

# Questions on data and the input to GEN\*

LUIS LÓPEZ  
*University of Illinois at Chicago*

(Received: January 13, 2016; final revision received: January 17, 2016; accepted: January 18, 2016; first published online 19 February 2016)

Keywords: code-switching, blending, GEN, optimality

The keynote article (Goldrick, Putnam & Schwartz, 2016) discusses doubling phenomena occasionally found in code-switching corpora. Their analysis focuses on an English–Tamil sentence in which an SVO sequence in English is followed by a verb in Tamil, resulting in an apparent VOV structure:

(1)(=7) They gave me a research grant *kadutaa*.  
give.3.PL.PAST

Crucial to the analysis is that the second verb is extracted from the grammar of a SOV language. The authors propose an optimality-inspired approach in which the two verbs are co-activated and placed in the input. The idea is intriguing and original. My questions revolve around the data and the architecture proposed.

First the data. The article discusses doubling of functional items and lexical items together, although they are in fact very different phenomena. The first occur often in code-switching and subjects accept them in judgment tasks. They should probably be analyzed in terms of morphological well-formedness requirements. The second type is really the main focus of this article. They are very rare in corpora and subjects reject them in acceptability tasks. Consequently, one should ask if the phenomenon under discussion actually exists. The articles that report on the corpora studies provide very little information regarding the intonation or the context of the data. Code switching is very often preceded by hesitation, repetition or change of strategy. Could it be the case that what is involved in these doubling constructions is a change of strategy in forming an utterance? If that is the case, it should hardly be studied as a phenomenon of linguistic competence.

The lack of information on the circumstances regarding the production of blended forms combines with a shallow or absent grammatical analysis of the languages involved in the article's cited sources. Notice that in all cases of VOV reported in the literature the OV language allows for

silent objects (Turkish, Tamil, Chinese). This opens the door for an alternative analysis: instead of VOV we might have VO<sub>pro</sub>V. If there is a pause after the O, then we might have a dislocated VP, a less extraordinary phenomenon than it seemed.

In order for us to ensure that blends are in fact features of the grammar of bilingual code-switchers we would need to have experimental data that taps speakers' linguistic competence. Additionally, the OV language involved should not allow empty objects.

Let's move onto the actual proposal and proceed on the assumption that lexical doublings are indeed part of the linguistic competence of bilingual code-switchers. The central idea is that doubling is the result of co-activation. Using a soft-constraint approach, co-activation is formalized as a consequence of having lexical items from two languages in the input to GEN. This approach raises a number of interesting points that I hope the authors will explore in future work or in an extended version of this contribution. The first question is: how is an input constituted? In Grimshaw's (1997) original proposal, which the authors invoke as their source, an input is a complete argument structure. This imposes strict limits to the kinds of things that can be an input to GEN (although Grimshaw leaves open what mechanism ensures that lexical arrays that do not constitute an argument structure are prevented from forming an input to GEN.) Although Grimshaw does not elaborate, it is easy to see why she placed this requirement: it is the only way to prevent the generation of sentences like “\*Pat breaks”, which no constraint in her system rules out. Goldrick et al. weaken Grimshaw's requirement: the input for doubling structures can include two predicates and one argument structure. Notice that the two predicates are independent lexical items, not copies of one and the same item. Once we have weakened Grimshaw's condition, a question arises: what constitutes an input to GEN? Should there be no restrictions? How do we prevent outputs like “\*Pat breaks”? The questions multiply when we consider blending data. The description in Goldrick et al. includes an implicit condition on the input, namely, that only

\* I would like to thank the members of the UIC Bilingualism Research Laboratory for discussion and continuous inspiration.

Address for correspondence:

Luis López, Department of Hispanic and Italian Studies (MC 315), 601 South Morgan St, University of Illinois at Chicago, Chicago IL 60607  
[luislope@uic.edu](mailto:luislope@uic.edu)

synonyms can be co-activated. How can we make sure that this condition is kept? For instance, in (1): could *kudataa* double ‘donate’ instead of ‘give’? We can wade a little wider. How about a completely different verb that has the same argument structure and assigns similar theta roles? An example of this last possibility is presented in Muysken (2000, p. 105, example 44). Moreover, the instances of co-activation that are cited in the literature involve phonetic co-activation (and it is an interesting feature of Goldrick et al.’s proposal that they have expanded the range to semantic co-activation), which can be enhanced if the items involved are cognates. This leads to the following question: Could phonetic resemblance also be a factor in creating a doubling input?

Finally, I also hope the authors will spend some time tying up loose ends of the analysis in future work. Their claim that Chomsky’s (1981) Theta criterion blocks structures with two objects and one verb (OVO) is correct, but the Theta criterion also blocks the structures that they discuss (VOV), since it forces the unique object to receive theta roles from two verbs. In traditional generative grammar this is impossible and forms the foundation of the traditional raising/control analyses of infinitives.

Additionally, their claim that their main datum cannot be accounted for via verb movement is debatable. A fairly standard analysis of predicate phrases involves a higher verbal head *v* that takes a lexical verb phrase as complement. The two heads *v* and *V* can be linked via movement. Their crucial datum (1) could be accounted for with the following assumptions (i) overt movement of the lexical verb to *v* in English, (ii) covert *V*-to-*v* movement in Tamil, (iii) the object is merged in or moved to a position to the left of *V*, (iv) doubling consists of spelling out of both copies. This would also have the advantage of accounting for the impossibility of OVO structures directly.

## References

- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- 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](https://doi.org/10.1017/S1366728915000802).
- Grimshaw, J. (1997). Projection, heads and optimality. *Linguistic Inquiry*, 28, 373–422.
- Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.