RESEARCH

Transderivational relations and paradigm gaps in Russian verbs

Katya Pertsova

Dept. of Linguistics, UNC, CB 3155, UNC-CH, Chapel Hill, NC, GB
pertsova@unc.edu

In this paper I argue that the notorious case of paradigm gaps, the first person singular gaps of Russian verbs, are not synchronically arbitrary as is often assumed (Graudina et al. 1976; Daland et al. 2007; Baerman 2008), but are predictable and connected to the opaque morphophonological alternations affecting stem-final consonants in the 1st person singular (1SG) present tense. I present evidence for a new empirical generalization showing that the problematic alternations in verbs are subject to a *lexical conservatism* effect (Steriade 1997). Namely, stems that appear in other derivationally or inflectionally related forms with the same alternation as the one expected in 1SG generally do not have gaps, while stems that have no attested related forms with alternations do. Overall, a larger set of verbs are problematic for the speakers than indicated in dictionaries, and there are degrees of "gappiness" with a lot of variation across speakers. Additionally, I consider how different theoretical proposals for handling ineffability fare in accounting for these findings. I propose to augment the framework of Harmonic Grammar (Legendre et al. 1990) with an additional post-competition step during which outputs can be compared to each other based on their Harmony scores. This proposal is not tied to violations of specific constraint and it has potential to account for both paradigm gaps and gradient grammaticality judgments.

1 Introduction

One of the puzzles in inflectional morphology is the phenomenon of *paradigm gaps*, a situation when speakers are uncertain about some wordform(s) in a paradigm of a word. An example of such a "defective" word in English noted in Pinker (1999) is the verb *stride* whose past participle is difficult for many speakers (*have stridden? have strided? have strode?*). This gap cannot be explained by the low frequency of *stride* since other low-frequency verbs do not present similar problems for the speakers. Paradigm gaps pose challenges for many theories of grammar, in particular theories that have defaults, including Optimality Theory (Prince & Smolensky 1993) and its variants. On the other hand, gaps are relatively easy to explain as a consequence of an unproductive rule or pattern, provided we have the right measure of productivity. In general, understanding what causes an otherwise extremely productive portion of grammar such as inflection to break down can shed light on the nature and scope of morpho-phonological phenomena.

Paradigm gaps often affect wordforms that are subject to irregular alternations or allomorphy e.g., gaps in Spanish verbs (Albright 2003; Maiden & O'Neill 2010), French verbs (Morin 1987), German diminutives (Fanselow & Féry 2002), Russian nouns (Pertsova 2005), and others. The English verbal participle example above is also of this type. In English, verbs like *stride* with the vowel /aj/ in the stem either change this vowel to

[ɪ] and appear with the suffix *-en* in the past participle (e.g., *hide – have hidden*), or have the same shape as the irregular past tense form (e.g., *grind – have ground*), or appear with the regular *-ed* suffix (e.g., *arrive – have arrived*). One natural hypothesis is that irregularity can lead to gaps under certain circumstances (Albright 2003; 2009b; Yang et al. 2012).

In this paper I examine a notorious case of paradigm gaps in Russian verbs which are also connected to alternations, but for which irregularity cannot be the explanation. All these verbs have stems ending in dental consonants which are expected to alternate with a palatal or alveo-palatal fricative in the first person singular present tense (henceforth 1SG). An example of a defective verb is *pylesosit'* [pɨlʲɪsósʲ-ɪtʲ][1] 'to vacuum' whose 1SG form 'I vacuum' is expected to be realized as [pɨlʲɪsóʂ-u] with the stem final [sʲ] ~ [ʂ] alternation characteristic of other stems ending in /sʲ/. However, for some reason speakers are often not comfortable producing this wordform and have a hard time deciding between [pɨlʲɪsóʂ-u] and [pɨlʲɪsósʲ-u] (a prescriptively unacceptable variant without the alternation). Many speakers tend to avoid 1SG forms of defective verbs altogether, resorting to circumlocutions. Some speakers nevertheless produce the problematic forms (with or without alternations), but their productions are often either ironic (when they know the form they produced is not prescriptively accepted) or are marked by hesitation or a feeling of discomfort (for more on this see section 4.1.2).

Linguists and dictionary writers have been aware of defective Russian verbs for a long time. Many such verbs are marked in Russian dictionaries as either lacking 1SG forms or having a "difficult" 1SG form. The researchers who have written on this topic (Graudina et al. 1976; Sims 2006; Daland et al. 2007; Baerman 2008; Albright 2009b; Moskvin 2015) agree that there is no single set of semantic and/or phonological properties that could explain the defectiveness of these verbs. Moreover, although it is tempting to connect defectiveness to uncertainty about morphophonological alternations in the 1SG, most of these alternations are regular and, thus, are not expected to cause the same problems as the verb *stride* mentioned above. This fact makes Russian gaps particularly challenging to explain. The most popular view is that Russian gaps are holdovers from a previously existing irregularity and are simply lexicalized (Graudina et al. 1976; Sims 2006; Baerman 2008). Daland et al. (2007) show how lexicalized gaps can spread to new words by analogy in a model which learns the frequencies of different wordforms in a paradigm.

Contrary to this view, I argue that the majority of Russian defective verbs are not lexically specified as lacking the 1SG form (in fact, new defective verbs continue to arise through borrowing, see section 4.3).[2] I also show that at least some problematic 1SG dental-fricative alternations are sufficiently well represented in the lexicon to be learned by native speakers, so these gaps also cannot be explained as resulting from sparse data leading to an unproductive rule. The crucial new empirical observation that sheds light on the Russian gaps is the fact that defective verbs differ from the otherwise similar non-defective verbs in lacking the same alternations in other related forms. In particular, in non-defective verbs the 1SG alternations also apply in other related forms, typically in the past passive participles or derivationally related secondary imperfectives. This finding presents evidence for the phenomenon of *lexical conservatism* (Steriade 1997; 2008; Yanovich & Steriade 2010) or *multiple correspondence* (Burzio 1998), which is sometimes

---

[1] In this paper I use IPA when phonetic detail is important, otherwise transliterations are used.
[2] The majority of defective verbs in Russian are of low-frequency. However, there are a few frequent defectives for which paradigm gaps could indeed be memorized. Speakers could do this by noticing the disparity between the expected and observed frequencies of 1SG wordforms.

described as dislike on part of the speakers of creating novel allomorphs of listed morphemes (this phenomenon is discussed in more detail at the beginning of section 4 and in section 5).

The broad implications of this work are that it provides further evidence for existence of transderivational relations among words suggesting that phonological computation is not strictly local and may involve reference to paradigmatic or derivational relatives other than the derivational bases. In the words of Burzio (2005), "phonological outputs are in attraction relation to neighboring representations." Existence of such relations is controversial and is often questioned on the grounds that they lead to a less constrained model of grammar (Bailyn & Nevins 2008; Bobablijik 2008; Bermúdez-Otero 2012).

Additionally, this work contributes to the conversation on the appropriate ways to model grammaticality and well-formedness. Paradigm gaps, conceived of as absence of an output, are challenging for main-stream optimization models in phonology because these models predict an output for every input. However, this problem can be avoided if we think of gaps as arising from outputs of very low relative well-formedness (see section 2.4). I consider ways in which one model of phonology, Harmonic Grammar (Legendre et al. 1990), can be modified to incorporate this assumption.

2 General approaches to paradigm gaps

Proposals for explaining paradigm gaps can be grouped into four broad categories summarized below. In this section I briefly review these approaches and comment on their applicability to the Russian verbal gaps. The last approach, termed here the *threshold* approach, will be applied to the Russian verbal gaps in section 5.

2.1 Lexicalization of gaps

Perhaps the simplest way to explain paradigm gaps is to assume that absence or avoidance of certain paradigmatic forms can be noticed by speakers and become lexicalized. Lexicalization of a gap can happen at the wordform level or at a more abstract stem level as proposed in Boyé & Hofherr (2010), who observe that the defective verbs in Spanish and French all share the same stem allomorph. This approach raises the question of how the wordform- or the stem-gaps arise in the first place. The answers may differ on a case by case basis. For the Russian case considered here, the most commonly cited reason for the original appearance of gaps has to do with a previous irregularity in the 1SG present forms stemming from co-existence of Russian and Old Church Slavonic alternations.

Baerman (2008) traces down the history of this phenomenon through historical grammatical accounts and comes to the conclusion that "problems that arose under prior morphological conditions were ultimately lexicalized to yield the defectiveness of contemporary paradigms." In particular, the [dʲ]-final stems used to have two different alternation patterns in 1SG: the currently surviving d ~ ʐ alternation and the Old Church Slavonic d ~ ʐd alternation (still used in past passive participles of some verbs, but no longer used in the 1SG). Additionally, according to some grammarians the non-alternation also used to be accepted for a selected set of lexemes, although this claim is more controversial. Sims (2006) suggests that non-alternations were characteristic of lexemes of "lower style". The historical explanation then can be roughly summarized as follows: when the minority 1SG patterns were lost, several stems which used to have these patterns failed to level out and retained a gap in the paradigm. Baerman, however, acknowledges that there is no perfect correspondence between the prior acceptability of minority alternations and defectiveness. Thus, the question remains why some of the etymologically Old Church Slavonic verbs adopted the majority alternation while others became defective.

Additionally, the general criticism of lexicalization approaches is that they assume that speakers can rely on negative evidence, that is, they can notice the absence of something. This assumption is only reasonable for high-frequency lexemes for which speakers might see all inflectional forms with a high enough probability that absence of such forms should become suspicious. However, lexemes that have gaps are often among the least frequent lexemes in a language. For such lexemes, absence of attested wordforms is the norm rather than the exception, and the lexicalization in such cases seems implausible. If gaps in low-frequency verbs appeared by analogy with a few high-frequency defective verbs (as proposed in Daland et al. 2007), one would expect them to occur in words that phonologically resembled the bases of the analogy, the frequent verbs of Old Church Slavonic origin. Instead, as we will see, the gaps in low-frequency verbs are systematically connected to presence of dental-palatal alternations in derivationally related forms.

2.2 The filtering approach

Another possible way to analyze defectiveness is to assume that some constraints or generalizations in language are inviolable and can act as a filter on the outputs of the grammar.

Within OT this type of approach has been used before for phonotactically-motivated gaps. The two implementations of filtering I am familiar with are the *null parse* theory (Prince & Smolensky 1993; McCarthy 2002) and the *Control* theory (Orgun & Sprouse 1999). The former theory presupposes that every competition includes a special candidate *null parse* (the gap), which violates a single constraint, M-PARSE. This candidate can win only if all other candidates violate constraints ranked above M-PARSE. Thus, ranking constraints above M-PARSE is equivalent to assuming that they are inviolable. The second theory presupposes that the outputs of the regular phonological grammar are run through a filter (CONTROL) of inviolable constraints. That is, the sequence of grammatical operations is [input → optimization → filter → output]. Violations of any constraints in CONTROL are fatal, which means some inputs will have no outputs.

The filtering approach could be technically applied to the gaps discussed in this paper. However, it has one potentially fatal flaw–it severely restricts morphological productivity. It assumes that any form violating the constraints in the filter must be ungrammatical. In the case examined here, if one wanted to make a general claim about the kinds of constraints that are in the filter in Russian, the most natural assumption would be that they include faithfulness constraints making reference to the lexically stored forms (for more discussion of this point see section 5). In that case, a morpho-phonological alternation would never be predicted to apply to a new stem unless a learner has already witnessed this stem with that alternation. Such state of affairs raises the question of how alternations spread through the lexicon in the first place. We know, for example, that certain types of alternations are rather automatic and apply even in low-frequency and novel contexts. For example, the reduction of [o] to [a] in pretonic syllables in many dialects of Russian[3] is very productive and is typically extended to borrowings (e.g., [matór] 'motor,' [manstrʲ-ónək] 'baby monster').[4]

2.3 The competition approach

One natural conclusion to draw from the fact that speakers often disagree on how to "fill the gap" and produce several different outputs, is that there are multiple patterns that are in close competition with each other. In the case of Russian defective verbs, when forced, speakers will usually produce either the form with the expected alternation (e.g.,

---

[3] Note that this process is phonological and not phonetic (Barnes 2007).
[4] Still, even vowel reduction is not fully automatic. It fails to apply in a small set of borrowed words, such as [rokokó], and in some ecclesiastical vocabulary (Avanesov 1968 via Gouskova 2012).

[pɨlʲɪsós-u] 'I vacuum') or the grammatically illicit form without the alternation (e.g., [pɨlʲɪsósʲ-u]). Perhaps, competition can lead to indecision on part of the speakers, which would in turn manifest itself as a paradigm gap. Nevins (2014) compares this situation to a "well-known paradox of a 'Burdian's Ass,' in which a donkey theoretically starves to death deciding between two equally appealing (or unappealing) bales of hay." Unlike the filtering approach, this approach has an advantage in that it does not require ascribing a special status to certain grammatical constraints as inviolable (a status which may be difficult to prove). However, the competition approach has also been criticized as implausible.

Albright (2009b) argues against competition-based approaches to defectiveness on the grounds that competition among equally matched forms has been assumed in other work to lead to variation, co-existence of multiple grammatical outputs for a single input (but see Pierrehumbert 2001 and Antilla 2007 for alternative views on variation). Assuming that close competition instead leads to paradigm gaps goes against the core assumptions about grammaticality in OT, in particular the assumption that a winner is the most optimal candidate. This assumption is conceptually natural in competition models. Consider what happens in an actual competition: when two competitors are equally good, they both win.

Another objection to a competition approach above mentioned in Albright (2009b), is that there are cases of paradigm gaps which do not seem to involve close competition with another pattern. The gaps in the genitive plural in Rusian discussed in Pertsova (2005) are of this sort. For example, for the noun *mechta* 'dream,' speakers do not entertain multiple options. They know that the genitive plural should be *mecht*, but they don't like it.

2.4 The threshold approach

I call the fourth approach to ungrammaticality the *threshold* approach. Taking up the competition metaphor again, imagine the following scenario: in addition to finishing first, a winner in each race is required to have a winning time that falls within some predefined window of times considered to be acceptable based on the distribution of winning times for all other equivalent races. In other words, if a winner were anomalously slow, his/her status as a winner may be questioned. The Minimum Generalization Learner of Albright discussed in section 3.2 could be viewed as a kind of a threshold model because it attaches confidence scores to rules. In this model, productivity of a pattern is proportional to its confidence score. Given such a scalar notion of well-formedness, one could define a threshold beyond which winners are considered to be defective.

