The strength of morphophonological schemas: Consonant mutations in Polish

BARTŁOMIEJ CZAPLICKI (iD)

]u[ubiquity press

ABSTRACT

Selective elimination of consonant mutations in Polish provides evidence supporting construction-based sublexicons and morphophonological schemas extracted from them. Morphophonological schemas exhibit various strength depending on their type frequency, they refer to morpheme-specific classes of segments and their impact is continuously mediated by paradigm uniformity pressures. A low-frequency and a high-frequency pattern are analyzed: agent nouns in *-ist-a/-yst-a* and diminutives in *-ek*. Two kinds of frequency are important predictors of pattern modification: type frequency and token frequency. As for the impact of type frequency, the less frequent a pattern is in the lexicon, the more susceptible it is to modifications promoted by paradigm uniformity. Schemas are ranked based on the type frequency of the morphological patterns they encode and interact with paradigm uniformity constraints. Schemas representing less frequent patterns are outranked by paradigm uniformity constraints and are thus more likely to be modified than those representing more frequent patterns. Regarding token frequency, the greater stability of high-frequency words than comparable low-frequency words is linked to their strong representations and a constraint promoting the use of stored representations. The continuous effect of paradigm uniformity effects gets support from the results of a nonce-word experiment.

CORRESPONDING AUTHOR:
Bartłomiej Czaplicki
University of Warsaw, PL
bczaplicki@uw.edu.pl

1 INTRODUCTION

Usage-based linguistics has provided ample evidence of the pervasiveness of frequency effects in language processing (see Bybee 2001 and Ellis 2002 for reviews). In contrast, frequency of use did not play a significant role in mainstream generative phonology until recently. It was partly due to the common assumption that words are generated on-line from their component parts (i.e. morphemes) (e.g. Chomsky & Halle 1968; Kiparsky 1982). Token frequency (i.e. the usage frequency of particular words) is an inherent property of whole words and the assumption that the words are generated through the concatenation of abstract morphemes makes it impossible to associate frequency values with stored representations. In classical generative approaches, phonological analyses lack the right tools for differentiating frequent from rare words. As a result, the observation that some (morpho)phonological processes apply or fail to apply to words based on their frequency either had to go unnoticed or be treated as due to extragrammatical factors. The advent of transderivational approaches using output-output constraints (Burzio 1996; Kenstowicz 1996; Benua 1997; Steriade 2000) within Optimality Theory (OT, Prince & Smolensky 1993/2004) made the task of representing frequency in the grammar easier. In these approaches, frequency values, which are ascribed to independently stored words (including morphologically complex words), can be used to account for the application or blocking of a process in a particular word or group of words.

In usage-based models morphological and phonological patterns (including segmental alternations) are represented in terms of schemas (Bybee 2001, Dąbrowska 2004). An example of a schema specifying English Past Tense in verbs like *stopped, begged* and *wanted* is given in (1) (Bybee 2001: 126).

(1)   a Past verb ends in [t], [d], or [ɨd]

Schemas are morphologically conditioned and their strength is a function of the frequency of the pattern they encode. The finding that the type frequency of a morphological pattern, that is, the number of words the pattern applies to, determines its productivity is referred to as gang effects (McClelland & Elman 1986; Stemberger & MacWhinney 1988; Alegre & Gordon 1999). The higher the number of words that adhere to a given pattern (i.e. the larger the gang), the more likely the pattern is to become extended to novel words. The type frequency of the schema in (1) corresponds to the number of English verbs (the size of the gang) that form their past tense using [t], [d], or [ɨd].

In addition to evidence pointing to the important role of type frequency, there is also evidence that token frequency, that is, the frequency with which a word is used, is relevant in pattern generalization.[1] High-frequency words are more likely to undergo phonetic reduction, while low-frequency words are more prone to analogical leveling (Mańczak 1980, Bybee 2001). Frequency plays a crucial role in dual-route models of lexical access (McQueen and Cutler 1998; Hay 2003; Plag 2012). A morphologically complex word can be accessed via the whole-word route (i.e. by accessing its stored whole-word representation) or the decomposed route (i.e. by accessing its component morphemes). The choice has been shown to depend on the relative frequencies of the derivative and its base. For example, in English the derivative *business* has a much higher frequency than its base *busy*. This entails that *business* is more likely to be accessed via the whole-word route.[2] Conversely, *blueness* is used less frequently than its base *blue* and, therefore, is predicted to show an advantage for the decomposed mode during access (Plag 2012; but cf. Hahn & Nakisa 2000, who, on the basis of German plurals, argue that dual-route models make the wrong predictions).

In the present analysis of consonant mutations in Polish, both type and token frequency are shown to have an impact on paradigm uniformity effects. Mutations are eliminated from a low-frequency morphological pattern, agent nouns in *-ist-a/-yst-a,* while a high-frequency pattern, diminutives in *-ek,* remains stable. It is argued that high-frequency patterns resist paradigm uniformity pressures due to their robust representations in the grammar. Low-frequency

---

1   I follow Bybee (2001) and others in distinguishing type from token frequency effects. Arguably, type frequency effects can in principle be derived from the sum of token frequencies.

2   The fact that *business* is synchronically only weakly associated with its base *busy* is fully consistent with the predictions of dual-route models of lexical access.

patterns, on the other hand, are represented with low-ranked schema-constraints and, thus, are more susceptible to modifications. It is also demonstrated that high frequency words (i.e. words with a high token frequency) show increased stability even if the pattern (schema) they represent displays a low type frequency.[3]

The paper is structured as follows. Section 2 provides the main data and introduces the basic elements of the analysis: frequency in phonology, allomorphy, schemas and cophonologies. It also discusses some representational aspects of consonant mutations in Polish. In Section 3, the formation of agent nouns in *-ist-a/-yst-a,* a low-frequency pattern, is discussed. This section focuses on the selectivity of consonant mutations. First, we look at the phonological conditioning responsible for the emergence of the pattern. Second, modern-day complexities resulting in the elimination of mutations are accounted for using a combination of phonological constraints and morphophonological schemas. This section also reports the results of a nonce-word experiment, which point to an on-going change driven by paradigm uniformity. In Section 4, we consider diminutives in *-ek,* a high-frequency pattern which results in stable mutations. Section 5 offers a discussion of the main implications of the analysis and Section 6 provides the conclusions.

2 DATA AND BASIC ASSUMPTIONS

2.1 A LOW-FREQUENCY PATTERN: AGENT NOUNS IN *-IST-A/-YST-A*

The suffix *-ist-a/-yst-a* [ista/ɨsta] is used to form agent nouns in Polish. It is of Latin/Greek origin and entered Polish at the turn of Old and Middle Polish, that is, around the 15th century (Długosz-Kurczabowa & Dubisz 2006: 369). The suffix is productive and no longer restricted to Latinate words. In spite of its long history in Polish, the suffix remains a low-frequency pattern. The quantitative data to back this claim will be given in Section 2.3. We start out with some facts about the distribution and consonant mutations triggered by the two variants of the suffix: *-ist-a* and *-yst-a.* I begin with examples illustrating the usage of the variant *-ist-a.* The data are drawn from Gussmann (2007: 157–161). For the purposes of this analysis, a consonant mutation is defined as a featural change in a consonant that results in a phonologically distinct (i.e. contrastive) segment. There is convincing evidence that palatalization of labials, dentals and velars before [i], [pʲ bʲ fʲ vʲ mʲ tʲ dʲ sʲ zʲ c ɟ ç], is a non-contrastive, coarticulatory effect of the following vowel. First, labials, dentals and velars do not contrast for palatality preconsonantally and word finally. Second, Święciński's (2014) acoustic study shows that [pʲ bʲ fʲ vʲ mʲ tʲ dʲ sʲ zʲ c ɟ ç] appear exclusively before the vowel [i]; before other vowels palatalization is manifested on a distinct segment, the palatal glide [j], e.g. *piasek* [pjasɛk] 'sand', *diabeł* [djabɛw] 'devil' and *kiosk* [kjɔsk] 'kiosk'. In light of this, [pʲ bʲ fʲ vʲ mʲ tʲ dʲ sʲ zʲ c ɟ ç] should not be regarded as mutated variants of [p b m f v t d s z k g x]. Such coarticulatory palatalization is omitted from the transcription in this paper, as it does not represent an instance of mutation in the relevant sense.[4]

---

3   A reviewer points out that it is in principle possible to extend the classical generative model in such a way as to associate frequency values directly with the morphemes which constitute the respective words. While an extension along these lines definitely deserves more attention, it does not discount the basic insight. Token frequency is a property of individual words, whereas type frequency is a property of an aggregate of words complying with a particular morphological pattern. In addition, while effectively modeling the effects of type frequency of affixes, such a solution might encounter problems with modeling the effects of the token frequency of complex words. In *blueness* both *blue* and *-ness* have a high token frequency (separately), but *blueness* as a complex word has a low token frequency. Insofar as *blueness* patterns with other words of low frequency, its behavior cannot be derived from the frequencies of its component morphemes. It must be derived from the token frequency of the entire word. Thus, in addition to the frequency of the component morphemes, such a model would still have to refer to the token frequency of the word.

4   A reviewer claims that labials *do* in fact contrast for palatality in Polish and refers to the well-known morphological regularity that involves the usage of two different case endings for stem-final labials. For example, the LOC SG has two possible endings *-u* and *-e.* The LOC SG of *karp* [karp] 'carp' selects *-u,* i.e. *karpi-u* [karpj-u], while the LOC SG of *sęp* [sɛmp] 'vulture' selects *-e,* i.e. *sępi-e* [sɛmpj-ɛ]. While it is true that the choice of the endings was *historically* motivated by the contrastive specification of labials (palatalized vs. non-palatalized), the synchronic grammar of contemporary speakers need not reflect this contrastive specification. In the usage-based approach adopted here the solution is simple. The two patterns of behavior of labials in morphology are represented by two competing schemas. For example, for base-final [p] the schemas representing LOC SG are [...p] ↔ [[...pj]u] and [...p] ↔ [[...pj]ɛ]. For novel words, the competition between the two schemas is resolved in favor of the more frequent one. For example, the LOC SG of *laptop* [laptɔp] 'laptop' is *laptopi-e* [laptɔpj-ɛ], not **laptopi-u,* suggesting that [...p] ↔ [[...pj]ɛ] is the more frequent and hence productive one of the two. As for well-established words, given that frequent words are stored whole, the competition between the two endings is resolved in favor of the stored one. See Section 2.5 for the details of the proposal.

(2) a. *labials*
   służba [swuʐb-a] 'service'   służbista [swuʐb-ist-a] 'martinet'
   program [prɔgram] 'program'   programista [prɔgram-ist-a] 'programmer'
   rezerwa [rɛzɛrv-a] 'reserve'   rezerwista [rɛzɛrv-ist-a] 'reservist'
   finał [finaw] 'end'   finalista [final-ist-a] 'finalist'

   b. *palatalized labials*
   kopia [kɔpj-a] 'copy'   kopista [kɔp-ist-a] 'scribe'
   utopia [utɔpj-a] 'utopia'   utopista [utɔp-ist-a] 'utopian'
   biografia [bjɔgrafj-a] 'biography'   biografista [bjɔgraf-ist-a] 'biographer'

   c. *coronals*
   tenis [tɛɲis] 'tennis'   tenisista [tɛɲiɕ-ist-a] 'tennis-player'
   krajobrazu [krajɔbraz-u]   krajobrazista [krajɔbraʑ-ist-a]
   'landscape' GEN SG   'landscape painter'
   plan [plan] 'plan'   planista [plaɲ-ist-a] 'planner'
   flet [flɛt] 'flute'   flecista [flɛtɕ-ist-a] 'flautist'
   ballada [ballad-a] 'ballad'   balladzista [balladʑ-ist-a] 'ballad writer'
   baseball [bɛjzbɔl] 'baseball'   baseballista [bɛjzbɔl-ist-a] 'b. player'

   d. *palatalized coronals*
   hokej [xɔkɛj] 'hockey'   hokeista [xɔkɛ(j)-ist-a] 'hockey player'

   e. *velars*
   Franco [frankɔ] 'Franco'   frankista [frank-ist-a] 'Frankoist'
   czołgu [tʂɔwg-u] 'tank' GEN SG   czołgista [tʂɔwg-ist-a] 'tank-driver'
   szachy [ʂax-ɨ] 'chess'   szachista [ʂax-ist-a] 'chess player'

The items in (2) show that the variant *-ist-a* is used after labials, (a), palatalized labials, (b), coronals [s z t d], (c), and velars, (e).[5] There are no cases of *-ist-a* after base-final palatalized coronals except after [j], shown in (d). The front glide is variably elided before *-ist-a*. Coronals (except [l]) show mutations before the suffix, while velars and labials (except [w]) do not. The patterning of base-final [l] in (c) and [w] in (a) requires an explanation. The coronal [l] does not alternate because it does not have a distinct mutated variant before [i]. The alternation [w ~ l] in (a) is a reflex of a historical alternation between a velarized /ł/ and a palatalized /lʲ/. It is omitted from further consideration for several reasons. First, it arose relatively recently due to the context-free change /ł/ > /w/. More time is necessary for paradigm uniformity pressures to take an effect on this alternation. Second, there are no words that end in [...w-ist-a] or [...w-ɨst-a]. In contrast, there is a large number of words in [...l-ist-a]: 135 (Bańko et al. 2003). There are two sources of such words: words that have a base in [...l], e.g. *motocykl-ist-a* 'motorcyclist' (< *motocykl*), and words that have no independently existing base and are not fully decomposable (i.e. they contain bound roots), e.g. *popul-ist-a* 'populist'. The existing words in [...l-ista] likely form an attractive bias for words in *-ist-a* with base-final [...w], perpetuating the [w ~ l] alternation. There is no comparable attractive bias from [...w-ista] or [...w-ɨsta], as these patterns are not attested.[6] The [w ~ l] alternation, thus, persists for lack of an alternative and for that reason it is not compatible with the other alternations.

The other variant, *-yst-a* [ɨst-a], is selected after base-final [r]. The rhotic alternates with the fricative /ʐ/ in the derived form, as illustrated in (3a). A context-free diachronic change /rʲ/ → /ʒʲ/ → /ʐ/ gave rise to this alternation in modern Polish. As exemplified in (3b), the variant *-yst-a* [ɨsta] also appears after alveolar and postalveolar (retroflex) affricates [ts dz tʂ dʐ] and the postalveolar (retroflex) fricatives [ʂ] and [ʐ], the latter fricative showing a similar behavior to the /ʐ/ ← /rʲ/.

---

5   The words *lobb-yst-a* 'lobbyist' and *hobb-yst-a* 'hobbyist' might seem exceptional to the distribution of the suffix alternants after labials. In fact, they are cases of "morphological absorption" (the term is commonly attributed to Mikołaj Kruszewski), where the stem-final vowel was reanalyzed as part of the suffix, i.e. *lobby* > *lobb-yst-a* and *hobby* > *hobb-yst-a.*

6   Admittedly, the influence of such product-oriented schemas as "agent nouns end in [...l-ista]" (see the discussion in Section 2.5) merits a closer look, but it would take us too far afield.

(3)　　a.　afera [afɛr-a] 'scandal'　　　　aferzysta [afɛʐ-ist-a] 'schemer'
　　　　b.　klasycyzm [klasɨts-izm] 'classicism'　klasycysta [klasɨts-ist-a] 'classicist'
　　　　　　brydża [brɨdʐ-a] 'bridge' GEN SG　brydżysta [brɨdʐ-ist-a] 'bridge player'
　　　　　　fetysz [fɛtɨʂ] 'fetish'　　　　　fetyszysta [fɛtɨʂ-ist-a] 'fetishist'

In addition to the words with coronals [t d r] that show mutations before *-ist-a*, illustrated in and (2) and (3), there are also those which do not evidence consonant mutations and choose the other shape of the suffix, as exemplified in (4). The distribution of the suffix alternants in (4) diverges from the distribution shown in (2) and (3).

(4)　　Bonaparte [bɔnapartɛ]　　　　　bonapartysta [bɔnapart-ist-a] 'supporter of
　　　　　　　　　　　　　　　　　　　　　　　　Bonaparte'

　　　　Conrada [kɔnrad-a] GEN SG　　　　konradysta [kɔnrad-ist-a] 'specialist in the
　　　　　　　　　　　　　　　　　　　　　　　　works of J. Conrad'

　　　　stypendium [stipɛndjum] 'stipend'　stypendysta [stipɛnd-ist-a] 'stipend holder'
　　　　parodia [parɔdj-a] 'parody'　　　　parodysta [parɔd-ist-a] 'parodist'
　　　　rygor [rɨgɔr] 'rigor'　　　　　　　rygorysta [rɨgɔr-ist-a] 'rigorist'

A number of words with the suffix *-ist-a* vacillate in current usage. Their bases end in [t d r]. This different behavior of words in [t d r] in (2) and (3), on the one hand, and (4) and (5), on the other, will be linked to differences in their token frequency in Section 3.3.2.

