

AUTHORS' RESPONSE

The future of code mixing research: Integrating psycholinguistic and formal grammatical theories*

MATTHEW GOLDRICK

Northwestern University

MICHAEL PUTNAM

Pennsylvania State University

LARA SCHWARZ

Pennsylvania State University

(Received: February 15, 2016; accepted: February 19, 2016; first published online 22 March 2016)

Keywords: Code mixing, Gradient Symbolic Computation

Our keynote article “Coactivation in bilingual grammars: A computational account of code mixing” (Goldrick, Putnam & Schwarz) aimed to provide a framework that would begin to unify psycholinguistic and formal grammatical approaches to code mixing. We situated our account within a large body of psycholinguistic and phonetic evidence suggesting that, under many conditions, multiple representational elements simultaneously occupy (to varying degrees) a single position within a linguistic structure. The presence of such blends in multilingual cognition is not compatible with many formal grammatical approaches that assume mental representations are necessarily discrete.

Using the Gradient Symbolic Computation architecture (Smolensky, Goldrick & Mathis, 2014), we developed a formal framework that addresses this challenge. It is based on 3 principles, which we summarize and clarify below:

1. **PROBABILISTIC GRAMMARS WITH WEIGHTED CONSTRAINTS:** The grammar defining variation in the mapping between meaning and syntactic structures is specified by numerically weighted constraints. These weights are acquired, reflecting both the speaker's experience and the prior constraints imposed by the learning algorithm.
2. **GRADIENT BLENDS OF GRAMMARS:** Any particular instance of code mixing reflects a weighted blend of the grammars that results from the process of acquiring multilingual competence (including experience with code mixing).
3. **GRADIENT BLENDS IN LINGUISTIC REPRESENTATION:** Multiple representational elements can, under certain

conditions, simultaneously occupy (to varying degrees) a single position within a linguistic structure.

We are pleased that our proposal stimulated such a vigorous discussion from researchers representing a variety of methodological and theoretical approaches to code mixing. In our reply, we discuss the core issues raised in the commentaries. We first discuss two issues that have general implications for the study of code mixing; we then turn to specific points related to our proposal.

General issues in theories of code mixing

The role of blends in theories of code mixing

Coactivation is a central claim of well-established psycholinguistic models of language processing in monolinguals (Dell, 1986) and bilinguals (Dijkstra & van Heuven, 1998). Key to such accounts is that in blended representations multiple elements can be present to varying degrees. In some cases, two representations are both strongly present – as in the doubling constructions we review, or in phonological mixtures of two words as discussed by Deuchar and Biberauer (Deuchar & Biberauer). These extreme cases of coactivation are but the tip of the iceberg. In most instances of bilingual language processing, it appears that one representation very strongly dominates processing; other representations are much more weakly activated. Such coactivation manifests in reaction times and phonetic data, rather than visibly as in blended productions. Critically, these reaction time (Starreveld, De Groot, Rossmark & Van Hell, 2014) and phonetic (Goldrick, Runnqvist & Costa, 2014) effects are dynamic and specific to particular novel processing contexts, suggesting they do not solely reflect stored patterns of usage (*pace* Poplack & Torres Cacoullos).

* This research was supported by NSF grant BCS1344269.

Address for correspondence:

Matthew Goldrick, Department of Linguistics, Northwestern University, 2016 Sheridan Rd., Evanston, IL 60208 USA
matt-goldrick@northwestern.edu

Our discussion focused on doubling constructions, as these provided a clear instance of the interaction between coactivation and discrete representations in grammar. But our account generalizes from such extreme cases to the most extreme situations where one representation strongly dominates processing (as illustrated in Figure 4 of our paper and the surrounding text). Our proposal is therefore not grounded solely in doubling constructions (*pace* Deuchar & Biberauer; Gullberg & Parafita Couto; López), but is situated in a rich tradition of psycholinguistic work that any theory of bilingual language knowledge and use must account for.

Note as well that our proposal does not claim that strongly activated blend states are typical or common. The need to constrain blends is a computational problem that must be addressed across many cognitive domains (Smolensky et al., 2014). The ubiquity of blends in (bilingual) cognition suggests that cognitive computations are not fully discrete, yet the highly restricted distribution of the degree of blending suggests that cognitive computations are not fully gradient. Our framework – and Gradient Symbolic Computation more generally – addresses this challenge head on, aiming to find the appropriate balance between the gradient and the discrete.

Finally, while strongly activated blend states are rare, the resulting doubling constructions are not anomalous (*pace* Poplack & Torres Cacoullos); they exhibit reliable, structurally driven patterns that theories must account for (Chan, 2015; Hicks, 2010).