Scalar or gradient well-formedness has been adopted in several recent proposals in phonology to explain speakers' gradient acceptability of phonotactic patterns (Berent et al. 2007; Coetzee & Pater 2008; Albright 2009a). For example, English speakers consistently judge certain attested onset clusters like [pr] as being better than other equally attested, but rare clusters like [pw]. One framework that has been proposed to model both gradient and categorical grammaticality is Harmonic Grammar (HG) (Keller 2006; Coetzee & Pater 2008). Below, I consider possibilities for how to extend such proposals to account for paradigm gaps.

Harmonic Grammar, like OT, is a constraint-based optimization model. It is similar to statistical models such as multiple linear regression because it relies on a linear function for computing well-formedness, Harmony. The constraints have weights representing their relative strength, and each violation of a constraint incurs a penalty expressed as a coefficient –1. The potential outputs for each input compete with each other, and the output with the highest Harmony score wins. Harmony of an output y is computed as the sum of

the weighted constraint violations incurred by *y*. Assuming that the violations are negative and constraint weights are positive, Harmony scores range from 0 to negative infinity, and the candidate with the harmony score closest to 0 is most optimal for a given input. Thus, grammaticality in this model is categorical much like in OT. For more background on HG see Pater (2009). As a brief illustration of how Harmony is computed consider the hypothetical tableau in (1), with three constraints and three hypothetical output forms for some input. Note that output 1 is the winner because it has the lowest Harmony score.

(1)   An illustration of a competition in HG.

	weight	1	4	3	
	/input/	C1	C2	C3	H
☞	output1	−1			−1
	output2		−1		−4
	output3		−1	−1	−4 + (−3) = −7

One way to incorporate gradient well-formedness into HG is to use the ordering imposed on the outputs by the Harmony scores. To implement the idea of a threshold for defectiveness, one could further require that any form that is a certain number of standard deviations away from the average Harmony score of the attested outputs be considered below the threshold. This threshold does not have to be thought of as a hard boundary; forms approaching the threshold from one side would be deemed as less and less defective and forms approaching it from the other side would be deemed as more and more defective. Defining a threshold in this way would correspond to the notion that speakers are skeptical of any output that looks like an outlier relative to the distribution of Harmony scores over all grammatical forms, yet they still recognize it as a potential output in contrast to truly ungrammatical forms.[5]

To test the idea that defective outputs have abnormally low Harmony scores in a particular language, one would need to know what the distribution of Harmony scores over a representative sample of winners looks like. Note that only winners can be compared to each other on the well-formedness scale defined by the Harmony scores. Likewise, losers can only be compared to losers. This assumption avoids the problem pointed out by Boersma (2004): namely, that a loser in some competition may be more harmonic than a winner in another competition. One can think of the process for determining acceptability of a form as having two steps:

(2)   a. Input-relative categorical optimization in HG: a most harmonic output for a given input is chosen as the winner
         b. Global scalar comparison: the harmony of a winner is evaluated relative to the distribution of Harmony scores for all winners

The second step provides a way of ordering winners on a well-formedness scale, and hence a way to account for gradient grammaticality judgments and relative degrees of "gappiness." If we assume that speakers' productions are determined by these two factors, grammaticality (which is categorical) and well-formedness (which is gradient), we can explain what forms are selected as outputs for specific inputs, why some of these forms are judged as more or less acceptable, and why some are considered so marginal, that they are avoided. In particular, for two grammatical outputs *a* and *b*, *a* is relatively more

---

[5] This idea raises many additional questions including the question of when during the learning process the threshold is established and whether it keeps being adjusted over time. Empirical investigation into whether children are more or less prone to paradigm gaps could shed light on this question.

acceptable than *b* if it has a higher Harmony score. Additionally, given some threshold of defectiveness, as the Harmony of a form approaches this threshold, speakers begin to avoid producing that form. Defectiveness in this sense is not the same thing as absence of an output or absolute ungrammaticality, since speakers know how to "fill the gap" (we know this because we see that they occasionally go ahead and fill it). Also note that having low harmony is not tied to violations of a specific set of constraints. In different situations different factors may conspire to produce "weak" competitions (i.e., competitions in which all winners have relatively low well-formedness).

One possible objection to this proposal has to do with cumulativity of constraint violations. Namely, it is possible that no threshold could in principle be defined if we want our grammar to generate arbitrarily long words which could accumulate constraint violations, but nevertheless be judged as acceptable by the speakers perhaps because each substring of a word is relatively well-formed. In that case, well-formed structures could have arbitrarily low Harmony scores. However, this objection is not a fatal one because the computation of relative well-formedness could be relativized to word-size. Additionally, it is not clear whether speakers do in fact judge really long strings as acceptable. For a recent discussion of issues related to word-size and grammaticality see Daland (2015).

There is an alternative approach to gradient grammaticality within Harmonic Grammar by Coetzee & Pater (2008). They propose to keep the notion of well-formedness relativized to a local competition. Namely, they measure well-formedness of a winner in terms of how far it is removed from the next most harmonic competitor for the same input. This approach has one potential problem: it implies that the least well-formed outputs are those that have neck-and-neck competitors, but as I already mentioned this is also the scenario that is assumed to produce free variation. Therefore, one would predict that free-variants should never be judged as more well-formed than outputs that are not in free variation. For example, if [pl] and [pr] were in free variation in English, they would be expected to be judged as less acceptable than [pw]. This problem could perhaps be avoided with an alternative method for measuring input-relativized well-formedness that considered not just the two most harmonic outputs but the overall spread of the Harmony scores in a local competition. Even then, such a measure is fundamentally different from defining well-formedness globally. The local approach appears to equate "fierce" competition (dense distributions of harmony scores over local competitors) with low well-formedness, which is conceptually similar to the competition approach to defectiveness. On the other hand, the global approach allows us to distinguish between two conceptually different cases: a competition that involves several equally strong or well-formed competitors vs. a competition that involves several equally weak or ill-formed competitors, where well-formedness is based on the weight of the constraints that are violated. One can imagine that both scenarios would lead to variation, but in the first case we want the model to predict speakers to be equally satisfied with the variants, while in the second case they should be equally unhappy with the variants, a situation that results in a paradigm gap.

3 Defective verbs in Russian: description and issues

Russian verbs have two morphological subparadigms: past and non-past. Semantically, the non-past paradigm is interpreted as present tense for imperfective verbs and as future tense for perfective verbs. But for simplicity, I will refer to this paradigm as the present tense paradigm. The gaps occur in the first person singular cell of the present tense paradigm of second conjugation verbs (verbs with the theme vowel -i in present tense forms) whose stems end in dental palatalized consonants. There are no obvious semantic or

phonological properties that set defective verbs apart from the non-defective ones (except possibly for a handful of verbs).[6]

Examples of the present tense paradigms for defective and non-defective verbs are given in Table 1. Notice that the 1SG form of the non-defective verb contains a morpho-phonological alternation: the stem-final /zʲ/ is realized as [ʐ]. Such alternations are typical for verbs of second conjugation. Below I discuss these alternations in more detail, and raise some questions in connection to the hypothesis that the gaps in 1SG are in some way caused by the difficulty in applying these alternations.

	SG	**PL**		**SG**	**PL**
1p.	/voʐ-u/	/vozʲ-im/		?	derzʲ-im
2p.	/vozʲ-iʂ/	/vozʲ-ite/		derzʲ-iʂ	derzʲ-ite
3p.	/vozʲ-it/	/vozʲ-at/		derzʲ-it	derzʲ-at

Table 1: Present tense paradigms of a non-defective verb /vozʲitʲ/ 'to transport' and a defective verb /derzʲitʲ/ 'to be rude.'

3.1 Morpho-phonological alternations in 1SG

A group of second conjugation verbs undergo consonantal alternations in the 1st person singular form. These verbs fall into two major classes: those whose stem ends in a dental consonant (henceforth *dental stems*), and those whose stem ends in a labial consonant (henceforth *labial stems*). These alternations historically derive from the process of jotation, changes that affected some consonants before the glide /j/ which was part of the 1SG suffix but was subsequently lost. The labial alternation involves insertion of [lʲ] before the 1SG suffix -*u*. The dental alternations involve a mutation of stem-final palatalized dental obstruents to post-alveolar or palatal continuants of the same voicing (see Table 2 for details). It is debatable whether these alternations are currently purely morphological or whether they could be analyzed as phonologically conditioned.[7] All verbs with gaps belong to the class of dental stem verbs with the exception of a single labial stem, *zatm-itʼ* 'to eclipse.'[8]

Notice that these consonantal alternations are completely regular (exceptionless) for all stems except the [tʲ]-final stems. That is, for each affected consonant there is only one alternant; non-alternation is not grammatical. For stems ending in [tʲ] (but not the

---

[6] For example, one defective verb, *pret-itʼ* 'to revolt,' is almost exclusively used in third person. But even then, hypothetical 2nd person forms are easy to come up with, while a hypothetical 1SG form is problematic. Two defective verbs, [at͡ʃutʲ-i-t͡sa] 'to appear' and [aɕːutʲ-itʲ] 'to sense' are claimed to be avoided in 1SG because the obligatory alternations create repeating syllables [t͡ʃut͡ʃu] and [ɕːuɕːu] respectively (Moskvin 2015). Although several linguists express doubt that this is the real reason for gaps in these words (Halle 1973; Sims 2006), I found some confirmation for avoidance of repeating consonants while searching for alternations in recent borrowings (reported in section 4.3). In particular, [t͡ʃatʲ-i-t͡sa] 'to chat' is used almost exclusively without alternations ([t͡ʃatʲ-u-sʲ] and only rarely [t͡ʃat͡ʃʲ-u-sʲ]), while other borrowings whose stem ends in [tʲ] either prefer the alternating pattern or are more evenly split between alternation and non-alternation.

[7] One possible phonological conditioning has to do with the phonotactic markedness of the sequence /Cʲu/ in Russian. In native words this sequence is rare except at a morpheme boundary, but even there it is rare if C is an obstruent (Padgett 2010). A phonologically-conditioned account of these facts might also be possible if one assumes that the 1SG suffix is /ju/, with an underlying glide that triggers the alternations. However, such an account would not only be opaque, but it would have difficulties explaining some pockets of irregularity in the dental alternations and the identity of the epenthetic segment in the labial alternation.

[8] Moskvin (2015) discusses how the defectiveness of this verb is related to a possible phonotactic problem, the ill-formedness of the [tmlʲ] cluster. Searching for words with this cluster in the Russian National Corpus (RNC) yields no results. The dictionaries list two past passive participle for this verb, one without the alternation (*zatmënnyj*), and one with the alternation (*zatmlʼënnyj*) 'eclipsed,' but both of them are quite rare.

cluster [stʲ]) there are two possibilities: the majority of such stems follow the [tʲ] ~ [tʃʲ] alternation, but a few historically Old Church Slavonic verbs follow the [tʲ] ~ [ɕː] alternation. The regularity of Russian alternations alone already rules out simple explanations in which gaps are a natural consequence of highly irregular patterns. For instance, Yang et al. (2012) propose that gaps in Russian [tʲ]-final stems are due to the [tʲ] ~ [tʃʲ] alternation having too many exceptions (i.e., the [tʲ] ~ [ɕː] verbs). However, this explanation cannot be extended to other dental stems which only have one alternant.

		Alternations		
IPA	**Orthography**	**3SG**	**1SG**	**Gloss**
[dʲ ~ ʐ]	(d ~ ž)	[sʲɪdʲ-ít]	[sʲɪʐ-ú]	'sit'
[zʲ ~ ʐ]	(z ~ ž)	[vanzʲ-ít]	[vanʐ-ú]	'pierce, stab
[sʲ ~ ʂ]	(s ~ š)	[kósʲ-ɪt]	[kaʂ-ú]	'scythe'
[tʲ ~ tʃʲ]	(t ~ č)	[lʲɪtʲ-ít]	[lʲɪtʃʲ-ú]	'fly'
[tʲ ~ ɕː]	(d ~ šč)	[sakratʲ-ít]	[sakraɕː-ú]	'reduce in number'
[stʲ ~ ɕː]	(st ~ šč)	[vɨrastʲ-ɪt]	[vɨraɕː-u]	'raise, cultivate smth.'
[+lab] ~ [+lab]lʲ	(C ~ Cl')	[lʲúbʲ-ɪt]	[lʲublʲ-ú]	'love'
(pʲ,bʲ,vʲ,mʲ,fʲ)				

Table 2: Alternations in 1SG present tense verbs of conjugation II.

Albright proposes that low productivity can result not only from irregularity, but also from paucity of the data supporting a rule (Albright 2003; 2009b). In the next section I test whether this idea can explain Russian gaps by examining lexical statistics related to the alternations in question. Before I do that, I draw readers' attention to another factor about 1SG forms which will become relevant in the subsequent discussion. Besides the consonanantal alternations, some verbs have stress alternations in the 1SG cell of the paradigm. Consider, for example, the verbs 'to scythe' and 'to love' in Table 2. These verbs have stress on the stem in 3SG and in all other present tense forms except for 1SG where the stress shifts to the suffix. Overall, stress alternations affect a relatively small number of stem-stressed verbs. Unaccented verbs (those with no stem-stress) have no stress alternations.

3.2 Are gaps due to paucity of the data?

Albright (2003; 2009b) advances a theory of ineffability based on a particular learning model, the Minimal Generalization Learner (Albright & Hayes 2002). He argues that paradigm gaps can result from unreliable generalizations, that is, generalizations that have a lot of exceptions and/or apply to a relatively small number of items. In general, each rule or generalization is associated with a confidence score that can be thought of as a measure of how confident speakers feel in applying this generalization.[9] Gaps arise as a result of low confidence in all potential outputs (cf. the discussion of the threshold approach in section 2.4).

---

[9] The confidence of a rule is calculated as the number of items that follow the rule divided by the number of items that could potentially follow the rule (i.e., items that fall within the structural description of the rule), and this ratio is further adjusted by taking the lower confidence limit to penalize rules that apply to too few items (for details see Albright 2002). For example, the raw confidence of a rule that applies to 6 out of 7 items is $6/7 = 0.85$ and the adjusted confidence is 0.79 with the lower confidence limit $\alpha = .75$.

