Bilingualism: Language and Cognition 19 (5), 2016, 887–888 © Cambridge University Press 2016 doi:10.1017/S1366728916000092

Constraints on variation, reduplication of semantics, and degrees of freedom: some notes on the computational account of code mixing

ROBERT J. HARTSUIKER
Ghent University

(Received: January 15, 2016; final revision received: January 18, 2016; accepted: January 18, 2016; first published online 3 February 2016)

Goldrick, Putnam and Schwarz (2016) propose a very explicit account of code mixing, which nicely draws implications from the well-established findings of coactivation during bilingual language production, and that code mixing is constrained by grammatical principles. This precise characterization will undoubtedly serve as a useful basis for further research on code mixing, second language sentence production, and syntactic learning. However, there are three issues with the account that require further elaboration.

Does having multiple word orders entail flexibility?

When defining the first principle of their account, Goldrick et al. argue that English requires the word order SVO, whereas Dutch allows "flexibility between SVO, SOV, and VSO." This is incorrect. All three orders indeed occur in Dutch, but they cannot be chosen flexibly. SVO is the order of the main clause and SOV is the order of a subordinate clause starting with "omdat" (because) (1a). But subordinate clauses starting with "want" (as) (2) have SVO order. VSO is required when a unit (e.g., a Wh-word) is moved sentence-initially (see Koster, 1975, for a generative analysis of Dutch word order). In contrast, order of auxiliary verb and participle within SOV clauses (1a-b) is flexible. Perhaps such facts of Dutch grammar can be captured by a more elaborate list of constraints. But it does seem that an analysis that simply ranks these word orders according to their overall probability misses an essential point about this variation: surely SVO, VSO, and SOV occur with certain probabilities of x, y, and $(1 - (x+y))$, but if we know the syntactic and lexical context, we can predict the order with certainty.

(1a) Zij belde de politie omdat de zakkenroller haar portemonnee had gestolen

Lit.: She called the police because the pickpocket her wallet had stolen

(1b) Zij belde de politie omdat de zakkenroller haar portemonnee gestolen had

Lit.: She called the police because the pickpocket her wallet stolen had

(2) Zij belde de politie want de zakkenroller had haar portemonnee gestolen

Lit.: She called the police as the pickpocket had her wallet stolen

"She called the police <because/as> the pickpocket had stolen her wallet"

Is semantics shared or separate across languages?

The account of doubled elements rests on the very strong assumption that there is co-activation, to varying degrees, of two separate semantic representations, one for English *gave* and one for Tamil *kodutaa* (see Goldrick et al.'s example 15). This assumption is crucial for the analysis of doubled elements, because compliance with the faithfulness constraint now promotes the rank of the doubled construction so that it acquires some probability. However, the authors do not provide an independent motivation for the assumption of a doubled representation in the semantics. In fact, the more fundamental question can be asked whether semantic representations in L1 and L2 are shared or separate. Note that psycholinguistic theories (e.g., Kroll & Stewart, 1994) traditionally assume a common semantic system for both languages or a feature-based system with high overlap between L1 and L2 semantic representations (Van Hell & De Groot, 1998). There is (masked) priming of non-cognate translation equivalents across languages (Schoonbaert, Duyck, Brysbaert & Hartsuiker, 2009) suggesting shared or overlapping representations. Relatedly, cross-language structural priming studies showed stronger priming if head verbs in prime and target sentence were translation equivalents (Schoonbaert, Hartsuiker & Pickering, 2007).

Address for correspondence: Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium
robert.hartsuiker@ugent.be

In light of such theories and findings, the assumption of discrete and fully separate semantic representations for each language is surprising.

Is the account overparametrized?

The paper refers to a learning algorithm, but does not describe the algorithm, parameters, training data, or fitting procedures in detail. This leaves it unclear how many free parameters there are in the unilingual case. But additionally, the adaptation to bilingualism introduces further parameters that for instance determine the strength of each language, the strength of each of the discrete semantic representations, the strength of the 'boost' to a constraint depending on the head word's language, and so on. Furthermore, all terminal and non-terminal nodes have activation values that interact with the contraints. If the goal is to provide a general, descriptive framework that demonstrates how code mixing COULD be analyzed, such aspects of training and fitting might be viewed as uninteresting details of implementation. But if the more ambitious goal is to develop a precise account that makes testable and quantitative predictions, clarification is needed of what data sets or experiments can be used to test the account and how the number of free parameters compares to the number of to be fitted data points. For now, the seemingly large number of degrees of freedom raises questions about the falsifiability of the account.

References

Goldrick, M., Putnam, M., & Schwartz, L. (2016). Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition*. doi: dx.doi.org/10.1017/S1366728915000802.

Koster, J. (1975). Dutch as an SOV language. *Linguistic Analysis, 1,* 111–136.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33,* 149–174.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language, 56,* 153–171.

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition, 37,* 569–586.

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory. Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193–211.