Testing tolerance for lexically-specific factors in Gradient Symbolic Computation

JANET G. VAN HELL, CLARA COHEN and SARAH GREY

Bilingualism: Language and Cognition / *FirstView* Article / April 2016, pp 1 - 3
DOI: 10.1017/S1366728916000122, Published online: 11 February 2016

Bilingualism: Language and Cognition: page 1 of 3 © Cambridge University Press 2016 doi:10.1017/S1366728916000122

COMMENTARY

Testing tolerance for lexically-specific factors in Gradient Symbolic Computation

JANET G. VAN HELL
Department of Psychology and Center for Language Science, Pennsylvania State University
CLARA COHEN
Department of Psychology and Center for Language Science, Pennsylvania State University
SARAH GREY
Department of Psychology and Center for Language Science, Pennsylvania State University

(Received: January 17, 2016; final revision received: January 18, 2016; accepted: January 18, 2016)

In their keynote article, Goldrick, Putnam and Schwarz (2016) present a computational account of code-mixing. Although they review literature on the co-activation of lexical representations and cognate facilitation effects in bilingual language processing, their model remains silent on how it interfaces with lexical factors, and how lexical factors impact code-switching. One such lexical factor is cognate status, which has been found to affect code-switching, as demonstrated in corpus analyses (e.g., Broersma & De Bot, 2006) and psycholinguistic experiments (Kootstra, Van Hell & Dijkstra, 2012). For example, using the structural priming technique to examine the role of lexical factors in code-switching, Kootstra et al. asked Dutch–English bilinguals to repeat a code-switched prime sentence (starting in Dutch and ending in English) and then describe a target picture by means of a code-switched sentence (also from Dutch into English). They observed that bilinguals' tendency to switch at the same position as in the prime sentence was increased when the prime sentence and target picture contained cognates.

To examine whether Goldrick et al.'s model is flexible enough to tolerate lexically-specific information, we extended their sample computation to include the lexical property of cognate status. We considered four variants of input to Goldrick et al.'s Table A1 (*gave/ koḍutaa [3rd plural, grant*]). Specifically, we replaced the object *grant* with the English–Tamil cognate pair *mango-mangai*. To model the activation of both cognate forms in bilinguals, we adapted Goldrick et al.'s approach with variable joint input activation, as illustrated in Goldrick et al.'s Tables 8 and 9. The four variants we considered had the following relative input activations:

Case 1: *mango* = *1*, *mangai* = 0 (English direct object)
Case 2: *mango* = 0*,* *mangai* = 1 (Tamil direct object)
Case 3: *mango* = 0.5, *mangai* = 0.5 (Cognate with equal activation of English and Tamil)
Case 4: *mango* = 0.4, *mangai* = 0.6 (Cognate with Tamil-biased activation)

The preparation of this commentary was supported by National Science Foundation grants OISE-0968369, BCS-1349110, and SMA-1514276.

To test the effect of variable weighting on these candidates, we built a table corresponding to Goldrick et al.'s Table A1, in which the output candidates of the grammar fragment are revised to include both *mango* and *mangai*. Thus, the first candidate in Goldrick et al.'s Table A1, *they gave grant koḍutaa*, has two counterparts in our tables: (1a) *they gave mango koḍutaa,* and (1b) *they gave mangai koḍutaa.* In Tables 1 and 2, the (a) candidate is always equivalent to the version in Goldrick et al.'s Table A1, substituting only *mango* for *grant*, while the (b) candidate always used *mangai*.

Case 1 behaved as expected: The probabilities predicted for the (a) candidate, containing the English direct object *mango*, matched their equivalents in Goldrick et al.'s Table A1. In Case 2, with the weights reversed to model a Tamil-only direct object, the model's predicted probabilities did not simply shift the probabilities to the (b) candidates, which differed only in using *mangai* rather than *mango* (Table 1, with Case 1 probabilities included for comparison). Rather, they showed a distinctly different distribution, along with the penalties for each constraint. As expected, the only two output candidates with non-zero probability are (b) candidates, containing *mangai*. However, the differences between the two distributions result from the asymmetrical weightings of compLeft between English and Tamil. The position of *mango* in the (a) candidates never incurs English-specific violations, while the position of *mangai* incurs Tamil-specific violations in every (b) candidate that places it anywhere beyond the left edge. Under the language-specific weightings proposed by Goldrick et al., doubling constructions are only possible when the direct object is English.