This theory crucially rests on a number of other assumptions which restrict the scope of possible rules. For example, if the rule for forming English past participles was formulated as "add -ed to the present tense stem" such a rule would be fairly reliable given the preponderance of regular verbs in English. However, if the participles were instead projected from past tense forms by more narrow rules, such as "if the past tense contains the vowel [o] then the participle will end in *-en* and [o] will change to [ɪ]" (e.g. *wrote – written*), such a rule would have low reliability due to narrower scope and many exceptions. An immediate question arises: what determines the starting point of the derivation? One restricting assumption that Albright makes is that inflectional forms are projected from other more informative inflectional forms, the *bases* of the paradigm, so that paradigmatic organization of a language will have an effect on which cells in a particular paradigm are more likely to be projected from other cells and, therefore, more likely to have gaps. Now, suppose we accept that the past participles in English are derived with reference to the past tense forms. It is still possible to formulate a general and quite reliable rule referencing the past tense: "the past participle is identical to the past tense form." This generalization is true for all regular and some irregular verbs, in particular irregular verbs ending in coronals. Do speakers compute such general rules in addition to the more specific ones, and if so, can they fall back on them when the more specific rules are unreliable? I take Albright's proposal to imply that speakers learn to ignore general rules in certain situations.[10]

Albright (2009b) suggests that the Russian gaps may also result from parochial rules and paucity of the data. That is, if we assume that the consonantal and stress alternations described in the previous section break up the dental-stem verbs into smaller subclasses and that speakers compute the rules for alternations separately for each consonant-stress type, there might not be enough data within each subclass to establish a reliable rule. Based on some rough preliminary calculations, Albright finds support for this hypothesis. He also notes that the labial stems are different from the dental ones in that they are less fragmented (e.g., they all have the same alternation, and are possibly more uniform with respect to stress alternations). Therefore, speakers probably have one general labial rule which explains the lack of gaps in this class of verbs. To test this proposal, the next section presents more in-depth analysis of lexical statistics for both dental and labial stems.

3.2.1 Lexical statistics: how robust are the 1sg alternations?

The question of robustness of the 1sg alternations is important for any potential explanation of Russian gaps. In particular, we have seen that the alternations in question are regular, but are they frequent enough for speakers to notice and to learn? To answer this question we need to know what counts as "frequent enough." First, it is important to look at type rather than token frequency since type frequency is better correlated with productivity (MacWhinney 1978; Bybee 1985). However, the exact number of different types of items required for establishing a rule may depend on a particular theory of rule-formation.

---

[10] For example, when discussing Spanish paradigm gaps connected to the mid-vowel alternations (e.g., gaps in 1sg of *abolir* 'to abolish' which are hypothesized to be due to uncertainty about the diphthongization of [o]), Albright assumes that speakers form narrow rules specific to each verb class and preceding consonant. He says "the division of the grammar into separate generalizations for different segments and different classes may be due to the fact that their statistics simply are so different across these different contexts, and, for accidental historical reasons, always have been" (Albright 2009b). Assuming that Spanish speakers do not have one general rule about mid-vowels helps explain why Spanish gaps appear in class III verbs only. In other classes the majority of verbs do not diphthongize, but in class III about half of the verbs with [e] diphthongize. Additionally, all verbs with [o] in class III diphthongize, but there are only three of them. If speakers considered [o]-verbs of all classes together, we would expect that they would follow the majority pattern of non alternation. But if they consider only class III [o]-verbs, there is not enough evidence for them to decide what to do.

Many theories are not specific enough to distinguish between rules that apply to a few versus many examples. According to Albright's metric of confidence, rules with narrow scope are punished and confidence falls too low – to the level of absolute uncertainty – when there are fewer than ten observations (Albright 2009b: 141). Taking this number as a rule of thumb, we now ask the following question: if speakers break up the lexicon into very specific subclasses by consonant and by stress, is it the case that each such subclass contains more than ten different stems?

To answer this question, I extracted all conjugation II verbs from the on-line list of Zaliznyak's paradigms.[11] I further divided these verbs into two groups: those whose lemma frequency is below or above one instance per million (ipm) in the *New Frequency Dictionary of the Russian Lexicon* based on the Russian National Corpus of 100 million wordforms (Lyashevskaja & Sharov 2009). The first set of words (frequency above 1 ipm) can be assumed to be familiar to an average native speaker and can be used for estimating type frequency of different alternations in the lexicon. The second set of words (frequency below 1 ipm) will be used to test the productivity of these alternations under the assumption that wordforms of verbs of this frequency are not memorized but derived by the grammar.[12] When compiling this database, I excluded those verbs from consideration that had the same stem and differed from each other only in the presence or absence of productive aspectual prefixes or the reflexive suffix. Typically, all derivatives of a stem have the same status with respect to defectiveness (e.g., *ubedit'* 'to convince,' *ubedit'sja* 'to convince oneself,' and *pereubedit'* 'to change someone's mind' are all defective). For better accuracy, I removed multiple derivatives of the same stem by hand since an automated procedure based on phonological shape would have difficulties with verbs that are distantly related and therefore appear to have the same stem (e.g., *s-razit'* 'to sleigh,' *vy-razit'* 'to express a thought,' *za-razit'* 'to infect'). There is evidence that stems which are merely historically related do not necessarily pattern in the same way. For example, *pretit'* 'to revolt' is defective, but the related *zapretit'* 'to prohibit' is not; the verb *obratit'* 'to turn into, to turn attention to' has the Old Church-Slavonic [tʲ] ~ [ɕ:] alternation, while the related *oborotit'* 'to turn around' has the [tʲ] ~ [tʃʲ] alternation. When considering verbs derived from the same root via productive aspectual prefixes (e.g., *katit'* 'to roll', *po-katit'* 'to start rolling', *za-katit'* 'to roll up'), I kept only one derivative in the database, namely the most frequent one which was not necessarily the least morphologically complex one. Note, that if we do not exclude multiple derivatives of the same root from consideration, then the sparsity of the data explanation would have no chance to succeed since there are definitely more then ten verbs per each consonant-stress type.

Next, each verb in the resulting database was classified based on its stress pattern, consonant alternation, and defectiveness. There were three categories of stress: "stem" for fixed stress on the stem, "end" for fixed stress on the suffix, and "alternating" for stress on the stem in all forms except 1SG. A verb was judged to be defective if it was on the list of defective verbs compiled by Sims (2006) with a few of my own modifications, which are summarized in Appendix A. The raw counts in Table 3 show the number of distinct verbal stems in each subclass of verbs cross-classified by stem-final consonant and stress. This table also shows the number of defective verbs in each subclass. Figure 1 presents the subset of this data for closer examination of the relationship between the number of stems in each subclass and defectiveness in dental stems.

---

[11] This list was downloaded from http://www.speakrus.ru/dict/all_forms.rar

[12] As a point of comparison, consider a few English verbs that have frequency < 1 ipm in Celex: *calk, bushwack, implode, lionize, miaw, nosh, inspan, junket, leather, reseat.*
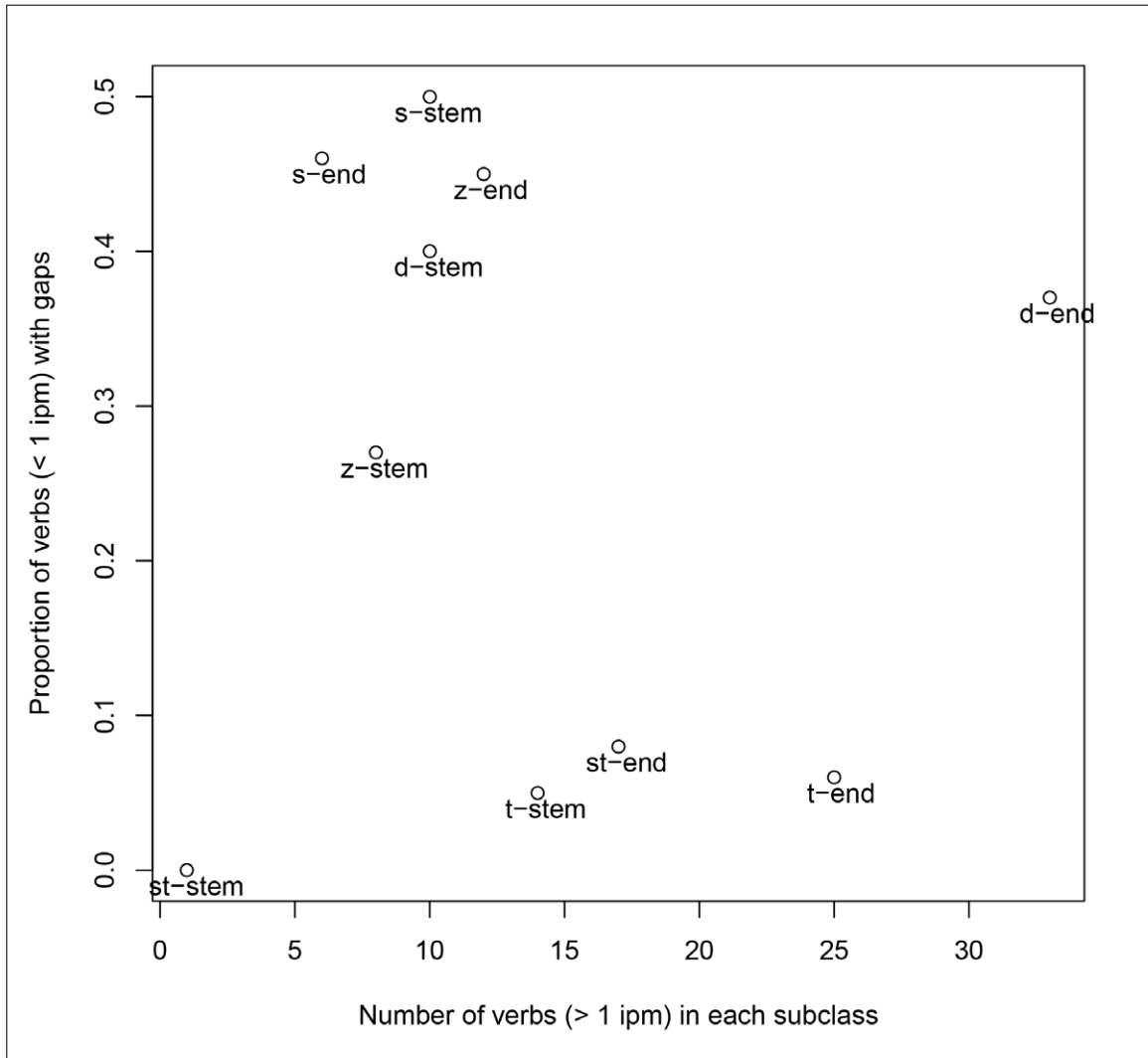
Figure 1: Number of dental stems with frequency > 1 ipm on the x-axis, proportion of defective verbs with frequency < 1 ipm on the y-axis.

Each point in Figure 1 corresponds to a subclass of dental stems.[13] The x-axis shows the number of verbs in each subclass that have frequency greater than 1 ipm and that are not defective – that is, verbs which can serve as the basis for forming the generalizations about the 1SG alternations. The y-axis shows the proportion of infrequent verbs (< 1 ipm) with gaps per each subclass, which can be thought of as a measure of defectiveness for a particular subclass. The reason why only low frequency verbs are considered for this measure, is that their defective status could not be memorized by the speakers. On the one hand, this graph confirms Albright's hypothesis that the data from narrow subclasses is somewhat sparse – the number of stems in each subclass is less then fifteen for many subclasses. However, on the other hand, there is no clear correlation between defectiveness and the number of stems per subclass. Likewise, there is no correlation between defectiveness and irregularity. For example, the only irregular stems, [tʲ]-final stems, in particular "t-end" stems which are highly irregular, have a relatively low proportion of defective verbs. On

---

[13] The graph does not show stress-alternating verbs. For defective verbs, it is impossible to determine whether they are stress-shifters or not since the stress shift occurs in the 1SG form. However, since practically all stress-shifting verbs have frequency greater than 1 ipm, it is safe to assume that defective verbs with frequency less than 1 ipm with stress on the stem in the rest of the present tense paradigm belong to the fixed stress type.

Alternation	stress type				
	alternating	stem	end	Total	Prop. of gaps
Dental Stems					
$t^j \sim t\int^j$	11 (11)	31 (12)	26 (11)	68 (34)	0.1 (0.12)
$t^j \sim \ç{:}$	0 (0)	2 (2)	14 (14)	16 (16)	
$t^j \sim ?$		4 (3)	5 (4)	9 (7)	
$st^j \sim \ç{:}$	5 (5)	8 (1)	29 (17)	42 (23)	0.07 (0.08)
$st^j \sim ?$		1 (1)	2 (1)	3 (2)	
$d^j \sim z̧$	14 (13)	19 (10)	42 (33)	75 (56)	0.23 (0.16)
$d^j \sim ?$		8 (2)	14 (9)	22 (11)	
$z^j \sim z̧$	2 (2)	16 (8)	18 (12)	36 (22)	0.25 (0.15)
$z^j \sim ?$		6 (3)	6 (1)	12 (4)	
$s^j \sim ş$	9 (9)	16 (10)	13 (6)	38 (25)	0.28 (0.1)
$s^j \sim ?$		6 (0)	9 (3)	15 (3)	
Labial Stems					
$f^j \sim fl$	0 (0)	5 (1)	1 (1)	7 (2)	0
$m^j \sim ml$	2 (2)	15 (9)	21 (17)	38 (28)	0.02 (0.03)
$m^j \sim ?$			1 (1)	1 (1)	
$b^j \sim bl$	3 (3)	18 (10)	22 (16)	44 (29)	0
$p^j \sim pl$	12 (12)	6 (1)	20 (17)	38 (30)	0
$v^j \sim vl$	5 (5)	34 (22)	23 (16)	62 (43)	0

Table 3: Number of stems per consonant-stress subclass. C ~ ? rows are used for defective verbs. The numbers in parenthesis indicate counts for verbs whose lemma frequency is greater than 1 ipm. For stem-stressed defective verbs, it is impossible to determine whether they shift the stress to the suffix in 1SG since they do not have 1SG forms. Therefore, there is no distinction between alternating and non-alternating stem-stress for defective verbs.

the other hand, the regular and most numerous types of stems, [d^j]-final stems with stress on the suffix ("d-end"), have a much higher proportion of defective verbs. A proportion test shows that the difference between the proportion of gaps in "t-end" and "d-end" infrequent stems is statistically significant (P(t-end) = 0.06, P(d-end) = 0.35, X^2 = 4.05, p = 0.04), and so are the differences between several other classes. For instance, the two equally frequent types of stems, "t-stem" and "z-end," have very different proportions of gaps (P(t-stem) = 0.05, P(z-end) = 0.45, x^2 = 5.07, p = 0.02). Thus, the lack of correlation between the size of a subclass and its "gappiness" is not just due to a floor effect.

