[pre-publication draft]

PEER COMMENTARIES

What's in a grammar? Modeling dominance and optimization in contact

DEVYANI SHARMA
Queen Mary, University of London
d.sharma@qmul.ac.uk

Muysken's article is a timely call for us to seek deeper regularities in the bewildering diversity of language contact outcomes. His model provocatively suggests that most such outcomes can be subsumed under four speaker optimization strategies. I consider two aspects of the proposal here: the formalization in Optimality Theory (OT) and the reduction of contact outcomes to four basic strategies.

Muysken warns us that his use of OT "constitutes a radical departure from its use in phonology or syntax. The principles are quite different and fairly general, and their application is stochastic rather than absolute" (Section 2.1).

A closer assessment does indeed suggest that this is not an OT grammar, but rather a typology of broad processes that, somewhat confusingly, uses OT terminology. This becomes clearer if we pose two questions standardly asked of OT analyses:

(i)   Are the constraints well-formulated and well-motivated?
(ii) Does their interaction generate the observed linguistic behavior?

Let us consider the first of these questions. Constraints in OT are typically grounded in properties relating to articulation, processing, or perception (Kager, 1999). In framework-neutral (e.g. learnability) terms, grammars are only composed of elements that we accept as universal building blocks of human language. Interesting extensions, such as Bhatt and Bolonyai (2011), raise the question of what is part of the grammar but crucially are still grounded in well-established universals.

Two of the four constraints Muysken proposes, FAITHFEAT ("Features of the input must be reflected in the output") and *CSL ("Don't switch between separate languages"), are from Hogeweg (2009) and conform to the above desiderata: the first is grounded in faithfulness to the input, the second in economy of processing.

Muysken introduces two new constraints – SL1 and SL2 ("Select L1", "Select L2") – designed to capture the effects of broad language dominance. These are less well-motivated. First, selecting L1 or L2 are forms of REPAIR that can satisfy *CSL, suggesting that they may be properties of candidates, not constraints (Kager, 1999, p. 52). Second, accounts of code-switching (CS) should ideally avoid CS-specific mechanisms (Mahootian, 1993; Woolford, 1983), and in particular should avoid reference to extrinsic language labels (MacSwan, 2005), as these are political rather than grammatical constructs. In fact, SL1 and SL2 even rely on a further extra-grammatical construct: they are formulated as "Select L1" but the tableaus indicate that their intended effect is "Select L1 as the matrix language". Again, a matrix

language is a disputed construct that, if it can be described consistently at all, should be generated by, not contained within, the components of a grammar. Furthermore, Muysken specifies language choice in the input ("$W_1W_1W_1W_2W_2W_2$"), so SL1 and SL2 awkwardly require language selection to occur twice, at different levels of computation.

Existing research in OT has dealt more intuitively and parsimoniously with the question of language selection and language transfer in contact, avoiding any reference to L1, L2, or to matrix languages. In Hogeweg's (2009) model, a single activated language along with a set of target MEANINGS constitutes the input, and the optimal output (consisting of forms from either language) is the one that best matches those meanings, given a specific ranking of faithfulness constraints and *CSL. Wiltshire (2006, forthcoming) models Indian English phonologies as a shift from L1 to L2 rankings, relying exclusively on well-established, universal phonological constraints. Wiltshire uses the Gradual Learning Algorithm (GLA; Boersma & Hayes, 2001) to model language dominance. The GLA matches outputs of the grammar to outputs encountered in the environment, so more exposure to the L2 leads incrementally to less L1-like rankings and thus less evidence of L1 influence. The syntax of contact varieties can similarly be generated through re-ranking (Bhatt, 2000; Koontz-Garboden, 2004).

All of this work has shown that language contact can be efficiently modeled simply as contact between two constraint rankings, with no reference to languages. Three of Muysken's strategies – L1, L2, and L1/L2 – would appear under these approaches as properties of the environment, not of the grammar. Stochastic OT models the effect of the environment on grammatical change not as a constraint, but as incremental re-ranking to match rankings encountered in the environment.

In short, language dominance is a socio-cognitive state that might be better modeled as influencing, rather than constituting, grammars. The same may be true of other socio-cognitive correlates of language, such as lexical frequency and speech accommodation. These too affect language optimization but through an individual's social exposure to data, rather than as statements in the grammar, and are modeled as such in Stochastic OT.

