Rethinking rule reliability: Why an exceptionless rule can fail

Vsevolod Kapatsinski
Indiana University

1 Introduction
One of the most puzzling phenomena in language change is the disappearance of previously productive processes. A morphophonological alternation may be productive at one point in time and lose productivity at a later point. In the historical record, the loss of productivity manifests itself as the accumulation of exceptions, where new words coming into the language are no longer required to obey a constraint that was previously exceptionless. The interesting questions, then, are: (1) how do the first exceptions appear, and (2) why does the number of exceptions continue to grow as time goes on?

One influence on productivity is the phonetic naturalness of the alternation. Several recent studies have suggested that phonetically natural alternations are easier to learn than phonetically unnatural ones (Finley & Badecker to appear, Wilson 2006). However, it is clear that an alternation may lose productivity while remaining the same in phonetic substance and that the same alternation may lose productivity in one language and remain productive in another. Thus, the loss of productivity must be conditioned by changes in how the alternation is exemplified by the lexicon.

Albright and Hayes (2003) proposed that at least some of the generalizations that a speaker of a language extracts from his/her experience with the language’s lexicon come in the form of rules. A rule is a generalization that specifies a mapping between an input (type) and an output (type) where the input and output types are classical categories, having necessary and sufficient conditions of membership.1 The likelihood that a rule will apply to a given input (i.e., the rule’s productivity with that input) is determined by the rule’s reliability relative to other rules competing for the same input. The reliability of a rule is taken to be the ratio of the number of input forms that obey the rule to the number of input forms that could obey the rule. For instance, the reliability of the past tense rule \[ \rightarrow -\text{ed}/\text{C}_- \] in English would be the number of bases that end in a consonant and form the past tense according to the rule divided by the number of bases that end in a consonant.

For the purposes of the present paper, the essential features of the model proposed by Albright & Hayes (2003) are that (1) input-output mappings compete for inputs, and (2) the outcome of the competition is determined by differences in reliability between the competing mappings. It is worth noting that competition

---

1The input and the output do not necessarily correspond to underlying and surface levels of representation respectively. In the present paper, the terms will be used simply to denote the directionality of the generalization, i.e., what is formed from what.
between input-output mappings is not assumed by all models of productivity. For instance, Plag (1999), following a long tradition in linguistics, argues that once the phonological and morphological characteristics of inputs that undergo a certain rule are specified, there is very little if any competition between rules. Bybee (2001:129) questions the reality of source-oriented generalizations.

The hypothesis of reliability-driven competition between input-output mappings predicts that a mapping will lose productivity as another mapping that competes for the same inputs and produces a different output gains in productivity. Rules can compete for the same inputs by virtue of overlap between their input categories. Suppose that we have a grammar containing only the rules C→Ci and k→ťʃi. These rules compete for inputs ending in /k/. However, inputs that end in other consonants can undergo only the more general rule (C→Ci). Thus if the language in question gains a large number of new input forms that end in a consonant other than /k/, the reliability of the general rule (C→Ci) will rise while the reliability of the specific rule (k→ťʃi) stays constant. Since the two rules compete for inputs ending in /k/, the increase in the productivity of the general rule (C→Ci) would result in the more specific rule (k→ťʃi) losing productivity. In the present paper, I provide support for this hypothesis using data from velar palatalization in Russian and artificial grammar learning.

2 Evidence from Russian

Russian morphophonology features a process of velar palatalization, whereby /k/ becomes [ťʃ], /g/ becomes [ʒ], and /x/ becomes [ʃ] in front of certain suffixes that begin or have historically begun with a front vowel, including the verbal stem extension –i and the diminutive suffixes –ok, and –ek/ik. If one examines a dictionary of modern Russian, even one that is designed to reflect recent borrowing (e.g., The big dictionary of youth slang, Levikova 2003), one would note that the velar palatalization process is exceptionless in front of all suffixes that trigger it. However, if one examines nonce borrowing, as seen in online communication of Russian Internet users, one notes that the process is more productive in front of some suffixes than in front of others. In particular, as we shall see later in this section, it is more productive in front of –ok and –ek than in front of –i and –ik. Since /i/ is a more natural trigger of palatalization than /e/ and /o/ (Wilson 2006), this pattern cannot be explained by differences in phonetic naturalness. The aim of the present section is to show that it is predicted by rule reliability.