(5)　　altysta　　　　~　alcista　　　　　　'alto singer'
　　　　[alt-ist-a]　　　　[altɕ-ist-a]
　　　　propagandysta　~　propagandz-ist-a　'propagandist'
　　　　[prɔpagand-ist-a]　　[prɔpagandʑ-ist-a]
　　　　manierysta　　~　manierzysta　　　'follower of mannerism'
　　　　[maɲɛr-ist-a]　　　[maɲɛʐ-ist-a]

As the summary of the distribution of the suffix *-ist-a/-yst-a* in (6) shows, the *-ist-a* variant appears after labials and velars, while *-yst-a* occurs after [ts dz tʂ dʐ ʂ ʐ]. Base-final coronals [t d r] show a more complex pattern: both allomorphs are attested. Based on these distributional facts, it is assumed that *-ist-a* is the basic variant of the suffix, while *-yst-a*, whose usage is restricted to retroflexes and alveolar affricates (categorically) and [t d r] (variably), is a positional variant of the suffix. In usage-based models, it is assumed that relations between morphological units, including identity relations, are established on the basis of phonological and semantic similarity. The suffixes *-ist-a* and *-yst-a* are both phonologically and semantically similar. A rationale for this assumption is provided in Section 2.5, where the properties of schemas are fleshed out.

(6)

consonant	*suffix*	*outcome*	
labials: [p b m f v]	-ist-a	no mutations	[p b m f v]
coronals:			
[s z n]	-ist-a	mutations	[ɕ ʑ ɲ]
[t d r]	-ist-a/-yst-a	mutations/no mutations	[tɕ dʑ ʑ]/[t d r]
[ts dz ʂ ʐ tʂ dʐ]	-yst-a	no mutations	[ts dz ʂ ʐ tʂ dʐ]
[l j]	-ist-a	no mutations	[l], [j] ~ ∅
velars: [k g x]	-ist-a	no mutations	[k g x]

We now turn to evidence indicating that consonant mutations before front vowels are historically and synchronically motivated in Polish. As for the historical motivation, there is compelling evidence that CV coarticulation preceded and induced the emergence of distinctive palatalization in Slavic languages. Jakobson (1929/1962: 71ff.) and Andersen (1978: 11–15) argue that in Common Slavic palatalized consonants arose due to coarticulation with the following front vowel. The developments of the reconstructed Common Slavic *věra* 'faith' and *klětŭka* 'cage' in Polish are used as illustrative examples (ě traditionally stands for yat', a long open front vowel). At the stage of [vʲɛra] and [klʲɛtka], the palatality of the consonant was interpreted as coarticulatory and attributed to the following front vowel. A series of changes in the quality of the vowel, depicted in (7), was partly responsible for the phonologization of a palatalized

consonant.[7] When the vowel became back, as in [vʲara] and [klʲatka], the palatality could no longer be attributed to the vowel and was phonologized on the consonant, which gave rise to the distinctively palatalized /vʲ/ in /vʲara/ and /lʲ/ in /klʲatka/ (Andersen 1978: 11–15). A similar mechanism was likely involved in the development of mutations before *-ist-a/-yst-a*. Thus, phonetic reconstruction and comparative evidence indicate that palatalization is historically motivated before front vowels in Polish.

(7)   Common Slavic věra > vʲɛra > vʲɛara > vʲara > Modern Polish vjara
        Common Slavic klětŭka > klʲɛtka > klʲɛatka > klʲatka > Modern Polish klatka

Synchronic alternations in Modern Polish also provide motivation for mutations before front vowels. Suffix-initial front vowels commonly trigger consonant mutations (Rubach 1984). In (7) mutations before the adjectival suffix *-ist-y* are shown. Two types are directly relevant for the discussion at hand: Coronal Mutation and Velar Mutation. Coronal Mutation is applicable to coronals [t d s z n r] and results in alveolopalatals and a postalveolar [tɕ dʑ ɕ ʑ ɲ ʐ], as illustrated in (8a). Velar Mutation applies to velars [k g x] and results in postalveolars [tʂ ʐ dʐ ʂ], as illustrated in (8b).[8] In addition to *-ist-y*, there is a host of other morphological patterns that generate mutations before coronals and velars (Rubach 1984; Gussmann 2007). Both types of mutation appear before /i/ and /ɨ/, as shown in (8), as well as before [ɛ] (e.g. *u*[d]-*o* 'thigh' – *u*[dʑ]-*ec* 'haunch', *kro*[k] 'step' – *kro*[tʂ]-*ek* DIM). The mutations result in a change of the featural specification of the consonant. The latter aspect will be fleshed out in Section 2.6.

(8)   a.   Coronal Mutation: [coronal, +anterior, +back] → [coronal, –anterior, –back]
               t ~ tɕ   złoto [zwɔt-ɔ] 'gold'   złocisty [zwɔtɕ-ist-ɨ] ADJ
               d ~ dʑ   gwiazda [gvjazd-a] 'star'   gwiaździsty [gvjaʑdʑ-ist-ɨ] ADJ
               s ~ ɕ   las [las] 'forest'   lesisty [lɛɕ-ist-ɨ] ADJ
               z ~ ʑ   wyrazu [vɨraz-u] 'expression' GEN SG   wyrazisty [vɨraʑ-ist-ɨ] ADJ
               r ~ ʐ   wzór [vzur] 'punishment'   wzorzysty [vzɔʐ-ist-ɨ] ADJ
               n ~ ɲ   bagno [bagn-ɔ] 'marsh'   bagnisty [bagɲ-ist-ɨ] ADJ

        b.   Velar Mutation: [dorsal, +back] → [coronal, +back]
               k ~ tʂ   wiek [vjɛk] 'age, century'   wieczysty [vjɛtʂ-ist-ɨ] ADJ
               g ~ ʐ   piargu [pjarg-u] 'colluvium' GEN SG   piarżysty [pjarʐ-ist-ɨ] ADJ
               g ~ dʐ   miazga [mjazg-a] 'mush'   miażdżysty [mjazdʐ-ist-ɨ] ADJ
               x ~ ʂ   puch [pux] 'ghost'   puszysty [puʂ-ist-ɨ] ADJ

Returning to the formation of agent nouns in *-ist-a/-yst-a* and the behavior of the base-final consonant, three facts require explanation in the context of Coronal and Velar Mutations. First, in considering the words in *-ist-a/-yst-a*, labials and velars do not show any significant consonant mutations before the suffix, while coronals (variably) do. This stands in contrast to the data exemplified in (8) showing that mutations apply to both coronals and velars in other morphophonological patterns.[9] The failure of velars to undergo mutations before *-ist-a/-yst-a* is in this light surprising and will be addressed in Section 3.1. Second, why do some words with base-final [t d r] show mutations before *-ist-a/-yst-a*, while others do not? Looking at the data illustrating Coronal Mutation, it seems safe to assume that mutations of coronals before front vowels are historically and phonetically motivated in Polish. There is evidence that palatalization before *-ist-a* in words with base-final [t d r] used to be a fully regular process. Words of this type without palatalization appeared later (Rubach 1984: 65–68). Therefore, the lack of mutations in this context for some words with base-final [t d r] before *-ist-a/-yst-a* must be seen as a later development, whose emergence is in need of explanation. This issue is tackled in Section 3.2.3. The third fact that requires explanation is the gradient elimination

---

7   The loss of jers, ultra-short vowels, is another factor implicated in the phonologization of palatalization in Late Common Slavic (Jakobson 1929/1962).

8   I refrain from using the terms Coronal Palatalization and Velar Palatalization (see, for instance, Rubach 1984), as the outcome of the processes is not always easily classifiable as involving a palatal articulation (see Section 2.5). There are several types of Velar Mutation (Rubach 1984).

9   Bateman (2007) in her survey of 45 languages or dialects found that palatalizations of coronals and dorsals are common, although the former type occurs more frequently than the latter: 54% for coronals vs. 18% for dorsals in her sample. Insofar as typological asymmetries in segmental alternations reflect common pathways of change, this indicates that both coronals and dorsals are susceptible to palatalization when followed by front vowels.

of mutations for base-final [t d r] and their preservation for base-final [s z n]. The gradient elimination of mutations for base-final [t d r] is dealt with in Section 3.2.3. In the summary in (9), the effects of *-ist-a/-yst-a* on the base-final consonant are juxtaposed with the effects of other morphological patterns, both inflectional and derivational (such as the adjectival suffix *-ist-y*). For compactness, in (9) and below I use *-ist-a* to refer to both alternants of the suffix: *-ist-a* and *-yst-a,* unless the choice of the alternant is somehow relevant.

(9)

consonant	-ist-a	other patterns
labials: [p b m f v]	no mutations	no mutations
coronals:		
[s z n]	mutations	mutations
[t d r]	mutations/no mutations	mutations
[l j]	no mutations	no mutations
[ts dz ʂ ʐ tʂ dʐ]	no mutations	no mutations
velars: [k g x]	no mutations	mutations

2.2 A HIGH-FREQUENCY PATTERN: DIMINUTIVES IN *-ek*

Let us take a closer look at the patterning of the diminutive suffix *-ek* in relation to different base-final consonants (Czaplicki 2013a; 2014a).

(10) a. labials
     słup [swup] 'pole'   słupek [swup-ɛk]
     grzyba [gʐɨb-a] 'mushroom' GEN SG   grzybek [gʐɨb-ɛk]
     syf [sɨf] 'syphilis'   syfek [sɨf-ɛk]

  b. coronals
     świat [ɕfjat] 'world'   światek [ɕfjat-ɛk]
     spodu [spɔd-u] 'bottom' GEN SG   spodek [spɔd-ɛk]
     wino [vin-ɔ] 'wine'   winek [vin-ɛk] GEN PL
     nos [nɔs] 'nose'   nosek [nɔs-ɛk]
     wozu [vɔz-u] 'cart' GEN SG   wózek [vuz-ɛk]
     wór [vur] 'sack'   worek [vɔr-ɛk]

  c. palatal(ized) labials and coronals
     gołębia [gɔwɛmbj-a] 'pigeon' GEN SG   gołąbek [gɔwɔmb-ɛk]
     liść [liɕtɕ] 'leaf'   listek [list-ɛk]
     kość [kɔɕtɕ] 'bone'   kostek [kɔst-ɛk] GEN PL
     niedźwiedźa [ɲedʑvjedʑ-a] 'bear' GEN SG niedźwiadek [ɲɛdʑvjad-ɛk]
     dzień [dʑɛɲ] 'day'   dzionek [dʑɔn-ɛk]
     gęś [gɛĩɕ] 'goose'   gąsek [gɔw̃s-ɛk] GEN PL

  d. velars
     krok [krɔk] 'step'   kroczek [krɔtʂ-ɛk]
     progu [prɔg-u] 'doorstep' GEN SG   prożek [prɔʑ-ɛk]
     duch [dux] 'ghost'   duszek [duʂ-ɛk]

Labials and coronals in (10a) and (10b) fail to mutate. In the case of palatalized labials and coronals in (10c), the palatal element disappears before *-ek*, which can be described in traditional terms as depalatalization.[10] Velars in (10d) mutate and appear as their postalveolar (retroflex) reflexes before the [ɛ] of the suffix.

There is inconsistency in the applicability of consonant mutations before the suffixes *-ek* and *-ist-a* (the latter discussed in the previous section). In contrast to the behavior of the *-ist-a* suffix, which triggers mutations of coronals but not of velars, the diminutive suffix *-ek* results in mutations of velars but not of coronals. There is ample evidence that the high front /i/ is typologically more likely to trigger mutations than the mid front /ɛ/ (Bateman 2007; Rubach 2007). In fact, phonological conditioning alone predicts that *-ist-a/-yst-a* should be responsible

---

10 The alternation in [liɕtɕ] NOM SG and [list-ɛk] DIM could be explained by positing an abstract palatalizing morpheme in the former case, but not in the latter. This solution can be taken to task for being abstract and ad hoc, as the morpheme would have to be deleted, once it has triggered palatalization.

for *higher* rates of mutations than *-ek*, as mutations are more likely to occur in the context of /i/ than in the context of /ɛ/. In a similar vein, Czaplicki's (2019) quantitative analysis of the effects of 27 suffixes in Polish has shown that /i/ is far more likely to trigger mutations than /ɛ/. The fact that the suffix *-ek* triggers palatalization before velars, but depalatalization before labials and coronals has been interpreted by Czaplicki (2019) as evidence for the irrelevance of phonological naturalness in the conditioning of consonant mutations. What is relevant for our purposes is that words in *-ek* do not exhibit variability, as opposed to words in *-ist-a*. There are no variants without mutations for base-final velars before *-ek*. Mutated coronals before *-ek* are equally uncommon.[11] It is shown that an element that is crucial in an explanatory analysis of morphophonological patterns is their frequency. Specifically, it will be shown that type frequency determines the stability of a pattern.

(11)   Relation between type frequency and morphological stability
         -ist-a – low type frequency – decreased stability
         -ek – high type frequency – increased stability

2.3 FREQUENCY

There is growing evidence suggesting that frequency plays an important role in morphophonology (Mańczak 1980; Bybee 2001; Ellis 2002; Albright & Hayes 2003; Baayen et al. 2003; Dąbrowska 2008; Czaplicki 2013a; 2013b; 2014a; 2014b). As already mentioned, the generalizability of a pattern has been shown to crucially depend on the number of stored words that exhibit the pattern (gang effects). More specifically, when two (or more) patterns are available in a particular context, the pattern with a higher frequency is the one most likely to become generalized to novel words. In other words, pattern extension deploys the most robust of the several patterns used in a particular morphological context. An explanation along these lines is available for the change of classes of the English verb *help*, which in Old English belonged to Strong Verbs, but now forms its past tense by means of the more robust *-ed* suffixation. Frequency enters into interactions with other factors. Dawdy-Hesterberg & Pierrehumbert (2014) have found that, while the generalization of the various patterns of Arabic broken plurals in large part depends on prosodic templates, gang effects are important predictors as well.

Another factor that is often implicated in language change and pattern extension is token frequency (Bybee 2001). High-frequency words are predicted to resist pattern extension, as evidenced by the preservation of such suppletive forms in English as *go – went*. As already mentioned, the relative token frequency of the derivative and its base plays a critical role in dual-route models of lexical access. This issue becomes relevant in Section 3.2, where the strength of lexical representations is considered.

Anttila (2006) discusses an interesting case of assibilation in Finnish. The bimoraic verbs which meet the structural description of assibilation (context of a following /i/) show three types of behavior that can be directly linked to their frequency: the most frequent verbs exhibit assibilation, the least frequent verbs fail to undergo assibilation and verbs of medium frequency show variation. Czaplicki (2016) argues that these regularities can be insightfully explained by reference to frequency, on the one hand, and avoidance of mutations between the base and the derivative (output-output correspondence, formulated below), on the other.

In an attempt to explain the different behavior of words in *-ist-a* and *-ek* with respect to mutations, I make reference to the frequency of the two patterns in the lexicon. It is argued that for robust patterns (i.e. those showing a high type frequency) identity pressures are overridden. Let us compare the type frequencies of the words in *-ist-a* and *-ek* using dictionary and corpus data. **Table 1** presents the counts of words in *-ist-a* and *-ek* drawn from a reverse dictionary of Polish (*Indeks a tergo do Uniwersalnego słownika języka polskiego pod redakcją Stanisława Dubisza*) (Bańko et al. 2003).

	-ek	*-ist-a*
Type frequency	1400	740

Table 1 Type frequency of words in *-ek* and *-ist-a*.

---

11   There is a handful of diminutives in *-ek* that do in fact show mutated coronals, e.g. [miɕ] 'teddy bear' – [miɕ-ɛk], [jaɕ] 'proper name' – [jaɕ-ɛk] and [ɔgjɛɲ] 'fire' – [ɔgjɛɲ-ɛk] (alongside [ɔgjɛn-ɛk]). Czaplicki (2014b) attributes the usage of palatals in such diminutives to the impact of Expressive Palatalization, a sound symbolic device commonly used in expressive morphology to signal smallness and affection.

The number of dictionary entries in *-ek* is nearly twice as high as the number of dictionary entries in *-ist-a*. There is good reason to believe that the difference in robustness between the two patterns is even greater. First, the list of words with the suffix *-ek* is not exhaustive, as it does not include recent borrowings and many diminutives whose semantics is fully predictable. Grzegorczykowa & Puzynina (1999: 425) mention that diminutives in *-ek* and *-ik* constitute an open class and that most nouns can form diminutives using one of these suffixes.[12] Second, a considerable number of words in *-ist-a* (but definitely not all) belongs to the learned stratum of vocabulary and their usage is often restricted to formal and technical registers. Therefore, the factor that is missing from **Table 1** is token frequency, i.e. the frequency with which each of the words is used in the discourse.

