

Polish yers are epenthetic: an argument from lexical statistics

Introduction. It is difficult to identify whether a rule is default or exceptional when a language supplies many examples of rule undergoers and non-undergoers. Polish vowel-zero alternations are such a case. The tendency has been to give Polish and Russian yer vowel alternations a unified analysis. We supply analytic and quantitative arguments for analyzing Polish alternations as general and epenthetic, but subject to exceptions, in contrast to Russian's exceptional alternation caused by deletion (Gouskova 2012, Gouskova and Becker 2013).

Polish yers. In Polish, [e] alternates with zero in the final syllables of some words but not others, shown in (1). Hayes (2009, ch. 12) points out that whether the alternation is treated as deletion (Gussmann 1980, Bethin 1992, Jarosz 2008, Rubach 1986, 2013) or epenthesis (Czaykowska-Higgins 1988), there must be lexical exceptions: there are vowels that do not alternate (1a), and there are contexts (e.g., [t_r]) in which vowels appear in some words but not others (cf. (1c) and (1e)). Regardless of a word's pattern with case suffixes, [e] appears in the last stem syllable with the diminutive suffix [-ek], as in the 'Diminutive' column in (1)—except in obstruent-obstruent clusters, which may be unbroken in the diminutive (see (1e)).

(1) *Six types of patterns in Polish*

	UR	Unaffixed	Case Suffix	Diminutive	Gloss
(a) Nonalternating V	/seter/	seter	seter-i	seter-ek	'setter'
	/kalek/	kalek	kalek-i	kalet-ek	'cripple'
(b) Epenthesis: 1σ	/dŋ/	džen	dŋ-i	džen-ek	'day'
	/mgw/	mgjew	mgw-i	mgjew-ek	'fog'
(c) Epenthesis CC# > 1σ	/sfetr/	sfeter	sfetr-i	sfeter-ek	'sweater'
	/lalk/	lalek	lalk-i	lalet-ek	'doll'
(d) Variable epenthesis	/bit-v/	bitf, bitef	bitv-i	bitev-ek	'battle'
	/vew-n/	vewn, vewen	vewn-i	vewen-ek	'wool'
(e) Exceptional blocking I	/v̥atr/	v̥atr	v̥atr-i	v̥ater-ek	'wind'
	/katedr/	katedr	katedr-i	kateder-ek	'cathedral'
(f) Exceptional blocking II	/most/	most	most-i	most-ek	'bridge'
	/swuʒb/	swuʒp	swuʒb-i	swuʒb-ek	'service'

Analysis. We argue that Polish vowel-zero alternations should be analyzed as epenthesis, using lexically indexed constraints (Pater 2008 *inter alia*). In non-alternating words such as (1a), the vowel is present in the UR. Alternating words such as (1b–d) differ in which constraint triggers epenthesis: in monosyllables, it is HEADEDNESS, the pressure to have a vowel nucleus (Szpyra 1992, Hayes 2009). In longer alternating words, the vowel breaks up a CC# cluster, so *CC# >> DEP (see (2a)). Sonorant-final clusters are especially common in this category. The third category of words has optional alternations at the morpheme boundary, usually affecting the same suffixes (-v, -n). The cases in which there are no alternations between unaffixed and case forms (see (1e, f)) are specified as exceptions to epenthesis: indexed CONTINUITY_{Ex} is ranked above *CC#, see (3). For such morphemes, CONTINUITY_{Ex} may be dominated, since there is obligatory epenthesis for CR-final stems in the context of diminutives: compare (3a) and (3b) for evidence that *CRC >> CONTINUITY_{Ex}. We attribute this to the selectional requirements of the [-ek] suffix, which favors bases that do not end in CR clusters (see Gouskova and Newlin-Łukowicz 2014 for a similar account of Russian [-ok]). *CC# also determines the site of epenthesis in /CCC/ words: in /mgw-a/ [mgw-a] 'mist' and /pxw-a/ [pxw-a] 'flea,' the vowel always appears after the first two consonants: [mgjew] 'mist (gen. pl.).' Finally, some words do not exhibit any alternations, such as (1f), in which epenthesis is blocked by CONTINUITY and not triggered by *CRC.

