1999): *Positional faithfulness and voicing assimilation in Optimality Theory*. „Natural Language & Linguistic Theory” 17, pp. 267–302.
- McCarthy J. & A. Prince (1995): *Faithfulness and reduplicative identity*. [In:] *University of Massachusetts Occasional Papers in Linguistics 18*. J.N. Beckman, L.W. Dickey & S. Urbanczyk (eds). Amherst, MA, pp. 249–384.
- McQueen J. & A. Cutler (1998): *Morphology in word recognition*. [In:] *The handbook of morphology*. A. Spencer & A. Zwicky (eds). Oxford, pp. 406–427.
- Mester A. (1994): *The quantitative trochee in Latin*. „Natural Language & Linguistic Theory” 12, pp. 1–61.
- Pater J. (2000): *Nonuniformity in English stress: the role of ranked and lexically specific constraints*. „Phonology” 17, pp. 237–274.
- Pierrehumbert J. (2001): *Exemplar dynamics: Word frequency, lenition and contrast*. [In:] *Frequency and the Emergence of Linguistic Structure*. J.L. Bybee & P.J. Hopper (eds). Amsterdam, pp. 137–158.
- Prince A. & P. Smolensky (2004): *Optimality Theory: Constraint interaction in Generative Grammar*. Oxford. [Revision of 1993 technical report, Rutgers University Center for Cognitive Sciences. Available on Rutgers Optimality Archive, ROA-537.]
- Rubach J. (1984a): *Cyclic and lexical phonology: The structure of Polish*. Berlin–New York.
- Rubach J. (1984b): *Segmental rules of English and Cyclic Phonology*. „Language” 60, pp. 21–54.
- Rubach J. (1997): *Extrasyllabic consonants in Polish: Derivational Optimality Theory*. [In:] *Derivations and Constraints in Phonology*. I. Roca (ed.). Oxford, pp. 551–581.
- Rubach J. (2000): *Backness switch in Russian*. „Phonology” 17, pp. 39–64.
- Rubach J. (2003a): *Duke-of-York derivations in Polish*. „Linguistic Inquiry” 34, pp. 601–629.
- Rubach J. (2003b): *Polish palatalization in derivational optimality theory*. „Lingua” 113, pp. 197–237.

- Rubach J. (2017): *Derivational meanders of High Vowel Palatalization in Polish*. „Lingua” 199, pp. 1–26.
- Rubach J. & G. Booij (2001): *Allomorphy in Optimality Theory: Polish Iotation*. „Language” 77, pp. 26–60.
- Sanders R. (2003): *Opacity and sound change in the Polish lexicon*. [PhD dissertation, UC Santa Cruz].
- Steele R.D. (1973): *The segmental phonology of contemporary Standard Polish*. [PhD dissertation, Harvard University].
- Zuraw K. (2000): *Patterned exceptions in Phonology*. [PhD dissertation, University of California].
- Zuraw K. (2010): *A model of lexical variation and the grammar with application to Tagalog nasal substitution*. „Nat Lang Linguist Theory” 28, pp. 417–472.
- „Sketch Engine”, <www.sketchengine.eu>, accessed: 15.03.2025.