In analyzing the frequency of the two suffixes, I use data extracted from the corpus *plTenTen: Corpus of the Polish Web*, available in *Sketchengine*, which is made up of texts collected from the internet in 2012 and comprises more than 7.7 billion words. **Table 2** provides the type and total token frequency of words in *-ist-a* and *-ek* in the corpus (accessed 12 October 2020). Words below the frequency of 50 have not been considered, as they turned out to be mostly proper names and spelling errors.[13]

	TYPE FREQUENCY	TOTAL TOKEN FREQUENCY
-ek	2303	17,119,953
-ist-a	302	1,869,247
sum	2605	18,989,200

Table 2 Frequency of words in *-ist-a* and *-ek* in *plTenTen: Corpus of the Polish Web*.

The type frequency of words in *-ek* in the corpus is 7.6 times higher than the type frequency of words in *-ist-a*. Predictably, the difference between the number of words in *-ist-a* and *-ek* in the corpus (i.e. their type frequency) is greater than for the data taken from the dictionary shown in **Table 1**. **Table 2** also shows the total token frequency, which is the sum of the token frequencies (the frequency with which a given word is used) of all the words in *-ist-a* and *-ek* in the corpus. The total token frequency of words in *-ek* is 9.2 times higher than the total token frequency of words in *-ist-a*.

The token frequency data are not normally distributed, so they have been log-scaled as in **Figure 1**. The histograms in **Figure 1** show that words in *-ek* outnumber words in *-ist-a* for all token frequency ranges. A non-parametric Mann-Whitney test run on the data reveals that the token frequency of words in *-ek* (Mdn = 459, M = 7433.76) is significantly lower than the token frequency of words in *-ist-a* (Mdn = 884, M = 6189.56), U = 390640.5, z = 3.49, p < .001. This result is likely due to the highly positively skewed distribution for words in -ek (skewness = 29.44). As can be seen from the histogram on the left in **Figure 1**, low-frequency words in *-ek* outnumber comparable words of medium and high frequency. It follows that type frequency is a better measure of the strength of a pattern than token frequency.

---

12   A reviewer draws attention to the multi-functionality of the suffix *-ek*. Apart from deriving diminutives, it is used to form feminine counterparts of masculine nouns, e.g. *aktor* 'actor' – *aktor-k-a* 'actress'. It may also function as a semantically opaque nominal marker, e.g. *tył* 'back' – *tył-ek* 'somebody's behind' and *sałat-a* 'lettuce' – *sałat-k-a* 'salad'. In some words it has no obvious semantic function at all, e.g. *grusz-k-a* 'pear' and *cór-k-a* 'daughter'. In line with the assumption that schemas are inferred from phonological and semantic similarity (see Section 2.5), only semantically transparent diminutives in *-ek* have been used to assess frequency.

13   The necessary condition for words to be included in the data was the existence of a base as an independent word. Words like *art-yst-a* 'artist', *dent-yst-a* 'dentist' and *stat-yst-a* 'extra' did not qualify because *art, dent* and *stat* are not actual words in Polish. This decision was partly dictated by the gradient and often debatable decomposability of such words with bound roots. With relevance to the present analysis, the lack of a base makes it impossible to assess the relative token frequency of the derivative and the base. Relative frequency is instrumental in determining the route of lexical access (See section 3.2.2). Furthermore, considering the assumptions of usage-based models, it would be meaningless to analyze the presence or lack of mutations in a derivative that has no independent base, such as *stat-yst-a*. Though, admittedly, such words would be central in an analysis of the gradience of semantic decomposability. Complex words without a shared independently occurring base like *klasyc-yzm* 'classicism' and *klasyc-yst-ycz-n-y* 'classicist' (where *klasyc* is not an actual word) are associated by means of second-order schemas (Booij & Audring 2017: 12–15). In this way, complex words form relations with multiple other complex words by virtue of the shared elements. A reviewer worries that the schema-approach might have problems with the fact that while words with bound roots like *ar*[t]*-yst-a* 'artist' and *ar*[t]*-yst-ycz-n-y* 'artistic' exist, a word like *ar*[tɕ]*-ist-ycz-n-y* (with a mutation) is unlikely. The proposed explanation makes use of (i) whole-word storage and (ii) second-order schemas that join elements of related complex words. Bound roots like *art-* are not independently stored, so a source-oriented schema such as: "a final [t] in the base corresponds to [tɕ] in the adjective in -ist-ycz-n-y" is not applicable. There is no independently stored base (see Section 2.5).
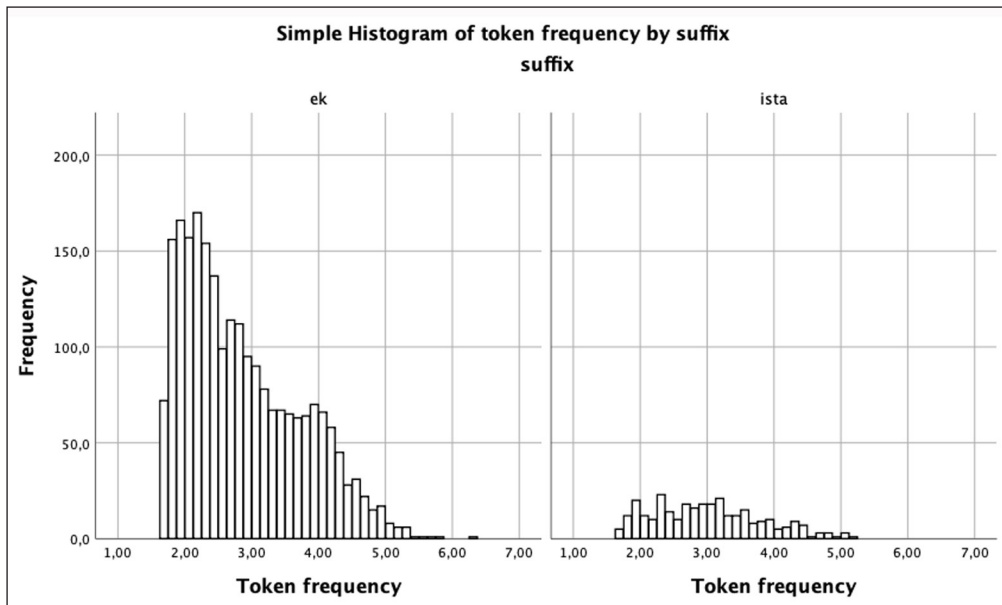
To sum up, a close analysis of type and token frequency confirms the difference in the robustness of the two patterns: the pattern *-ek* is significantly more robust than the pattern *-ist-a* in the grammar. In addition, low-frequency words in *-ek* are overrepresented relative to words in *-ek* of medium and high frequency.

The two suffixes are useful in assessing the role of frequency, as they are compatible in the relevant aspects of distribution and phonological behavior. First, they are both derivational. Second, they are both fully productive and readily extended to new words. Third, neither of them is restricted to loanwords.[14] Finally, they both begin with front vowels and trigger mutations. In fact, as mentioned in Section 2.2, phonological conditioning alone predicts that *-ist-a* should be responsible for *higher* rates of mutations than *-ek*, as /i/ is more likely to trigger mutations than /ɛ/ both in Polish and more generally (Bateman 2007; Rubach 2007; Czaplicki 2019). It is claimed that the differences in the phonological and morphological behavior of the two constructions are derivable from the differences in their frequency.

2.4 LEXICAL STORAGE OF ALLOMORPHS

Lexical storage of allomorphs is not a new idea. There is ample evidence from usage-based research that morphologically complex words are stored whole (McQueen & Cutler 1998; Bybee 2001; Baayen et al. 2003; Hay 2003; see also Section 3.2.2). Similarly, certain analyses representative of generative phonology demonstrate that specific alternations cannot be derived using purely phonological operations and the relevant allomorphs need to be listed. In phonologically conditioned suppletion two phonologically dissimilar alternants need to be stored, however, their distribution is phonologically conditioned (Carstairs 1988; 1990). Anderson (2008) discusses a pattern from Surmiran and argues that vowel reduction, once a phonologically conditioned process, has become opaque. A solution proposed by Anderson involves reference to two listed alternants of the stem whose distribution is governed by prosodic considerations (i.e. stress). The distribution of alternants may also be regulated by phonologically neutral considerations. Paster (2006) and Embick (2010) use subcategorization frames which make reference to phonological and lexical information. However, allomorph selection does not as a rule result in phonologically optimized structures. In Kaititj the ergative suffix appears as [-ŋ] after disyllabic stems and as [-l] after trisyllabic stems. Although the formula that captures the generalization makes reference to phonological vocabulary, in this

---

14   For example, out of 28 words showing the alternation [r ~ ʐ] found in a Polish dictionary (Bańko et al. 2003), 18 have no structural equivalents in English, e.g. *chałturz-yst-a* 'person doing odd jobs', *fakturz-yst-a* 'invoice clerk', *maturz-yst-a* 'high school graduate', *kamerz-yst-a* 'cameraperson' and *pokerz-yst-a* 'poker player'. English was chosen for comparison, as it is currently the main source of borrowings in Polish.

case the syllable, this instance of allomorph selection does not in any way improve phonological well-formedness (Paster 2006). The existence of such patterns shows two things: allomorph selection need not be phonologically optimizing and some alternants, including different shapes of a stem, have to be listed (i.e. stored).

2.5 MORPHOPHONOLOGICAL SCHEMAS AND COPHONOLOGIES

In order to shed light on the differences in the assumptions of traditional generative and usage-base approaches, it is vital to compare the relevant aspects of rules (Current generative analyses generally employ constraints instead of rules, but certain properties of rules remain implicit.) and schemas (usage-based approaches). Schemas in contradistinction to rules emerge from the lexicon, that is, from the stored representations of words and phrases. As a consequence, they have "no existence independent of the lexical units from which they emerge" (Bybee 2001: 27). Rules, on the other hand, exist independently of the stored items and form part of a module that is separate from the lexicon. The productivity of a schema is a function of the number of participant items (gang effects). Put differently, the more words comply with a given schema, the more productive the schema is predicted to be.[15] In this view, productivity of a schema is gradient and probabilistic, which is a consequence of the close connection between schemas and stored words. While in early generative models rules did not show a direct relationship with the number of words they apply to, more recently, modeling non-categorical effects, including the effects of frequency, has been facilitated by the use of gradient and probabilistic OT constraints (e.g. MaxEnt, Hayes & Wilson 2008).

Since the publication of *The Sound Pattern of English* (Chomsky & Halle 1968), the well-formedness of rules has typically been associated with the notion of markedness or phonological naturalness. Chomsky & Halle (1968) observe that their model largely overpredicts the types of processes that occur in natural languages and propose to constrain the set of possible rules by appealing to phonological markedness (Chapter 9). Rules that lead to the reduction in markedness are preferred over those that do not. More recently, Hayes and Steriade (2004: 1) claimed that markedness constraints are gleaned from phonetic knowledge, the latter being somewhat vaguely defined as "the speakers' partial understanding of the physical conditions under which speech is produced and perceived". In this view, final devoicing of obstruents is possible, while final voicing is predicted to be impossible, as it would lead to more marked structures (Kiparsky 2006). The position that synchronic universals (i.e. markedness) constrain diachronic change and shape linguistic patterns is advocated in, for instance, de Lacy (2002; 2006), Kiparsky (2006; 2008) and de Lacy & Kingston (2013).

On the other hand, there is accumulating evidence that undermines the role of markedness as an active bias in synchronic grammars. Processes that result in more marked structures are pervasive (Bach & Harms 1972; Anderson 1981; Blevins 2004; Hale & Reiss 2008; Czaplicki 2013a; 2014a; 2019). Examples of such processes include final voicing in Lezgian (Blevins 2006) and unnatural patterns of consonant epenthesis in various languages (Blevins 2008). Typological asymmetries can be explained by extragrammatical factors (e.g. common trajectories of phonetically based sound change) (Ohala 1983; Blevins 2004). Insofar as schemas are based on stored representations, they are language specific and not necessarily dependent on naturalness, understood as a universal learning bias. Thus, schemas are *a priori* markedness-free. However, it should be noted that markedness constraints are *not* logically incompatible with schemas. In principle, the emergence of markedness constraints is independent of schemas. Yet, one of the predictions of usage-based models is that language-specific considerations should override markedness constraints when a conflict between the two pressures arises (a preference for morphological conditioning over phonological conditioning, see below). The role of markedness in schema-based approaches certainly deserves more attention.

---

15 Bybee (2001: 124) discusses evidence indicating that the minimum number of items necessary to form a gang is greater than three. However, the productivity of a pattern is not solely determined by the size of the gang.

Rules refer to distinctive features (Chomsky & Halle 1968). If segments are referred to, this is done as a shorthand for featural specification that underlies a particular segment. Schemas are not similarly restricted in their vocabulary.

(12)   Lexical organization provides generalizations and segmentation at various degrees of abstraction and generality. Units such as morpheme, segment, or syllable are emergent in the sense that they arise from the relations of identity and similarity that organize representations. Since storage in this model is highly redundant, schemas may describe the same pattern at different degrees of generality (Langacker 2000; Bybee 2001: 7–8).

As mentioned in (12), schemas can refer to various organizational units, such as segment, syllable and feature, as long as these units emerge from stored representations. The claim cited in (12) points to yet another important difference between rules and schemas. Rules are preferably stated using phonological vocabulary (this is a requirement of modularity, Scheer 2012). In contrast, schemas require that different types of information – phonological, syntactic, morphological and semantic – be simultaneously accessible. This is the fundamental property of Parallel Architecture, a theory of grammar developed by Ray Jackendoff (cf. Jackendoff 2002). Schemas are formed on the basis of phonological and semantic similarity between stored words. Morphological structure emerges from these identity relations. On this view, *-ist-a* and *-yst-a* are predicted to be identified as alternants of a single suffix because of their phonological and semantic similarity as well as their near complementary distribution.

Schemas contain information about morphological structure (e.g. English past tense formation using *-ed*) (Bybee 2001: 23–24). In fact, Bybee (2001: 97–100) argues that segmental alternations display *a preference for morphological conditioning* over phonological conditioning. She also claims that "once morphological conditioning becomes dominant, it follows that phonological principles, such as patterning based on natural classes, will no longer be applied in the same way as for phonetically conditioned processes" (2001: 105). Put differently, markedness considerations are less important for morphologized patterns than for patterns that are fully phonetically transparent. Two phonologically similar contexts can give rise to different segmental alternations, as long as the morphological conditioning is different (i.e. two different morphemes). For example, Velar Mutation applies before the suffix *-ek* but not before *-ist-a* (see Sections 2.1–2.2), even though the context of a following front vowel is present in both cases.

Schemas are compatible with approaches that assume the existence of multiple cophonologies within one grammar (Booij 2010; Inkelas 2014; Booij & Audring 2017; Czaplicki 2020), where each morphological construction is associated with its own phonological subgrammar. In other words, each morphological construction has its own phonological properties. Representative of early approaches that assume the existence of multiple cophonologies within one grammar is Itô & Mester's (1995) Core-Periphery Model of the lexicon, which identifies the core area of the lexicon, governed by a maximum set of markedness constraints (markedness dominates faithfulness). Their markedness constraints are syllable- and segment-related and penalize, for example, voiced obstruent geminates and non-geminate [p]. Structures that occupy less and less central areas of the lexicon show increasingly more violations of markedness constraints (faithfulness dominates markedness). By extension, the Core-Periphery Model predicts that the less nativized (more peripheral) a structure, the more faithful it should be to its input. Crucially, the model predicts that there are multiple layers within the lexicon and each layer is differentiated from the others in that it has its own specific constraint ranking (phonological grammar). In Japanese four layers are distinguished: Yamato, Sino-Japanese, Assimilated Foreign and Unassimilated Foreign. An important claim of Itô & Mester (1995) is that the differences between the core area of the grammar and more peripheral areas relate to the reranking of faithfulness constraints. The ranking of markedness constraints remains constant for the whole grammar. For example, the hypothetical input /paka/ is realized differently depending on the layer of the grammar in which it is generated. If it is processed in the Sino-Japanese stratum it surfaces as [haka], due to the ranking of No-P above Faith. In the Assimilated Foreign stratum, the output is [paka], due to the reranking of Faith above

No-P.[16] Morphological conditioning can also be handled by indexing constraints to specific morphological constructions (Itô & Mester 1999).