In order to test the productivity of velar palatalization in front of –i, –ik, –ek, and –ok, I took all velar-final verbs and nouns found in the British National Corpus (http://corpus.byu.edu/bnc/). The resulting verbs were transliterated into Cyrillic. For each verb, possible Russian infinitival forms were derived. For instance, if the English verb is ‘lock’, some possible Russian infinitives are /loťʃitj/, /lokitj/, /lokatj/, /lokovatj/ and /lokisovatj/. Verbs for which an established Russian form already existed (e.g., format > /formatirovatʃ/) were excluded.
Existence was determined by the occurrence in either the Reverse Dictionary of Russian (Sheveleva 1974), Big Dictionary of Youth Slang (Levikova 2003), or the present author’s memory. This yielded 472 different verbs. For 56 of them, the final consonant of the English form was a velar, for 99 it was a labial, and for 317 it was a coronal. In the case of the nouns, all possible English monosyllables ending in /k/ or /g/ were created and transliterated into Russian manually. Then possible diminutive forms were created from them and submitted to Google. An additional sample of non-velar-final nouns was then created by matching the distribution of final consonant types in terms of manner and voicing and preceding vowels in the sample of velar-final nouns. Finally, to have a reasonably reliable estimate of the likelihood of failure of velar palatalization before –i for each verb, velar-final verbs and nouns that had 10 or fewer tokens with at least one of the palatalizing suffixes were excluded from the sample. This yielded 36 velar-final verbs and 19 velar-final nouns that could undergo velar palatalization and had a reasonably large number of tokens containing the relevant suffixes.

A Russian verb must have a stem extension, thus one must be attached to an extensionless English verb being borrowed into the language. However, –i is not the only stem extension competing for the English inputs. The stem extension –a, which does not trigger velar palatalization, is the other prominent option, being the second most productive stem extension (after –i) in nonce probe tasks (Kapatsinski 2005). Figure 1 shows that most velar-final inputs are highly unlikely to take –i, preferring to take –a instead, while most labial-final and coronal-final inputs are very likely to take –i (the difference between velar-final and non-velar-final inputs is highly significant, p<.0005 according to Monte Carlo resampling). Thus, the stem extension that triggers a stem change in the lexicon is disfavored by the stems that can undergo the change.

Figure 1: Histograms showing that most velar-final stems are unlikely to take –i while most labial-final and coronal-final stems tend to take –i.

The rate of failure of velar palatalization before the suffix –i is the number of tokens in which a /k/ or /g/ in the input corresponds to a /ki/ or a /gi/ in the output divided by the number of tokens in which the stem extension –i is used with an input ending in /k/ or /g/. Thus, for the input ‘lock’, the rate of failure of velar palatalization would be the frequency of the infinitive /lokiti/ divided by the sum
of the frequency of /lokitʃ/ and the frequency of /lotʃitʃ/. Velar palatalization fails before –i 44% of the time.

In the case of the diminutives, the suffixes –ek and –ok are favored by velar-final inputs while the suffix –ik is favored by non-velar-final inputs. The suffix –ik is chosen 26% of the time with velar-final inputs and 91% of the time with non-velar-final inputs. The suffix –ok is chosen 55% of the time with velar-final inputs and 7% of the time with non-velar-final ones. Finally, –ek is chosen 19% of the time with velar-final and 2% of the time with non-velar-final inputs, all p<.001 according to Monte Carlo resampling. Thus, the palatalizing suffixes –ek and –ok are favored by the inputs that can undergo velar palatalization whereas –ik is disfavored by such inputs. This pattern reflects the distribution in the lexicon, as depicted by dictionaries, where –ek is restricted to velar-final inputs, –ok is heavily favored by velars, and –ik almost never occurs with velar-final inputs.

