Svetlana Gorokhova (St. Petersburg, Russia)

The Selection of Grammatical Features: A Corpus-based Analysis of Speech Errors

Abstract

Russian is a highly inflected language in which grammatical features such as gender, number, case, tense and aspect usually surface as bound morphemes. To study the ways grammatical features are selected during sentence production, Russian speech errors (slips of the tongue) involving inflected word forms were analyzed and compared to the data from the National Corpus of Spoken Russian.
The analysis involved 200 context-free substitutions of a grammatical feature yielding a wrong word form, e.g.:

Ty	otvezi	ix	v	Moskvu	k	**rodstvennik-am**
You	take	them	to	Moscow	to	relative-PL.**DAT**
				→ …	k	**rodstvennik-ov**
					to	relative-PL.**GEN**

Take them to your relatives in Moscow
(the genitive case form of the noun 'relatives', *rodstvennikov*, is selected instead of the target dative case form, *rodstvennikam*).

Assuming that context-free substitutions of a grammatical feature result from competition among the grammatical forms of a target word during lexical processing, it seems reasonable to expect that within the inflectional paradigm of a given word, some forms may be more likely to be selected.
A comparison between the relative frequencies of the target and error word forms within the word's inflectional paradigm in the National Corpus of Spoken Russian shows that the frequency of the error word form is generally higher than that of the target word form.
The above data lead to a number of conclusions on how grammatical features are selected during lexical processing. Firstly, the evidence from speech errors reveals that the selection of grammatical features is generally a competitive process. This is incompatible with the claim that although the selection of lexical nodes may be competitive, the selection of their grammatical properties is an automatic consequence of lexical selection (e.g. Caramazza et al. 2001). In addition, context-free grammatical feature substitutions suggest the existence of a distinct grammatical encoding level of sentence production.
Secondly, the comparison with the corpus data indicates that during sentence production, some grammatical features may have a priority within a word's inflectional paradigm; consequently, the more robust (higher-frequency) inflected forms of the word may be more readily accessible than the weaker forms to serve as potential substitutes for the latter. This is consistent with the finding that lexical processing is affected by the frequency of an inflected word form (e.g. Kostić, Mirković 2002).

References

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F. X. (2001), A crosslinguistic investigation of determiner production. In E. Dupoux (Ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler*. Cambridge, MA: MIT Press.

Kostić, A., Mirković, J. (2002), Processing of inflected nouns and levels of cognitive sensitivity. *Psihologija*, 35, 287-297.