To sum up, schemas are dependent on stored representations, their productivity (strength) depends on the number of words they derive (type frequency), they are *a priori* markedness-free, they can be stated at various degrees of generality (e.g. segment, feature, syllable) and they include morphological information. Supporting evidence for these properties of schemas can be found in, for example, Bybee (2001), Ellis (2002) (frequency effects), Booij & Audring (2017), Czaplicki (2013a; 2019; 2020) (morphological conditioning) and Blevins (2004) (markedness-free generalizations).

Schemas can be product- or source-oriented. An example of a product-oriented schema specifying English Past Tense in *walked, begged* and *wanted* was given in (1) and is repeated in (13) (Bybee 2001: 126). Source-oriented schemas mention input (base) as well as output. Several source-oriented schemas representing Polish consonant mutations are exemplified in (14).

(13)   a Past verb ends in /t/, /d/, or /ɨd/

(14)   a.   […k] NOUN ↔ [[…tʂ]ɛk] NOUN DIM
                […g] NOUN ↔ [[…ʐ]ɛk] NOUN DIM
                […x] NOUN ↔ [[…ʂ]ɛk] NOUN DIM
                […t] NOUN ↔ [[…t]ɛk] NOUN DIM
                […ɕtɕ] NOUN ↔ [[…st]ɛk] NOUN DIM

         b.   […k] NOUN MASC ↔ [[…kj]ɛm] NOUN MASC INSTR SG
                […g] NOUN MASC ↔ [[…gj]ɛm] NOUN MASC INSTR SG
                […x] NOUN MASC ↔ [[…x]ɛm] NOUN MASC INSTR SG

The schemas in (14) express the formation of diminutives, (a), and instrumentals, (b), from nouns whose stems end in various consonants. The former pattern was illustrated in Section 2.2, the latter can be applied to *kro*[k] 'step', *pro*[g]-*u* 'threshold' GEN SG and *du*[x] 'ghost'. Source-oriented schemas are preferable for describing consonant mutations, as the applicability and type of mutation in the derivative crucially depend on the final consonant *in the base*. In (13a) base-final velars appear as their mutated alternants before -*ek*. However, at the same time base-final [t] fails to mutate and the cluster [ɕtɕ] depalatalizes to [st] in the same context. So we could not refer to a product-oriented schema requiring that a consonant appear in its palatalized (or mutated, or retroflex) form before -*ek*. Rather, the output of the concatenation of a suffix depends on a particular base-final consonant. Before the instrumental suffix -*em*, illustrated in (14b), velar plosives mutate but the fricative remains unchanged. Such arbitrary suffix- and consonant-specific alternations, which abound in Polish, would be difficult to conceptualize as product-oriented schemas (see Becker & Gouskova 2016 for more evidence that source-oriented schemas are necessary).

Schemas defined along these lines constitute the core of the proposed analysis. Type frequency determines the strength of linguistic patterns, in the sense that the more frequent the pattern, the more likely it is to be extended to novel words. In the formalization of the analysis, type frequency finds reflection in the ranking of constraints representing morphophonological schemas. Constraints representing more frequent schemas are ranked higher than constraints corresponding to less frequent schemas. In the case at hand, schemas representing -*ek* will be ranked higher than schemas pertaining to -*ist-a*, as illustrated in (15). Schema-constraints are interspersed with phonological constraints, as will be demonstrated in Section 3.2.3.

(15)   […k] NOUN ↔ [[…tʂ]ɛk] DIM > > … > > […k] NOUN ↔ [[[…k]ist]a] AGENT NOUN

---

16   While there are parallels between the Core-Periphery Model and the present analysis, an important difference pertains to the role assigned to frequency in the two analyses. In the Core-Periphery Model frequency is not invoked directly; it is rather an epiphenomenon of a structure's assignment to a sublexicon. Thus, it is basically an accident that structures occupying peripheral areas of the lexicon are also less frequent than those centrally located. By contrast, in the present analysis the frequency of a structure determines its sensitivity to paradigm uniformity pressures. Less nativized patterns are more susceptible to paradigm uniformity pressures than more entrenched patterns precisely because they are less frequent.

2.6 REPRESENTING CONSONANT MUTATIONS IN POLISH

In traditional terms, palatalization is a type of consonant mutation that is caused by a following front vowel or a palatal glide. Palatalization has its diachronic roots in the phonologization of coarticulatory effects between a consonant and the following front vowel or a palatal glide (Jakobson 1929/1962; Bateman 2007; Kochetov 2011). In phonological analyses, palatalization has been represented as agreement in certain features between a consonant and a following vowel. Several feature theories have been in use, for example, the Halle-Sagey model (Sagey 1986) and the Clements-Hume model (Clements & Hume 1995). In the Halle-Sagey model palatalization involves the agreement in the feature [–back] and in the Clements-Hume model the process is represented as agreement in the features [coronal, –anterior]. More recent approaches cast in Optimality Theory concentrate on defining feature classes, i.e. features that function together in phonology (see, for instance, Padgett 2002 and Halle 2005). In the present analysis front vowels are represented as [+coronal, –back] and, consequently, palatalization imposes agreement in the features [+coronal, –back] on adjacent consonants and vowels. The assumption that the feature [±back] can refer to both [±coronal] and [±dorsal] is in line with Padgett (2002) and Halle (2005) in the sense that feature sets (classes of segments) are defined on a language- and alternation-particular basis.

Coronal Mutation, e.g. [t] ~ [tɕ], can be viewed as a change of the feature [±back], alongside the change of [±anterior]. With relevance to this analysis, Coronal Mutation does not affect the major place of articulation ([±coronal]).

Velar Mutation, e.g. [k] ~ [tʂ], on the other hand, involves a change of the major place of articulation ([+dorsal, –coronal] > [–dorsal, +coronal]), with the feature [±back] remaining unaffected. The assumption that the output of Velar Mutation, postalveolars, is [+back] gets support from their phonetics and distribution. Hamann (2002) has found that Polish postalveolars meet the criteria of retroflex consonants and retroflexion is incompatible with palatalization. In addition, postalveolars never appear before the high front vowel /i/ (except in recent borrowings). Based on these facts, Velar Mutation is an instance of coronalization.

The analysis at hand focuses on one important aspect of these mutations: Velar Mutation results in a change of the major place of articulation (from dorsal to coronal), while Coronal Mutation generates no similar change (the coronal place remains). This difference will be central.

A family of output-output faithfulness (paradigm uniformity, PU) constraints will be relevant for the present purposes (Kenstowicz 1996; Benua 1997; Steriade 2000). Such constraints have an important function of improving the transparency of morphological relationships between words and, thus, may facilitate lexical access. Dressler (2003: 464) refers to this pressure as "morphotactic transparency" and adds that "the most natural forms are those where there is no opacifying obstruction to ease of perception". Research on morphological processing suggests that transparent phonology aids morphological decomposition (Frauenfelder & Schreuder 1992: 173). Output-output faithfulness is violated whenever a consonant undergoes a mutation. Both Velar Mutation and Coronal Mutation incur a violation of output-output faithfulness.

IᴅᴇɴᴛPl, an output-output faithfulness constraint given in (16), is used to represent the difference between the effects of Velar Mutation and Coronal Mutation.

(16)   IᴅᴇɴᴛPl$_{O-O}$:
       Corresponding consonants in the stem of the base and the output have identical values for [±labial, ±coronal, ±dorsal].

IᴅᴇɴᴛPl is violated when Velar Mutation applies (–coronal, +dorsal > +coronal, –dorsal). When Coronal Mutation occurs, the constraint is respected (+coronal, –dorsal > +coronal, –dorsal). This analysis relies on output-output faithfulness constraints, as opposed to input-output faithfulness constraints. Identity relations between a derivative and its base are evaluated.

3 A LOW-FREQUENCY PATTERN: AGENT NOUNS IN -*IST-A*

3.1 HISTORICAL CHANGE: WHY CORONALS UNDERWENT MUTATIONS BUT VELARS DID NOT

This section aims to account for the emergence of a general pattern, in which coronals show mutations before -*ist-a*, while velars do not. Thus, the proposed analysis is historical and refers to the time when the suffix -*ist-a* became productive in the language. At that time, there were no morphophonological schemas related to the suffix -*ist-a* that would be sufficiently entrenched in the lexicon to determine the output. As mentioned in Section 2.1, there is historical evidence that Coronal Mutation initially applied across the board and was phonetically conditioned. Therefore, the evaluation of candidates proceeds according to phonological constraints. The modern-day complexities surrounding coronals are addressed in Section 3.2, where it will be argued that currently morphophonological schemas play a greater role than at the inception of the -*ist-a* pattern.

In explaining why coronals underwent mutations before -*ist-a* and velars did not, I make use of the observation made in Section 2.6 that Coronal Mutation does not change the major place of articulation, while Velar Mutation does. Thus, IDENTPl is violated by Velar Mutation but not by Coronal Mutation.

In addition to IDENTPl, a faithfulness constraint that will be relevant is MAXV$_{[\pm back]}$, formulated in (17a). Given that [ista] is the principal variant of the suffix, while its allomorph [ista] is positionally restricted (see Sections 2.1–2.2), MAXV$_{[\pm back]}$ is violated when [ista] is used. MAXV$_{[\pm back]}$, rather than IDENTV$_{[\pm back]}$, is used because central vowels, including [ɨ], lack place features (Clements and Hume 1995). In consequence, the change of [i] to [ɨ] involves the deletion of [–back], which is penalized by MAXV$_{[\pm back]}$. IDENT$_{[\pm anterior]}$ in (17b) enforces faithfulness to [±anterior]. The constraints regulating consonant mutations in Polish are formulated in (17c, d) (based on Rubach 2007). This analysis rests on the assumption that all features are binary.[17]

(17)   a.   MAXV$_{[\pm back]}$: Input [±back] on a vowel must be preserved on an output correspondent of that vowel.

   b.   IDENT$_{[\pm anterior]}$: Corresponding consonants in the base and the output have an identical value for [±anterior].

   c.   AGREE$_{[\pm coron]}$: A consonant and a following vowel must agree in [±coronal] when the vowel is specified for it.

   d.   AGREECOR$_{[\pm back]}$: A coronal consonant and a following vowel must agree in [±back].

The AGREE constraints are used here as shorthands for restrictions on specific consonant + vowel sequences. AGREE$_{[\pm coron]}$ effectively bans velars before front vowels, e.g. [ki gi xi]. Following Rubach (2007), the formulation of AGREE$_{[\pm coron]}$ in (17c) ensures that the constraint is moot for the vowel [ɨ], as central vowels lack place features. AGREECOR$_{[\pm back]}$ in (17d) bans non-palatal coronal consonants before front vowels, e.g. [ʂi ʐi tʂi dʐi] and [ti di tsi dzi]. In line with the evidence mentioned in Section 2.1, the latter constraint requires *phonological* palatalization; presence of coarticulatory effects of the following vowel, e.g. [tʲi], is not sufficient to satisfy it. The formulation of AGREECOR$_{[\pm back]}$ illustrates the already-mentioned assumption that feature sets are established on the basis of the phonological behavior of segments in a particular language and possibly in a particular alternation, rather than on the basis of a universal hierarchy of features. In the proposed analysis the feature [±back] can co-occur with both [±dorsal] and [±coronal], which in essence is reminiscent of the claims of Padgett (2002) and Halle (2005).

In (18), an evaluation of a word in -*ist-a* with a base-final velar is shown. The high-ranked IDENTPl mandates that a candidate with no change of major place be selected (candidate a). Candidates (b) and (c) are eliminated due to a change from dorsal to coronal. Some remarks are in order about the AGREE constraints. Candidate (a) violates AGREE$_{[\pm coron]}$ because the consonant does not agree in [±coronal] with the following vowel. Candidate (b) violates AGREECOR$_{[\pm back]}$ because a postalveolar [ʐ] is [+back], while the following vowel is [–back]. AGREE$_{[\pm coron]}$ is respected because

---

17   Other analyses have argued for privative (unary) features. For example, Lombardi (1999) contends that all features are privative and Halle (2005) assumes that features designating articulators are privative, e.g. [coronal], while most other features are binary, e.g. [±back], [±anterior].

both the consonant and the vowel are [+coronal]. Candidate (c) fares well on AGREE$_{[\pm back]}$, as both the consonant and the vowel are [+back]. It also vacuously satisfies AGREE$_{[\pm coron]}$, because, as mentioned above, [ɨ], being a central vowel, is not specified for place features. Candidate (d) incurs a fatal violation of MAXV$_{[\pm back]}$, because the backness of the relevant vowel has been modified. This evaluation demonstrates that AGREE$_{[\pm coron]}$ must be ranked below IDENTPl and MAXV$_{[\pm back]}$.

(18)　　Evaluation of a derivative in *-ist-a* of [tʂɔwg-u] 'tank'

Input: tʂɔwg + ist-a	IDENTPl	AGREECOR$_{[\pm back]}$	MAXV$_{[\pm back]}$	AGREE$_{[\pm coron]}$
☞　a.　tʂɔwg-ist-a				*
b.　tʂɔwʐ-ist-a	*!	*		
c.　tʂɔwʐ-ist-a	*!		*	
d.　tʂɔwg-ist-a			*!	

The evaluation in (19) focuses on words in *-ist-a* with base-final coronals. The faithful candidate shown in (a) fails to respect agreement in backness (a violation of AGREECOR$_{[\pm back]}$) and loses to candidate (b), which satisfies AGREECOR$_{[\pm back]}$ by employing an alveolopalatal before a front vowel. IDENTPl and AGREE$_{[\pm coron]}$, the latter not shown, are respected by all the candidates. Surely, the change of [t] to [tɕ] in the winning candidate cannot be of no consequence for faithfulness. The constraint that is violated here is the low-ranked IDENT$_{[\pm anterior]}$. The tableau also confirms the relevance of MAXV$_{[\pm back]}$. Candidate (c) is not optimal because the quality of the vowel has been changed.

(19)　　Evaluation of a derivative in *-ist-a* of [flɛt] 'flute'

Input: flɛt + ist-a	IDENTPl	AGREECOR$_{[\pm back]}$	MAXV$_{[\pm back]}$	IDENT$_{[\pm anterior]}$
a.　flɛt-ist-a		*!		
☞　b.　flɛtɕ-ist-a				*
c.　flɛt-ist-a			*!	
d.　flɛtɕ-ist-a		*!	*	*

It has been shown that mutations of velars were avoided, as they involve a change of major place, the latter being detrimental to base recognition. In contrast, mutations of coronals were tolerated, because they do not result in a comparable modification. The ranking responsible for the emergence of the pattern *-ist-a* is given in (20).

(20)　　IDENTPl, AGREECOR$_{[\pm back]}$, MAXV$_{[\pm back]}$ >> AGREE$_{[\pm coron]}$, IDENT$_{[\pm anterior]}$

3.2 PRESENT-DAY DEVELOPMENTS: WHY (SOME) MUTATIONS ARE BEING ELIMINATED

In this section, we address the issue of the gradual elimination of consonant alternations from *-ist-a* words with base-final coronals [t d r]. It is worth noting that alternations with base-final coronals [s z n] show a weaker tendency towards elimination. I refer to an interplay of the type frequency of a pattern with the requirement that the base be transparent in the derivative (output-output faithfulness). In contrast to the analysis in Section 3.1, where phonological constraints determined the output, I assume that, the analysis in this section must refer to morphophonological patterns because at the present stage some patterns are more entrenched than others and this must be reflected in the grammar. This assumption is based on the finding that the frequency of a pattern in the lexicon (its type frequency) determines its productivity. As words showing a particular pattern accumulate, they form a gang. When the gang reaches a certain threshold, a schema emerges and the size of the gang determines the productivity of the schema (gang effects; see Section 2.3). In Section 3.2.1–3 we look at data drawn from a dictionary and a corpus, and Section 3.2.4 examines experimental data probing native speaker intuition.

3.2.1 Dictionary entries

Table 3 provides the counts of dictionary entries with the variants *-ist-a* and *-yst-a* for each base-final consonant drawn from the reverse dictionary of Polish (Bańko et al. 2003). Labials, velars and the coronals [s z n] show categorical behavior: labials and velars do not mutate, while the coronals [s z n] mutate. The coronals [t d r] show two competing patterns. Looking at the type frequency of the mutated and non-mutated variants for each derivative with base-final [t d r], it appears that the lexical entries without mutations are actually more numerous. A corpus-based analysis of words with base-final [t d r] yields similar results. Based on a search of *The National Corpus of Polish,* words showing stable alternating patterns exhibit an average frequency of M = 391.17, while words without alternations and fluctuating words have an average frequency of M = 38.13. The difference is significant at p < .05. An in-depth analysis of the corpus-based data is postponed until the next section, where the concept of listedness is elucidated.