This data also casts doubt on the hypothesis that speakers do not generalize over subclasses of verbs that have different stress patterns and that uncertainty about stress placement further contributes to defectiveness. For instance, end-stressed verbs for which the stress shift is never an option are as likely to be defective as stem-stressed verbs. The proportion of defective infrequent end-stressed verbs is 0.27, and the corresponding proportion for stem-stressed verbs is 0.24, not significantly different by the proportion test, p = 0.76. If we assume that speakers generalize over stem-stressed, end-stressed, and stress-alternating verbs – after-all, stress is irrelevant for determining the quality of the alternating consonant – then speakers would be able to rely on more than twenty stems per each consonant type to learn the dental alternations. On the other hand, if we assume

that speakers stick with narrowly defined subclasses even when further generalization is possible, then it is not clear why the labial alternations should be protected from gaps since they are as poorly represented within each narrow subclass as the dental alternations and are not obviously more uniform with respect to stress-alternations. Overall, the narrow generalization hypothesis leading to paucity of the data is insufficient to explain all facts connected to defectiveness of Russian dental-stem verbs.[14]

3.3 Further questions about the distribution of defective verbs

The previous section revealed two puzzles for any account linking defectiveness to alternations in Russian verbs: the asymmetry between labial and dental stems and the fact that the presumably problematic alternations are mostly regular and relatively well-attested in the lexicon. There are two additional puzzles for an account that connects defectiveness to consonantal alternations.

verb (inf.)	Gloss	1SG pres.	2SG pres.	lemma freq.
[kəlʲɪsʲ-ítʲ]	'wheel around'	?	[kəlʲɪsʲ-íṣ]	1.8
[grʲézʲ-ɪtʲ]	'(day)dream' (poet.)	?	[grʲézʲ-ɪṣ]	2.6
[zəltʲ-ítʲ]	"yellow (smth)'	?	[zəltʲ-íṣ]	0
[jɪrundʲ-ítʲ]	'speak nonsense'	?	[jɪrundʲ-íṣʲ]	0
[kvásʲ-itʲ]	'ferment'	[kváṣ-u]	[kvásʲ-ɪṣ]	0.8
[gəradʲ-ítʲ]	'enclose'	[gəraz̧-ú]	[gəradʲ-íṣ]	2.2
[skapʲítʲ-ɪtʲ-sʲa]	'to die' (colloq.)	[skapʲítʃ-u]	[skapʲít-ɪṣ]	0
[əpaskúdʲ-ɪtʲ]	'to debase'	[əpaskúz̧-u]	[əpaskúdʲ-ɪṣ]	0

Table 4: Dental stem verbs of low frequency ("?" indicates a gap).

The first puzzle has to do with explaining which specific dental stems are defective. If defectiveness is a systematic phenomenon due to low productivity of dental alternations, one would expect that all low-frequency dental stems would be equally defective, but this is not the case. Although it is true that almost all defective verbs are verbs of low frequency, it is not true that all low-frequency verbs with dental stems are defective. Examples in Table 4 show verbs that have similarly low lemma frequency but differ in whether they have a gap in 1SG.

Another puzzle has to do with the fact that the presumably problematic alternations apply elsewhere in the inflectional paradigm of the verb without causing widespread gaps. Namely, the same alternations apply in the past passive participle forms of verbs that allow such forms semantically. For example, a past passive participle of *kvas-it'* 'ferment' is *kva[ṣ]-ennyj* 'fermented'. A caveat worth noting here is that for a small set of frequent [dʲ]-final verbs of Old Church Slavonic origin, the participles have the [dʲ] ~ [z̧d] alternation instead of the [dʲ] ~ [z] alternation found in the 1SG forms (e.g., ro[dʲ]-it' 'to give birth', ro[z]-u 1SG, ro[z̧dʲ]-ënnyj PAST PASS. PART.). Furthermore, a handful of verbs lack the expected alternations in the past passive participles (e.g., *von[zʲ]it'*, *von[zʲ]ënnyj*,

---

[14] Another subclass of verbs in Russian presents a problem for this hypothesis. These are verbs of first conjugation that have similar alternations as those described here: stem-final consonant mutations (g ~ z̧ and k ~ tʃʲ) in several present tense forms (e.g., *izrek-u* 'utter, 1sg,' *izre[tʃʲ]-ëš* 'utter, 2nd person sg.'). No verb of this type is defective despite the fact that the number of verbs undergoing these alternations is very small. According to my estimates, there are only thirteen g-final stems and seven k-final stems that are expected to undergo these alternations, with one of the verbs, *tkat'* 'weave,' failing to alternate. It is possible that the reason why these verbs are not defective is because the velar alternations apply in four out of six forms in the present tense paradigm, including the most frequent 3rd person SG. form.

von[z̦]*ënnyj* 'to stab'). However, the overwhelming majority of verbs have the same alternations in past passive participles as in 1SG forms. Thus, any account linking gaps to alternations must also explain why these alternations do not typically lead to paradigm gaps in the participles.[15]

3.4 Local summary

To summarize so far, the 1SG dental alternations in Russian are regular with a single exception of [tʲ]-final stems. There are also sufficiently many stems with these alternations for speakers to learn the pattern even if we consider stems with different stresses separately (at least according to a model on which a pattern may be productive if it holds of at least 10 examples). Additionally, although the overwhelming majority of defective verbs are of low frequency, it is not the case that all low frequency dental verbs are defective. And finally, the 1SG alternation also appears in past passive participles, where it does not lead to wide-spread gaps. All of these facts challenge an account linking defectiveness in Russian verbs to morpho-phonological alternations.

In the following sections I will argue that despite these challenges, gaps in Russian verbs are indeed synchronically connected to the morpho-phonological alternations affecting the 1SG form with the exception of a few frequent verbs whose defective status may be lexicalized. The challenges above can be addressed once a relevant, but previously missed, generalization is observed. This generalization is discussed next.

4 The transderivational dependency

In this section I will argue that all verbs with dental-stems are defective unless they have listed stem allomorphs with the same alternation as the one expected in the 1SG. This observation calls to mind the phenomenon of *lexical conservatism*, a term coined in Steriade (1997) to describe reluctance on the part of the speakers to produce novel allomorphs of existing morphemes. Such reluctance can manifest itself as a transderivational dependency of the following sort: a specific pattern only applies in a novel context if the morpheme to which it applies already has a listed allomorph where this pattern is attested. For example, Steriade (2008) shows that in Romanian a derivational suffix -/ik/ causes palatalization of stem-final /k/ only for stems which show the same palatalization alternation (k → tʃ before /i/) in the inflectional paradigm. In English the -*able* derivative *custódiable* (from *cústody*) with a stress-shift is preferred by the speakers to *cústodiable* because it better satisfies *LAPSE[16] and because there already is an existing derivative of this stem with the stress on the second syllable – *custódial*. For nouns which lack such derivatives, stress does not shift when -*able* is attached (e.g., *chállengable*, *challéngable) (Steriade 1997). How exactly such an effect should be understood and analyzed and in what way it can contribute to defectiveness is the question I take up in section 5. Here, I just focus on empirical evidence supporting the following transderivational relation:

(3)   Verbs of second conjugation that have other morphologically related forms with

---

[15] I say "typically" because a small number of verbs also have "difficult" past passive participles according to Zaliznyak's morphological dictionary. But the nature of most participial gaps appears to be different from the 1SG gaps judging from Zaliznyak's descriptions. He classifies gaps in participles into several categories mostly based on the morphological properties of the stem (e.g., one such category is imperfective verbs ending in unstressed -*ivat'* and derived via prefixation to a base that already contains a prefix: e.g., *na-vy-dum-iv-a-t'* 'to think up'). Only a handful of these gaps may be tied to the dental alternations and almost all occur in verbs which also have a gap in 1SG (*uvjazit'* 'to sink (smth.),' *šerstit'* 'to search vehemently,' *prijutit'* 'to shelter'). Further work is necessary to systematically investigate the gaps in past passive participles and their possible connection to dental alternations.

[16] *LAPSE is defined as assigning a violation for every occurrence of two consecutive unstressed syllables.

the 1SG dental-palatal alternation are not defective, and vice versa: verbs that lack such alternations in other forms are defective.

4.1 Evidence from the native verbs

Recall that the 1SG form stands out in the present tense paradigm of second conjugation as the only form with alternations. However, the same alternations affect dental and labial stems in a few other related forms outside the present tense. As I have already mentioned, within the same paradigm, this alternation occurs throughout the past passive participle paradigm. For example, compare the full paradigms of the two verbs given in Table 5, 'transport' and 'be rude'. The second of these verb is intransitive and lacks passive participles; it is also defective.

Infinitive	voz^j-i-t^j	derz^j-i-t^j
Past	voz^j-i-l, voz^j-i-l-a,	derz^j-i-l, derz^j-i-l-a,
	voz^j-i-l-o, voz^j-i-l-i	derz^j-i-l-o, derz^j-i-l-i
Non-past	voʐ-u, voz^j-iʂ,	–, derz^j-iʂ,
	voz^j-it, voz^j-im,	derz^j-it, derz^j-im,
	voz^j-ite, voz^j-at	derz^j-ite, derz^j-at
Imperative	voz^j-i, voz^j-ite	derz^j-i, derz^j-ite
Present part.	voz^j-aɕ:-ij, voz^j-aɕ:-ije,	derz^j-aɕ:-ij, derz^j-aɕ:-ije,
(nom.)	voz^j-aɕ:-eje, voz^j-aɕ:-aja	derz^j-aɕ:-eje, derz^j-aɕ:-aja
Past part.	voz^j-ivʂ-ij, voz^j-ivʂ-ije,	derz'-ivʂ-ij, derz^j-ivʂ -ije,
(nom.)	voz^j-ivʂ -oje, voz'-ivʂ-aja	derz^j-ivʂ-oje, derz'-ivʂ-aja,
Present pass. part.	voz^j-im-ij, voz^j-im-ije,	
(nom.)	voz^j-im-aja, voz^j-im-oje	
Past pass. part.	voʐ-enn-ij, voʐ-enn-ije,	
(nom.)	voʐ-enn-aja, voʐ-enn-oje	

Table 5: Paradigms of /voz^j-/ 'transport' and /derz^j-/ 'be rude' (defective).

Outside of the inflectional paradigm, the dental and labial alternations appear throughout the verbal paradigm of secondary imperfectives which are productively derived from perfective verbs via imperfective suffixes such as -(y/i)va, -va, -a (e.g., rod-it'$_{PFV}$ – rož-a-t'$_{IPFV}$ 'give birth') or prefix-suffix combinations (e.g., xod-it'$_{PFV}$ 'to walk' – po-xaž-iva-t'$_{IPFV}$ 'to walk occasionally'). Finally, the same alternations appear less systematically in some nominalizations and in adjectival forms. Below is an example of a verb that has the 1SG alternation appear in multiple morphologically related forms.

(4)   Derivatives of the root -vstre[t^j]- 'meet' with the stop-affricate alternations:
1. vstre[tʃ^j]-u   'meet' PFV 1SG PRES.
2. vstre[tʃ^j]-enn-yj   'one who was met,' PAST PASS. PART., NOM.SG.M.
3. vstre[tʃ^j]-a-t'   'meet' IPFV INFINITIVE
4. vstre[tʃ^j]-a   'meeting' N. NOM.SG.F.

However, for various reasons some verbs lack derivational relatives with alternations. For example, it is semantically impossible for intransitive verbs to have past passive participles. Imperfective verbs also often lack such forms. Likewise, not all verbs have secondary imperfectives. The alternations in nominal and adjectival forms have largely leveled out, so only a few verbs still have related non-verbal forms with alternations. Thus, for many

low-frequency verbs speakers have only encountered the non-alternating variant of the stem. I hypothesize that it is precisely these verbs that are defective.

4.1.1 The distribution in the lexicon

As a first test of the hypothesis that defective verbs lack other inflectional and derivational relatives with consonant alternations, I checked the compiled corpus of verbs described in section 3.2.1 to see if verbs that were marked in dictionaries as defective lacked alternations in their wider morphological family. Table 6 summarizes the findings.

	defective verbs	non-defective verbs
alternations elsewhere	6 (10%)	215 (78%)
no alternation elsewhere	55 (90%)	60 (22%)

Table 6: Dependency between defectiveness and presence of 1SG alternations in related forms in dental stems (data from a corpus of second conjugation verbs described in section 3.2.1).

While there is no perfect correspondence between existence of expected alternations elsewhere and defectiveness, the statistical dependency between these two factors was highly significant by Fisher's exact test (p = 0.0001). Ninety percent of defective verbs lack related forms with the expected alternation, and 78% of non-defective verbs have related forms with the 1SG alternation. This dependency by itself does not establish a causal relationship, but it is consistent with the hypothesis that existence of alternations in at least one stem allomorph protects the stem from defectiveness or, alternatively, that lack of such alternations can push the stem towards being defective.

The largest exception to the generalization in (3) are the sixty verbs that lack attested alternations in other forms and yet are not defective. It may be that these verbs are frequent enough so that speakers can simply remember their 1SG forms. However, at least twenty-seven of these verbs have lemma frequency less than 1 ipm and are unlikely to be memorized. Another possibility is that the criteria I used for determining defectiveness, observations of dictionary writers, is inaccurate. In the next section I take a closer look at verbs of this type in the context of another measure of defectiveness, interspeaker agreement in the productions of 1SG forms.

The other exceptions in Table 6 are the six defective verbs which have past passive participles (ppp) or other forms with alternations listed in the online version of Zaliznyak's Dictionary (Starostin 1998a). These are *obessmert-it'* (ppp: *obessmer[tʃʲ]-enn-yj*) 'immortalize,' *oščut-it'* (noun: *oščú[ɕː]enije*) 'sense,' *prigvozd-it'* (ppp: *prigvo[ʐ]ënnyj*) 'nail down,' *umiloserdit'* (ppp: *umiloser[ʐ]ennyj*) 'have mercy,' *želtit'* (ppp: *žel[tʃʲ]ënnyj*) 'yellow,' *obezlošadit'* (ppp: *obezloša[ʐ]ennyj*) 'dehorse.' However, note that *oščut-it'* has the phonological problem mentioned in footnote 6 and the defective status of *obessmert-it'* is questionable. The 1SG form *obessmerču* appears in Zaliznyak (1980); Ozhegov (1988), and is used in literary writing, for example in poems of Brodsky ("Ja obessmerču plot' – ty obessmertil dušu" 'I will immortalize the body – you have immortalized the soul.' *Goodbye Ode*, 1964). All the other verbs above have very infrequent ppp's, some of which may in fact not be used with alternations by the speakers. For example, the 2007 horror film *Nailed* was translated by some Russian websites as *Prigvoz[dʲ]-ënnyj* without the alternation, and others as *Prigvo[ʐdʲ]-ënnyj* with the irregular Old Church Slavonic alternation; the latter form also appears in some dictionaries.

4.1.2 Interspeaker agreement as a measure of defectiveness

The data discussed in section 4.1 reveals that there is a strong statistical correlation in the Russian lexicon between defectiveness and absence of the expected alternation in other

related forms. However, since the list of defective verbs used in establishing that correlation is based on dictionaries, a different picture may emerge when we look at actual language use. So, to verify the status of defective verbs further and to test the correlation between defectiveness and presence of expected alternations elsewhere, I examine speakers' uses of 1SG forms on the web. Before that, however, I turn to the question of what happens when speakers produce problematic forms, that is, when they "fill the gap".

