Chapter 8

Rethinking the reach of categorical constraints: The final-over-final constraint and combinatorial variability

Neil Myler
Boston University

This squib argues that categorical rules and constraints of the sort traditionally found in generative syntax can, in principle, make interesting and testable quantitative predictions about surface frequencies in language use, despite occasional claims to the contrary. Specifically, the final-over-final constraint (FOFC, Biberauer et al. 2014; 2009; Holmberg 2000; Walkden 2009; many others) is predicted to exert a specific influence on the likelihood of OV vs. VO word order in the language use of a speaker that allows both, given a COMBINATORIAL VARIABILITY approach to intra-speaker syntactic variation (Adger 2006 et seq.).

1 Introduction

Generative linguistics has traditionally employed categorical rules and constraints in its quest to understand the properties of the syntax of particular languages and the properties of the syntactic component of the language faculty more generally. For this reason, its theoretical postulates have often been taken to be either irrelevant to or at odds with the inherent variability of language use (see Guy 2005; Newmeyer 2005; inter alia).

In this squib, I will argue that categorical constraints can, in fact, make interesting and testable quantitative predictions about surface frequencies, given a certain theory of how intra-speaker syntactic variation is to be modeled. More

Neil Myler. 2020. Rethinking the reach of categorical constraints: The final-over-final constraint and combinatorial variability. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 137–147. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280641

specifically, I will show that the *final-over-final constraint*[1] (FOFC – Biberauer et al. 2014; 2009; Holmberg 2000; Walkden 2009, many others) should exert a specific influence on the likelihood of OV vs. VO word order in the language use of a speaker that allows both, given a *combinatorial variability* approach to intra-speaker syntactic variation (Adger 2006 et seq.).

The squib is structured as follows. In §2, I introduce the combinatorial variability approach, showing how it might be used to generate predictions concerning the expected baseline surface frequencies of OV vs. VO order in the speech of Quechua–Spanish bilinguals, focusing on DP complements and the head-directionality of VP and TP. In §3, I introduce FOFC and demonstrate that the surface frequencies predicted by the combinatorial variability approach change if FOFC is held to be valid. In §4, I outline the prospects and challenges for testing these predictions in a sociolinguistic study of actual Quechua–Spanish bilinguals in Cochabamba, Bolivia. §5 is a brief conclusion.

2 Quechua–Spanish contact and combinatorial variability

To make the discussion of combinatorial variability more concrete, I will frame this section around the specific example of language contact between speakers of Quechua and Spanish. Speakers of these two languages are in contact in Peru, Bolivia, Ecuador, parts of Colombia, and parts of northern Chile and northern Argentina. Many Quechua speakers in these places are bilingual in Spanish. As is well-known, Quechua and Spanish are almost typological opposites in terms of their basic word order. Quechua is predominantly head-final, as shown in the example from Cochabamba Quechua (a Bolivian variety) in (1). Spanish, on the other hand, is a head-initial language, as shown in (2).

(1) Cochabamba Quechua
Kay runa Cochabamba-man ri-q ka-rqa.
This man Cochabamba-to go-NMLZ be-PST
'This man used to go to Cochabamba.'

(2) Spanish
Este hombre ha ido a Cochabamba.
This man has gone to Cochabamba
'This man has gone to Cochabamba.'

[1]Note that FOFC is referred to as the *final-over-final condition/constraint* in some more recent work, including Sheehan et al. (2017).

Pre-theoretically, one might expect contact between Quechua speakers and Spanish speakers to give rise to mutual influence on word order, such that head-initial orders increase in Quechua usage, and/or head-final ones increase in Spanish usage, depending on the degree of bilingualism of the speaker, attitudes towards each language, and so on. Indeed, such has been reported in the literature on Andean Spanish (e.g., Muntendam 2008; Muysken 1984; Sánchez 2003) and in studies of the influence of Spanish on Quechua (Camacho 1999; Hintz 2009; Sánchez 2003, 2012). Let us now turn to the combinatorial variability approach, and how it might analyze such variation.