Velar palatalization almost never fails before –ek and –ok (failure rate = 0.3%) while failing 40% of the time before –ik (p<.0005, paired-samples Wilcoxon signed ranks test). Thus, velar palatalization is likely to fail before –i and –ik, the suffixes that are disfavored by inputs that can undergo velar palatalization, but not before –ek and –ok, which are favored by inputs that can undergo velar palatalization.

3 The Rule-Based Learner: A simulation

The Rule-Based Learner (Albright & Hayes 2003) is a computational model of rule induction and weighting. The model starts with a set of morphologically related word pairs as in (1).

(1) mot motatʃ
    tʃmok tʃmokatʃ
    drug druʒitʃ
    krug kruʒitʃ
    golos golosovatʃ

For each word pair, the model creates a word-specific rule as in (2).

(2) []→a/mot_
    []→a/tʃmok_
    g→ʒi/dru_
    g→ʒi/kru_
    []→ova/golos_

Then, rules that involve the same change are combined. Contexts in which the same change, e.g., []→i, happens are compared by matching segments starting from the location of the change. If segments match, they are retained in the
specification of the context for the change and the pair of segments further away from the change is compared. When this comparison process reaches the nearest pair of segments that do not match, the phonological features they share are extracted and retained in the specification of the context. Segments that are further away from the location of the change than the closest pair of non-matching segments are not compared and are replaced by a free variable in the specification of context.

For instance, the rules in (3) are combined into the rule in (4). Since the change involves the end of the stem, comparison starts from the end. The last segments in the context are both /u/, so they are retained and preceding segments are compared. Since the preceding segments are both /r/, they are retained as well and comparison proceeds to the preceding segment. These segments do not match but they are the closest pair of segments to the change that doesn’t match, so the matching features are retained in the rule.

(3)  g→ʒi/dru_
     g→ʒi/kru_
(4)  g→ʒi/ [+cons; -cont; -son; -Labial]ru_

The resulting more general rules are then compared to each other and even more general rules derived if the same change occurs in multiple contexts, eventually resulting in quite general rules, such as []→i/C_. However, all rules are retained in the grammar. Instead of removing non-maximally-general rules from the grammar, the Rule-Based Learner weights each rule by its reliability. A reliable rule is more likely to apply to a novel word than a less reliable rule. For instance, if Vg→Vʒi is more reliable than Vg→Vga, and these are the only rules that can apply to the novel verb /dig/, the verb should be more likely to be borrowed as /diʒi/ than as /diga/.

The Rule-Based Learner requires a lexicon of pairs of morphologically related words, on which the model is trained, and a test set, on which it is tested. The test sets consisted of the borrowed words analyzed above. To estimate the probability of a given verb undergoing velar palatalization given that a particular suffix is chosen the reliability of the most reliable rule that requires palatalization is divided by the sum of its reliability and the reliability of the most reliable applicable rule that does not require palatalization but still attaches the same suffix. For instance, suppose the verb is /dig/ and the model has extracted the rules in (5) with reliability estimates shown in parentheses. The only rules that can apply to /dig/ are (a), (c), (d), and (f). Of these, the only rules that require velar palatalization are rules (c) and (d). Rule (c) is more reliable than rule (d), so it would get to apply. Its reliability is .272. The rule that attaches –i without palatalizing the stem-final /g/ is rule (f). Its reliability is .232. Therefore, the predicted probability that the final consonant of /dig/ will be palatalized, given that –i is selected as the stem extension, is \(0.272 / (0.272 + 0.232) = 54\%\) (cf. Albright & Hayes 2003:128).
The training lexicon consisted of stem-verb and noun-diminutive pairings found in the *Reverse dictionary of Russian* (Sheveleva 1974) and/or the *Big dictionary of youth slang* (Levikova 2003). The main results presented below held regardless of whether the *Reverse dictionary*, the *Slang dictionary*, or both were used. No suffixes were excluded from the training set. Thus, aside from verbs featuring the highly productive –i and –a, verbs having –ova, –irova, –izirova, and –e were also included. The full training set for the model of stem extension choice consisted of 2396 verb-stem pairs. The model of diminutive formation was trained on a set of 1154 diminutive nouns.