As mentioned in the introduction, speakers do not always avoid problematic forms of defective lexemes, especially in informal settings. For frequent defective verbs, productions of 1SG forms can be intended as ironic.[17] For infrequent defective verbs such productions are marked by uncertainty. This uncertainty manifests itself in writing in several ways: the problematic forms may be enclosed in quotes, or followed by smilies or parenthetical remarks. In general, speakers' willingness to fill the gap most likely reflects differences in their personalities rather than differences in grammar.

It is interesting to examine what speakers do when they fill a gap. It has been observed in experimental studies that productions of problematic forms are marked by low interspeaker agreement, which on the surface looks like variation (Albright 2003; Sims 2006). This is also true of Russian defective verbs. The two most frequent variants produced by the speakers are a variant with the expected alternation (e.g. *pyleso[ş]-u*) and a variant without the alternation (e.g., *pyleso[sʲ]-u*). This variation is reported in descriptive works such as Graudina et al. (1976) and is confirmed in production studies (Sims 2006; Pertsova & Kuznetsova in press), as well as in Google searches presented shortly. The fact that speakers often produce the non-alternating pattern is striking given that this pattern is virtually unattested in standard speech, although it can be found in colloquial speech, slang, certain dialects, and foreign borrowings (Slioussar & Kholodilova 2013). I discuss borrowings in section 4.3.

Besides the expected alternations and non-alternations, speakers occasionally produce some other variants of the 1SG, although very rarely. In particular, in her experiment Sims (2006) found that some speakers produced 1SG forms with the "incorrect" alternation [dʲ] ~ [ʐd] for the two verbs of Old Church Slavonic origin in which this alternation appears in the past passive participles. Speakers did not use this alternation with any other dental verbs. This fact is interesting because it directly shows that the other allomorphs of the stem can influence how the 1SG is realized. When it comes to novel stems, in particular novel borrowings, Slioussar & Kholodilova (2013) report that speakers sometimes produce the [dʲ] ~ [dz] alternation which is different from both the Old Church Slavonic [ʐd] and from the expected [ʐ]. This alternation preserves the stem segments in their original order and contains the expected fricative. Even more surprisingly, speakers occasionally apply the labial alternations to dental stems, producing forms like *zafrend-lʲ-u* (as the 1SG of *zafrend-itʲ* 'friend (on social media)'), or use the wrong dental alternant, optionally keeping the stem consonant (e.g., *zafren(d)[tʃʲ]-u*). Although such "mistakes" are exceedingly rare,[18] they do crop up in dental stems but virtually never occur in labial stems. These facts further confirm the asymmetry between labial and dental stems and suggest that the computation of 1SG alternations involves competition of many alternatives, the two most prominent ones for the dental stems being the form with the expected alternation and the form without the alternation.

---

[17] For example, here is an excerpt from Winnie the Pooh's song which uses the defectiveness of the famous defective verb, *pobed-itʲ* 'win, conquer,' for a comic effect: *Xorošo bytʲ medvedem, ura! Pobežu (net, pobedʲu!) Pobedʲu ja žaru i moroz!* 'It's good to be a bear, hurray! I will conquer *(pobežu?* no, *pobedʲu!)* I will conquer *(побежу)* the heat and the frost!'

[18] I found only ten instances of the last two patterns online, while the expected alternation appears in hundreds of thousands.

The variation observed in the production of defective words, however, is different from the canonical cases of variation, when multiple forms are equally acceptable for the speakers. In particular, experimental studies mentioned earlier show that the variation observed in the production of defective forms is tied to low confidence (measured by subjective ratings) and goes hand in hand with other signs of defectiveness, such as circumlocutions, hesitations, ironic usage, and so on. There is a general correlation between low confidence and low interspeaker agreement for defective lexemes, while no such correlation is found in cases of true variation. This is illustrated in Figure 2 from Sims (2006) who elicited 1SG forms of verb doublets (free variants such as *kudaxt-aj-u* and *kudaxč-u* 'cluck 1SG')[19] and verbs with dental stems of second conjugation. Note that confidence remains high for verb doublets regardless of whether intersubject agreement is high or low. Thus, they pattern differently from defective verbs.

4.1.3 Web searches

Given that low interspeaker agreement and low confidence in productions of problematic wordforms are indicative of defectiveness, we can use these measures to diagnose the status of verbs instead of, or in addition to, relying on dictionaries. To this end, I examined speakers' productions of 1SG forms of dental and labial stems on the web and in two elicitation experiments. Here, I focus on the data from the web. The experiments are discussed in section 4.2.

I used the web because the 1sg forms of defective verbs are virtually unattested in standard corpora. The web, in addition to its size, has an advantage of containing texts representing the informal, spoken-version of the language. The disadvantage of course is that data from the web is notoriously noisy and search engine counts are generally not precise. Nevertheless, if a clear pattern emerges despite all the noise, it would be extremely unlikely that such a pattern were an accidental artefact. We will see that a very stark pattern emerges when we look at the productions of 1SG forms of dental and labial stem verbs attested on the web.

To test the productivity of 1SG alternations in dental versus labial verbs, defective versus non-defective verbs, and verbs with or without expected alternations in related forms, I examine four type of verbs:

(5)   Types of verbs included in the searches
   a.   Group A: no alternations elsewhere, defective, dental. These are verbs with recognized or prescribed gaps and no alternations in related forms. We expect to find low interspeaker agreement for the 1SG forms of these verbs.
   b.   Group B: no alternations elsewhere, suspected defective, dental. These are verbs with no prescribed gaps, but no alternations in any of the related forms. These verbs are among the sixty apparent exceptions in Table 6. If they are indeed non-defective, they should show higher interspeaker agreement than the defective verbs. If, on the other hand, they are defective, they should behave similarly to Group A.
   c.   Group C: alternations elsewhere, regular, dental. These verbs have no recognized or prescribed gaps and have attested alternations in some other related form. These verbs are expected to show high inter speaker agreement and have a high rate of 1SG forms with the expected alternation.

---

[19] Verb doublets present a well-known case of morphological variation in Russian due to the *-a* to *-aj* suffix-shift (Andersen 1980; Shvedova 1980; Nesset 2010, and many others). This shift can be viewed as leveling at the level of the stem, rather than the wordform. Namely, speakers interpret first conjugation stems with the suffix *-a* as stems with the suffix *-aj* because the distinction between the *-a* and the *-aj* suffixes is neutralized in many forms.
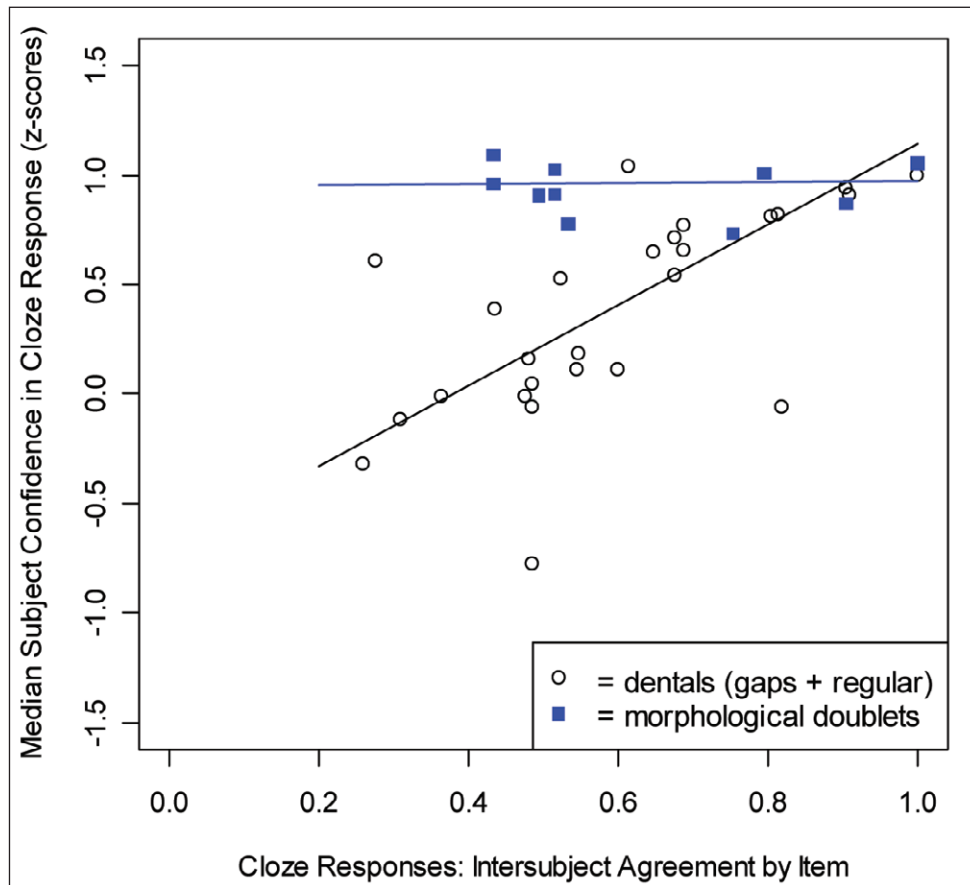
Figure 2: Median confidence vs. intersubject agreement rate in the production of 1SG forms (from Sims 2006, Fig. 26. "No correlation between intersubject agreement and confidence ratings for doublets.")

　　d.　Group D: labial. These verbs are also expected to show high inter speaker agreement and a high rate of 1SG forms with the expected alternations.

Methods: In my queries, I focused on native verbs in the lower frequency spectrum (frequency range between 0 and 6 ipm). To estimate interspeaker agreement, I searched for all logically possible 1SG variants of each verb: one without any alternations, one with the expected alternation, and those with prescriptively unacceptable but occasionally attested alternations identified in similar searches in Slioussar & Kholodilova (2013). To reduce the amount of noise, I limited the language to Russian, the region to Russia, and searched for phrases where the 1SG form is immediately preceded by the 1st person singular pronoun "I" using advanced Google search (in July of 2014). Upon examining the returns of searches, I excluded verbs which yielded a high number of erroneous responses due to the fact that the form in question was homophonous with some other form (e.g., *sglažu* is ambigous between 'I will even out' and 'I will jinx'), or due to the fact that it was a frequent misspelling of some other form. Verbs that yielded fewer than five hits were also excluded since interspeaker agreement could not be reliably estimated from such a low number. I ended up with twenty one to twenty six verbs per group (see Appendix B for the full list).

　Results: Figure 3 shows the overall interspeaker and inter-item agreement for the four groups of verbs defined above. Each dot in the graph corresponds to a single verb. The

x-axis is the log of the total number of hits for all 1SG forms, and the y-axis shows the proportion of hits in which the 1SG form contained an expected alternation. The unexpected alternations were extremely rare for these verbs, so the overwhelming majority of responses were binomially distributed as having the expected alternation or having no alternation.

The resulting pattern is apparent to the naked eye: groups C and D, labial stems and dental stems that have 1SG alternations elsewhere, have high interspeaker and inter-item agreement. That is, almost all of these stems almost always alternate in the expected way. On the other hand, dental stems that do not have alternations elsewhere, groups A and B, show great variability in their alternation behavior. The two groups look very similar to each other, providing support for the generalization in (3). To see whether proportion of alternations for a verb significantly depends on its frequency and membership in a specific group (A, B, C, D), I ran a logistic regression whose results are summarized in Table 7. Effects of both group and frequency were significant as well as the interaction between them. The interaction shows that the effect of frequency is different for different groups.
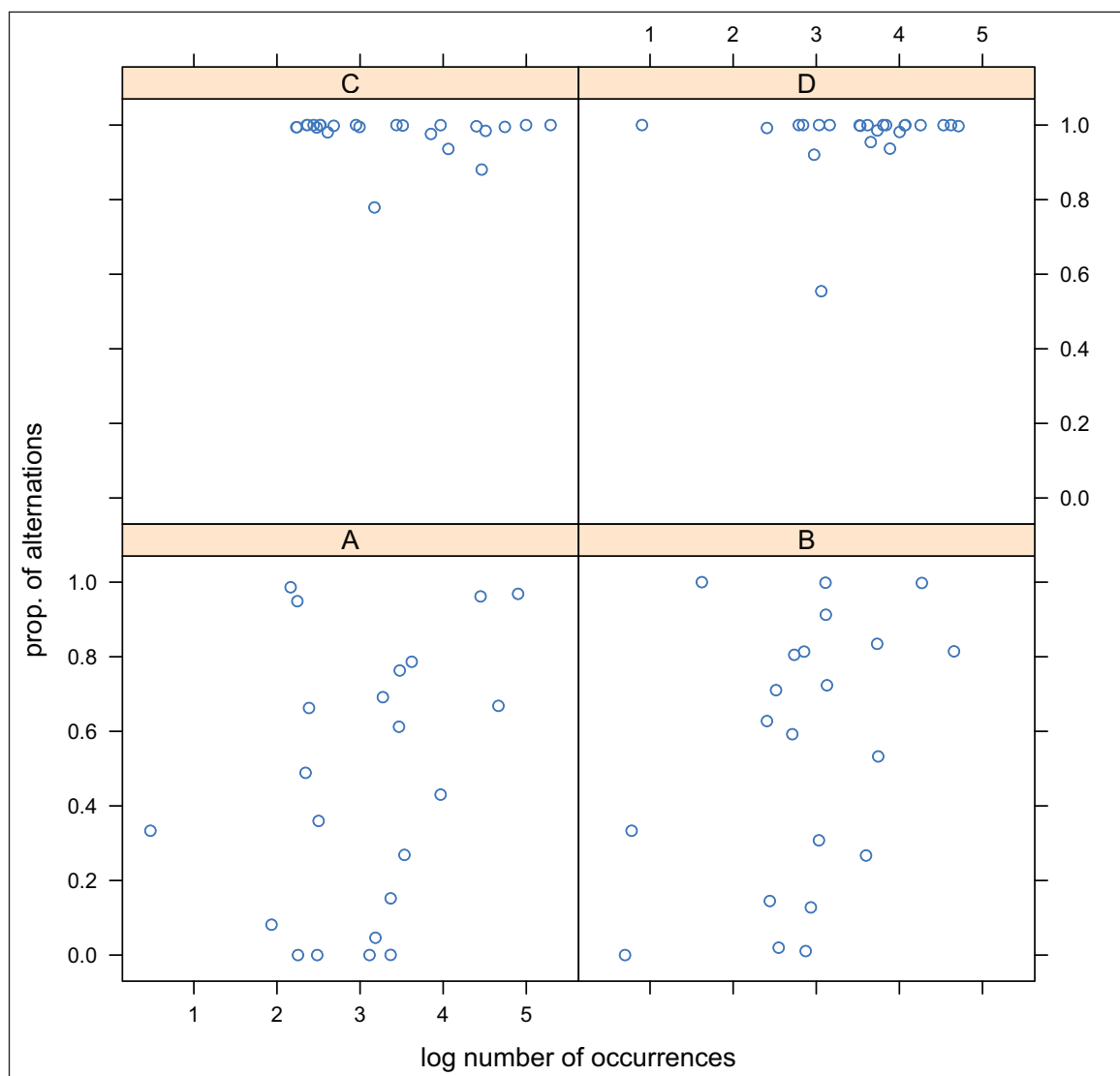
Figure 3: x-axis: log frequency of the total number of 1SG occurrences; y-axis: proportion of 1SG occurrences that contained the expected alternations.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−6.46	0.05	−133.45	0.0000
group B	3.83	0.07	52.62	0.0000
group C	4.78	0.09	55.51	0.0000
group D	1.99	0.16	12.22	0.0000
freq.	1.77	0.01	160.70.	0.0000
group B * freq.	−0.82	0.02	−47.74	0.0000
group C * freq.	−0.49	0.02	−25.38	0.0000
group D * freq.	0.55	0.04	12.69	0.0000

Table 7: Coefficients of the logistic regression model predicting proportion of alternations in the 1SG forms. Group A is the reference category for "group" (log likelihood = −36719, df = −89).