Comparative syntax research within the Minimalist program has pursued the idea that syntactic variation across languages/dialects should be analyzed only in terms of variation in the featural needs of functional items (the so-called Borer-Chomsky conjecture, as it is dubbed by Baker 2008; see Borer 1984; Chomsky 1995). This presents a generativist pathway to *orderly heterogeneity* in the sense of Weinreich et al. (1968): Suppose that an individual's lexicon contains function morphemes with the same categorial feature and the same contribution to truth conditions (and thus roughly the same distribution), but which differ in one or more of their morphosyntactic features. Then, the choice of one or the other lexical item in a derivation will result in somewhat different outputs, but with no difference in meaning. Thus, there will be an appearance of syntactic optionality, but in reality the only optionality is in lexical choice: once particular lexical items have been chosen, the syntactic derivation is fully determined. This is the essence of Adger's (2006 et seq.) proposed reconciliation of Minimalist syntax with sociolinguistic variation.

As Adger (2006) points out, it is possible to calculate quantitative predictions about variability which arise from the combinatorics of the relevant syntactic elements (hence the name *combinatorial variability* for the overall approach). Take lexical items A, B, and C; all with identical truth-conditional meaning but with distinct syntactic features. A and B, when chosen, give rise to a series of derivational steps S_1. C, on the other hand, differs in some aspect of its feature content from A and B, and thus gives rise to a distinct derivation S_2, whose output differs on the surface from S_1. This will give the appearance of syntactic variability. All else held equal, a prediction is made about the nature of that variability. Since two out of a possible three lexical choices give rise to S_1, but only one choice yields S_2, the prediction is that the output corresponding to S_1 should appear in usage two thirds of the time, and the output of S_2 should appear one third of the time.[2]

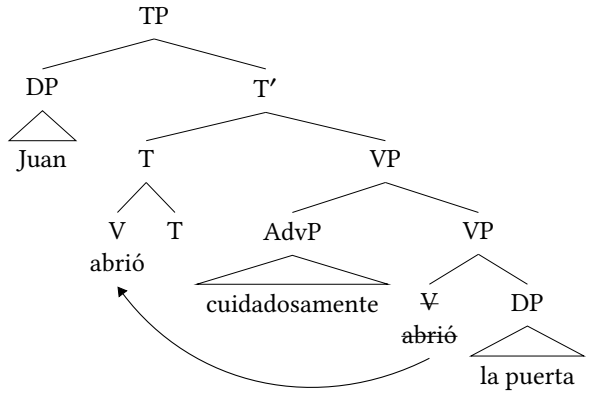
[2]This follows only if no other factors favor A, B, or C over the others, so that the choice is determined by chance. In actual use, of course, the probability distribution predicted by purely syntactic combinatorics will be modulated by sets of factors influencing lexical choice itself, including sociolinguistic factors. I return to this issue below.

Returning to our example from Quechua–Spanish contact, we will now examine the baseline frequencies of OV and VO word order that a combinatorial variability approach would predict. First, we need an inventory of the syntactic microparameters that are relevant to analyzing word-order differences between the two languages.

The first is HEAD-DIRECTIONALITY OF THE VP.[3] In Spanish, the head of VP is on the left (this value will be denoted "L" for short). In Quechua, the head of the VP is on the right ("R" for short).

The second parameter is HEAD-DIRECTIONALITY OF THE TP. This parameter, of course, is directly analogous to the first. Spanish T is on the left, and Quechua T is on the right. This parameter has a direct influence on where the verb surfaces relative to its complement, because T in these languages attracts the verb (i.e., there is V-to-T movement). V-to-T movement is known to apply in Spanish because of the placement of VP-peripheral adverbs relative to the verb and the direct object (Pollock 1989; Zagona 2002).[4]

(3) Spanish
Juan abrió cuidadosamente la puerta.
Juan opened carefully the door
'Juan carefully opened the door.'



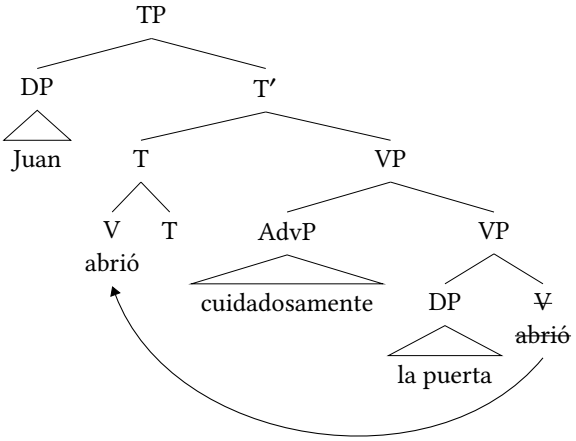
[3]For simplicity I will assume the traditional head parameter in the ensuing discussion, but nothing I have to say is incompatible with an antisymmetric approach to the relationship between structure and linearization (see Kayne 1994). Since Kayne's linear correspondence axiom is a key component of many existing approaches to deriving FOFC, this is good news.

