A Contrastive Study of Verbless Sentences in Russian and English Parallel-Corpora

Bondarenko A.

*Université Paris Diderot, Paris, France*

tonyabondarenko@gmail.com

Although the absence of the verb has fascinated linguists for ages, the difficulty of automatic processing has meant that most analyses have relied on fragmented data. Overcoming the typical problems of fixed annotation and verb-centric syntactic modeling (Landolfi et al. 2010), this paper presents a new method of verbless sentence extraction. Verbless sentences are investigated through parallel-text corpora (Stolz 2007) in English and Russian with the aim of uncovering the semantic and pragmatic factors associated with the absence of the verb.

Cross-linguistic differences make it particularly relevant to compare Russian, a language that permits the most liberal use of verbless sentences among the Indo-European family, with English, known for its dependency on the verb phrase (Kopotev 2007). The present study combines corpus methods with the parallel-text approach to contrastive linguistics developed by Guillemin-Flescher (2003). Following the latter, we rely on the principle that re-occurring translation patterns reveal linguistic constraints that would otherwise remain hidden. The results are based on a 150,000-word specially created corpus, which includes Dostoyevsky's Russian dialogue centered *Brothers Karamazov*, Pinter's English play *The Caretaker*, and several translations. Automatic verbless sentence extraction was developed using Trameur (Fleury & Zimina 2014) and the data was aligned by paragraph. Characteristic elements of verbless sentences for each language were found using statistical corpus tools. Translation correspondences were manually annotated for antecedent-based verbal ellipsis (McShane 2000), information structure (Lambrecht 1994), predication (Hengeveld 1992).

While establishing statistical significance of expected verbless sentence frequency differences between Russian and English, the results surprisingly reveal English overrepresentation of antecedent-based ellipsis. Verbless sentences are overwhelmingly non-elliptical, associated with direct speech and statistically characterized by informal markers. Information structure analysis of translation correspondences, shows an English trend of activating contextually implied topics, particularly over distances. Predication transformation findings imply that the verbal/non-verbal dichotomy is inadequate for a cross-linguistically stable definition of predication.

*Keywords*: verbless sentences, parallel-corpora, Russian-English.