4.2 Experimental evidence

These corpus findings were further confirmed in two experiments, in which native speakers of Russian performed a fill-in-the blank cloze reading task by providing first, second, and third person singular forms of dental and labial verbs of second conjugation. Participants also rated confidence in their responses on a 5-point Likert scale. The stimuli consisted of low frequency dental and labial stems of the four types described in (5). The reader is referred to Pertsova & Kuznetsova (in press) for a detailed description of the experimental design and stimuli. The first experiment showed that verbs which did not have alternations elsewhere elicited significantly fewer alternations in the participants' responses and had lower confidence scores compared to verbs which did have 1SG alternations in other forms. More specifically, we found that the confidence ratings were significantly higher for verbs of group C compared to verbs of groups A and B, which were not significantly different from each other. Looking at the proportion of expected alternations produced in 1SG, we found significant differences among all three groups, with group A having the lowest percentage of expected alternations (59%), followed by group B (74%), followed by group C (94%). Labial stems also had high percentage of alternations, 94%. This experiment confirms that the pattern revealed in the corpus is not simply due to lexical variation because true variation is not associated with low confidence (see section 4.1.2).

A second experiment reported in Pertsova & Kuznetsova (in press) was designed to test whether the transderivational effect was gradient or categorical. Namely, we tested whether speakers' confidence in the 1SG alternations increased proportionally to the token frequency of the expected alternations present in the derivational nest of the verb. The stimuli in this experiment included dental verbs which varied with respect to the size of their derivational nest, the token frequency of alternating forms in the nest (calculated from the Modern subcorpus of RNC), and the total frequency of all the forms in the nest. We found that the transderivational effect was not gradient, but categorical. That is, speakers' confidence and proportion of produced alternations in 1SG did not gradually increase with the increase in token frequency of these alternations in the derivational nest of the verbs. Instead, verbs were more or less bimodally distributed into likely alternators and less likely alternators (i.e., verbs with mean proportion of alternations < 0.8). The less likely alternators were almost exclusively verbs that lacked any forms with the alternating variant of the stem in their nest. There was no difference among the rest of the verbs; speakers almost always produced the expected alternation in 1SG regardless of whether this alternation was frequent or not for a specific verb.

The corpus and the experimental data on native verbs confirm the lexical conservatism hypothesis and suggest that the number of defective verbs is currently underestimated.

Verbs like *frantit'* 'dress like a frant,' *gundosit'* 'speak nasally,' *strekozit'* 'be too lively,' and other verbs of low frequency of type B also appear to be defective despite not being marked as such in dictionaries.

4.3 Evidence from recent borrowings

Recently borrowed verbs present a good opportunity to test the hypothesis that defectiveness is synchronically tied to dental (but not labial) alternations, and that it is subject to the transderivational dependency found in the native verbs. When a word is just borrowed into a language, it will lack many derivational relatives. However, it is possible that more frequent borrowings quickly develop such relatives, which eventually become lexicalized and mutually reinforce each other. Thus, we might also expect a difference between high and low frequency borrowings.

In recent years, the Russian language has absorbed a great number of borrowings from English. A large number of borrowings are specialized terms that are being used within smaller communities and that do not become wide-spread, such as gaming terms. There are also many more widely-used computer and social media borrowings such as *frendit'* 'friend,' *fludit'* 'flood with meaningless information, to spread nonsense online,' *bekapit'* 'back-up,' and so on. Here is an example of a sentence with recent borrowings from English (from http://forum.ee/t158721/aaa-pomogite/page-2).

(6) Skaž-u   prosto čto **serf'-u**   pod **furrifoks-om**   s
    say-1SG.PRES simply that surf-1SG.PRES under Firefox-INSTR.SG. with
    kuč-ej   **plagin-ov…**
    bunch-INSTR.SG. plug-in-GEN.PL.…
    'Let me just say that I surf with Firefox with a bunch of plug-ins…'

As this example shows, foreign borrowings usually acquire Russian morphology. For verbs, this means that they are assigned to the productive second conjugation class and, therefore, are expected to alternate in 1SG.

To collect data on 1SG productions of novel borrowings, I used the same methods as those for the existing verbs described previously. Table 8 lists the borrowings used in my searches. Many of them come from Slioussar & Kholodilova (2013) with the addition of several of my own examples, including a much greater number of [fʲ]-final stems. These stems are of particular interest because they present a good opportunity to test the productivity of the labial alternation and to see at what level of generality it applies. There are very few f-stem verbs in the lexicon and practically all of them come from borrowed vocabulary due to the fact that /f/ is not a native Russian phoneme.[20] However, since there is a general productive rule that applies to all labial consonants, one would expect that the labial alternations would be freely extended to f-final stems. On the other hand, if speakers use narrow consonant-specific rules, f-stem borrowings might behave differently from the other labial stem borrowings.

Figure 4 shows the overall difference between labial and dental borrowings in terms of proportion of 1SG forms attested with alternations on the web. It is apparent that the two groups of verbs behave differently. Labial stems for the most part have the expected alternations in 1SG (except for a few verbs discussed shortly). Dental stems, on the other hand, vary greatly from each other, and many of them show a low degree of interspeaker agreement. As a group, the dental borrowings resemble the existing defective verbs in Figure 3. The logistic regression on proportion of alternations with dependent variables being type of stem (labial vs. dental) and frequency shows both factors and the interaction between

---

[20] /f/ came into Russian relatively late (after Christianization).

f^j	*serfit'* 'surf,' *grifit'* 'grief' (gaming term), *defit'* 'defend', *nerfit*^j 'nerf' (reduce effectiveness),' *blafit'* 'bluff,' *morfit*^j 'morph,' *debafit*^j 'debuff,' *sejfit'* 'save', *prufit'* 'proof'
v^j	*sejvit'* 'save,' *folovit'* 'follow,' *kreativit'* 'be creative,' *muvit'* 'move,' *testdrajvit'* 'test-drive,' *karvit'* 'carve'
_pj	*stopit'* 'stop,' *bekapit'* 'backup,' *oftopit'* 'be off-topic' *fotošopit'* 'photoshop,' *tajpit'* 'type,' *šejpit'* 'shape'
m^j	*gejmit'* 'play games,' *flejmit'* 'flame,' *zumit'* 'zoom' *spamit'* 'spam,' *farmit'* 'farm'
b^j	*krabit'* 'crab (ruin someone's game)', *nubit'* 'noob (to annoy)'
s^j	*frilansit'* 'freelance,' *klasit'* 'give a high mark,' *baunsit'* 'bounce'
z^j	*rejzit'* 'raise,' *brauzit'* 'browse,' *relizit'* 'release'
t^j	*sabmitit'* 'submit,' *rebutit'* 'reboot,' *saportit'* 'support,' *kopirajtit'* 'copywrite,' *printit'* 'print,' *kajtit'* 'kite' (gaming) *resetit'* 'reset,' *kammentit'* 'comment,' *invajtit'* 'invite,' *čatit'-sja* 'chat'
st^j	*kopipastit'* 'copy-paste', *bustit'* 'boost'
d^j	*snoubordit'* 'snowboard,' *vardit'* 'ward,' *vikipedit'* 'use wikipedia' *aplodit'* 'upload,' *daunlodit'* 'download,' *trejdit'* 'trade,' *apgrejdit'* 'upgrade,' *rejdit'* 'raid,' *frendit'* 'friend,' *fludit'* 'flood with comments,' *zafrendit'* 'friend' PFV

Table 8: Recent borrowings with dental and labial stems used in the web searches.

them to be highly significant (see Table 9). This model fits the data better than the null model (likelihood ratio test: $\chi^2 = 39330$, df $= 3$, p $= 0$), but the overall fit is relatively poor due to high dispersion in the dental stems. The frequency effect and the interaction between frequency and type of stem are extremely small.

Notice a number of outliers in the labial stems, verbs with low interspeaker agreement. It turns out that almost all of them are f-final stems (see Figure 5, which breaks down proportion of alternations in labial stems by consonant). Interestingly, there appears to be a negative correlation between frequency and proportion of alternations in f-stems such that more frequent verbs are less likely to alternate. For example, the most frequent f-stem with more than 3,000 hits, *serf-it'* 'surf (the web),' was attested with alternations (as *serfl'-u*) only 35% of the time. Although this correlation is based on a very small sample (nine verbs), it is consistent with the data from the three more established f-stems included in the native verb searches: the more frequent verb *drejf-it'* 'be afraid of' was less likely to alternate than the other two less frequent f-stems, *razgraf-it'* 'make columns' and *proštraf-it'sja* 'get a ticket, to do something wrong'. Verbs with f-final stems also behaved differently from other labial stems in the first production experiment discussed in section 4.2. Speakers had significantly lower proportion of alternations in 1SG forms of f-stems (0.83) compared to other low-frequency labial stems (0.94) and significantly lower confidence in these alternations. Overall, these numbers place the f-stems somewhere in between defective and non-defective stems. This finding raises questions about the granularity level at which labial alternations are learned (see the next section for more discussion).

To sum up, the results for novel borrowings confirm the difference between labial and dental stems and show that dental borrowings have a high degree of variation or low degree of interspeaker agreement. I take this to be evidence for their status as defective. However, further empirical work is necessary to establish whether speakers are as uncomfortable producing 1SG forms of these verbs as they are of native defective verbs. For what it's worth, there are many metalinguistic discussions on the web about the 1SG forms of
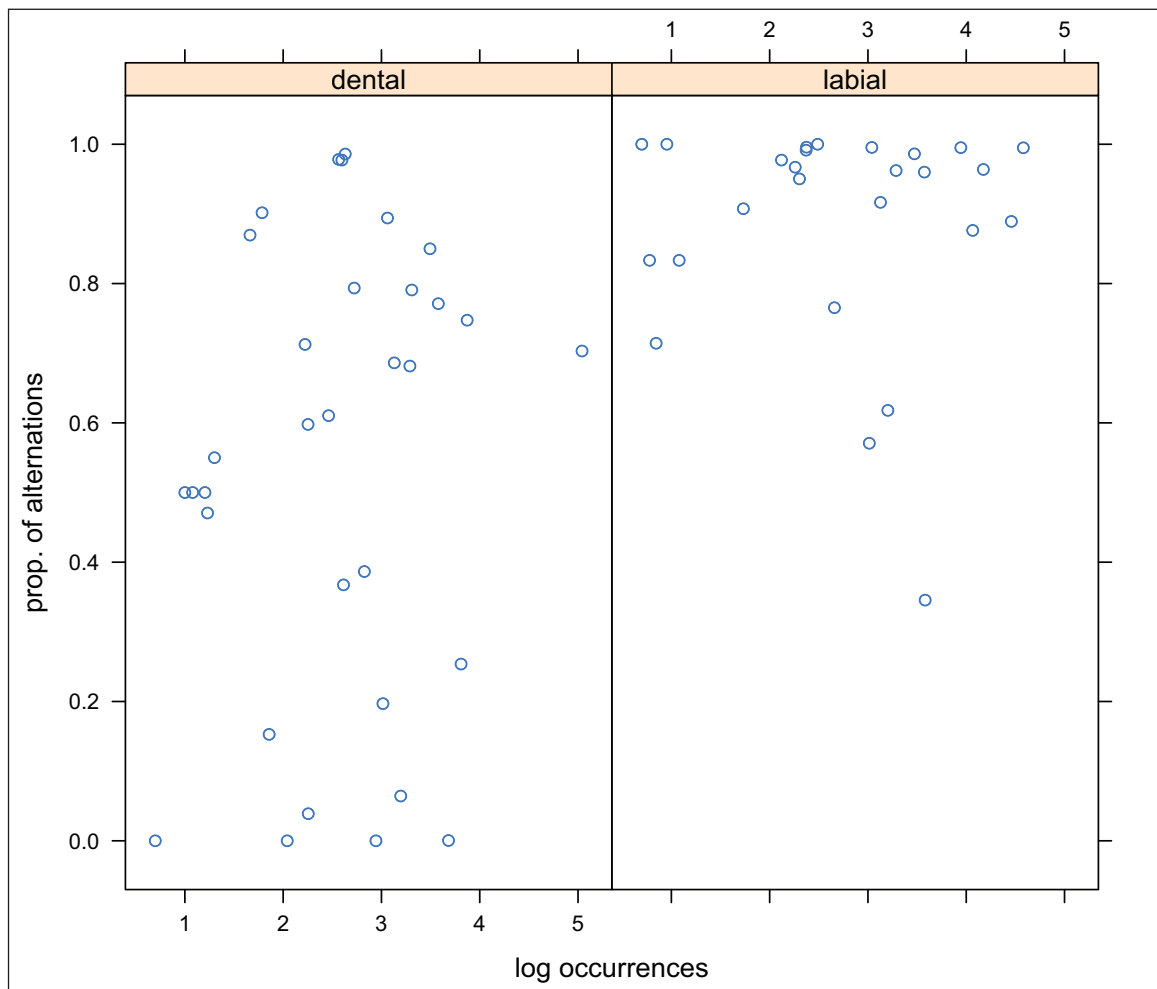
Figure 4: x-axis: log frequency of the total number of 1SG occurrences; y-axis: proportion of 1SG occurrences that contained the expected alternations.

borrowings. Speakers ask each other about the right way to say 'I paste,' 'I post,' 'I friend,' etc. Here is a particularly vivid example:

(7)   A discussion on *Wikislovar'* (Wiktionary) about *zafrendit'* 'friend':
        Contributor 1:   I think that in Russian *zafrendl'u* sounds better than *zafrend'u, zafrenžu.*
        Contributor 2:   But less natural. We probably shouldn't invent a "better sounding" solution. We should let *zafrend'u/zafrenžu* be the colloquial form, I don't know what conjugation number that would be. [. . .]
        Contributor 1:   You have talked me into it, but what is the passive participle *zafrendlennyj, zafrendennyj* or *zafrenžennyj?*

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−0.0039	0.0107	−0.37	0.7145
labial	−1.6184	0.0199	−81.36	0.0000
frequency	−0.0000	0.0000	−67.42	0.0000
labial:frequency	−0.0000	0.0000	−45.92	0.0000

Table 9: Coefficients of the logistic regression model predicting proportion of alternations in 1SG of borrowed verbs (log likelihood = −17716, df = 4).

