

Cynthia Vakareliyska, University of Oregon

Processing [N[N]] Constructions in Macedonian vs. Bulgarian: A Corpus Linguistics Study

Abstract

[N[N]] constructions, illustrated by English “action hero” or “state senate”, have entered the various modern Slavic languages since the end of the Soviet period in 1989, as English loanblends consisting of an English qualifier noun followed by a native head noun. Most Slavic loanblend [N[N]]’s, such as Russian _SIM-kartočka ‘_SIM card’, are semi-calques representing new concepts and products that have been introduced through English. The introduction of this grammatical construction is significant, because it violates traditional Slavic grammatical rules by leaving the qualifier without adjectival and agreement morphemes.

The study on which this talk is based is a comparison of corpora of Serbian, Bulgarian, Macedonian, Russian, Ukrainian, and Belarusian online newspaper issues (the last three collected and analyzed by Valeriia Tretiak) over a single period. The results shows that the acceptability of [N[N]]’s varies from language to language, ranging from Bulgarian, which has adopted English loanblend [N[N]]’s into its official prescriptive grammar, new word dictionary, and orthographic dictionary as part of the standard language, to Belarusian, which tends to avoid them. The Bulgarian corpus also has by far the largest number of productive English-loanblend [N[N]]’s, i.e., innovative and often metaphorical [N[N]]’s formed from English qualifier nouns together with new native head nouns that are not attested together in English: e.g., _parti zona ‘party zone’, _džaz emocii ‘jazz emotions’.

It has been argued informally by some Slavic linguists that the easy adoption and processing of [N[N]]’s in Bulgarian is simply a consequence of the language’s analytic case system (as opposed to the inflectional system of most of the Slavic languages). If this were the case, then Macedonian, which too has an analytic case system, as well as grammatical and lexical systems nearly identical to Bulgarian’s, should also exhibit significant [N[N]] productivity, but comparison of the Macedonian and Bulgarian newspaper corpora shows that it does not.

This talk shows that the reason for the lack of [N[N]] productivity in Macedonian is a consequence of parsing restrictions that make them more difficult to process as NPs as opposed to indivisible compound nouns. And, based partly on evidence from the Serbian newspaper corpus, the talk explains what causes these restrictions in Macedonian.