xuj-Reduplication in Russian

Russian exhibits a pattern of expressive echo reduplication in *xu(j)-*: *sélfi-xuélfi* 'selfie-RED', *startápy-xujápy* 'startups-RED' with derogative / depreciative meaning. The word *xuj* 'penis' is strongly tabooed in Russian, and, for this reason, the phenomenon is rarely discussed in the Russian-speaking scientific community. It is telling that most papers mentioning this type of reduplication appeared in *Russian Linguistics*, a journal published outside Russia (Dreizin and Priestly 1982, Plähn 1987, Belikov 1990, Voinov 2012).

The most extensive description of the pattern is provided by Belikov (1990):

1) if the base does not end in an open stressed syllable, its pretonic part is replaced with *xuj-* in the reduplicant (*úlica-xujúlica* 'street-RED', *avtóbus-xujóbus* 'bus-RED', *tumán-xuján* 'fog-RED');
2) if the base ends in an open stressed syllable, *xue-* is added to the last syllable of the base (*oslý-xueslý* 'donkeys-RED').

These rules were inferred from introspection and disparate anecdotal examples. In order to test their validity on a large sample, we used the Araneum Russicum Maximum corpus of 13.7 billion tokens (Benko 2014). We ran a CQL query [lc=".*(.{1,})-*xu*.?\1"], looking for hyphenated words whose first component contains (i) any substring of 0 or more characters and (ii) a substring of 1 or more characters repeated in the reduplicant after *xu* and possibly one more character. The query yielded 3,801 hits. After manually filtering out words like *čixua-xua* 'chihuahua' or *San-Xuan* 'San Juan', which amount to more than 40% of results, we were left with 397 types and 439 tokens of *xuj*-reduplication. Among these 439 examples, there are 434 nouns, two verbs, one adverb, one pronoun and one question word.

Belikov's rules are broadly confirmed with several exceptions that mostly involve repetition of the onset of the stressed syllable (*profkóma-xujkóma* 'trade.union.GEN.SG-RED' instead of the expected *profkóma-xujóma*, which in this case might be due to the fact that *profkom* is a compound itself). However, words with ultimate stress show substantial variation. Out of 15 non-monosyllabic ultimately stressed words, only seven conform to Belikov's rules. In five words, the reduplicant follows the *xuj*$VC_1^nVC_0^1$ pattern (*m*[a]*ntáž-xujantáž* 'assembling-RED' instead of *montáž-xujáž*, *ljubví-xujubví* 'love.GEN.SG-RED' instead of *ljubví-xuebví*); in three words, the pattern is *xuj*$C_1^nVC_0^1$ (*lar'kí-xujkí* 'booth-RED', *xujkí* coinciding with an existing derivative of *xuj* meaning 'little pricks').

Interestingly, the reduplication of ultimately stressed words seems to be avoided. The distribution of stress positions in reduplicated words is clearly different from the general distribution of stress in nouns. A control group of 100 nouns selected randomly from blog posts shows a substantially higher proportion of ultimately stressed words (see Table).

The avoidance of *xuj*-reduplicated forms in words with ultimate stress is best explained by a constraint on the length of the repeated part of the word: it should not be too short, which is typical of morphological doubling (Inkelas and Zoll 2005, Inkelas 2008). Otherwise it would violate the MAX-BR constraint requiring base-reduplicant identity (McCarthy and Prince 1995; see also Alderete et al. 1999). In our talk, we will discuss the application of competing approaches to Russian *xuj*-reduplication.

Stress	Examples of *xuj*-reduplication		Control group
	Number of words	Example	Number of words
ultimate	31 (7.1%)	*gúgl-xujúgl* 'Google-RED'	26 (26%)
penultimate	237 (54%)	*kásting-xujásting* 'casting-RED'	40 (40%)
antepenultimate	166 (37.8%)	*sánkcii-xujánkcii* 'sanctions-RED'	33 (33%)
preantepenultimate	5 (1.1%)	*ménedžery-xuénedžery* 'managers-RED'	1 (1%)

Table. Stress positions in the examples with *xuj*-reduplication as compared with the control group

References

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Suzanne Urbanczyk. 1999. Reduplication with fixed segmentism. *Linguistic Inquiry* 30(3). 327–364. doi:10.1162/002438999554101.

Belikov, Vladimir. 1990. Produktivnaja model' povtora v russkom jazyke: Material dlja obsuzhdenija [A productive model of repetition in Russian]. *Russian Linguistics* 14(1). 81–86. doi:10.1007/BF02743719.

Benko, Vladimír. 2014. Aranea: Yet another family of (comparable) web corpora. In Petr Sojka, Aleš Horák, Ivan Kopeček, and Karel Pala (eds.). *Text, Speech and Dialogue. 17th International Conference, Brno, Czech Republic, September 8–12, 2014*. Cham: Springer International Publishing Switzerland. 247–256.

Dreizin, Felix and Tom Priestly. 1982. A systematic approach to Russian obscene language. *Russian Linguistics* 6(2). 233–249.

Inkelas, Sharon. 2008. The dual theory of reduplication. *Linguistics* 46(2). 351–401. doi:10.1515/ling.2008.013.

Inkelas, Sharon and Cheryl Zoll. 2005. *Reduplication: Doubling in morphology*. Cambridge: Cambridge University Press.

McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory*. 249–384.

Plähn, Jürgen. 1987. Xujnja-mujnja i tomu podobnoe [Xujnja-mujnja and the like]. *Russian Linguistics* 11(1). 37–41. doi:10.1007/BF02743729.

Voinov, Vitaly. 2012. Rhyming reduplication in Russian paired words. *Russian Linguistics* 36(2). 175–191. doi:10.1007/s11185-012-9091-5.