BASE-FINAL CONS.	MUTATIONS	NO MUTATIONS	% OF NO MUTATIONS
[p]		13	100
[b]		10	100
[m]		37	100
[f]		11	100
[v]		30	100
[t]	20	33	62.3
[d]	18	21	53.8
[r]	28	51	64.6
[n]	109		0
[s]	19		0
[z]	2		0
[ts]		6	100
[z�é]		11	100
[ʂ]		6	100
[dz�é]		1	100
[k]		7	100
[g]		11	100
[x]		5	100

Table 3 Number of alternating and non-alternating patterns for words in *-ist-a/-yst-a*.

It is argued that PU pressures are responsible for the elimination of morphophonological alternations between the base and the derivative. The schemas in **Table 4** represent some of the attested morphophonological patterns based on dictionary entries. They are divided into two categories according to the behavior of the base-final consonant: the schemas in the first column show alternations (mutations), unlike the schemas in the second column.

For the base-final consonants [t d r] in (d–f), both the alternating and the non-alternating schemas are present. In fact, in accordance with the quantitative data in **Table 3**, the non-alternating patterns seem to be more frequent in the lexicon than the alternating patterns. Given that the frequency of a pattern determines its strength, the strength of the non-alternating schemas for [t d r] is higher than the strength of the corresponding alternating schemas in modern usage and is reflected in the relative ranking of the schema-constraints in (21) ("》" indicates dominance).

(21) a. [...t] ⟷ [[[...t]ɨst]a] 》 [...t] ⟷ [[[...tɕ]ist]a]
     b. [...d] ⟷ [[[...d]ɨst]a] 》 [...d] ⟷ [[[...dʑ]ist]a]
     c. [...r] ⟷ [[[...r]ɨst]a] 》 [...r] ⟷ [[[...ʐ]ist]a]

	ALTERNATING SCHEMAS	NON-ALTERNATING SCHEMAS
a.		[…p] ↔ [[[…p]ist]a]
b.		[…m] ↔ [[[…m]ist]a]
c.		[…f] ↔ [[[…f]ist]a]
d.	[…t] ↔ [[[…tɕ]ist]a]	[…t] ↔ [[[…t]ist]a]
e.	[…d] ↔ [[[…dʑ]ist]a]	[…d] ↔ [[[…d]ist]a]
f.	[…r] ↔ [[[…ʐ]ist]a]	[…r] ↔ [[[…r]ist]a]
g.	[…s] ↔ [[[…ɕ]ist]a]	
h.	[…z] ↔ [[[…ʑ]ist]a]	
i.	[…n] ↔ [[[…ɲ]ist]a]	
j.		[…ts] ↔ [[[…ts]ist]a]
k.		[…dʑ] ↔ [[[…dʑ]ist]a]
l.		[…ʑ] ↔ [[[…ʑ]ist]a]
m.		[…k] ↔ [[[…k]ist]a]
n.		[…x] ↔ [[[…x]ist]a]

Table 4 Selected alternating and non-alternating schemas.

Returning to **Table 4**, the base-final consonants [s z n] in (g–i) show only alternating patterns, while for the base-final consonants [p b f] in (a–c) and [ts dʑ ʑ k x] in (j–n) only non-alternating patterns can be identified. The patterns in the first column impinge on PU because they introduce segmental alternations in the derivatives. In other words, the transparency of the base is diminished.

The issue that needs to be tackled first is the selectivity of the bias against alternations within coronals.

(22) a.   Alternating patterns   Non-alternating patterns
        [t]   [t ~ tɕ]   [t ~ t]
        [d]   [d ~ dʑ]   [d ~ d]
        [r]   [r ~ ʑ]   [r ~ r]

      b.
        [s]   [s ~ ɕ]
        [z]   [z ~ ʑ]
        [n]   [n ~ ɲ]

As for the base-final coronals [t d r] in (22a), mutations can be avoided thanks to the patterns on the right, which have emerged recently and are affecting more and more words (evidence for the latter claim is given below). Looking at (22b) it is clear that for the coronals [s z n] the older alternating patterns have not been replaced.

How do we account for the asymmetry in the treatment of the two groups of coronals in (22)? Crucial to the analysis is the fact that, while all of the patterns in (22) satisfy IDENTPl, the alternating patterns in (22a) incur a violation of IDENT$_{[\pm strid]}$, a faithfulness constraint defined in (23).

(23)   IDENT$_{[\pm strid]O\text{-}O}$
         Corresponding consonants in the stem of the base and the output have an identical value for the feature [±strident].

IDENT$_{[\pm strid]}$ is violated by the alternating patterns in (22a) because [t d r] are specified as [-strident] and when they mutate in -ist-a words they become [tɕ dʑ ʑ], which are [+strident]. In contrast, the specification of the [s z n] in (22b) remains the same in this respect: both [s z] and [ɕ ʑ] are [+strident] and both [n] and [ɲ] are [-strident]. IDENT$_{[\pm strid]}$ is respected.

Similarity can be measured in terms of features, as shown in (24). The contrast [s] vs. [ɕ] involves a difference in one feature, while the contrasts [t] vs. [tɕ] and [k] vs. [tʂ] each involve a change

of two features, the latter mutation crucially involves a change in major place.[18] As a result, mutations of [t d r] and [k g x] involve more contrast (i.e. they result in less similar segments) between the base and the derivative than mutations of [s z n]. This observation can be used to explain the relative acceptability of the mutations for [s z n] and their avoidance for [t d r] (gradient) and [k g x] (categorical).

(24)

	expected mutations	features that change	mutations before *-ist-a*
[s z n]	[ɕ ʑ ɲ]	[±anter]	accepted
[t d r]	[tɕ dʑ ʐ]	[±anter], **[±strid]**	avoided (gradiently)
[k g x]	[tʂ ʐ ʂ]	**Place, [±strid]**	avoided (categorically)

3.2.2 Listedness

This analysis makes crucial reference to the availability of a lexical representation of a particular word. During lexical retrieval, the lexical representations of some morphologically complex words are stored and available. The mental representations of other words are not available and the words need to be processed on-line from their component parts, e.g. base + affix. I refer to a dual-route model of lexical access (McQueen and Cutler 1998; Hay 2003; Plag 2012), which proposes that the availability of a mental representation of a word depends on the word's token frequency. A complex word may be accessed via the whole-word route or the decomposed route. The choice between the two ways of access is determined by the relative frequency of the derivative and the base, as well as the phonotactic constraints of the language. If the derivative is more frequent than the base, the derivative is more likely to be accessed via the whole-word route. If, on the other hand, the base is more frequent than the derivative, we expect the derivative to be accessed via the decomposed route. The latter type of relationship occurs more commonly in general and has been identified in all the cases of *-ist-a* words discussed here. I extend the model and propose that some words in *-ist-a* are accessed via the whole-word route and others via the decomposed route. The choice depends on their absolute frequency. In other words, frequency impacts the morphological decomposability of the derivative. The second factor that has been found to determine the decomposability of words is phonological: the phonotactic probability of segmental sequences (Plag 2012). Plag's (2012) phonotactic probability in some crucial ways parallels the phonological restrictions (formulated in terms of features) on consonant-vowel sequences, e.g. *[ʂi ʐi tʂi dʐi], enforced by Agree constraints in Section 3.1.

The constraint UseListed promotes the selection of a stored lexical representation as the input. If such a form is unavailable (for instance, as for rare and novel words), the constraint is moot, as all the potential outputs violate it (Zuraw 2000). For example, the word [altɕ-ist-a] can be accessed via the listed form /altɕ-ist-a/ or via its component morphemes /alt/ + /ist-a/. UseListed promotes the former type of access, whenever available.

(25)   UseListed
        The input portion of a candidate must be a single lexical entry.

The strength of a word's lexical entry has been found to depend on its frequency of use (token frequency) (Hay 2003). Therefore, it is necessary to gauge the strength of the lexical representations of words in *-ist-a*. In (26) I give the frequencies of the relevant bases with stem-final [t d r] and their derivatives in *-ist-a*. The words were compiled on the basis of the data extracted from the reverse dictionary (Bańko et al. 2003) and later used to investigate token frequency in a corpus. The frequencies of the lemmas were extracted from the balanced *National Corpus of Polish* (12 January 2018), which contains 250 million words from various sources from 1988 to 2010, including literature, newspapers, journals, conversations and internet texts. The Pelcra search engine (Pęzik 2012) was used. The items in (a) illustrate stable alternating patterns. The items in (b) show the only four vacillating words found in the corpus. The two values stand for the number of extracted alternating/non-alternating lemmas. Stable

---

18   In fact, mutations of the coronals [s z n t d] additionally involve a change in the feature [±back], while mutations of the velars do not. It follows that a change in major place is penalized more strongly than a change of [±back] (and [±anterior]). Ranking IdentPl above Ident$_{[±back]}$ generates this pattern.

non-alternating patterns are exemplified in (c). *Figure 2* shows boxplots for frequency of all the words in *-ist-a* with and without mutations identified in the corpus.
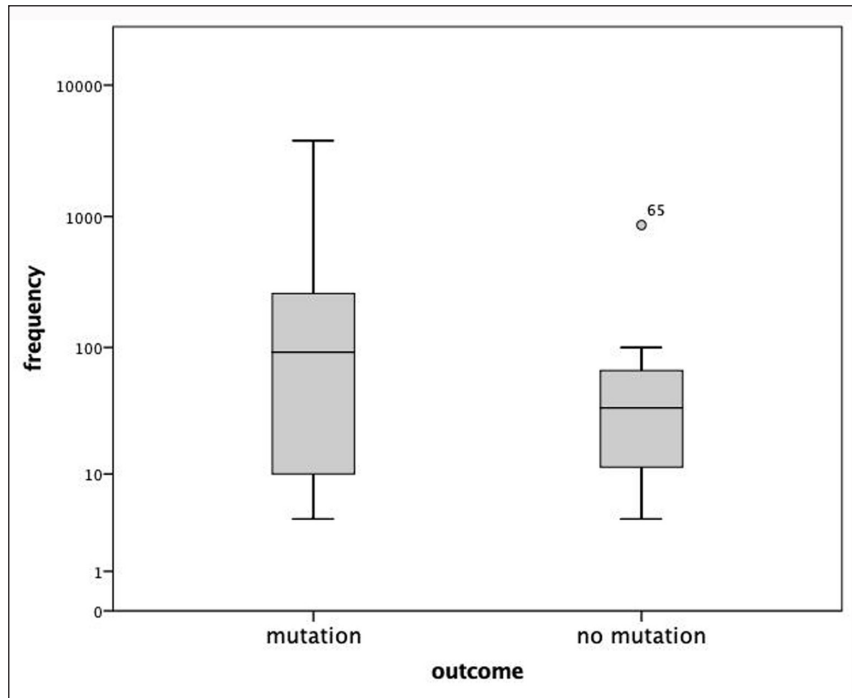
(26)   Token frequencies of representative derivatives in *-ist-a* and their bases

a.	fle[t] 'flute' 757 fle[tɕ]-ist-a 96	ren[t]-a 'disability benefit' 10593 ren[tɕ]-ist-a 3780	bryga[d]-a 'brigade' 3278 bryga[dʑ]-ist-a 255	rekor[d] 'record' 7993 rekor[dʑ]-ist-a 1650
	kontraban[d]-a 'contraband' 283 kontraban[dʑ]-ist-a 7	afe[r]-a 'scandal' 8131 afe[ʐ]-yst-a 353	rowe[r] 'bicycle' 10936 rowe[ʐ]-yst-a 3662	matu[r]-a 'high-school certificate' 6872 matu[ʐ]-yst-a 2333
b.	al[t] 'alto' 649 al[tɕ]-ist-a/ al[t]-yst-a 4/1	epizo[d] 'episode' 2758 epizo[dʑ]-ist-a/ epizo[d]-yst-a 1/6	balla[d]-a 'ballad' 1379 balla[dʑ]-ist-a/ balla[d]-yst-a 8/1	propagan[d]-a 'propaganda' 3852 propagan[dʑ]-ist-a/ propagan[d]-yst-a 81/91
c.	Bonapar[t]e 143 bonapar[t]-ysta 29	Freu[d] 4657 freu[d]-yst-a 22	folklo[r] 'folklore' 1488 folklo[r]-yst-a 28	stypen[d]ium 'stipend' 5579 stypen[d]-yst-a 861

All the derivatives extracted from the corpus show a lower frequency than their corresponding bases. An interesting tendency is discernible in the boxplots in *Figure 2*. The derivatives with mutations, such as those in (26a), on average exhibit a higher frequency than the derivatives without mutations, exemplified in (26b, c) (with the notable exception of *stypen*[d]-*yst-a*, marked as an outlier). This agrees with the predictions of usage-based models. Well-established words

are expected to be more stable, hence more resistant to PU pressures, than are rare or novel words, because of their strong mental representations. Rare and novel words, on the other hand, are more susceptible to the influence of PU pressures because their representations are weak or unavailable. Thus, we expect a difference in the frequency of words with and without mutations.

In analyzing the frequency of all the words with base-final [t d r] available in the corpus, two things deserve a mention. First, the overall number of the words is not very high: 47 with mutations and 24 without mutations. Second, among the words with mutations the degree of variance of frequency is very high. For example, there are 11 words in this group with the frequency below 10. At the same time, 6 words in this group exceed the frequency of 1000 (two of them exceed 3000). The mean frequency of all the derivatives showing stable alternating patterns, (26a), is M = 391.17, SE = 123.94 and the median is Mdn = 92. In the group of vacillating and non-alternating words, (26b) and (c), the degree of variance is not that high. Only one word, *stypen*[d]-*yst-a* (visible as an outlier in **Figure 2**), shows a frequency higher than 100; the frequency of most of the remaining words is well below 100. Crucially, the frequencies of the words that vacillate, all of them shown in (26b), are low. In fact, *propagan*[dʐ]-*ist-a/ propagan*[d]-*yst-a* is the only one among them whose frequency exceeds 100. Note, however, that the two variants of this word show very similar frequencies, which might explain their persistence. In the remaining cases in (b), one of the variants is significantly more frequent than the other (even though the overall frequency does not exceed 10). The mean frequency of the vacillating and non-alternating derivatives (excluding *stypen*[d]-*yst-a*) is M = 38.13, SE = 6.21 and the median is Mdn = 33. A non-parametric Mann-Whitney test was run on the data. On average, the words with mutations are more frequent than the words without mutations or vacillating words, U = 356.00, z = –2.308, p < .05.[19]

Finally, how can we explain the unusual behavior of *stypen*[d]-*yst-a*? Being of recent origin, the word was formed when the non-alternating pattern was more frequent than the corresponding alternating pattern. It began to be used very frequently and its representation stabilized (without mutations). Words like *stypen*[d]*ysta* show that high-frequency words can represent two patterns in modern usage: an alternating one when the word was formed earlier and a non-alternating one when it is of a more recent origin.

Some of the important predictions of usage-based models are upheld. Words of the highest frequency (i.e. above 1000 in the analyzed data) are phonologically stable. We would not expect a word like *ren*[tɕ]-*ist-a*, whose frequency is the highest among the analyzed words (3780), to appear in current usage in its corresponding form without a mutation, i.e. as **ren*[t]-*yst-a*. Its strong memory trace prevents such an outcome and UseListed promotes the selection of its stored representation as input. In contrast, words showing a relatively low value of frequency are susceptible to PU pressures, for example, *bonapar*[t]-*yst-a* with the frequency of 29. In addition, low-frequency words show more variation than high-frequency words, e.g. *al*[tɕ]-*ist-a/al*[t]-*yst-a* (4/1). A less expected discovery is a substantial number of words with mutations whose frequency is low – 11 of them have a frequency below 10, for example, *kontraban*[dʐ]-*ist-a* (7). Their persistence may be due to the gradual and probabilistic nature of change (lexical diffusion). What is more, some of the words with mutations (e.g. *kontraban*[dʐ]-*ist-a* 'contrabandist') are relatively old and predictably comply with the phonological requirements which were regular at the time of their formation. Their low frequency today may well reflect changes in the society. In low-frequency vacillating words such as *al*[tɕ]-*ist-a* ~ *al*[t]-*yst-a* the variant with a mutation is most likely to be older than the variant without a mutation. Another reason for the persistence of low-frequency words with mutations might have to do with the fact that the corpus represents mostly written language, which means that the data reflect conservative usage and may be an imperfect representation of spoken language. In order to address this issue, in Section 3.2.4, we will examine data elicited in an experiment involving native speakers of Polish.

3.2.3 Analysis

We begin with an analysis of words with base-final [s z n], which show mutations. IdentPl is not shown, as it is moot for coronals. In the evaluation of *bas-ist-a* [baɕ-ist-a] 'bass player'

---

19   If we include *stypendysta* in the sample, the mean frequency of non-alternating and vacillating words is M = 72.42, SE = 34.8 and the median is Mdn = 34. The difference between the two groups remains significant, U = 397.00, z = –2.03, p < .05.