The learner models competition between input-output mappings. Therefore it is crucial to define what is meant by the input and the output. For the present paper, we are interested in modeling competition between input-output mappings in which some mappings require velar palatalization. The input form for these mappings may or may not have the stem extension already specified. The output of the competition can either be a phonetic form, specifying the allophone of /i/ ([ɨ] or [ɨ]) used, or a phonemic form, which does not include this specification. Finally, the diminutive suffixes –ek and –ik can be considered allomorphs since –ek is restricted to velar-final inputs and –ik is largely restricted to non-velar-final ones (with a handful of exceptions in which –ik can be attached to a velar-final input according to the dictionary but –ok is preferable). Therefore, I tested a version of the model in which the –ik/ek choice is made after the choice on whether to palatalize the stem velar and a version in which the choice of the suffix allomorph preceded the choice of the stem shape. All versions of the model behaved in the same way with regard to the hypothesis tested in the present paper.

All versions of the model were able to predict that –i is less productive with velar-final stems than with coronal-final and labial-final stems, predicting –i around 20% of the time with velar-final inputs, 50% of the time with labial-final inputs and 65% of the time with coronal-final inputs. The model also correctly predicted that –ek and –ok are favored by velars while –ik is disfavored by them, predicting –ik 79% of the time with non-velar-final inputs and only 30% of the time with velar-final ones. The suffix –ok was predicted to be used 20% of the time with non-velar-final inputs and over 40% of the time with velar-final ones. Finally, –ek was predicted to be used only with velar-final inputs, accounting for 20% of outputs derived from velar-final inputs. Most importantly, the model correctly predicted that velar palatalization is more likely to fail before –i and –ik than before –ek or –ok. In the model, velar palatalization never failed before –ek, always failed before –ik, failed about half the time before –i and only around 20% of the time before –ok.
For both /k/-final and /g/-final verbs, the model extracts only one applicable rule that favors adding –i and leaving the final consonant of the stem unchanged. For /g/-final roots, this is the rule C_{[+voiced]} → C_{[+voiced]i} and for /k/-final roots this is the rule C → Ci. Thus, in order for the more specific rules requiring /k/ to change into /ʧ/ or /g/ to change into /ʒ/ to fail, they must lose to a very general rule. In the Russian lexicon used to train the model, coronal-final and labial-final stems tend to take –i while velar-final stems tend to take –a. Since most stems in the lexicon end up taking –i, the model extracts the very general rule C → Ci and C_{[+voiced]} → C_{[+voiced]i}, assigning them a moderate reliability. On the other hand, the fact that velar-final stems disfavor –i drives down the reliabilities of rules that add –i to velar-final stems. This includes the rules that add –i and change the root-final consonant. As a result, these rules will sometimes lose the competition for application to the more general rules C → Ci and C_{[+voiced]} → C_{[+voiced]i}. Thus, the model predicts that velar palatalization will often fail before an affix if and only if the affix is more productive after non-velars than after velars. This holds for the stem extension –i and the diminutive suffix –ik but not for the diminutive suffixes –ek and –ok. Therefore, the model correctly predicts that velar palatalization should fail often before –i and –ik and rarely before –ek and –ok. This prediction follows directly from the hypothesis that input-output mappings compete, with the outcome determined by reliability.

4 Evidence from artificial grammar learning

The data from Russian strongly suggest that the productivity of velar palatalization is connected to whether the palatalizing affix is used mostly with inputs that can undergo velar palatalization or with inputs that cannot. However, the data are ambiguous with regard to the directionality of this effect. It is possible that the low productivity of velar palatalization before –ik and –i, whatever its cause, makes speakers of Russian avoid using –i and –ik with velar-final inputs. If changing an input consonant in front of a certain suffix is difficult while keeping the consonant unchanged results in a suboptimal output, the best course of action may be to avoid using the suffix altogether.