[4]I assume here that T is the relevant landing site in all cases, but this is certainly an oversimplification. See Schifano (2015; 2018) for evidence that considerably more granularity is needed, with verb movement targeting different positions in the Cinquean extended IP (Cinque 1999 et seq.) in different languages. This does not affect the main point here, so long as verb movement is to a landing site higher in the structure than the final position of the direct object. Thanks to an anonymous reviewer for raising this issue.

It is much more difficult to ascertain whether or not there is V-to-T movement in Quechua, since both VP and TP are head-final in that language, and this makes it impossible to check whether the verb "crosses over" adverbs at the edge of VP. The empirical evidence we have to hand is therefore compatible with V-to-T movement being present or absent in Quechua. However, there is one typological consideration which weighs in favor of assuming that Quechua does have V-to-T movement. The syntactic literature has found that VO languages with rich agreement inflection on the finite verb always have V-to-T movement (Kosmeijer 1986; Pollock 1989; see Koeneman & Zeijlstra 2012 for a recent reaffirmation of this correlation). Since Quechua has extremely rich agreement inflection on its finite verbs, we may assume it has V-to-T movement also.[5]

To see why this matters for surface word-order, consider the case of a derivation in which VP-headedness has the Quechua "R" value, but TP-headedness has the Spanish "L" value. In such a case, the surface word order will be VO in spite of the fact that the structure is "underlyingly" OV, because of V-to-T movement.

(4) V-to-T movement obscures head-finality of VP



Given these basic assumptions about clause structure and the points of parametric variation which differentiate Spanish and Quechua, we can now ask about the predictions of combinatorial variability for the baseline frequencies of OV vs. VO order.

[5] An anonymous reviewer points out that there remain a number of potential problems for this conclusion (referring to Vikner 2005; Han et al. 2007; 2016). This must be borne in mind, because if it turns out that Quechua lacks V-to-T, then another test-bed for the quantitative predictions of FOFC would need to be found. The broader point of this squib, that such predictions are formulable and testable in principle, stands regardless.

Let us assume that a bilingual speaker is able to represent syntactic objects from each language in much the same way as a monolingual speaker. That is, a bilingual speaker has access to a left-headed VP structure much as a monolingual Spanish speaker does, and also has access to a right-headed VP structure in the same way that a monolingual Quechua speaker does. Similarly, the bilingual's functional lexicon will contain a lexical item T which takes its complement to its right, Spanish-style, and another lexical item T which takes its complement to the left, Quechua-style, and so on for other syntactic objects. Of course, in making utterances, bilingual speakers will have to make a choice between these options. It turns out that the different parameter settings discussed above, simply through the nature of their logically possible combinations, give rise to quantitative predictions about what the baseline frequencies of these different choices should be.

For the purposes of simplicity, I will concentrate on DP direct objects only. The calculations below would have to be somewhat different for QP and CP complements. In the case of QPs, the fact that Quechua allows overt scrambling for scope would somewhat increase the chance of OV order surfacing, relative to non-quantificational DPs. For CPs, the possibility of clausal extraposition in both languages would boost the predicted baseline frequency of VO order.

There are $2 * 2 = 4$ possible combinations of parameter settings relevant here, shown below.

(5) Combinations of parameter settings: DPs

Combination A	Output: VO
Parameter	Setting
VP-headedness	L
TP-headedness	L

Combination B	Output: OV
Parameter	Setting
VP-headedness	R
TP-headedness	R

Combination C	Output: VO
Parameter	Setting
VP-headedness	R
TP-headedness	L

Combination D	Output: OV
Parameter	Setting
VP-headedness	L
TP-headedness	R

Hence, the logically possible combinations predict a 50/50 split between VO orders and OV orders for DPs.