For Case 3, a cognate with equal activation in both languages, the model output is not some blend of Case 1

Address for correspondence: Janet G. van Hell, Department of Psychology, 414 Moore Building, Pennsylvania State University, University Park, PA 16802 USA
jgv3@psu.edu

Table 1. *Model output for Case 1 and 2.*

candidate	Spec Left	Head Left	Comp Left	Case 2 parse	Case 2 quant	Case 2 harmony	Case 2 pr	Case 1 pr
1a they gave mango kodutaa	0	−30	−12	−37.5	−16	−95.5	0.000	0.039
1b they gave mangai kodutaa	0	−30	−24	0	−16	−70	0.000	0.000
2a they kodutaa mango gave	0	−42	−12	−37.5	−16	−107.5	0.000	0.000
2b they kodutaa mangai gave	0	−42	−24	0	−16	−82	0.000	0.000
3a they gave mango	0	−12	−12	−75	0	−99	0.000	0.001
3b they gave mangai	0	−12	−24	−37.5	0	−73.5	0.000	0.000
4a they kodutaa mango	0	−6	−12	−75	0	−93	0.000	0.480
4b they kodutaa mangai	0	−6	−24	−37.5	0	−67.5	0.002	0.000
5a they mango gave	0	−24	−6	−75	0	−105	0.000	0.000
5b they mangai gave	0	−24	−12	−37.5	0	−73.5	0.000	0.000
6a they mango kodutaa	0	−12	−6	−75	0	−93	0.000	0.480
6b they mangai kodutaa	0	−12	−12	−37.5	0	−61.5	0.997	0.000

Table 2. *Model output for Case 3 and 4. Values for specLeft, headLeft, and compLeft do not change from Case 2, and are therefore omitted here.*

candidate	Case 3 parse	Case 3 quant	Case 3 harmony	Case 3 pr	Case 4 parse	Case 4 quant	Case 4 harmony	Case 4 pr
1a they gave mango kodutaa	−18.75	−16	−76.75	0.039	−22.5	−16	−80.5	0.012
1b they gave mangai kodutaa	−18.75	−16	−88.75	0.000	−15	−16	−85	0.000
2a they kodutaa mango gave	−18.75	−16	−88.75	0.000	−22.5	−16	−92.5	0.000
2b they kodutaa mangai gave	−18.75	−16	−100.75	0.000	−15	−16	−97	0.000
3a they gave mango	−56.25	0	−80.25	0.001	−60	0	−84	0.000
3b they gave mangai	−56.25	0	−92.25	0.000	−52.5	0	−88.5	0.000
4a they kodutaa mango	−56.25	0	−74.25	0.479	−60	0	−78	0.151
4b they kodutaa mangai	−56.25	0	−86.25	0.000	−52.5	0	−82.5	0.002
5a they mango gave	−56.25	0	−86.25	0.000	−60	0	−90	0.000
5b they mangai gave	−56.25	0	−92.25	0.000	−52.5	0	−88.5	0.000
6a they mango kodutaa	−56.25	0	−74.25	0.479	−60	0	−78	0.151
6b they mangai kodutaa	−56.25	0	−80.25	0.001	−52.5	0	−76.5	0.676
12 they kodutaa	−75	0	−81	0.001	−75	0	−81	0.008

and Case 2, as we would expect, but nearly identical to Case 1, differing only in the now non-zero probability of candidate (6b; see Table 2).

In Case 4, where the weighting of the two cognate forms is biased towards Tamil, quite a different picture emerges (Table 2). Here, the predicted probabilities seem to be a more intuitive blend of Case 1 and Case 2. The non-zero probability of the doubling construction and the symmetrical probabilities of candidates (4a) and (6a) are consistent with Case 1, while the preference for (6b) is consistent with Case 2.

This demonstration shows that Goldrick et al.'s model is indeed flexible enough to incorporate cognate status, which we know can affect code-switching. Whether the predictions that emerge are accurate is an empirical question. Our demonstration has yielded two such predictions. First, the English and Tamil-specific weights for compLeft are different, while they are equivalent for specLeft. This predicts that the different probability distributions for possible outputs should be sensitive to whether the cognate is in object position, as we modeled here, or in subject position. Second, the modeling of equal activation of both cognate forms in Case 3 showed that the output probabilities are quite similar to those produced in Case 1, the English-only direct object. It required a distinct bias for *mangai* in Case 4 to produce a more intuitive combination of the probabilities predicted for English-only and Tamil-only direct objects. It would be interesting to see whether a similarly non-linear relationship between relative language (cognate) activation and code-mixing can be observed in experimental data.

References

Broersma, M., & de Bot, K. (2006). Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition, 9,* 1–13.

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.

Kootstra, G. J., Van Hell, J. G., & Dijkstra, T. (2012). Priming of code-switches in sentences: The role of lexical repetition, cognates, and language proficiency. *Bilingualism: Language and Cognition, 15,* 797–819.