Another difficulty with the Russian data is that the dictionary is not a perfect model of the Russian lexicon as it exists in the mind of a Russian speaker. Therefore, the data on whose basis velar palatalization is acquired by the model are different from the data on whose basis velar palatalization is acquired by Russian speakers. A way to address both of these issues is provided by the paradigm of artificial grammar learning. By training the subjects and the model on the same language featuring velar palatalization, we can maximize the similarities between their relevant learning experiences. Furthermore, by varying the lexical distribution of the palatalizing suffix and keeping all other aspects of the competing rules constant, we can determine whether the distribution of the palatalizing suffix can influence the productivity of palatalization.
Native English speakers were randomly assigned to two groups. Both groups were presented with an artificial language featuring two plural suffixes, –a and –i, and an exceptionless rule that palatalized velars before –i, turning /k/ into [tʃ] and /g/ into [dʒ]. Since the same input-output mappings are used in both languages, phonetic naturalness is controlled. In both languages, velar-final singulars always corresponded to plurals ending in –tʃi or –dʒi. Both subject groups were presented with 30 singular-plural pairs in which the singular ended in a velar. Therefore, the palatalizing rule has the same type (and token) frequency in both languages. The difference between the two languages was that in Language I –i was not very productive with non-velar-final singulars, being used in only 25% of the cases with –a being used 75% of the time. In Language II, the rates were reversed: –a was used 25% of the time with non-velar-final bases while –i was used 75% of the time. In both cases, 40 non-velar-final bases were used. Just like velar-final bases, the non-velar-final bases ended in oral stops (/p/, /b/, /t/, and /d/). Non-velar consonants were always constant across the singular and the plural.

The experiment consisted of a training stage and a testing stage. In the training stage, subjects repeated singular-plural pairs presented to them auditorily over headphones. The aural presentations of the words were accompanied by visual presentations of the referents on a computer screen. The participants’ productions were used to assess whether the training stimuli were perceived correctly. If a participant made perception errors on more than 5% of singular-plural pairs, s/he was excluded from the experiment. In the testing stage, subjects were presented with novel singular forms (not presented during training) and asked to orally produce the plural. Participants who used –i fewer than 10 times with velar-final singulars were excluded from the present analyses. Thus the analyses below are based on thirty participants, half of whom were exposed to each language.

The only rules that are applicable for novel velar-final bases and are extracted by the Rule-Based Learner upon exposure to the two languages are presented in Table 1. As the table shows, the two languages differ only in the reliabilities of the rules that do not require velar palatalization. The rule attaching –i without changing the preceding consonant is much more reliable in Language II than in Language I. Therefore, velar palatalization is predicted to fail before –i in Language II more often than in Language I.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Language I</th>
<th>Language II</th>
</tr>
</thead>
<tbody>
<tr>
<td>k → tʃi/V_</td>
<td></td>
<td>0.85</td>
</tr>
<tr>
<td>g → dʒi/V_</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] → i/C_</td>
<td>0.18</td>
<td>0.57</td>
</tr>
<tr>
<td>[ ] → a/C_</td>
<td>0.57</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 1: Rules and the corresponding reliability values extracted by the Rule-Based Learner when it is exposed to the same artificial languages presented to human participants.
Figure 2 shows that –i was much more likely to be used with singulars ending in a non-velar consonant if the subjects were exposed to Language II than to Language I (t(28)=4.4, p<.001). The frequency of use of –i with non-velar-final inputs in the two languages closely matches frequency in the input (30% vs. 25% for Language I, 67% vs. 75% for Language II). This result shows that subjects were sensitive to the frequency of –i vs. –a after non-velars and therefore that this variable can potentially influence the productivity of velar palatalization.