5 Towards an analysis

The question I now turn to is how the facts discussed in the previous section help us explain the underlying cause of defectiveness in Russian verbs. There are several possibilities. One possibility suggested by a reviewer is that this defectiveness might be purely morphological. That is, if 1SG were derived from one of the other surface forms with the same alternation, the imperfective or the past passive participle forms, then in cases when such forms were lacking, the 1SG simply could not be derived. However, the 1SG present could not plausibly be derived from a past passive participle or a secondary imperfective, which itself is derived from the perfective stem. It suffices to note that some verbs have 1SG forms but lack secondary imperfectives or past passive participles. For example, the verb *spat'* 'sleep' is intransitive and, therefore, has no past passive participles. It also has no secondary imperfective forms because it is already imperfective, but it has a regular 1SG form with the expected labial alternation, *spl'-u* 'I sleep'.

The other two possibilities I consider here are as follows (1) defectiveness is simply a symptom of extreme lexical conservatism (refusal to produce novel allomorphs), and
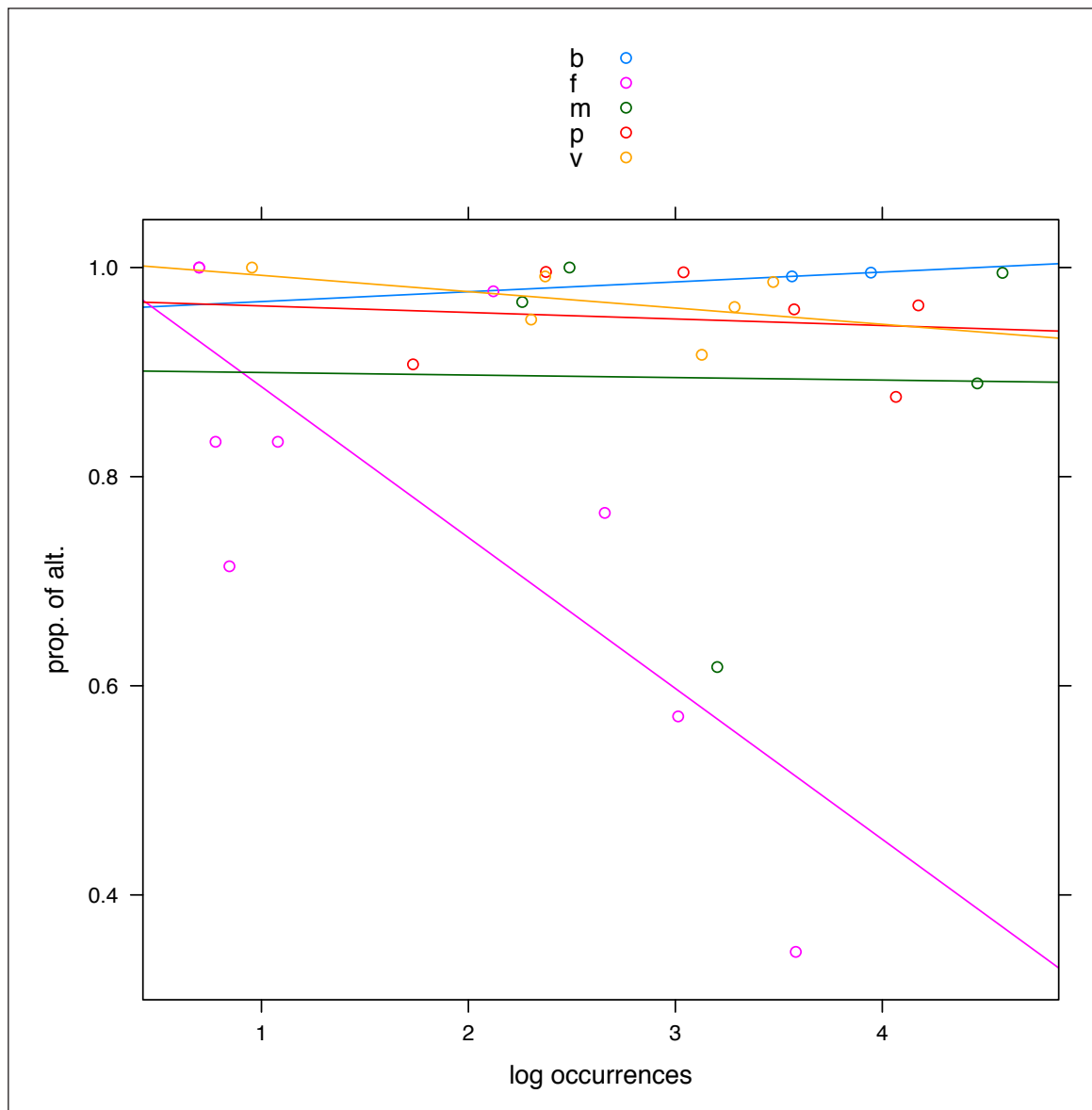
Figure 5: Proportion of alternating 1SG forms in labial borrowed verbs.

(2) several different factors together with lexical conservatism conspire to produce defectiveness. Assuming a constraint-based optimization model, the first possibility could be modeled using the filtering approach to defectiveness discussed in section 2.2, while the second possibility could be modeled with the threshold approach to defectiveness in HG discussed in section 2.4. In either case we have to allow constraints that can reference other allomorphs for a specific morpheme. To that end I use the following constraint from a family of LEXP constraints proposed in Steriade (2008):

(8)   IDENT$_{lex}$[αF]: For any segment s in a subconstituent C of an expression under evaluation, if s is [αF] then s has an [αF] correspondent in a listed allomorph of C.

The LEXP constraint above consults a set of listed forms in the input (not just a single base or UR) to look for a correspondent of some property in the output. If the property in question is found in *at least one of those forms,* the constraint is satisfied. "Listed" can be understood as frequent enough to be stored in the lexicon. Observe that this constraint presupposes existence of several different UR's or several different "bases" for a single morpheme. Alternatively, it could be formulated as a type of an Output-Output Correspondence constraint, referencing surface forms, assuming that they are lexically listed. For my purposes it won't be important to decide which of these possibilities is more appropriate.

Adopting a filtering approach can be done by placing the IDENT$_{lex}$[αF] in the Filter component of the grammar, e.g., CONTROL in the Orgun & Sprouse (1999) model. Thus, the regular OT grammar would produce outputs with the expected alternations in the 1SG, but those forms in which these alternations are never observed anywhere else would be filtered out in CONTROL. One advantage of this approach is that it affords a simple explanation of the complex facts presented in this paper and unifies them with another case of paradigm gaps in Russian, the gaps in the genitive plural of some feminine nouns (Pertsova 2005). These gaps also show a lexical conservatism effect. Namely they occur in nouns that are always stressed on the suffix but where stress is expected to shift to the stem in the genitive plural. Nouns which do not have any other stem-stressed allomorphs are the ones that are defective.

However, a problem with this approach is that there are no principled restrictions on what kind of constraints belong in the filter. If all LEXP constraints were in the filter, for example, this would amount to saying that Russian speakers are extremely lexically conservative and do not tolerate any alternations in a morpheme which they have not already seen with that alternation. This appears to be too strong of a prediction, since there are cases when speakers generalize morphophonological alternations to new stems as previously discussed. Thus, without a worked-out theory about what types of alternations are lexically conservative and what types are not, this approach to paradigm gaps remains flawed.

The other alternative is that the violation of the lexical conservatism constraint combined with violations of other constraints tips the scale towards defectiveness. This is the threshold approach to defectiveness discussed in section 2.4. Recall that I proposed that the Harmony scores of HG grammars be used to rank outputs on a well-formedness scale to account for both gradient grammaticality judgments and defectiveness. Namely, defectiveness could be understood as unwillingness or hesitation on the part of some speakers to produce outputs whose well-formedness scores are abnormally low. An obvious technical difficulty with testing this approach is that it requires computing the constraint weights of all constraints in a grammar which are needed to estimate the distribution of

Harmony scores over a large set of outputs. Since doing so is not feasible, my analysis below is given only as a "proof of concept" to make this idea more concrete and to illustrate the constraints that are involved in the computation of 1SG forms.

I assume that the morpho-phonological alternations involving dental consonants discussed in this paper are learned via language-specific constraints. In the tableau below, the constraint [sʲ → ş] stands for such a set of language-specific constraints supporting this particular morpho-phonological change in 1SG forms. Several proposals exist in the literature about how weights of language-specific constraints are determined during learning (Boersma 1997; Albright & Hayes 2006; Hayes & Wilson 2008; Allen & Becker 2015). It is commonly assumed that these weights must be proportional to the probability of observing constraint violations in the lexicon.

Recall that the only difference between a regular and a defective dental stem is that the former has an additional violation of the LEXP constraint. Thus, we must assume that this constraint has a significantly high weight, such that when this weight is added to the weights of other constraints violated by the alternating form, the cumulative Harmony score becomes too low enough for this output to be considered an outlier relative to the distribution of harmony scores over the whole language. The other constraints used in (9) are (1) the regular faithfulness constraint, IDENT(C), which must have a relatively high weight as well since consonant mutations are overall relatively rare, and (2) *[cʲu], which prohibits sequences of a palatalized consonant followed by [u]. Palatalized consonants before back vowels used to be illegal in Russian and are still very rare within native morphemes, in particular if C is a labial consonant.[21]

(9)   A hypothetical constraint weighting leading to gaps in dental stems assuming that the Harmony score of < –15 produces low confidence in outputs (gaps). The two shown verbs are the defective *pylesos-itʲ* 'vacuum' and the non-defective *zames-itʲ* 'mix (the dough).'

weight	10 $[s^j \rightarrow ş]$	6 $*[s^j u]$	8 $\text{ID}_{lex}[\alpha\text{PLACE}]$	7 IDENT(C)	H
defective dental [pɨlʲɪsósʲ-u]	–1	–1			–16
(☞) [pɨlʲɪsóş-u]			–1	–1	–15
regular dental [zamʲɪsʲ-ú]	–1	–1			–16
☞ [zamʲɪş-ú]			([zamʲɪş-atʲ])	–1	–7

Notice that according to this account, a dearth of s-final stems in the data would not necessarily lead to a gap. Given less evidence for [sʲ ~ ş] alternation, the constraints enforcing this alternation would have much lower weights, and the non-alternating form would become the winner. Kapatsinski (2010) describes a similar situation in which Russian speakers often fail to apply velar palatalization before -i (that is, prefer non-alternations)

---

[21] [Cʲu] sequences can occur at the morpheme boundary, before the 1SG suffix -/u/ and two other -/u/ suffixes in Russian, the accusative suffix of feminine nouns ending in -Cʲa and the dative suffix of masculine nouns ending in -Cʲ(Padgett 2010). As we have seen, palatalized labials do not occur before the 1SG suffix because of the labial alternations, and they are extremely rare before the nominal -/u/ suffixes because there are only two nouns with a labial stem in the relevant declensions (*golub'* 'pidgin' and *červ* 'worm' according to my search of Zaliznyak's paradigms.) On the other hand, it is much easier to find nouns with palatalized dental stems that appear before a nominal case suffix -/u/ (e.g., [dʲadʲ-u] 'uncle ACC.SG.', [stʲɪzʲ-u] 'road ACC.SG' (poet.)., [gusʲ-u] 'goose DAT.SG.', [katʲ-u] 'Katja ACC.SG.'), and hence the sequence [Cʲu] is more common when C is a dental rather than a labial consonant.

because there are relatively few examples in the lexicon of velar-final stems occurring before i-initial suffixes. On the other hand, velar palatalization almost never fails to apply in borrowings before the diminutive suffix -ok even though palatalization in this context is less phonologically natural, though more frequent in the lexicon. Thus, the degree of robustness of language-specific alternations is crucial for determining whether an alternation will apply, fail to apply, or lead to defectiveness.

The difference between the labial and dental stems can be attributed to the fact that labial stems do not violate the LexP constraint because labial alternation involves insertion at a morpheme boundary rather than changing morpheme internal segments.[22]

(10)   No gaps are predicted in labial stems: the verb *rubit'* 'chop'

weight	11 $b \rightarrow bl^j$	7 *[b^ju]	8 $\text{ID}_{lex}[\alpha\text{PLACE}]$	8 DEP	H
[rub^j-u]	−1	−1			−18
☞ [rub-l^j-u]				−1	−8

One could also argue that the LexP constraint is not violated by nonce-words because these words lack lexical representations. Thus, we would expect speakers to be less conservative when it comes to wug tests.

Violation of a LexP constraint alone will not necessarily lead to a gap under the threshold approach. Gaps are only expected when the alternating form also violates some other constraint with the effect that the cumulative weight of all violations pushes the well-formedness of the winner too far down. For instance, as mentioned earlier, vowel reduction in borrowings can also lead to the violation of a LexP constraint, but this violation alone will not lower the well-formedness of the alternating form in question below the threshold if the faithfulness for unstressed vowels is ranked low (which must be the case since unstressed vowels are reduced in Russian almost without exception). This approach would also work equally well for cases that do not involve neck-and-neck competition or violations of LexP constraints; the only requirement for a gap is that the most optimal form for some input has a very low relative well-formedness score.

