Regularity is overrated:
Stochastic competition in grammar and the primacy of the lexicon

Vsevolod Kapatsinski
University of Oregon

Linguistic theory draws a distinction between regular and irregular systems of generalizations. Regular systems are to be preferred over irregular ones by the analyst, and thus presumably are also preferred by the learners. In rule-based theory, a regular system is one in which rules predicting different outputs do not compete with each other for application to any (class of) inputs. Thus, if rule A states that an input X corresponds to an output Y, there can be no rule B stating that X corresponds to a different output, Z. If the learner of a language has a preference to learn regular rule systems, s/he should minimize competition between rules in the grammar s/he induces from linguistic data (Plag 2003), the way a linguist does when analyzing a dataset. In Optimality Theory, a regular system is one that can be described by a single constraint ranking where the choice regarding which constraint to obey is deterministic, rather than stochastic. Again, other things being equal, strict rankings are preferred as analytic solutions and thus are (implicitly) suggested to be preferred by learners.

We present a case in which learners of a language do not appear to be minimizing competition between generalizations. Table 1 shows an artificial language with two plural suffixes, -i and –a, where –i always palatalizes the preceding velar. A competition-minimizing learner would be expected to learn that [k] becomes [tʃ] before –i while other consonants do not change when –i is attached. This could be achieved by either extracting the rules shown on the left in Table 1 or a constraint ranking like *ki >> Ident-Velar. If this were the case, velar palatalization would be equally productive in Language 1 and Language 2 because the palatalizing rule or constraint ranking has no competitors that can produce errors. Saying that rule/constraint application is noisy does not change this prediction, since the amount of noise would be presumably the same in both languages.

On the other hand, a learner that does not attempt to minimize competition could extract a generalization like Ident-Place or $C_{[\alpha place]} \rightarrow C_{[\alpha place]}i$. The resulting generalization would then compete with the palatalizing generalization and would be supported by more examples in Language 2 than in Language 1. This greater strength of the anti-palatalizing generalization in Language 2 matters iff competition between generalizations is resolved stochastically rather than always being resolved in favor of the "best" generalization because the best generalization is palatalizing in both languages. With stochastic resolution of competition, palatalization would be expected to be more productive in Language 1 than in Language 2. This is in fact what we observe with English speakers (Figure 1; $p<.01$). This account also explains the puzzling fact that in nonce borrowing from English into Russian as seen on Google, velar palatalization (exceptionless in the native lexicon) is shown to be less productive before the verbal stem extension –i (blok→ blotʃ+i+t^j) than before the diminutive suffix –ok (blok → blotʃ+ok): 44% vs. 100% ($p<.0005$), which can be attributed to the Russian lexicon providing greater support for $C_{[\alpha place]} \rightarrow C_{[\alpha place]}i$ than for $C_{[\alpha place]} \rightarrow C_{[\alpha place]}ok$ because –i tends to attach to non-velars while –ok tends to attach to velars.

Thus many grammatical systems that look regular to an analyst might not be regular to the speakers of the language. Retrieval of known complex expressions from the lexicon in preference to grammatical generation appears almost inevitable in an irregular grammatical system for a speaker to be very certain about the forms of words s/he knows. The grammar must account for the speaker's stochastic behavior with novel lexical items, since these items by definition cannot be retrieved from memory. However, the grammar cannot at the same time account for lexically-specific deterministic behavior exhibited with familiar lexical items (Frisch et al. 2004: 220; Zuraw 2000). The speaker's certainty about the form of a familiar word must come from retrieving the information about the familiar word from the lexicon. To further support our contention that seemingly regular systems are sometimes not, we will present a case where a seemingly regular grammatical system (the spelling of Russian prefixes, e.g., Figure 2) is nonetheless shown to be very reliant on lexical retrieval in precisely the cases when the prefix is highly confusable with a differently-spelled preposition (e.g., for the adjectival prefix *bez-* but not the verbal prefix *raz-*) resulting in the potential for competition between spelling rules in production that is not apparent from a linguistic description of the seemingly regular spelling system.

References:
Frisch, S. A., J. B. Pierrehumbert, & M. B. Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, 22, 179-228.
Plag, I. 2003. *Word formation in English*. Cambridge: Cambridge University Press.
Zuraw, K. 2000. Patterned exceptions in phonology. Ph.D. Dissertation, UCLA.

Table 1: The two artificial languages presented to learners. The variables M and N show the numbers of word pairs exemplifying a particular rule in each of the four languages. M and N can be unequal, and are greater than zero.

	Language 1	Language 2
{k;g} → {tʃ;dʒ}i	M	
{t;d;p;b} → {t;d;p;b}i	N	3N
{t;d;p;b} → {t;d;p;b}a	3N	N

Figure 1: Adding –i without changing the stem-final consonant is more productive in Language 2 than in Language 1 (the dashed line shows the median).

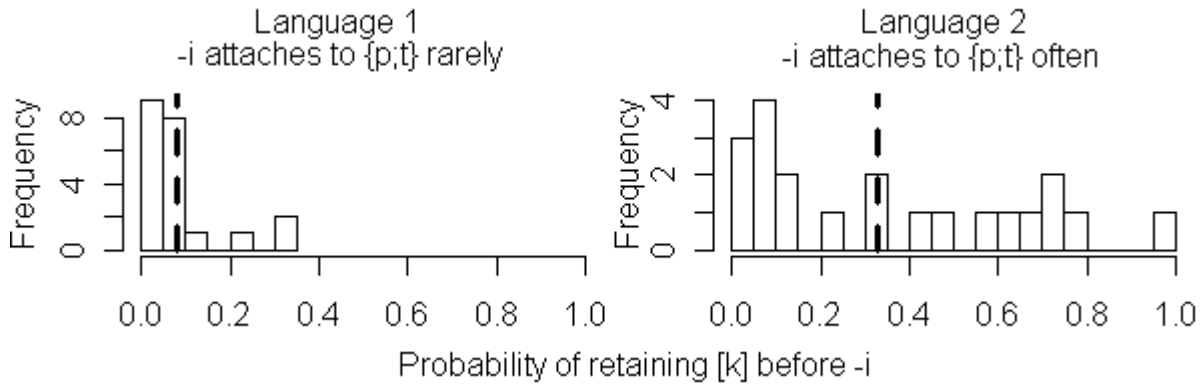

Figure 2: The (regular) spelling of *bez-* in regularly inflected Russian adjectives is driven by lexical retrieval of whole words (unlike for prefixes of verbs, which are less like differently spelled prepositions)