(6) VO vs. OV order with DP complements
VO = 2/4 outputs = 50%
OV = 2/4 outputs = 50%

3 Bringing in the final-over-final constraint (FOFC)

The *final-over-final constraint* of Biberauer et al. (2014: 171) has an interesting effect on this calculation.

(7) *The final-over-final constraint* (FOFC)
A head-final phrase αP cannot dominate a head-initial phrase βP, where α and β are heads in the same extended projection.

This constraint will, of course, make the categorical prediction that V-O-Aux orders will be absent from compound tenses in the Spanish and the Quechua of bilinguals. In addition, however, FOFC has a quantitative effect. In particular, it rules out combination D in (5), because that combination involves a head-final TP dominating a head-initial VP. In terms of the predicted baseline surface frequencies, we thus obtain the following results instead of the ones we saw in (6):

(8) VO vs. OV order with DP complements (if FOFC is valid)
VO = 2/3 outputs = 67%
OV = 1/3 outputs = 33%

This is an exciting finding, because it shows that categorical constraints can give rise to stochastic effects, meaning that such constraints *are* of potential relevance to variationist work after all. This result emerges from the fact that combinatorial variability derives quantitative predictions by looking at the interaction of different parameter settings, and universal constraints like FOFC take certain combinations of parameter settings out of the picture. Another intriguing consequence of this result is that it becomes possible, in principle, to use variationist data to test the predictions of such universal constraints. Since the baseline frequencies predicted are different if FOFC holds than they are if it does not, in principle it becomes possible to test FOFC by seeing how the variationist data pan out. In the next section, I examine the prospects for doing this.

4 Testing the predictions: Prospects and challenges

It is clear what the signature of FOFC should be in quantitative data: because FOFC bars one of the logically possible routes to OV word order, OV should be less common than VO all else held equal if FOFC is valid. If FOFC is not valid, then OV and VO should be equally frequent, all else held equal.

The challenge in testing predictions of this sort, of course, is that all else is seldom equal, and a range of social factors that have been discussed in the sociolinguistics literature will also influence the actual surface frequencies of the orders. These must be controlled for or accommodated somehow if the signature of FOFC is to be detected. Most obviously, although the literature reports mutual influence between Spanish and Quechua word orders, it still might be the case that speakers have some (presumably subconscious) sense that Quechua exhibits more head-finality. If so, language mode would be expected to favor OV when the speaker is talking in Quechua, and VO when the speaker is talking in Spanish. Such an effect would be especially likely if the VO vs. OV difference turned out to be a socially salient linguistic variable.

The issue of social salience raises the possibility that speakers might use OV vs. VO order as a way of indexing particular identity categories, including attitudes to Quechua and Spanish, orientation towards or away from indigenous culture, and so on. Since exposure to standard Spanish will favor VO order, degree of education is another factor to be considered. In addition, of course, degree of bilingualism/proficiency in each language would be expected to be relevant.

Finally, there is a presupposition of the combinatorial variability approach which itself has yet to be tested; namely, the idea that the probability that a given variable will be used is determined by chance if no other factor intervenes. This assumption is not unreasonable, but nor is it certain to be correct – we still await an empirical demonstration that it is on the right track.

In an ongoing collaboration, the sociolinguist Daniel Erker and I have carried out a pilot study involving demographic/attitudinal surveys, sociolinguistic interviews, reading passage data, and grammaticality judgments on both Spanish and Quechua as spoken in Cochabamba, Bolivia. The data set includes 19 speakers: 4 monolingual Spanish speakers, and 15 Quechua–Spanish bilinguals. For the bilinguals, we have interview data, reading passage data, and grammaticality judgment data on both languages. The analysis of this data is still in progress. As well as addressing a number of issues in the sociolinguistics of language contact, we hope that a full version of this study (including monolingual Quechua speakers, and many more speakers overall) will allow us to test the quantitative predictions of FOFC, and the predictions of the combinatorial variability approach more generally.

5 Conclusion

This squib has shown that categorical principles and constraints can make predictions about apparently non-categorical phenomena. Testing those predictions, however, is a difficult and delicate task, one that is not yet within our reach from a practical standpoint. Bringing it within our reach will require the collaboration of formal linguists and sociolinguists.

Abbreviations

FOFC	final-over-final condition/constraint	NMLZ	nominalizer/nominalization
		PST	past

Acknowledgements

For Ian Roberts, the syntactic Eric Cantona.