Beyond blends: learning and control

Although blended representations play a key role in explaining bilingual language processing in general, and code mixing in particular, they by no means constitute a complete theory. Below we discuss two additional components that must be incorporated into any complete theory of code mixing.

Learning multilingual grammars

Our approach focused on how a bilingual would integrate the grammatical knowledge they have already acquired. We were relatively agnostic on the precise structure of this acquisition process, simply emphasizing the role that experience plays in shaping such knowledge. However, it is clear that greater specificity is required if we are to account for the enormous intra- and inter-speaker variation in code mixing (Bhatt; Gullberg & Parafita Couto; Hartsuiker; Poplack & Torres Cacoullos).

An issue common to both monolingual and multilingual learning that any theory must address is the contribution of prior constraints on the structure of the learning algorithm to the outcome of learning (Verissimo; see Goldrick & Larson, 2010, for discussion in the context of learning of stochastic phonological patterns).

With respect to multilingual learning specifically, theories of multilingual grammars must specify how learning reflects the dynamic, bidirectional interaction between the grammatical systems of multiple languages (Bobb & Hoshino; Sorace). This is a clear challenge that current formal learning theories do not address in any systematic way.

Control

While our account integrates some aspects of psycholinguistic and formal grammatical approaches to code mixing, it is far from a complete integration. Like other grammatical accounts, our proposal is stated at a high level of abstraction, specified in terms of a probabilistic mapping between inputs and outputs. It does not specify HOW this mapping is computed, which represents a central focus of psycholinguistic and neurolinguistic theories of code mixing. Clearly, the mechanisms by which speakers control the language of production make a strong contribution to patterns of code mixing (Green; Sorace). Theories of code mixing have, up to this point, failed to precisely articulate the relative contribution of these mechanisms vs. grammatical knowledge to code mixing patterns. In our estimation, theories incorporating control have not sufficiently specified the role of highly structured grammatical knowledge; theories incorporating such grammatical knowledge have not articulated the contribution of control. This is a clear challenge that the field must address.

Issues for the Gradient Symbolic Computation Account

Input to the grammar

Our proposal left relatively unspecified the nature of the input to the grammatical component that was the focus of our analysis of doubling constructions. Our focus in this analysis was linearization: determining the surface syntactic structures specifying word order. In the input to linearization, language-specific representations of lexical items are associated with grammatical functions (e.g., argument structure). In the output, lexical items are placed within surface syntactic structures. This roughly corresponds to the mapping between functional and positional levels of grammatical encoding in psycholinguistic theories (Garrett, 1975).

What is the source of the input to this linearization process? Building on models of speech production, we assume that the input is determined by at least two processes. First, encoding of the message the speaker wants to communicate (defined over semantic representations that are largely shared between languages, as discussed by Hartsuiker); and second, a mapping from the elements of this message to the lexical items

and grammatical functions. Clearly, additional work is needed to specify the structure of these components of the grammatical system. If, as we propose, blend representations are allowed to be the output of these processes (and hence input to linearization), it is important that we clarify the restrictions on the grammatical properties of such blend representations (as discussed in López).

Conceptual and empirical extensions of our account

Our intent was to introduce a framework for researchers to test out their own proposals for code mixing; a means by which proposals could be formalized and their predictions clearly articulated. We are delighted that van Hell, Cohen and Grey (van Hell, Cohen & Grey) and Veríssimo (Veríssimo) took the opportunity to explore our approach. Veríssimo's analysis examines how probabilistic grammars can be used to model distinctions between native- and non-native compound formation. Van Hell et al. explore the predictions made by gradient variations in input activation. This illustrates the complex interactions that can occur between lexically-specific and structural factors in code mixing. We look forward to the further development of such accounts and empirical assessment of their predictions.

Other commentaries identified particular areas that warrant deeper investigation within our framework. Gullberg and Parafita Couto (Gullberg & Parafita Couto) discuss cases that may appear to challenge our account – when conflicting word order does not give rise to doubling (as discussed in our paper) but rather to MIXED phrases, containing elements from both languages but no doubling. As noted above, while blended representations are possible in our account, they are strongly dispreferred. In fact, as shown in the Appendix of our paper, our analysis predicts that mixed constructions will have much higher probability than blends. In ongoing work (Putnam & Goldrick, submitted), we explore such mixed structures in greater detail, examining how our framework accounts for the broad empirical patterns Gullberg and Parafita Couto have observed in mixed determiner phrases.