The second question noted above is whether the constraints derive the observed behavior. Unfortunately, Muysken's sample tableaus do not list actual inputs or candidates, so they are difficult to assess. This may be due to his aim of characterizing broad socio-historical outcomes, such that the tableaus do not represent standard, generative OT grammars. We are therefore obligated to set aside the OT framework and instead assess in more general terms whether the four proposed strategies can generate the linguistic behavior observed in contact settings.

Before doing this, let us note one final detail of the OT formalization. Although not described as such, the proposal appears to be a "competing grammars" model with respect to intra-speaker variation. Intra-speaker variation can take the form of multiple types of contact outcomes within a single clause (e.g., Muysken's Papiamentu example) or within a single interaction (e.g., a British Asian speaker simultaneously using insertion, backflagging, and alternation; Sharma, 2011). The four proposed constraints are so general that only one outcome is generated for mixed language input for a given speaker (footnote 7 raises this issue). Muysken mentions stochastic application of constraints, but it cannot help here: stochastic ranking between two constraints can generate micro-variation (e.g. Bresnan, Deo &

Sharma, 2007), but for all four constraints to be rerankable in a single speaker, the grammar would have to have no significant difference in ranking, leading to unattested randomness in outputs. So every shift in CS type appears to require the activation of a distinct grammar (ranking). This brings with it the various pros and cons of a competing grammars approach.

The OT formalization aside, Muysken draws our attention to some interesting broad parallels across contact language outcomes. He groups these into four types of cognitive optimization: "maximize structural coherence of the first language (L1); maximize structural coherence of the second language (L2); match between L1 and L2 patterns where possible; and rely on universal principles of language processing" (abstract).

Are contact outcomes sufficiently parallel to be reducible to these four "speaker optimization strategies"?

Let us first consider the reduction to four optimization types, starting with the example of strong L2 influence. Muysken draws a parallel between the dominant role of the (erstwhile) L2 in backflagging and its similarly dominant role in Creoles with strong lexifier input, immigrant pidgins, L2-oriented mixed languages, and ethnolectal varieties. The high rank of SL2 ("Select L2 as the matrix language") can generate backflagging, as proposed. It cannot, however, also generate an ethnolectal utterance. For instance, in an ethnolectal Dutch utterance, we might find that every lexical form is Dutch, but that the final-obstruent devoicing found in indigenous Dutch is absent. Such an utterance is generated by a standard Dutch constraint ranking and lexicon, just with a lower ranking of *VDOBSCODA that matches its low rank in the speaker's L1 (see Wiltshire, 2006). This matched ranking helps reduce the speaker's cognitive load of managing two different phonologies. The optimization is at the level of PHONOLOGY, not language selection or social meaning as in the case of backflagging.

The same issue arises for strong L1 influence. Muysken groups insertional CS alongside relexification and other processes that favor a dominant role for the L1. Again, SL1 ("Select L1 as the matrix language") may generate insertional CS, but it cannot also generate relexification. This is because relexification involves the use of entirely L2 surface forms while retaining faithfulness to an abstract L1 semantic feature. Here, the optimization reduces the speaker's cognitive load of maintaining two featural representations. It occurs at the level of SEMANTICS, not language choice or social meaning as in the case of insertion.

These examples show that, although closely linked, language dominance cannot be equated with optimization. Muysken is right in noting shared properties of language dominance in the sets of outcomes grouped in Tables 1–3, but optimization is not equally shared within these sets. Language dominance is a socio-cognitive state that causes form A to be favored over form B when a speaker executes diverse optimization processes. Language dominance may even TRIGGER optimization processes at times, but the two are still distinct. Optimization can target very diverse and specific mental representations; Muysken's own extensive contributions to the field have shown that the idealized boundaries of "a language" disintegrate when we examine the fine details of contact-based restructuring in syntax, semantics, phonetics, and pragmatics. It is at these micro-levels of cognitive representation that optimization takes place. Indeed, the diversity that Muysken sees as a shortcoming in the field might simply reflect this diversity of dimensions along which human language can be optimized.

A revised model could be composed of two modules: language dominance (reducible to a few types along the lines of Muysken's proposal) and optimization (not currently reducible to a few types). The two are closely linked, in that dominance may influence the form favored in optimization processes and may trigger some such processes. We could then explore a structured typology of optimization itself (e.g. processes such as replacement, convergence, and regularization, occurring at different levels, including semantic features, phonetic features, phonological contrast, and syntactic function), and how this interacts with language dominance and other ecological factors.