This squib is an outgrowth of ongoing joint work with Daniel Erker on Quechua–Spanish contact, including a poster we jointly presented at the annual meeting of the Linguistic Society of America in 2016. I would like to thank the audience of that poster presentation, a colloquium audience at UPenn, anonymous reviewers from the National Science Foundation, David Adger, Byron Ahn, Carol Neidle, Cathy O'Connor, and Ian Roberts for their feedback on various aspects of my and Daniel's joint work. Thanks also to two anonymous reviewers and to Daniel Erker for their comments on an earlier version of this squib. Errors and infelicities are on my head alone.

References

Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42(3). 503–530. DOI: 10.1017/S002222670600418X.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–373. Amsterdam: John Benjamins. DOI: 10.1075/la.132.16bak.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.

Biberauer, Theresa, Glenda Newton & Michelle Sheehan. 2009. Limiting synchronic and diachronic variation and change: The Final-over-Final Constraint. *Language and Linguistics* 2009(10). 701–743.

Borer, Hagit. 1984. *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris.

Camacho, José. 1999. From SOV to SVO: The grammar of interlanguage word order. *Second Language Research* 15(2). 115–132. DOI: 10.1191/026765899673532714.

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

Guy, Gregory. 2005. Grammar and usage: A variationist response. *Language* 81(3). 561–563. DOI: 10.1353/lan.2005.0132.

Han, Chung-hye, Jeffrey Lidz & Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38(1). 1–47. DOI: 10.1162/ling.2007.38.1.1.

Han, Chung-hye, Julien Musolino & Jeffrey Lidz. 2016. Endogenous sources of variation in language acquisition. *PNAS* 113(4). 942–947. DOI: 10.1073/pnas.1517094113.

Hintz, Daniel J. 2009. Reordenamiento sintáctico en construcciones verbales analíticas del quechua por el contacto con el castellano. *Lingüística* 22. 187–201.

Holmberg, Anders. 2000. Deriving OV order in Finnish. In Peter Svenonius (ed.), *The derivation of VO and OV*, 123–152. Amsterdam: John Benjamins. DOI: 10.1075/la.31.06hol.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Koeneman, Olaf & Hedde Zeijlstra. 2012. One law for the rich and another for the poor: The Rich Agreement Hypothesis rehabilitated. lingbuzz/001462.

Kosmeijer, Wim. 1986. The status of the finite inflection in Icelandic and Swedish. *Working Papers in Scandinavian Syntax* 26. 1–41.

Muntendam, Antje. 2008. Crosslinguistic influence in Andean Spanish: Word order and focus. In Melissa Bowles, Rebecca Foote, Silvia Perpiñán & Rakesh Bhatt (eds.), *Selected proceedings of the 2007 Second Language Research Forum*, 311–322. Somerville, MA: Cascadilla Proceedings Project.

Muysken, Pieter. 1984. The Spanish that Quechua speakers learn: L2 learning as norm-governed behavior. In Roger Andersen (ed.), *Second languages: A crosslinguistic perspective*, 101–119. Rowley, MA: Newbury House.

Newmeyer, Frederick J. 2005. *Possible and probable languages*. Oxford: Oxford University Press.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Sánchez, Liliana. 2003. *Quechua-Spanish bilingualism: Interference and convergence in functional categories*. Amsterdam: John Benjamins. DOI: 10.1075/lald.35.

Schifano, Norma. 2015. *Verb movement: A Pan-Romance investigation*. University of Cambridge. (Doctoral dissertation).

Schifano, Norma. 2018. *Verb movement in Romance: A comparative study*. Oxford: Oxford University Press.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts & Anders Holmberg. 2017. *The Final-over-Final Condition: A syntactic universal*. Cambridge, MA: MIT Press.

Vikner, Sten. 2005. Immobile complex verbs in Germanic. *The Journal of Comparative Germanic Linguistics* 8(1). 85–117. DOI: 10.1007/s10828-004-0726-9.

Walkden, George. 2009. Deriving the Final-over-Final Constraint from third factor considerations. *Cambridge Occasional Papers in Linguistics* 5. 67–72.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 95–195. Austin, TX: University of Texas Press.

Zagona, Karen. 2002. *The syntax of Spanish*. Cambridge: Cambridge University Press.