Why not deletion? In our analysis, [v̥atr] and [most] are exceptions to epenthesis. The alternative is that [seter] resists deletion. A deletion analysis does not explain why only the [e] vowel alternates, or why [e] is predictably present in the context of the diminutive suffix even in morphemes that resist the

alternation elsewhere. To explain that pattern, a Russian-style deletion account (Gouskova 2012) would still have to posit epenthesis in diminutives.

(2) *Analysis of Polish words with alternations*

/sfetr/ 'sweater'	HEADEDNESS	*CRC	CONTIG _{EX}	*CC#	DEP
a. sfeter~sfetr				W	L
/sfetr-i/ 'sweater'					
b. sfetri~ sfeteri					W
/mgw/ 'mist'					
c. mqiew~ mgw	W				L

(3) *Analysis of Polish words without alternations*

/v̥atr/ 'wind'	HEADEDNESS	*CRC	CONTIG _{EX}	*CC#	DEP
a. v̥atr~v̥ater					W
/ v̥atr-k-a/					
b. v̥aterka~v̥atrka		W	L		L

Alternation is the general rule in Polish. The POLEX lexicon of Polish (Vetulani et al. 1998) contains 41,742 nouns. Of these, 6.3% contain a non-alternating [e] (see (4b)), 15.8% exhibit vowel alternation (see (4a)), and 16.1% end in CC# in some grammatical case (see (4c)). Of nouns that end in CC#, which may be considered exceptions to alternation, the majority end in the suffixes [-oœte], [-izm], [-ist], [-stv], [-ovŋ] and [-itm], see (4ci). These suffixes categorically never host alternating vowels, tend to be part of a more formal register in language use, and represent 11.1% of the lexicon. Thus, 5.0% of the lexicon ends in CC#, but does not contain these particular suffixes, see (4cii). If Polish speakers know that the above suffixes are unacceptable contexts for vowel insertion, and so rank faithfulness to them above *CC#, then the number of CC# words that must be treated as idiosyncratic exceptions to the epenthesis rule (5% of the lexicon) is smaller than the number of words that undergo it (15.8% of lexicon) in Polish, compared to Russian's 17% unbroken CC# and ~9% alternation (Gouskova and Becker 2013).

(4) *Corpus statistics*

	Count of forms	Of lexicon	Example
(a) Alternating [e]	6,581	15.8%	sfeter~sfetri
(b) Non-alternating [e]	2,624	6.3%	seter~seteri
(c) Ends in CC# cluster	6,729	16.1%	
i. <i>Suffixed</i>	4,630	11.1%	markeizm~markeizmu
ii. <i>All unsuffixed</i>	2,099	5%	swuʃp~swuʒba
(d) CCV# or non-[e] CVC	25,808	61.8%	azja~azji
Total	41,742	100%	

Discussion. Russian and Polish vowel alternations are historically related, but they diverged: in Russian, they are exceptionally triggered deletion, but in Polish, they are the result of a productive epenthesis process subject to exceptions. Many differences between the languages follow from this. In Russian, alternation is not extended to loanwords (dizel̩~dizel̩a 'diesel'), vowel quality is only semi-predictable (mid [e] and [o]), and there are paradigm gaps (e.g., [mgl̩] 'mist' does not have a genitive plural). In contrast, Polish readily extends alternation to loanwords (dizel~dizl-a 'diesel'), predictably alternates [e], and has no paradigm gaps (/mgw/ 'mist' is [mg̩ew] in the genitive plural).

(5) *Russian versus Polish vowel alternations*

	Russian	Polish
Extended to loans?	No	Yes
Vowel predictable?	No	Yes
Paradigm gaps?	Yes	No
Unbroken CC#	17%	5%
Alternation	9%	15.8%

Traditional analyses of Polish and Russian yers, which posit similar representations for both languages, do not address the different qualities of the alternations, and do not predict that speakers of the two languages should behave differently with novel words. In contrast, our analysis makes a testable prediction that Polish speakers' regular epenthesis should extend alternation to novel items more readily than Russian speakers' lexically restricted deletion.