Finally, what does this account predict for the past passive participles and the secondary imperfectives, the forms in which the same problematic alternations occur? The participles and other derivatives with alternations should not be problematic as long as the verb in question has an attested 1SG form, which will be the case for most verbs. Of course, this may not be true for verbs of very low frequency and for defective verbs, but as we have seen, defective verbs typically do not have participles or imperfective derivatives for semantic reasons. There is only a handful of defective verbs which could potentially have participles or secondary imperfectives. For example, verbs such as *uvjazit'* 'sink into,' *šerstit'* 'search vehemently,' and *prijutit'* 'shelter' are defective and could have past passive participles such as 'one who was sunk into,' 'one who was searched vehemently,' and 'one who was sheltered' respectively. Yet such participles are indicated in Zaliznyak (1980) as "difficult". Thus, the dental-palatal alternations in question are most likely problematic across the board and lead to some degree of ineffability in all cases unless an alternating form becomes lexicalized through frequent use. However, it's worth noting that past passive participles and imperfectives are different from 1SG forms in a way that can provide additional protection from gaps. Namely, the 1SG is the only form in the present tense sub-paradigm that is expected to alternate. On the other hand, both participles

---

[22] The candidate in which the epenthetic [l^j] is parsed as part of the stem would violate DEP_{lex} (C), but it would be harmonically bound by the candidate that doesn't parse [l^j] as belonging to any morpheme.

and secondary imperfectives have alternations in their stem throughout their respective paradigms (imperfectives in all verbal forms, and participles in all adjectival forms varying over gender, number, and case). Thus, if the Paradigm Uniformity (Benua 1997) constraint is also part of the grammar, it would be violated by the 1SG alternating forms, but not by the participle or imperfective forms, making them more well-formed than the 1SG forms. I will not, however, pursue this issue further here.

I leave it to future research to explore the feasibility of defining a relative well-formedness scale based on Harmony scores to account for both ineffability and gradient grammaticality, notions that appear related. This research would require a better understanding of the variability in Harmony scores within a language and an appropriate model for how the weights of language-specific constraints are computed. For instance, one issue with regard to the last point that needs to be resolved is the effect of pattern generality on constraint weights. In some models (e.g., the Maximum Entropy learner of Hayes & Wilson 2008 or the Minimum Generalization Learner of Albright & Hayes 2006) more general rules/patterns are preferred over more specific ones. However, in this paper I have discussed evidence suggesting that speakers may not be applying the general labial alternation rule to f-stems, which are very rare in the lexicon. Similar results are reported in an artificial language study by Linzen & Gallagher (2014) who found lower endorsement rates for novel unattested conforming stimuli over attested conforming stimuli; an unattested conforming stimulus is the stimulus that fits the general pattern being learned, but is not attested during training. This result seems to suggest that speakers rely more on specific, rather than general, constraints. However, Linzen & Gallagher (2014) also show that when speakers are exposed to just one token of each type of segment, they learn the general pattern and apply it in the same way to novel items with attested and unattested segments. Further work is required to figure out the balance between favoring general versus specific generalizations given different levels of exposure to the data.

6 Conclusion

To conclude, the gaps in the 1SG present tense of Russian verbs, thought to be arbitrary, are not arbitrary after all. I have presented evidence from lexical statistics and from speakers' productions of native and novel borrowed verbs to show that these gaps are connected to the 1SG morpho-phonological alternations. These synchronically opaque alternations appear in a few other forms outside the 1SG, most notably in past passive participles and in derived imperfectives. Verbs that have such related derivatives with the alternating variant of the stem do not have gaps in the 1SG. Additionally, 1SG forms of frequent verbs may be lexicalized, including lexicalization of a 1SG gap in a handful of frequent defective verbs. The existence of the transderivational relationships described above provide further evidence for the fact that phonological computations are not strictly local.

I have also considered several alternatives for formalizing defectiveness as a grammatical phenomenon. A promising approach, but one which requires further investigation, is to use the Harmony scores derived from constraint violations in Harmonic Grammar to rank all grammatical outputs on a well-formedness scale. This scale could be taken to reflect speakers' confidence in producing these outputs and in principle allows one to account for both gradient well-formedness and gradient ineffability. This approach distinguishes between grammaticality and well-formedness, with grammaticality being tied to optimization in a local competition and well-formedness being tied to the global distribution of Harmony scores over all grammatical outputs. Violation of many high-ranked constraints could lead to defectiveness on this view; so, defectiveness is not tied to violations of a specific set of constraints.

Supplementary Files

The supplementary files for this article can be found as follows:

Acknowledgements

Competing Interests

The author declares that she has no competing interests.

References

Albright, Adam. 2002. *The identification of bases in morphological paradigms*. Los Angeles, CA: UCLA dissertation.

Albright, Adam. 2003. A quantitative study of Spanish paradigm gaps. In Gina Garding & Mimu Tsujimura (eds.), *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL),* vol. 22, 1–14.

Albright, Adam. 2009a. Feature-based generalization as a source of gradient acceptability. *Phonology* 26(1). 9–41. DOI: http://dx.doi.org/10.1017/S0952675709001705

Albright, Adam. 2009b. A lexical and morphological condition of paradigm gaps. In *Modeling ungrammaticality in optimality theory* (Advances in Optimality Theory). London: Equinox Publishing.

Albright, Adam & Bruce Hayes. 2002. Modeling English past tense intuitions with minimal generalization. In Mike Maxwell (ed.), *Proceedings of the 2002 workshop on morphological learning,* 58–69. Association of Computational Linguistics Philadelphia: Association of Computational Linguistics. DOI: http://dx.doi.org/10.3115/1118647.1118654

Albright, Adam & Bruce Hayes. 2006. Modeling productivity with the gradual learning algorithm: the problem of accidentally exceptionless generalizations. In Gisbert Fanselow, Caroline Fery, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in grammar: Generative perspectives*, 185–204. Oxford: Oxford University Press.

Allen, Blake & Michael Becker. 2015. Learning alternations from surface forms with sublexical phonology. Lingbuzz/002503.

Andersen, Henning. 1980. Russian conjugation: acquisition and evolutive change. In Elizabeth Close Traugott, Rebecca Labrum & Susan Shepherd (eds.), *Papers from the 4th international conference on historical linguistics*, Series IV: Current Issues in Linguistic Theory 14, 285–301. Amsterdam: Amsterdam Studies in the Theory and History of Linguistic Science. DOI: http://dx.doi.org/10.1075/cilt.14.30and

Antilla, Arto. 2007. Variation and optionality. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 519–536. Cambridge: Cambridge University Press.

Baerman, Matthew. 2008. Historical observations on defectiveness: the first singular non-past. *Russian Linguistics* 32. 81–97. DOI: http://dx.doi.org/10.1007/s11185-007-9017-9

Bailyn, John & Andrew Nevins. 2008. Russian genitive plurals are impostors. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity*, 237–270. Oxford University Press.

Barnes, Jonathan. 2007. Phonetics and phonology in Russian unstressed vowel reduction: A study in hyperarticulation. Ms. Boston University.

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

Berent, Iris, Donca Steriade, Tracy Lennertz & Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104. 591–630. DOI: http://dx.doi.org/10.1016/j.cognition.2007.01.006; http://dx.doi.org/10.1016/j.cognition.2006.05.015

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The morphology and phonology of exponence,* vol. 41: Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

Bobablijik, Jonathan. 2008. Paradigms (optimal and otherwise): A case for scepticism. In Asaf Bachrach & Andrew I. Nevins (eds.), *Inflectional identity,* 29–54. Oxford: Oxford University Press.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. *IFA Proceedings* 21. 43–58.

Boersma, Paul. 2004. A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments. Ms. University of Amsterdam, ROA-648.

Boyé, Gilles & Patricia Cabredo Hofherr. 2010. Defectiveness as stem suppletion in french and spanish verbs. In Dunstan Brown, Matthew Baerman & Greville G. Corbett (ed.), *Defective paradigms: Missing forms and what they tell us.* Oxford University Press. DOI: http://dx.doi.org/10.5871/bacad/9780197264607.003.0003

Burzio, Luigi. 1998. Multiple correspondence. *Lingua* 103. 79–109. DOI: http://dx.doi.org/10.1016/S0024-3841(97)00025-9

Burzio, Luigi. 2005. Lexicon and grammar: unequal but inseparable. Ms. John Hopkins University.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Philadelphia: Benjamins. DOI: http://dx.doi.org/10.1075/tsl.9

Coetzee, Andries W. & Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in muna and arabic. *Natural Language and Linguistic Theory* 26(2). 289–337. DOI: http://dx.doi.org/10.1007/s11049-008-9039-z

Daland, Robert. 2015. Long words: the problem they pose for phonotactic grammars, and the solution for maxent harmonic grammars. *Phonology Phonology* 32(03). 353–383.

Daland, Robert, Andrea D. Sims & Janet B. Pierrehumbert. 2007. Much ado about nothing: A social network model of russian paradigmatic gaps. In *Proceedings of the 45th annual meeting of the association for computational linguistics,* 936–943. Prague, Czech Republic. http://groups.lis.illinois.edu/amag/langev/paper/daland07socialNetworkRussian.html.

Fanselow, Gisbert & Caroline Féry. 2002. Ineffability in grammar. In Gisbert Fanselow & Caroline Féry (eds.), *Resolving conflicts in grammar: Optimality Theory in syntax, morphology, and phonology* (Linguistische Berichte 11), 265–307. Hamburg: Helmut Buske Verlag.

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30(1). 79–133. DOI: http://dx.doi.org/10.1007/s11049-011-9142-4

Graudina Ljudmila K., Victor A. Itskovitch & Lija P. Katlinskaya. 1976. *Grammaticheskaja pravil'nost' russkoj rechi (opyt chastotno-stilisticheskogo slovar'a variantov).* Moskva: Nauka.

Halle, Morris. 1973. Prolegomena to a theory of word formation. *Linguistic Inquiry* 4. 3–16.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 27(7). 379–440. DOI: http://dx.doi.org/10.1162/ling.2008.39.3.379

Kapatsinski, Vsevolod. 2010. Velar palatalization in Russian and artificial grammar: Constraints on models of morphophonology. *Laboratory Phonology* 1. 361–393. DOI: http://dx.doi.org/10.1515/labphon.2010.019

Keller, Frank. 2006. Linear optimality theory as a model of gradience in grammar. In Gisbert Fanselow, Caroline Féry, Ralph Vogel & Matthias Schlesewsky (eds.), *Gradience in grammar: Generative perspectives*, 270–287. Oxford: Oxford University Press.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic grammar – a formal multi-level connectionist theory of linguistic wellformedness: An application. In *Proceedings of the twelfth annual conference of the Cognitive Science Society*, 388–395. Hillsdale, NJ.

Linzen, Tal & Jill Gallagher. 2014. The timecourse of generalization in phonotactic learning. *Proceedings of the Annual Meetings on Phonology* 1(1). 1–12. DOI: http://dx.doi.org/10.3765/amp.v1i1.18

Lyashevskaja, Olga & Sergej Sharov. 2009. Novyj chastotnyj slovar' russkoj leksiki [new frequency dictionary of the Russian lexicon]. Http://dict.ruslang.ru/freq.php.

MacWhinney, Brian. 1978. *The acquisition of morphophonology* (Monographs of the Society for Research in Child Development 43). Wiley. DOI: http://dx.doi.org/10.2307/1166047

Maiden, Martin & Paul O'Neill. 2010. On morphomic defectiveness: Evidence from the Romance languages of the Iberian peninsula. In Matthew Baerman, Greville G. Corbett & Dunstan Brown (eds.), *Defective paradigms: Missing forms and what they tell us* (Proceedings of the British Academy 163), 103–124. Oxford University Press.

McCarthy, John. 2002. *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.

Morin, Yves Charles. 1987. Remarques sur l'organisation de la flexion des verbes français. *ITL Review of Applied Linguistics* 77. 13–91.

Moskvin, Vadim 2015. Nedostatochnye glagoly: k utochneniju pon'atija (defective verbs: towards clarifying the term). *Russian Linguistics* 39(2). 209–232. DOI: http://dx.doi.org/10.1007/s11185-015-9146-5

Nesset, Tore. 2010. Suffix shift in Russian verbs: a case for markedness? *Russian Linguistics* 34. 123–138. DOI: http://dx.doi.org/10.1007/s11185-010-9057-4

Nevins, Andrew. 2014. Book review of "Defective Paradigms: Missing forms and what they tell us". *Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro* 10(1). 29–36.

Orgun, Cemil Orhan & Ronald Sprouse. 1999. From MParse to Control: deriving ungrammaticality. *Phonology* 16. 191–224. DOI: http://dx.doi.org/10.1017/S0952675799003747

Ozhegov, Sergey I. 1988. *Slovar' russkogo jazyka (the dictionary of the Russian language)*. Moskva: Russkij jazyk.

Padgett, Jaye. 2010. Russian /Cju/ and "perceptual" vs. "phonological" theories of borrowing: a reply to Paradis (and Thibeault). *Lingua* 120(5). 1233–1239. DOI: http://dx.doi.org/10.1016/j.lingua.2009.07.006

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

Pertsova, Katya. 2005. How lexical conservatism can lead to paradigm gaps. In Jeffrey Heinz, Andy Martin & Katya Pertsova (eds.), *UCLA working papers in phonology,* vol. 11.

Pertsova, Katya & Julia Kuznetsova. in press. Experimental evidence for lexical conservatism in Russian: Defective verbs revisited. In *Proceedings of FASL 24*. Michigan Slavic Publications.

Pierrehumbert, Janet. 2001. Stochastic phonology. *Glot International* 5(6). 195–207.

Pinker, Steven. 1999. *Words and rules (the ingredients of language)*. New York: Basic books.

Prince, Alan & Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. RuCCS-TR-2 ROA-537 Rutgers University.

Shvedova, N. 1980. *Russian grammar.* Academy of Sciences USSR, the Institute of the Russian Language. Moskva: Nauka.

Sims, Andrea. 2006. *Minding the gaps: inflectional defectiveness in a paradigmatic theory.* The Ohio State University dissertation.

Slioussar, Natalia & Maria Kholodilova. 2013. Paradigm leveling in non-standard Russian. In Alexander Podobryaev (ed.), *Formal approaches to Slavic linguistics: The second MIT meeting 2011.* Ann Arbor, MI: Michigan University Press.

Starostin, Sergei. 1998a. Zaliznyak's on-line dictionary. http://starling.rinet.ru/cgi-bin/main.cgi?flags = eygtmnl.

Steriade, Donca. 1997. Lexical conservatism. In *Linguistics in the morning calm,* Selected Papers from SICOL 1997, 157–179. Linguistic Society of Korea Hanshin Publishing House.

Steriade, Donca. 2008. A pseudo-cyclic effect in Romanian morpho-phonology. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity,* vol. 18. 313–360. Oxford studies in theoretical linguistics, Oxford University Press.

Yang, Charles, Kyle Gorman, Jennifer Preys & Margaret Borwczyk. 2012. Productivity and paradigmatic gaps. Talk presented at NELS 43.

Yanovich, Igor & Donca Steriade. 2010. Uniformity, subparadigm precedence and contrast derive stress patterns in Ukrainian nominal paradigms. Talk at OCP 6.

Zaliznyak, Andrej. 1980. *Grammatical dictionary of the Russian language.* Russkij jazyk.