![Figure 2: Subjects exposed to Language II are more likely to use –i to form the plural than subjects exposed to Language I.](image1)

Figure 3 shows that, as predicted by the model, velar palatalization was less likely to be acquired by subjects exposed to Language II, the language in which the palatalizing suffix, –i, is more productive with singulars that cannot undergo velar palatalization (t(28)=2.316, p<.05).

![Figure 3: Subjects exposed to Language II are less likely to palatalize the velar before –i than subjects exposed to Language I.](image2)

Figure 3 shows that, as predicted by the model, velar palatalization was less likely to be acquired by subjects exposed to Language II, the language in which the palatalizing suffix, –i, is more productive with singulars that cannot undergo velar palatalization (t(28)=2.316, p<.05).

Like speakers of Russian, subjects exposed to the artificial languages do not simply match the rate of velar palatalization to which they are exposed (100% for
all subjects, regardless of whether they were exposed to Language 1 or Language 2). Rather, learners appear to be sensitive to the reliability of the ‘just add –i’ rule relative to the palatalizing rule. Figure 4 shows that there is a strong and significant negative correlation ($r = -0.67, p < 0.001$) between how much a subject uses –i with non-velar-final inputs and how likely s/he is to palatalize a velar before –i.

![Figure 4: Subjects for whom –i is productive with inputs that cannot undergo velar palatalization are the subjects for whom velar palatalization is unproductive. Curves show the 95% confidence interval for the regression line.](image)

Once the correlation between rate of velar palatalization exhibited by a subject and his/her rate of –i use with non-velar-final inputs (Figure 4) is taken into account, the difference between subject groups no longer contributes towards explaining between-subject differences in velar palatalization productivity (according to an ANCOVA with rate of velar palatalization as the independent variable, the rate of –i use with non-velar-final inputs as a covariate, and Language as a fixed factor, rate of –i use is significant, $F(1,27) = 14.23, p < 0.001$, while Language is not, $F(1,27) = 0.82, p > 0.5$). Thus the difference in productivity of –i with non-velar-final inputs accounts for all differences in productivity of velar palatalization between subjects that can be attributed to the artificial language they are exposed to.

The present experiment manipulated the productivity of –i with non-velar-final inputs and assessed the influence of this manipulation on the productivity of velar palatalization before –i. As predicted by the hypothesis of reliability-driven competition between input-output mappings, instantiated by the Rule-Based Learner, the more the palatalizing suffix is used with inputs that are not subject to palatalization, the more it loses its palatalizing properties.
5 Discussion

The present results provide support for the hypothesis that alternative source-oriented generalizations compete for inputs and that the outcome of this competition is determined by differences in reliability between the competing generalizations. Bybee (2001:129) questions the reality of source-oriented generalizations: since ‘any morphological pattern that can be described by a source-oriented rule can also be described by a product-oriented one’, which would specify only the form of the output, rather than an input-output mapping, ‘we still await conclusive evidence for source-oriented schemas’, which would include rules.

In both studies reported in the present paper, palatalization was largely restricted to velar-final inputs. Pierrehumbert (2006) has argued that product-oriented generalizations cannot account for cases like these, in which an output can be derived only from a restricted class of inputs, since they place no restrictions on the shape of the input. However, if product-oriented generalizations are combined with faithfulness/paradigm-uniformity constraints (for possible implementations of paradigm uniformity see Albright 2005, Kenstowicz 1996, McCarthy 2005), such restrictions can in fact be derived.

Let us consider the difference between Language I and Language II. Both languages feature the same number of singular-plural pairs obeying the k→tʃi rule. Thus, both languages also have the same number of outputs ending in /tʃi/. Therefore, reliability of a simple product-oriented generalization like ‘plurals must end in /tʃi/’ cannot account for the fact that velar palatalization is more productive in Language I than in Language II. However, reliability of a conditional product-oriented generalization, like ‘if a plural ends in /i/, the preceding consonant must be /tʃ/’, can account for the present data. The reliability of the product-oriented generalization above would be the number of plurals ending in /tʃi/ divided by the number of plurals ending in /i/. While the numerator of this expression is the same in Language I and Language II, the denominator is much larger in Language II, where –i is productive after non-velars. Thus, the product-oriented generalization requiring an alveopalatal in the output is more reliable in Language I than in Language II.