Given this, Muysken's inclusion of universal principles (UP) alongside L1, L2, and L1/L2 seems awkward. A separation of dominance and optimization in the revised model outlined above would instead treat UP as part of the grammar, and the other three as dimensions of dominance. Universal principles would then always be underlyingly present, but only emergent under certain types of L1–L2 structural mismatch or limited L1 or L2 input (Sharma 2005, 2009; Wiltshire, forthcoming).

Muysken's use of the term "optimization" thus needs unpacking into distinct components. His use of the term "speaker strategies" also appears to conflate a range of community behaviors. If an Irish community has experienced language shift to English and only uses selective backflagging in Irish, there is no longer any choice on the part of the speaker to use the L1 or L2 as a matrix language (which is how the analysis is framed). Not all contact-driven outputs can be equally described as speakers' "rational decision making in interactions" or "strategic choices" (Section 6).

We are left, then, with something more akin to a new typology of contact types than a formal generative model of optimization strategies. Although the novelty of the proposal is reduced as a result, the typology makes a number of important points, drawing our attention to the systematic role of language dominance, linking this to optimization processes, and accounting for why certain groups of languages, e.g. Creoles, do not form a uniform typological class.

Finally, it is worth noting that although the model aims to incorporate social factors, the focus is on those that influence language dominance. It would be worthwhile to consider the proposal in light of a more robust social model that incorporates other social factors in contact-induced change, such as iconicity (e.g. Herbert, 2002) or indexicality (e.g. Roberts, 2004). For example, it may well be that an L1-based or an L2-based outcome develops for a particular grammatical construction because it is part of a particular social register, thus deriving not from the speaker's general language dominance but from the finer details of social arrangements.

References

Bhatt, R. M. (2000). Optimal expressions in Indian English. *English Language and Linguistics, 4,* 69–95.

Bhatt, R. M., & Bolonyai, A. (2011). Code-switching and the optimal grammar of bilingual language use. *Bilingualism: Language and Cognition, 14,* 522–546.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry, 32,* 45–86.

Bresnan, J., Deo, A., & Sharma, D. (2007). Typology in variation: A probabilistic approach to *be* and *n't* in the Survey of English Dialects. *English Language and Linguistics, 11,* 301–346.

Herbert, R. K. (2002). The sociohistory of clicks in Southern Bantu. In R. Mesthrie (ed.), *Language in South Africa*, pp. 297–315. Cambridge: Cambridge University Press.

Hogeweg, L. (2009). Word in process: On the interpretation, acquisition, and production of words. Ph.D. dissertation, Radboud University Nijmegen.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Koontz-Garboden, A. (2004). Language contact and Spanish aspectual expression: A formal analysis. *Lingua, 114,* 1291–1330.

MacSwan, J. (2005). Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language and Cognition, 8,* 1–22.

Mahootian, S. (1993). A null theory of code switching. Ph.D. dissertation, Northwestern University.

Roberts, S. J. (2004). The role of style and identity in the development of Hawaiian Creole. In G. Escure & A. Schwegler (eds.), *Creoles, contact, and language change: Linguistic and social implications*, pp. 331–350. Amsterdam: John Benjamins.

Sharma, D. (2005). Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition, 27,* 535–566.

Sharma, D. (2009). Typological diversity in New Englishes. *English World-Wide, 30,* 170–195.

Sharma, D. (2011). Return of the native: Hinglish in Britain. In R. Kothari & R. Snell (eds.), *Chutnefying English*, pp. 1–21. New Delhi: Penguin.

Wiltshire, C. (2006). Word-final consonant and cluster acquisition in Indian English(es). In D. Bamman, T. Magnitskaia & C. Zaller (eds.), *Online proceedings supplement of the 30th Boston University Conference on Language Development* (available online). Somerville, MA: Cascadilla Press

Wiltshire, C. (forthcoming). Emergence of the unmarked in Indian Englishes with different substrates. In M. Filppula, J. Klemola & D. Sharma (eds.), *The Oxford handbook of world Englishes*. Oxford: Oxford University Press.

Woolford, E. (1983). Bilingual code-switching and syntactic theory. *Linguistic Inquiry, 14,* 520–536.