Other types of blended constructions are a clear target for a Gradient Symbolic analysis. Bhatt (Bhatt) discusses portmanteau sentences, where a constituent in one language is shared with a structure in another language. Bhatt illustrates how this could be characterized using a blended representation (with the shared constituent simultaneously occupying positions in two distinct phrases). We suggest that while this blended structure violates well-established structural constraints, there may be contexts where other constraints compel violation of such considerations in favor of the blended structure. Muysken (Muysken) discusses more complex instances

of blending, where portions of a collocational expression are embedded (in a sometimes discontinuous fashion) within the structural frame of another language. Analyzing such complex blends will require a fuller understanding of the representation of multiword expressions (see Brehm & Goldrick, submitted, for discussion); this will allow more precise predictions to be made about how such constructions will interact with other structural constraints contributed from both conflicting grammars in the optimization process.

Conclusions

While there is a clear need for greater quantitative data (Poplack & Torres Cacoullos), psycholinguistic and formal grammatical investigations of code mixing have yielded tremendous insights into the computations and processes underlying this aspect of multilingual cognition. We hope that our framework has provided some initial steps towards integrating these traditions and pointed towards several interesting areas of future investigation.

References

- Bhatt, R. M. Coactivation: The portmanteau constructions in bilingual grammar. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000146](https://doi.org/10.1017/S1366728916000146).
- Brehm, L., & Goldrick, M. (submitted). *Distinguishing discrete and gradient category structure in language: Insights from verb-particle constructions*. Unpublished manuscript, Northwestern University.
- Bobb, S. C., & Hoshino, N. Fusing languages in the bilingual cognitive architecture. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000109](https://doi.org/10.1017/S1366728916000109).
- Chan, B.-H. (2015). Portmanteau constructions, phrase structure and linearization. *Frontiers in Psychology*, 6, 1851. doi:[10.3389/fpsyg.2015.01851](https://doi.org/10.3389/fpsyg.2015.01851)
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dijkstra, T., & Van Heuven, W. J. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.) *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Lawrence Erlbaum.
- Deuchar, M., & Biberauer, T. Doubling: An error or an illusion? *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000043](https://doi.org/10.1017/S1366728916000043).
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.) *The psychology of learning and motivation: Advances in research and theory* (pp. 133–177). New York: Academic Press.
- Goldrick, M., & Larson, M. (2010). Constraints on the acquisition of variation. In C. Fougeron, B. Kuhnert, M. D'Imperio & N. Vallee (Eds.) *Laboratory Phonology 10: Variation, phonetic detail and phonological representation* (pp. 285–310). Berlin: Mouton de Gruyter.

- Goldrick, M., Putnam, M. T., & Schwarz, L. (2016). Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition*, doi: [10.1017/S1366728915000802](https://doi.org/10.1017/S1366728915000802).
- Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less native-like. *Psychological Science*, *25*, 1031–1036.
- Green, D., & Li, W. Code-switching and language control. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000018](https://doi.org/10.1017/S1366728916000018).
- Gullberg, M., & Parafita Couto, M. C. An integrated perspective on code-mixing patterns beyond doubling? *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000080](https://doi.org/10.1017/S1366728916000080).
- Hicks, C. (2010). *Morphosyntactic doubling in code switching*. MA thesis, University of North Carolina-Chapel Hill.
- Hartsuiker, R. J. Constraints on variation, reduplication of semantics, and degrees of freedom: Some notes on the computational account of code mixing. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000092](https://doi.org/10.1017/S1366728916000092).
- López, L. Questions on data and the input to GEN. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000079](https://doi.org/10.1017/S1366728916000079).
- Muysken, P. Bilingual complexes: The perspectives of the Gradient Symbolic Computation framework. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000031](https://doi.org/10.1017/S1366728916000031).
- Poplack, S., & Torres Cacoullos, R. Data before models. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000067](https://doi.org/10.1017/S1366728916000067).
- Putnam, M., & Goldrick, M. (submitted). *The good, the bad, and the gradient: The role of 'losers' in code-switching*. Unpublished manuscript, Pennsylvania State University and Northwestern University.
- Smolensky, P., Goldrick, M., & Mathis, D. (2014). Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science*, *38*, 1102–1138.
- Sorace, A. Language and cognition in bilingual production: Will GPS show us the road ahead? *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000110](https://doi.org/10.1017/S1366728916000110).
- Starreveld, P. A., De Groot, A. M. B., Rossmark, B. M. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition*, *17*, 258–276.
- van Hell, J. G., Cohen, C., & Grey, S. Testing tolerance for lexically-specific factors in Gradient Symbolic Computation. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000122](https://doi.org/10.1017/S1366728916000122).
- Veríssimo, J. Extending a Gradient Symbolic approach to the native versus non-native contrast: The case of plurals in compounds. *Bilingualism: Language and Cognition*. doi: [10.1017/S1366728916000134](https://doi.org/10.1017/S1366728916000134).