(a recent word), the listed form is unavailable, as indicated below the Input in (27). The mutation in the winning candidate causes a violation of the low-ranked IDENT$_{[\pm anterior]}$, the latter requires faithfulness to the feature [±anterior]. The evaluation highlights two important issues. First, novel words do not have strong lexical representations, therefore, UseListed is violated by all the candidates, and the relevant schema-constraint determines the output. Second, once morphophonological schemas become entrenched (i.e. patterns are morphologized), markedness constraints regulating mutations (e.g. AgreeCor$_{[\pm back]}$) take a back seat. This is due to the preference for the morphological over phonological conditioning of patterns, discussed in Section 2.5. The entrenchment of a pattern is a function of the size of the gang it represents (gang effects, see Section 2.3). In OT formalism, the morphologization of a pattern may be viewed as the emergence and promotion of the relevant schema-constraint in response to the increasing number of words that observe the pattern.

(27)   Evaluation of a derivative in *-ist-a* of [bas].

Input: bas + ist-a Listed: ∅	UseListed	[...s] ↔ [[[...ɕ]ist]a]	IDENT$_{[\pm anter]}$
a. bas-ist-a	*	*!	
☞ b. baɕ-ist-a	*		*
c. bas-ist-a	*	*!	

As regards derivatives with base-final [t d r], the token frequency values in the corpus of the representative words *al*[tɕ]*-ist-a*, *bonapar*[t]*-yst-a* and *fle*[tɕ]*-ist-a* are 4, 29 and 96, respectively. The appearance of *al*[tɕ]*-ist-a* alongside *al*[t]*-yst-a* is an indication that the word is variably accessed via either the whole-word route or the decomposed route. *bonapar*[t]*-yst-a* is mainly retrieved via the decomposed route. As the mental records of both words are weak (due to their low frequency), transparent bases facilitate their retrieval. Finally, *fle*[t] ~ *fle*[tɕ]*-ist-a* does not show fluctuations because the derivative, having its own mental trace (due to a considerably higher frequency than *al*[tɕ]*-ist-a* and *bonapar*[t]*-yst-a*), is mainly accessed via the whole-word route.

In the evaluation of the derivative from *Bonapar*[t]*e* the listed form is unavailable due to a low token frequency and, therefore, UseListed (not shown) is violated by all the candidates. IDENT$_{[\pm strid]}$ and IDENT$_{[\pm anterior]}$ represent PU pressures.

(28)   Evaluation of a derivative in *-ist-a* of [bɔnapartɛ]

Input: bɔnapart + ist-a Listed: ∅	IDENT$_{[\pm strid]}$	[...t] ↔ [[[...t] ist]a]	[...t] ↔ [[[...tɕ] ist]a]	IDENT$_{[\pm anter]}$
a. bɔnapart-ist-a		*!	*	
☞ b. bɔnapart-ist-a			*	
c. bɔnapartɕ-ist-a	*!	*		*

Currently, the non-alternating schema for base-final [...t] has a higher type frequency, hence higher strength, than the alternating schema. This finds reflection in the ranking of the two schemas in tableau (28). Candidate (b) is selected because it satisfies both IDENT$_{[\pm strid]}$ and the dominant schema. Candidate (a) uses neither of the schemas available in the lexicon for the formation of words in *-ist-a* and hence violates the respective schema-constraints. Candidate (c) is eliminated due to a violation of IDENT$_{[\pm strid]}$.

To derive an earlier state of affairs when the alternating pattern prevailed in the lexicon, the ranking of the two schema-constraints must be reversed in (28). This would account for the expansion of non-alternating patterns at the stage when the alternating patterns were in fact more common. With IDENT$_{[\pm strid]}$ ranked high, the candidate with a transparent base, (b), is selected regardless of the ranking of the two schema-constraints with respect to each other.

In the case of *fle*[t] ~ *fle*[tɕ]*-ist-a,* the derivative *fle*[tɕ]*-ist-a* has a relatively high token frequency, which entails that the listed form is available. The morphologically complex

word is independently stored and hence is accessed via the whole-word route. The impact of PU pressures as well as schemas is outweighed by the impact of the high-ranked UₛₑLısᴛᴇᴅ.

(29)   Evaluation of a derivative in *-ist-a* of [flɛt]

Input: flɛt + ist-a Listed: flɛtɕ-ist-a	UₛₑLısᴛᴇᴅ	IDENT$_{[\pm strid]}$	[...t] ↔ [[[...t]ist]a]	[...t] ↔ [[[...tɕ]ist]a]	IDENT$_{[\pm anter]}$
a. flɛt-ist-a	*!		*	*	
b. flɛt-ist-a	*!			*	
☞ c. flɛtɕ-ist-a		*	*		*

Vacillating forms such as *al*[tɕ]*-ist-a ~ al*[t]*-yst-a* also indicate that the impact of lexical strength, established on the basis of token frequency, may reduce the influence of PU pressures. The token frequency of *al*[tɕ]*-ist-a* is low. Together with the attested variation, this is interpreted as evidence that the word may be alternately accessed via the whole-word route when the listed form is available, or the decomposed route when the listed form is unavailable. The availability of the listed form is determined probabilistically on the basis of token frequency.[20] When the word is accessed via the whole-word route, faithfulness constraints are outranked, and UₛₑLısᴛᴇᴅ effectively selects the output. The tableaux in (30) illustrate the variability of the output. When a stored derivative is available, as shown in the first tableau, candidate (c) is selected (with a mutation), as it is the only one that satisfies UₛₑLısᴛᴇᴅ. On the other hand, when the stored representation is unavailable and UₛₑLısᴛᴇᴅ is violated by all the candidates (the second tableau), the word is accessed via the decomposed route and derived on-line from its component parts. In such a situation, candidate (a) fails the evaluation because it does not respect either of the available schemas. Candidate (c) loses to candidate (b) because it contains a modification of stridency specification. To summarize, the variation between *al*[t]*-yst-a* and *al*[tɕ]*-ist-a* can be explained in a dual-route model of lexical access by assuming that the former variant is derived on-line from its component parts, while the latter is retrieved from memory via the whole-word route. The outcome depends on the probabilistic availability of the memorized representation.

(30)   Evaluation of a derivative in *-ist-a* of [alt]

Input: alt + ist-a Listed: altɕ-ist-a	UₛₑLısᴛᴇᴅ	IDENT$_{[\pm strid]}$	[...t] ↔ [[[...t]ist]a]	[...t] ↔ [[[...tɕ]ist]a]	IDENT$_{[\pm anter]}$
a. alt-ist-a	*!		*	*	
b. alt-ist-a	*!			*	
☞ c. altɕ-ist-a		*	*		*

Input: alt + ist-a Listed: ∅	UₛₑLısᴛᴇᴅ	IDENT$_{[\pm strid]}$	[...t] ↔ [[[...t]ist]a]	[...t] ↔ [[[...tɕ]ist]a]	IDENT$_{[\pm anter]}$
a. alt-ist-a	*		*!	*	
☞ b. alt-ist-a	*			*	
c. altɕ-ist-a	*	*!	*		*

Such vacillating forms suggest that the three pressures, i.e. PU, type frequency, and token frequency, are important in the language and that the actual usage reflects their combined effects. The differences between the impact of type and token frequency will be elaborated in Section 5. The attestation of *al*[tɕ]*-ist-a* alongside *al*[t]*-yst-a* may be explained by the different method of retrieval of the word. When it is accessed via the whole-word route, the stored representation is used, i.e. *al*[tɕ]*-ist-a*. When the word is accessed via the decomposed route, mutations in the base are detrimental to word recognition and PU becomes relevant, yielding *al*[t]*-yst-a*. In the case of *bonapar*[t]*-yst-a*, as the derivative with a mutation is not listed in the mental lexicon, a transparent base serves to facilitate word recognition. Finally, in the case of well-established words, such as *fle*[t] ~ *fle*[tɕ]*-ist-a*, the impact of PU constraints is mitigated by

---

20   Alternatively, the availability of listed representations could be categorical, while the ranking of UₛₑLısᴛᴇᴅ could be probabilistic and modeled using stochastic (Boersma 1998) or weighted constraints (Pater 2009). This aspect of the analysis is left for future research, as I can see no clear advantage of one approach over the other.

their strong lexical representations every time they are accessed. The ranking of the relevant constraints established in the course of the analysis is given in (31).

(31)   UseListed, IdentPl $>>$ Ident$_{[\pm strid]}$, [...t] $\leftrightarrow$ [[[...t]ist]a], [...s] $\leftrightarrow$ [[[...ɕ]ist]a]
         $>>$ [...t] $\leftrightarrow$ [[[...tɕ]ist]a], Ident$_{[\pm anter]}$

3.2.4 Experiment

In order to offer more evidence for the continuous role of PU pressures in the grammar, I use the findings of a nonce-word experiment conducted on 61 participants, speakers of Polish, aged 19–23. The participants were asked to fill out a written questionnaire probing their intuitions regarding the formation of words in *-ist-a/-yst-a*. The main area of interest was the formation of *-ist-a/-yst-a* words for base-final coronals [t d r] and [s z n]. In the first part of the questionnaire, the participants were given examples of *-ist-a/-yst-a* words along with their meaning. At this stage, real words were used and only those whose bases ended in a non-coronal. In the second part, the participants were given a list of two-syllable nonce words and asked to form words in *-ist-a/-yst-a*. Non-coronals were used as fillers, as we do not expect significant variation here. In order to get a more varied sample of the relevant patterns, twice as many words with base-final [t d r] and [s z n] were elicited than with the remaining consonants, that is, each participant had to form two words with base-final [t d r s z n] and one word with base-final [p b m f v ʂ ʐ tʂ dʐ k g x]. The words comply with the Polish phonotactic constraints. The base-final syllable is of the structure CVC(C), where the V is [a], [ɔ], [u], [ɛ] or [i]. Two versions of the questionnaire, differing in the choice of nonce words, were assigned randomly to the participants, each containing 24 stimuli. 30 participants completed Questionnaire A and 31 participants completed Questionnaire B. The list of words used in the questionnaires is provided in the Appendix. An example of a task (translated) is given in (32). The respondents were instructed to provide the word that sounds the most natural to them and that there are no bad answers. They were asked to provide the first word that comes to their mind. The responses of one participant were excluded, as they were largely incongruous.

(32)   **klunar** is a new scientific discipline.
         A person who studies or represents **klunar** is called _____.

The results of the experiment are given in **Table 5**. Predictably, the results are not as straightforward as the ones given for dictionary entries in **Table 3**. The category "other" subsumes all the cases in which the participants failed to form a derivative in *-ist-a/-yst-a* using one of the attested patterns (e.g. the expected derivatives of *miarte*[z] are *miarte*[ʑ]*-ist-a* and *miarte*[z]*-yst-a*, yet, 5 participants chose *miarte*[ʐ]*-yst-a* instead). For labials and retroflexes mutations were not elicited. Of particular interest are consonants which appear in both non-alternating (no mutation) and alternating (mutation) patterns in **Table 5**, i.e. velars and coronals (except retroflexes). While the former appear in alternating patterns sporadically, in the case of the latter, the variation reaches a noticeable level.

Table 5 provides grounds for several interesting observations. Mutations of velars are marginal and may be the result of an indirect influence of Velar Mutation (discussed in Section 2.1) on this pattern. Mutations occur more frequently in the case of coronals (except retroflexes). While non-alternating patterns occur in more than 80% of the cases for [t d r], for [s z] the rates of non-alternating patterns do not exceed 60%. The coronal nasal is the most susceptible to mutations; the rate of non-alternating patterns for [n] does not reach 6%.

A conditional inference tree analysis using the *party* package in R (Hothorn et al. 2006) has been run on the data, with *outcome* (mutation/no mutation) as the dependent variable and *consonant* as the predictor (18 levels). The category "other" has been omitted, as it is not particularly revealing. A conditional inference tree provides estimates of the likelihood of the value of the response variable (mutation/no mutation) on the basis of a series of binary questions about the values of predictor variable (consonant). This method has been chosen, as it is designed for binary dependent variables and deals well with unbalanced data and small or even zero cell counts (Tagliamonte & Baayen 2012).

The first partitioning in the decision tree in **Figure 3** shows that mutations of [s z n] are overwhelmingly more common than mutations of the remaining consonants (p < .001). Similarly, mutations of [t d r] are more common than mutations of all the other consonants on the right-hand side of the tree (p < .001). In addition, [n] is significantly more susceptible

BASE-FINAL CONS.	MUTATION	NO MUTATION	% OF NO MUTATION	OTHER
t	10	105	91.3	5
d	8	107	93	5
r	19	96	83.5	5
s	66	47	41.6	7
z	45	62	57.9	13
n	113	7	5.8	0
p		60	100	0
b		60	100	0
m		60	100	0
f		59	100	1
v		60	100	0
ş		58	100	2
z̧		59	100	1
tş		60	100	0
dz̧		59	100	1
k	2	58	96.7	0
g	3	56	94.9	1
x	1	59	98.3	0

Table 5 Number of alternating and non-alternating patterns in *-ist-a* words for each base-final consonant in the nonce-word experiment.
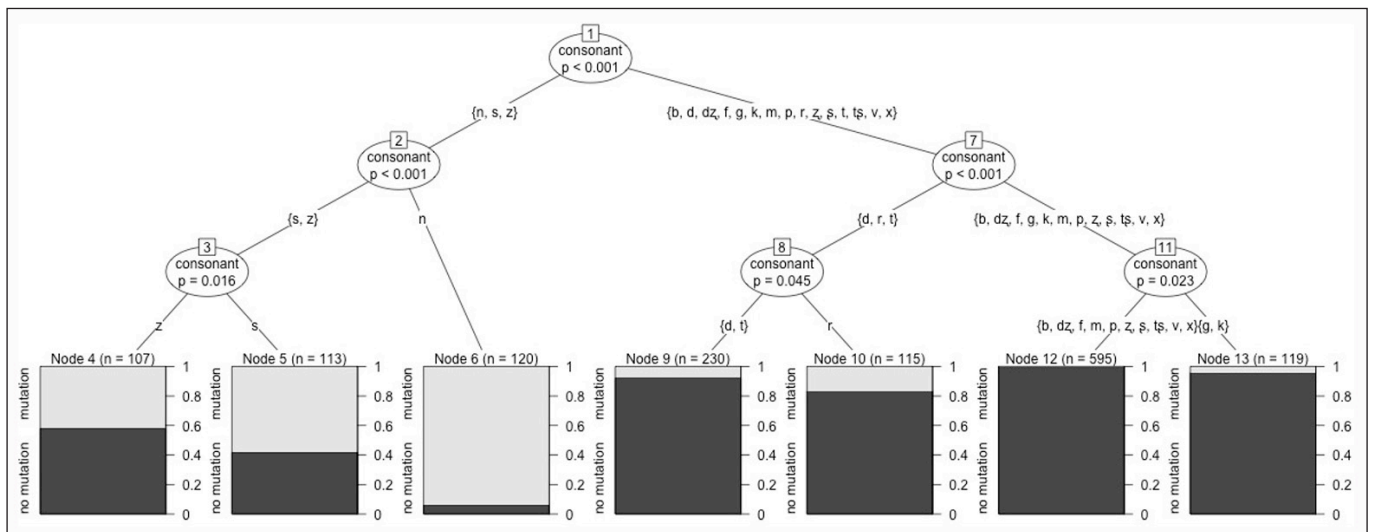
Figure 3 Conditional inference tree for *outcome* (mutation/no mutation) as the response variable and *consonant* as a predictor. Mutation is depicted as light gray, No mutation as dark gray.

to mutations than both [s] and [z] (p < .001); the latter two are also different from each other (p = .016). The likelihood of mutations for [r] is higher than for [t d] (p = .045). In a similar way, [k g] are different from all the other consonants. Crucially, mutations of [s z n] are significantly more common than mutations of [t d r], and the latter are more common than mutations of the remaining consonants.[21] These findings are compatible with the OT analysis provided in the previous sections.