On its own, the product-oriented generalization does not predict that only velars will become alveopalatals. However, suppose that the generalization competes with constrains like ‘a velar in the singular corresponds to a velar in the plural’ and ‘place values of the singular and the plural must be the same’. If the constraints against changing the consonant are stronger for non-velars than for velars, then only the velars will change to satisfy the product-oriented generalization. Thus the present results are also predicted by a purely product-oriented model with reliability-weighted conditional product-oriented constraints and paradigm uniformity constraints.

The present results are also consistent with the hypothesis that there is competition between source-oriented generalizations but it is resolved by differences in type frequency, rather than reliability (Bybee 1985, 2001). The
present study does not allow us to distinguish between the two statistics if the competing generalizations are assumed to be source-oriented. Under the product-oriented story, the relevant statistic must be reliability, since the generalization ‘if the plural ends in –i, the preceding consonant must be an alveopalatal’ has the same number of supporting examples, featuring an alveopalatal followed by –i, in both artificial languages.

Whether the relevant statistic is reliability or type frequency and whether the reliability-weighted generalizations are source- or product-oriented, the present results require the competition between generalizations to be resolved according to a stochastic, probability-matching decision rule (Luce 1959). Thus, the present model requires the speakers not to apply the most reliable input-output mapping available 100% of the time. If the most reliable mapping were applied 100% of the time, there would be no difference between Language I and Language II, since input-output mappings requiring velar palatalization are more reliable than those simply attaching –i in both languages (the same would also hold for the alveopalatal-favoring constraint vs. the paradigm uniformity constraints applying to velar-final singulars). Furthermore, probability-matching decisions are also required for the Rule-Based Learner to exhibit prototype effects. Using the same choice rule as in the present study, Albright & Hayes (2003) show that the Rule-Based Learner successfully accounts for the finding that novel verbs that are similar to many regular English verbs are more likely to take the regular past tense than novel verbs that are similar to neither regular nor irregular English verbs. The regular past tense is the preferred output for both classes of inputs. If the most reliable mapping always won out, the reliabilities of dispreferred mappings would be irrelevant, hence no difference between the two classes of inputs would be expected.

Probabilistic choice does not maximize the probability of producing the correct output. In part because of the probabilistic choice rule, the model of grammar instantiated by the Rule-Based Learner does not produce the correct output as often as humans do if tested on the ‘training lexicon’, the lexicon of the subject’s native language. Thus, as Albright & Hayes (2003:154) note, the Rule-Based Learner is “solely… a model of morphological productivity, and not a model of how existing words are stored and produced”. In order to account for the fact that known words are very likely to be produced correctly, it must be assumed that they are simply retrieved for production, rather than being derived online using the input-output mappings of the grammar. Lexical storage and retrieval during production is also necessary to account for the fact that an input-output mapping can become unproductive while words instantiating the mapping are retained in the lexicon and produced correctly (Bybee 1985, 2001; Zuraw 2000). In other words, a morphophonemic rule can lose productivity while remaining exceptionless only if whole-word storage and retrieval are assumed.

There are two types of individual variability evident in the data in Figure 4. First, it is possible to be exposed to Language II and behave like a subject who was exposed to Language I and vice versa. I have argued that languages vary
along a continuum between a state in which a morphophonological rule triggered by an affix is fully productive and the affix is used only with stems that can undergo the rule and a state in which the suffix is used only with stems that couldn’t undergo the rule even if it were productive and the rule is consequently unproductive. If exposure to a language necessarily caused correct acquisition of that language, a language would occupy a fixed point on the proposed continuum forever. It is imperfect acquisition that allows a language to move from one point of the continuum to another (occupying different points on the diagonal stretching from upper left to lower right in Figure 4).

