Rule Reliability and Productivity: 
Velar Palatalization in Russian and Artificial Grammar 

Vsevolod Kapatsinski 
Department of Linguistics and Cognitive Science Program, Indiana University, USA 
vkapatsi@indiana.edu 

Russian velar palatalization (k→tʃ, g→ʒ, x→ʃ) presents an intriguing case of a morphophonological alternation that is exceptionless for a given suffix, yet not fully productive in natural loanword adaptation with some of the suffixes. In particular, while the process is shown to be only partially productive before –i and –ik but fully productive before –ok and –ek, contrary to the naturalness of the suffixes as triggers of palatalization. A model of rule induction and weighting (the Rule-based Learner, developed by Albright and Hayes 2003) is trained on the lexicon of Russian verbs, in which velar palatalization is exceptionless (always applying before the stem extension –i), and tested on new borrowings and nonce probes from elicited production tasks. Despite the fact that velar palatalization is exceptionless in the training set, it is correctly predicted to often fail with novel words based on information in the lexicon.

When a foreign verb is borrowed into Russian, it must be assigned a stem extension, e.g., upload → /ʌploʌd+i+tʃ/, duck → /dʌk+i+tʃ/, lock → /lotʃ+i+tʃ/. The two most productive stem extensions in modern Russian are –i and –a. While –i always triggers velar palatalization in the native lexicon, -a does not, e.g., duck → /dʌk+a+tʃ/, lock → /lotʃ+i+tʃ/. Importantly, velar-final roots favor –a while –i is favored elsewhere.

When exposed to the native lexicon, the Rule-based Learner discovers the following crucial rules: C→C+a, C→C+i, k→tʃ+i, g→ʒ+i. In this model, the likelihood that a rule will apply is determined by its reliability relative to competing rules. Reliability is defined as the number of words to which the rule applies divided by the number of words to which it could apply. For instance, the reliability of the rule k→tʃ+i is the number of k-final roots that take –i and change the final consonant to the total number of k-final roots in the lexicon. Since velars favor –a over –i while most consonants favor –i over -a, the rules k→tʃ+i and g→ʒ+i, which involve velar palatalization, are not very reliable relative to the rule C→C+i, which leaves the consonant unchanged. Thus, velar palatalization is correctly predicted to fail before -i. On the other hand, the diminutive suffixes –ek and –ok usually attach to velar-final inputs, hence rules stipulating velar palatalization before these suffixes are much more reliable than the rules that stipulate that the suffix can be attached with no change to the preceding consonant.

These predictions hold regardless of whether the stem change and the suffix are chosen during a single decision stage or the choice of the suffix precedes the decision on whether to change the stem. However, only the former (single-stage) model is able to account for segmental context effects on palatalization. As Figure 1 shows, the single-stage model successfully predicts that velar palatalization should fail more often when the stem ends in a consonant cluster than when it ends in a VC sequence, and more if it contains a front vowel than if it contains a back one. Importantly, these predictions are made only if speakers choose the suffix and the stem change at the same time. If the decision to apply the stem change is made after the affix is chosen (Stage I: choose –i, Stage II: choose whether to palatalize), the model predicts no influence of penultimate segment identity. Thus, I argue that the suffix should not be seen as a ‘trigger’ of velar palatalization but rather that both the suffix and the stem change are markers of the function typically associated with the suffix and are chosen simultaneously.
The importance of rule reliability for productivity is confirmed by an artificial grammar learning experiment, in which native English speakers were exposed to one of two languages. Both languages contained two plural suffixes, -i and –a. In both languages, 30 words exhibited velar palatalization, which was exceptionless before –i and never occurred before –a. The same total number of training trials was presented. However, in Language I, -i occurred mostly with velar-final inputs whereas in Language II non-velar-final inputs also took –i most of the time. Thus, the languages differed only in the reliability of the no-change rule C → Ci, which was higher in Language II than in Language I. As expected, velar palatalization was much more likely to fail before –i in Language II than in Language I (63% vs. 29% respectively).

Rule reliability does not provide a complete account of productivity. Thus, in the Russian loanword adaptation data, it was observed that velar palatalization before –i was blocked if its application would produce a homonym with an existing Russian verb. This base recoverability effect is not predicted by the Rule-Based Learner. In the artificial-grammar learning experiment, subjects exposed to either language learned that singulars ending in /tʃi/ or /dʒi/ take –i, despite having seen no examples in which this happens (singulars ending in /tʃi/ or /dʒi/ were not presented during training). Evidently, subjects generalized over the plurals, which featured /tʃi/ and /dʒi/ but not /tʃa/ and /dʒa/. Since the Rule-Based Learner can only generalize over input-output pairings, it does not predict this product-oriented effect. Thus, I conclude that rule reliability is an important influence on productivity but not the only one, interacting with base recoverability/homonymy avoidance and reliability of product-oriented generalizations.

References