Figure 3 shows a highly significant difference between the likelihood of mutations for [s z] and [n] in the experimental results. Why is the tendency to eliminate mutations stronger for both [s] and [z] than for [n]: 41.6% and 57.9 vs. 5.8%? The answer might have to do with cue robustness and perceptibility. Wright (2004: 43–44) reviews the activity of auditory nerve

---

21 Pairwise comparisons for [t d r s z n] using a generalized mixed-effects logistic regression (glmer, binomial) with speaker and word as random intercepts and consonant as a fixed effect were also run. The regression could not be used for all the consonants, as the method does not handle zero counts. The results in essence confirm the results of the decision tree analysis. It was found that /n/ is significantly different from all the consonants (p < .001) and /s/ and /z/ are different from /t/, /d/ and /r/ (p. < 001). The other comparisons that yielded differences are: /s/ is different from /z/ (p = .026), /r/ is different from /t/ (p = .046) and from /d/ (p = .016). /t/ and /d/ are not significantly different from each other.

fibers and suggests that consonants receive an auditory boost in CV sequences. Crucially, the boost is particularly true of stops, fricatives and affricates but less so of nasals. Thus, the auditory difference between [n] and [ɲ] is smaller than the distance between [s] and [ɕ] (and [z] and [ʑ]). It follows that a higher rate of mutations for [n] than for [s z] can be explained by appealing to cue perceptibility. Steriade's (2008) approach using P-maps and contrast-based constraints offers a formal solution. PU constraints require that the relevant parts of derivatives be maximally similar to their bases in order to aid word recognition. Acceptability of mutations of the base in the derivative is in fact gradient, with mutations that diverge the most from their base correspondents being less acceptable than mutations of consonants that render them more similar to their correspondents. In other words, mutations that introduce less perceptual contrast between the derivative and the base are preferred. An analysis employing P-maps appeals to the ranking of the contrast-based constraints Δ (s~ɕ) >> Δ (n~ɲ) and their interaction with phonotactic constraints and morphophonological schemas. This ranking derives from the fact that the alternation (s~ɕ) universally involves more perceptual contrast than (n~ɲ) and, therefore, is more likely to be avoided, all else being equal. Perceptibility also explains the distribution of palatal coronals in Polish: [s] is possible before [i] in recent borrowings ([sinus] *sinus* 'sine') and is thus potentially contrastive with [ɕ] ([ɕivɨ] *siwy* 'gray'), while [n] is impossible before [i] across the board and neutralizes to [ɲ] in this context ([uɲifɔrm] *uniform* 'uniform').

The rates in *Table 5* can be directly compared with the results of the investigation of dictionary entries in *Table 3*. *Table 6* shows the rates of words without mutations in the two relevant groupings of consonants, that is, [t d r] and [s z n], and juxtaposes the results obtained in the experiment with those taken from the dictionary. A cursory look at *Table 6* warrants a claim that the non-alternating patterns (no mutations) are more common in the experimental data than in the dictionary data for [t d r s z n].

% OF WORDS WITH NO MUTATIONS	t	d	r	s	z	n
experiment	91.3	93	83.5	41.6	57.9	5.8
dictionary	62.3	53.8	64.6	0	0	0

Table 6 Rates of words without mutations in the experiment and the dictionary.

Figure 4 offers a conditional inference tree with *outcome* (mutation/no mutation) as the dependent variable, and *consonant* ([t d r s z n]) and *source* (experiment/dictionary) as predictor variables. *Table 6* and the decision tree in *Figure 4* show two things: first, the experimental results are not as categorical as the dictionary data and, second, the alternating patterns are significantly less commonly used by the participants of the experiment than in the dictionary ($p < .001$ for [t d r s z] and $p = .021$ for [n]).[22]
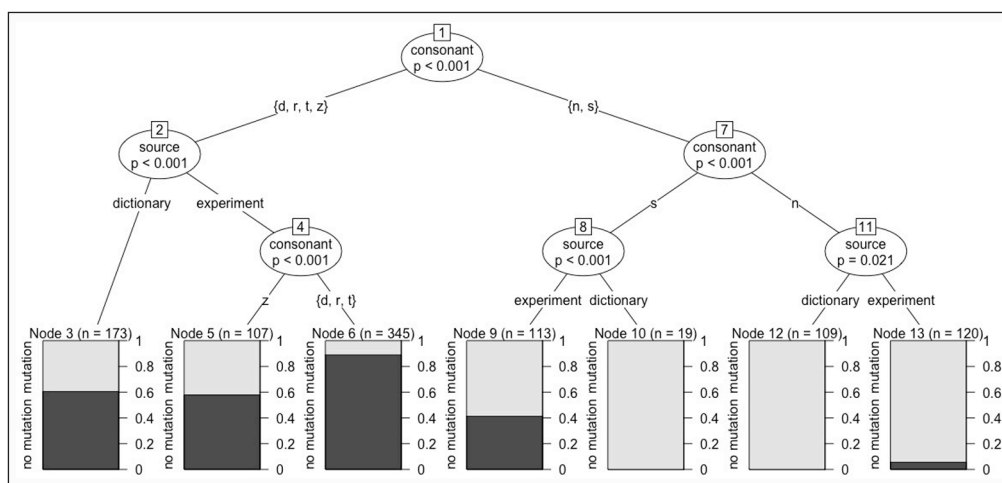

Figure 4 Conditional inference tree for *outcome* (mutation/no mutation) as the dependent variable and *consonant* ([t d r s z n]) and *source* (dictionary/experiment) as predictors. Mutation is depicted as light gray, No mutation as dark gray.

Admittedly, the comparison across the two conditions (dictionary/experiment) should be approached with caution because of the different methods of data collection. Dictionaries

---

22 The less expected patterning of [z] with [t d r] in the first partitioning in Figure 4 is likely associated with the paucity of *-ist-a* words with base-final [z] in the dictionary data (2 items), see Table 3.

offer prescriptive data, often ignoring variation found in actual language use. Data extracted from a corpus provide a better match for experimental data because the method of collecting data for a corpus is more compatible with the method of obtaining data from respondents in experimental conditions.[23] ***Table 7*** compares the number of mutations of base-final [t d r] in two conditions: experiment vs. corpus (*The National Corpus of Polish*). A logistic regression run on the data suggests that the *condition* variable (experiment/corpus) has a significant influence on the *outcome* variable (mutation/no mutation). The coefficient on the *condition* variable has a Wald statistic equal to 83.581, which is significant at the .001 level (df = 1). The overall model is significant at the .001 level according to the Model chi-square statistic ($\chi^2(1) = 92.607$).

	MUTATION	**NO MUTATION**
experiment	37	308
corpus	47	24

Table 7 Number of words in base-final [t d r] with and without mutations in the experiment and in the corpus.

There are two possible reasons for the different rates of paradigm uniformity effects in the experiment vs. in the corpus. First, while novel and rare words in the corpus correspond to nonce words in the experiment, well-established words in the corpus do not have analogues among the experimental stimuli. Given that novel words are more susceptible to paradigm uniformity pressures than established words, the rates of mutations are expected to be lower for the experimental results. Second, it is likely that the data indicate an on-going change that might lead to the gradual elimination of mutations from the *-ist-a* pattern. The fact that there is a difference between the dictionary and corpus data, on the one hand, and the experimental data, on the other, might suggest that PU pressures continuously affect mental grammars and are responsible for a gradual change in morphophonological patterns. In particular, the promotion of IDENT$_{[\pm anterior]}$ above schema-constraints is held accountable for the reduction in the number of mutations for base-final [s z n] in the experimental results, as compared with the dictionary and corpus data.

4 A HIGH-FREQUENCY PATTERN: DIMINUTIVES IN *-ek*

The purpose of this section is to demonstrate that morphophonological patterns exhibiting a high type frequency in the lexicon are stable and that the schema-constraints encoding them are ranked higher than schema-constraints representing patterns with a lower type frequency. With relevance to the analysis at hand, the schema-constraints representing the *-ek* suffix are ranked higher than the schema-constraints representing the *-ist-a* suffix. This reflects their relative robustness in the grammar. The schema-constraints for the suffix *-ek* are ranked above the faithfulness constraints preventing consonant mutations, i.e. PU IDENT constraints. In this way, well-established morphophonological patterns are respected at the expense of PU considerations. In order to evade the potential impact of strong lexical representations and ensure that the words are accessed via the decomposed route, let us consider some relatively new diminutives ending in *-ek*.[24]

(33)   drink [drink] 'drink'   drineczek [drinɛtʂ-ɛk], drinczek [drintʂ-ɛk]
        buldoga [buldɔg-a] 'bulldog' GEN SG buldożek [buldɔʐ-ɛk]

A search of the *plTenTen* corpus for diminutives in *-ek* without mutations of velars (e.g. *[drink-ɛk] or *[buldɔg-ɛk]) has given zero results. It appears that words in *-ek* show stable mutations of velars in spite of PU violations. Given the stability of this pattern confirmed by the lack of words representing this pattern without mutations in the corpus, a nonce-word experiment probing speaker intuition is unlikely to be insightful. The high type frequency of this pattern overrides phonological constraints (PU pressures). In tableau (34), the relevant schema-constraint is ranked above the constraints enforcing base transparency.

---

23   There is accumulating evidence suggesting that comparisons across different conditions provide valuable tools for assessing the psychological reality of corpus-based models. For example, Djivak et al. (2016) explicitly compare the performance of a statistical model based on data derived from a corpus with the performance of native speakers in selecting one of six Russian verbs meaning 'try'.

24   Two diminutives for [drink] are attested: one with e-insertion in the base, [drinɛtʂ-ɛk], and one without, [drintʂ-ɛk]. The conditioning of e-insertion in diminutives is beyond the scope of this paper, but see Czaplicki (2020). The token frequencies of the lemmas *drinecz-ek* and *drincz-ek* in the *plTenTen* corpus are low: 594 and 33, respectively. The token frequency of the lemma *buldoż-ek* is higher: 2,351.

(34)   Evaluation of a derivative in *-ek* of [drink]

Input: drink + ɛk Listed: ∅	[...k] ↔ [[...tʂ] ɛk]	IDENTPl	IDENT$_{[\pm strid]}$
a. drinɛk-ɛk	*!		
☞ b. drinɛtʂ-ɛk		*	*

UseListed (not shown) is violated by both candidates. The faithful candidate in (a) fails the evaluation because it does not respect the dominant schema-constraint encoding the formation of words in *-ek* with base-final [k]. Candidate (b) fares well on the schema-constraint and, in spite of violations of the PU identity constraints, comes out victorious. As a result, the high type frequency of the pattern *-ek* ensures that mutations in novel words are stable.[25]

5 DISCUSSION

This section aims to tease apart and compare the formal mechanisms of dealing with the effects of type and token frequency. As regards type frequency, constraints that represent morphophonological schemas are ranked according to the frequency of the patterns they encode. In this way, the impact of stronger patterns, i.e. those with a higher type frequency, is greater than the impact of weaker patterns. Ranking schemas allows us to model the competition between morphophonological patterns and PU constraints. To be more specific, more robust patterns are predicted to override PU considerations, which might result in the preservation of mutations and is manifest in the formation of words in *-ek*. The formation of words in *-ist-a*, on the other hand, shows that weaker morphophonological patterns are dominated by PU pressures, a ranking which opens the way for patterns respecting base transparency (i.e. patterns without mutations). The different impact of paradigm uniformity pressures on patterns (schemas) of high and low frequency can be represented as in (35). The stronger the pattern (schema), the more resistant it is to PU pressures.

(35)   Schema$_{high\ frequency}$ >> PU >> Schema$_{low\ frequency}$

In addition, individual words of high frequency are predicted to resist PU pressures, whether they represent a high-frequency or a low-frequency schema. This is due to the availability of the listed representations of high-frequency words and the bias for the whole-word route of lexical access. In the proposed analysis the impact of token frequency was mediated by the constraint UseListed, which promotes the use of listed forms (when available). The effects of token frequency were detectable on agent nouns in *-ist-a*, a low-frequency pattern. While mutations are stable for high-frequency words representing the pattern *-ist-a*, they tend to be eliminated from low-frequency words representing this pattern.

(36)   UseListed >> PU

The effects of token frequency (whether high or low) have not been identified for *-ek*, a high-frequency pattern (schema).[26] The ranking of the constraints established in the course of this analysis is given in (37).

---

25   It is in principle possible that within a construction such as diminutives in *-ek*, one of the schemas, for instance, [x] ~ [ʂ-ɛk], might have a significantly lower type frequency than the other schemas, and would, consequently, behave differently with respect to PU pressures. Czaplicki (2014b) discusses diminutives in *-ik/-yk* in Polish and demonstrates that a schema of a low type frequency (pertaining to [ʂ]) within this construction has been altered, whereas comparable stronger schemas (e.g. pertaining to [ʐ]) remain stable. However, schema modification of this type is not likely in the case of such high-frequency constructions as diminutives in *-ek*. For example, in the *plTenTen corpus*, the type (and total token) frequency values of words in *-ek* with base-final velars [k g x] are as follows: [k] ~ [tʂ-ɛk] – 528 (1,896,609), [g] ~ [ʐ-ɛk] – 28 (341,000), [x] ~ [ʂ-ɛk] – 81 (1,423,892). The corresponding values for words in *-ist-a* are: [k] ~ [k-ist-a] – 5 (9,536), [g] ~ [g-ist-a] – 13 (14,951) and [x] ~ [x-ist-a] – 4 (39,182). The frequency values of all the schemas for *-ek* are much higher than those for *-ist-a*, which makes modification of the schemas for *-ek* unlikely.

26   The working assumption is that the type frequency of the entire construction (the number of all the words that represent it) determines its strength. Pattern strength can also be measured by the sum of *token* frequencies of all the words that show the pattern. The hypothesis that the sum of token frequencies of all the words that represent a pattern is a more reliable determinant of the pattern's strength than the type frequency of the pattern certainly deserves verification.

(37)   Schema$_{\text{high frequency}}$, UseListed >> PU >> Schema$_{\text{low frequency}}$

More generally, the possibility of assigning different rankings to schemas means that each construction has its own subgrammar (cophonology), where schema-constraints are interspersed with phonological constraints.

The proposed account diverges from previous accounts of Polish consonant mutations (e.g. Rubach 1984; Szpyra 1989; Ćavar 2004 and Gussmann 2007) in two crucial aspects. First, it claims that morphological patterns (constructions) differ in terms of the different degrees of stability of consonant mutations, where the conditioning factor is type frequency. Token frequency is held accountable for the stability of alternations in individual words. Neither type nor token frequency was explicitly used in previous accounts, which means that the different propensity of particular constructions and words for segmental alternations was unaccounted for. Second, previous accounts argued that phonological context is the principal factor determining the targets and triggers of palatalization.[27] In contrast, the proposed account places emphasis on the morphological (i.e. construction-specific) conditioning of consonant mutations.

6 CONCLUSIONS

Two morphophonological patterns (constructions) showing consonant mutations in Polish have been analyzed: a low-frequency pattern – formation of agent nouns with *-ist-a* and a high-frequency pattern – diminutive formation with *-ek*. It has been shown that the tendency to avoid mutations has an effect on a construction with a low type frequency in the lexicon. The more frequent the construction, the more likely it is to remain unaffected by paradigm uniformity pressures. Put differently, less frequent patterns are more susceptible to modification due to phonological constraints than more frequent patterns. In order to account for this regularity, morphophonological schemas embodying these patterns must be represented in the grammar according to their type frequency. In Optimality Theory schema-constraints pertaining to the pattern *-ek* were ranked higher than those enforcing the pattern *-ist-a*. The impact of paradigm uniformity was represented as an interaction of output-output faithfulness constraints with morphophonological schemas. The discussion has offered evidence for construction-based cophonologies. In this approach, each morphological construction has its own phonological properties. It has been argued that schemas and the constructions they represent may show various degrees of susceptibility to PU pressures. The stability of a construction depends on its type frequency. The dynamic role of PU pressures in the grammar was confirmed by the results of an experiment which showed that mutations are being gradually eliminated from the *-ist-a* pattern.

In addition, the frequency of words (token frequency) has been shown to impact the drive for identity between the base and the derivative. The less frequent a word, the stronger the pressure to preserve its base and avoid mutations. The relationship between frequency and mutations is rooted in language processing. Words of higher frequency are accessed via the whole-word route, while low-frequency words are processed by accessing both the base and the affix. Thus, preservation of the base intact in the derivative is more important for less frequent words, as it speeds up their lexical access. Finally, the acceptability of mutations has been shown to depend on the degree of featural/perceptual similarity of mutated consonants to their base correspondents. A family of Output-Output faithfulness constraints targeting various features has been used to enforce base transparency. It has been demonstrated that the degree of featural similarity between derivative-base correspondents determines the acceptability of the derivative.

The discussion has shown that frequency is a key element of a predictive and explanatory phonological analysis. Both type and token frequency condition the stability of morphophonological patterns and are relevant in pattern maintenance and change. It follows that phonological and morphological theories must be designed in such a way as to allow frequency to play a major role. This is a departure from previous analyses of

---

27   Ćavar (2004) additionally attributes the effects of mutations to the pressures to mark morphological boundaries and maintain underlying contrasts.

segmental alternations in the generative tradition, where frequency was epiphenomenal and extraphonological.