However, not all learners line up perfectly along the diagonal. In particular, some subjects in the present experiment failed to acquire velar palatalization while also being unlikely to use the palatalizing suffix –i with non-velar-final inputs. These are the subjects on the lower left in Figure 4. While this learning outcome is relatively unlikely, as shown by the strong negative correlation in Figure 4, it is possible and appears to be more likely than a situation in which –i is used extensively and velar palatalization is productive. There are no points on the upper right in Figure 4 and the rate of velar palatalization in both of the artificial languages is substantially lower than is predicted by the model (83% predicted vs. 67% observed for Language I, 60% predicted vs. 38% observed for Language II). Unlike the model, the human learners appear to be biased against accepting the fact that singular and plural stem shapes are different. That is, learners appear to come to the experiment with a high weight on paradigm uniformity. Thus languages constituting exceptions to the proposed connection between high productivity of a rule-triggering affix with stems that cannot undergo the rule and low productivity of the rule are possible but are predicted to be less frequent than would be expected by chance and may be especially likely to satisfy a preference for fixed stem shapes.²

Finally, exceptions to the proposed connection may happen when the to-be-acquired rule is phonetically unnatural. The Russian data suggest that rule reliability overrides naturalness of the trigger: while /i/ is a more natural trigger of palatalization than /e/ or /o/ (Wilson 2006), stem extensions and masculine diminutive suffixes beginning with /i/ are least likely to trigger palatalization in Russian. However, rule reliability may not override naturalness of the change itself. In the data from loanword adaptation on the web, velar palatalization in Russian tends to fail with /g/, which changes to [ʒ], more often than it fails with /k/, which changes to [tʃ], a fact that is not predicted reliably by the Rule-Based Learner. The /k/→[tʃ] change is more phonetically natural than the /g/→[ʒ] change, involving more confusable segments (Guion 1998). Thus, the low

²A necessary caveat is that this argument relies on the assumption that paradigm uniformity is universal. The data of the artificial grammar learning experiment reported here can also be accounted for if the relevant assumption of paradigm uniformity arises as a result of the subjects’ exposure to English. In either case, the relevant paradigm uniformity constraint appears to be feature-specific, since Bybee & Newman (1995) find no preference for affixes over vowel changes in the stem as plural morphemes in artificial grammar learning with English participants.
productivity of /g/→[3] compared to /k/→[tʃ] is expected if learning a phonetically natural alternation is easier than learning a phonetically unnatural one (Finley & Badecker to appear, Wilson 2006). Finally, since rule reliability is quantitative in nature, it may fail to override trigger naturalness when differences in the reliabilitys of palatalizing rules before different triggers are not as great as they are in Russian.

6 Conclusion

We started out with the hypothesis that source-oriented generalizations, i.e., input-output mappings, compete for inputs with the outcome of this competition determined by differences in reliability between the competing mappings. This hypothesis, implemented within a computational model of rule induction and weighting (the Rule-Based Learner, Albright & Hayes 2003) was shown to produce a prediction for language change. It is predicted that the more often a rule-triggering suffix is used with stems that cannot undergo the rule, the less likely the rule is to apply to stems that can undergo it. This prediction was confirmed with both natural loanword adaptation in Russian, where rule reliability (or rule type frequency) was shown to override phonetic naturalness of the rule’s trigger, and artificial language learning, which allowed us to disambiguate the direction of causation. These data strongly suggest that, as predicted by the hypothesis of reliability-driven competition between rules and a product-oriented version in which conditional product-oriented generalizations compete with paradigm uniformity constraints in a reliability-driven manner, a morphophonemic rule can lose productivity if the affix(es) that trigger the rule come to be used mostly with stems that do not contain segments that could undergo the rule. If morphologically complex words are retrieved from the lexicon as wholes, an assumption required by the Rule-Based Learner, this process of productivity loss can occur while no lexical exceptions appear in the lexicon. Thus a morphophonemic rule can lose productivity while remaining exceptionless.

References