We have looked at two constructions at the opposite ends of the type frequency spectrum. Future research should also focus on investigating patterns of intermediate frequency. If the proposed analysis is on the right track, the effects of phonological constraints (e.g. paradigm uniformity constraints) should be less categorical for such patterns, that is, there should be more variation.

APPENDIX

Nonce words used in the experiment

BASE-FINAL CONS.	QUESTIONNAIRE A		QUESTIONNAIRE B	
/p/	skranɔp		dɔɲap	
/b/	truzarb		tɔnab	
/m/	plɛnɔm		sɔram	
/f/	spakɔf		ligarf	
/v/	pɔskrav		maskav	
/t/	ʂpankat	mjarkut	prɔnat	tʂukat
/d/	tʂɛwɔd	szpalkɛd	mravad	mɔntard
/s/	maklɛs	zdultus	pɔntas	bɔnɛrs
/z/	mjartɛz	vujtaz	lɔkrɛz	vɨntruz
/n/	spirɔn	pɔskrɛn	raspan	kɔltan
/r/	mizdar	blɛdɔr	klunar	buɲɛr
/ʂ/	mlanɛʂ		guntɛrʂ	
/ʐ/	vajtɛʐ		hɔmpɛʐ	
/tʂ/	kɔrpitʂ		bɔnkɛtʂ	
/dʐ/	vaspadʐ		dʐɛvjɛdʐ	
/k/	prantuk		surtak	
/g/	bramag		ʂɔmag	
/x/	mrunax		dɔvrix	

ABBREVIATIONS

ADJ = adjective, DIM = diminutive, GEN = genitive, INS = instrumental, LOC = locative, MASC = masculine, NOM = nominative, PL = plural, SG = singular

ACKNOWLEDGEMENTS

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR AFFILIATION

Bartłomiej Czaplicki   ⓘD   *orcid.org/0000-0003-2833-2421*
University of Warsaw, PL

REFERENCES

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90. 119–161. DOI: *https://doi.org/10.1016/S0010-0277(03)00146-X*

Alegre, Maria & Peter Gordon. 1999. Rule-based versus associative processes in derivational morphology. *Brain and Language* 68. 347–354. DOI: *https://doi.org/10.1006/brln.1999.2066*

Andersen, Henning. 1978. Perceptual and conceptual factors in abductive innovations. In Jacek Fisiak (ed.), *Recent developments in historical phonology,* 1–22. The Hague: Mouton. DOI: *https://doi.org/10.1515/9783110810929.1*

Anderson, Stephen R. 1981. Why phonology isn't "natural". *Linguistic Inquiry* 12(4). 493–539.

Anderson, Stephen R. 2008. Phonologically conditioned allomorphy in the morphology of Surmiran (Rumantsch). *Word Structure* 1. 109–124. DOI: *https://doi.org/10.3366/E1750124508000184*

Anttila, Arto. 2006. Variation and opacity. *Natural Language and Linguistic Theory* 24. 893–944. DOI: *https://doi.org/10.1007/s11049-006-0002-6*

Baayen, R. Harald, James M. McQueen, Ton Dijkstra & Robert Schreuder. 2003. Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In R. Harald Baayen & Robert Schreuder (eds.), *Morphological structure in language processing,* 355–390. Berlin: Mouton de Gruyter.

Bach, Emmon & Robert Harms. 1972. How do languages get crazy rules? In R. P. Stockwell & R. K. S. Macaulay (eds.), *Linguistic change and generative theory*, 1–21. Bloomington: Indiana University Press.

Bańko, Mirosław, Dorota Komosińska, & Anna Stankiewicz (eds.) 2003. *Indeks a tergo do Uniwersalnego słownika języka polskiego pod redakcją Stanisława Dubisza*. Warszawa: Państwowe Wydawnictwo Naukowe.

Bateman, Nicoleta. 2007. *A crosslinguistic investigation of palatalization.* Ph.D. dissertation, University of California, San Diego.

Becker, Michael & Maria Gouskova. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47(3). 391–425. DOI: *https://doi.org/10.1162/LING_a_00217*

Benua, Laura. 1997. *Transderivational identity: Phonological relations between words*. Ph.D. dissertation, University of Massachusetts.

Blevins, Juliette. 2004. *Evolutionary Phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press. DOI: *https://doi.org/10.1017/CBO9780511486357*

Blevins, Juliette. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32. 117–165. DOI: *https://doi.org/10.1515/TL.2006.009*

Boersma, Paul. 1998. *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives*. Ph.D. dissertation, University of Amsterdam.

Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press. DOI: *https://doi.org/10.1111/j.1749-818X.2010.00213.x*

Booij, Geert & Jenny Audring. 2017. Construction morphology and the parallel architecture of grammar. *Cognitive Science* 41(S2). 277–302. DOI: *https://doi.org/10.1111/cogs.12323*

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

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press. DOI: *https://doi.org/10.1017/CBO9780511612886*

Carstairs, Andrew. 1988. Some implications of phonologically conditioned suppletion. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1988*, 68–94. Dordrecht: Foris.

Carstairs, Andrew. 1990. Phonologically conditioned suppletion. In Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer & John R. Rennison (eds.), *Contemporary morphology,* 17–23. Berlin: Mouton de Gruyter.

Ćavar, Małgorzata E. 2004. *Palatalization in Polish*. Ph.D. dissertation, University of Potsdam.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, George N. & Elizabeth Hume. 1995. The internal organization of speech sounds. In John Goldsmith (ed.), *Handbook of phonological theory,* 245–306. Oxford: Blackwell.

Czaplicki, Bartłomiej. 2013a. Arbitrariness in grammar: Palatalization effects in Polish. *Lingua* 123. 31–57. DOI: *https://doi.org/10.1016/j.lingua.2012.10.002*

Czaplicki, Bartłomiej. 2013b. *R*-metathesis in English: An account based on perception and frequency of use. *Lingua* 137. 172–192. DOI: *https://doi.org/10.1016/j.lingua.2013.09.008*

Czaplicki, Bartłomiej. 2014a. *Lexicon based phonology: Arbitrariness in grammar*. Munich: Lincom Europa.

Czaplicki, Bartłomiej. 2014b. Frequency of use and expressive palatalization: Polish diminutives. In Eugeniusz Cyran & Jolanta Szpyra-Kozłowska (eds.), *Crossing Phonetics-Phonology Lines,* 141–160. Newcastle upon Tyne: Cambridge Scholars Publishing.

Czaplicki, Bartłomiej. 2016. Word-specific phonology: The impact of token frequency and base

transparency. In Grzegorz Drożdż (ed.), *Studies in lexicogrammar: Theory and applications,* 261–276. Amsterdam/Philadelphia: John Benjamins. DOI: *https://doi.org/10.1075/hcp.54.14cza*

Czaplicki, Bartłomiej. 2019. Measuring the phonological (un)naturalness of selected alternation patterns in Polish. *Language Sciences* 72. 160–187. DOI: *https://doi.org/10.1016/j.langsci.2018.10.002*

Czaplicki, Bartłomiej. 2020. Construction-specific phonology: Evidence from Polish vowel-zero alternations. In Krzysztof Jaskuła (ed.), *Phonological and phonetic explorations*, 77–93. Lublin: Wydawnictwo KUL.

Dąbrowska, Ewa. 2004. Rules or schemas? Evidence from Polish. *Language and Cognitive Processes* 19. 225–271. DOI: *https://doi.org/10.1080/01690960344000170*

Dąbrowska, Ewa. 2008. The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language* 58. 931–951. DOI: *https://doi.org/10.1016/j.jml.2007.11.005*

Dawdy-Hesterberg, Lisa & Janet Pierrehumbert. 2014. Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience* 29. 1268–1282. DOI: *https://doi.org/10.1080/23273798.2014.899377*

de Lacy, Paul. 2002. *The formal expression of markedness*. PhD dissertation, University of Massachusetts, Amherst. Amherst, MA: GLSA Publications.

de Lacy, Paul. 2006. Transmissibility and the role of the phonological component. *Theoretical Linguistics* 32. 185–196. DOI: *https://doi.org/10.1515/TL.2006.012*

de Lacy, Paul & John Kingston. 2013. Synchronic explanation. *Natural Language and Linguistic Theory* 31(2). 287–355. DOI: *https://doi.org/10.1007/s11049-013-9191-y*

Divjak, Dagmar, Ewa Dąbrowska & Antti Arppe. 2016. Machine meets man. Evaluating the psychological reality of corpus-based probabilistic models. *Cognitive Linguistics* 27. 1–34. DOI: *https://doi.org/10.1515/cog-2015-0101*

Długosz-Kurczabowa, Krystyna & Stanisław Dubisz. 2006. *Gramatyka historyczna języka polskiego.* Warsaw: Wydawnictwa Uniwersytetu Warszawskiego. DOI: *https://doi.org/10.31338/uw.9788323526025*

Dressler, Wolfgang U. 2003. Naturalness and morphological change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics,* 461–471. Oxford: Blackwell. DOI: *https://doi.org/10.1002/9780470756393.ch12*

Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24. 143–188. DOI: *https://doi.org/10.1017/S0272263102002024*

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. (Linguistic Inquiry monographs 60). Cambridge, MA: MIT Press. DOI: *https://doi.org/10.7551/mitpress/9780262014229.001.0001*

Frauenfelder, Uli H. and Robert Schreuder. 1992. Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1991*, 165–183. Dordrecht: Kluwer. DOI: *https://doi.org/10.1007/978-94-011-2516-1_10*

Grzegorczykowa, Renata & Jadwiga Puzynina. 1999. Rzeczownik. In Renata Grzegorczykowa, Roman Laskowski & Henryk Wróbel (eds.), *Gramatyka współczesnego języka polskiego: Morfologia,* 389–468. Warsaw: Wydawnictwo Naukowe PWN.

Gussmann, Edmund. 2007. *The phonology of Polish*. Oxford: Oxford University Press.

Hahn, Ulrike and Ramin Charles Nakisa. 2000. German inflection: Single route or dual route. *Cognitive Psychology* 41. 313–360. DOI: *https://doi.org/10.1006/cogp.2000.0737*

Hale, Mark and Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.

Halle, Morris. 2005. Palatalization/Velar softening: What it is and what it tells us about the nature of language. *Linguistic Inquiry* 36. 23–41. DOI: *https://doi.org/10.1162/0024389052993673*

Hamann, Silke. 2002. Postalveolar fricatives in Slavic languages as retroflexes. In Sergio Baauw, Mike Huiskes & Maaike Schoorlemmer (eds.), *OTS Yearbook 2002,* 105–127. Utrecht: Utrecht Institute of Linguistics.

Hay, Jennifer. 2003. *Causes and consequences of word structure*. London: Routledge. DOI: *https://doi.org/10.4324/9780203495131*

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440. DOI: *https://doi.org/10.1162/ling.2008.39.3.379*

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15. 651–674. DOI: *https://doi.org/10.1198/106186006X133933*

Inkelas, Sharon. 2014. *The interplay of morphology and phonology*. Oxford: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780199280476.001.0001*

Itô, Junko & Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh (eds.), *University of Massachusetts Occasional*

Papers in Linguistics Vol. 18: Papers in Optimality Theory, 181–209. University of Massachusetts, Amherst: GLSA.

Itô, Junko & Armin Mester. 1999. The Phonological Lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics,* 62-100. Oxford: Blackwell. DOI: *https://doi.org/10.1002/9781405166225.ch3*

Jackendoff, Ray. 2002. *Foundations of Language*. Oxford: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780198270126.001.0001*

Jakobson, Roman. 1929/1962. Remarques sur l'évolution phonologique du russe comparée à celle des autres langues slaves. *Traveaux du Cercle Linguistique de Prague 2*. Reprinted in *Selected writings, vol. I: Phonological studies*, 2nd exp. edn., 7–116. The Hague: Mouton.

Kenstowicz, Michael. 1996. Base identity and uniform exponence: Alternatives to cyclicity. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods* 1. 363–394. Salford: University of Salford.

Kiparsky, Paul. 1982. From cyclic to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations. Part I,* 131–175. Dordrecht: Foris Publications.

Kiparsky, Paul. 2006. The amphichronic program vs. evolutionary phonology. *Theoretical Linguistics* 32. 217–236. DOI: *https://doi.org/10.1515/TL.2006.015*

Kiparsky, Paul. 2008. Universals constrain change, change results in typological generalizations. In Jeff Good (ed.), *Linguistic universals and language change,* 25–53. Oxford: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780199298495.003.0002*

Kochetov, Alexei. 2011. Palatalization. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology,* 1666–1690. Malden: Wiley-Blackwell. DOI: *https://doi.org/10.1002/9781444335262.wbctp0071*

Langacker, Ronald. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language,* 1–63. Stanford, CA: CSLI.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17. 267–302. DOI: *https://doi.org/10.1023/A:1006182130229*

Mańczak, Witold. 1980. Laws of analogy. In Jacek Fisiak (ed.), *Historical morphology,* 283–288. The Hague: Mouton. DOI: *https://doi.org/10.1515/9783110823127.283*

McClelland, James & Jeffrey Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18(1). 1–86. DOI: *https://doi.org/10.1016/0010-0285(86)90015-0*

McQueen, James & Ann Cutler. 1998. Morphology in word recognition. In Andrew Spencer & Arnold A. Zwicky (eds.), *The handbook of morphology,* 406–427. Oxford: Blackwell. DOI: *https://doi.org/10.1002/9781405166348.ch21*

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech,* 189–216. New York: Springer. DOI: *https://doi.org/10.1007/978-1-4613-8202-7_9*

Padgett, Jaye. 2002. Feature classes in phonology. *Language* 78. 81–110. DOI: *https://doi.org/10.1353/lan.2002.0046*

Paster, Mary. 2006. *Phonological conditions on affixation*. Ph.D. dissertation, University of California at Berkeley.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33. 999–1035. DOI: *https://doi.org/10.1111/j.1551-6709.2009.01047.x*

Pęzik, Piotr. 2012. *Wyszukiwarka PELCRA dla danych NKJP. Narodowy Korpus Języka Polskiego*. Adam Przepiórkowski, Mirosław Bańko, Rafał L. Górski & Barbara Lewandowska-Tomaszczyk (eds.), 2012. Warszawa: Wydawnictwo PWN.

Plag, Ingo. 2012. *Word-formation in English*. Cambridge: Cambridge University Press.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Basil Blackwell. DOI: *https://doi.org/10.1002/9780470759400*

Rubach, Jerzy. 1984. *Cyclic and Lexical Phonology: The structure of Polish*. Dordrecht: Foris. DOI: *https://doi.org/10.1515/9783111392837*

Rubach, Jerzy. 2007. Feature Geometry from the perspective of Polish, Russian and Ukrainian. *Linguistic Inquiry* 38, 85–138. DOI: *https://doi.org/10.1162/ling.2007.38.1.85*

Sagey, Elizabeth. 1986. *The representation of features and relations in nonlinear phonology*. Ph.D. dissertation, MIT.

Scheer, Tobias. 2012. *Direct interface and one-channel translation: A non-diacritic theory of the morphosyntax-phonology interface.* Berlin: de Gruyter. DOI: *https://doi.org/10.1515/9781614511113*

Stemberger, Joseph P. & Brian MacWhinney. 1988. Are inflected forms stored in the lexicon? In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics,* 101–116. San Diego, CA: Academic Press.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Michael Broe & Janet Pierrehumbert (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon,* 313–334.

Cambridge: Cambridge University Press.

Steriade, Donca. 2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In Kristin Hanson & Sharon Inkelas (eds.), *The nature of the word,* 151–179. Cambridge: MIT Press. DOI: *https://doi.org/10.7551/mitpress/9780262083799.003.0007*

Święciński, Radosław. 2014. The phonological status of the palatal glide in Polish: Acoustic evidence. In Eugeniusz Cyran & Jolanta Szpyra-Kozłowska (eds.), *Crossing phonetics-phonology lines,* 365–402. Newcastle: Cambridge Scholars Publishing.

Szpyra, Jolanta. 1989. *The phonology-morphology interface: Cycles, levels and words.* London and New York: Routledge.

Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24. 135–178. DOI: *https://doi.org/10.1017/S0954394512000129*

Wright, Richard. 2004. A review of perceptual cues and cue robustness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology,* 34–57. Cambridge: Cambridge University Press. DOI: *https://doi.org/10.1017/CBO9780511486401.002*

Zuraw, Kie. 2000. *Patterned exceptions in phonology.* Ph.D. dissertation, University of California, Los Angeles.

INTERNET RESOURCES

National Corpus of Polish: www.nkjp.pl.

plTenTen: Corpus of the Polish Web, available in *Sketch Engine: www.sketchengine.eu.*