ON THE 'MAFIOSO EFFECT' IN GRAMMAR
Theresa Biberauer[1,2], Ian Roberts[1] & Michelle Sheehan[1]
University of Cambridge[1] and Stellenbosch University[2]

There is an obvious tension in the Minimalist Program between the desire to posit a minimally genetically specified syntactic component whilst maintaining the empirical insights of the GB era, notably the fact that variation is constrained and structured (in that certain logically possible options are simply never attested). In this talk, we address this challenge, building on Chomsky (2005) by proposing that this 'structured variation' emerges because of UG-external forces such as (i) the nature of the PLD, (ii) system-internal pressure, (iii) acquisition biases and (iv) processing pressures. Our central proposal is that these forces trigger 'Mafioso Effects' whereby only one of the options associated with a given (emergent) parameter is ever actually attested, i.e. this option is effectively one that cannot be refused. As such, certain GB principles can be rethought as 'no-choice parameters', with crosslinguistic gaps and skewings resulting from certain parameter settings being strongly preferred/dispreferred as a result of (i)-(iv).
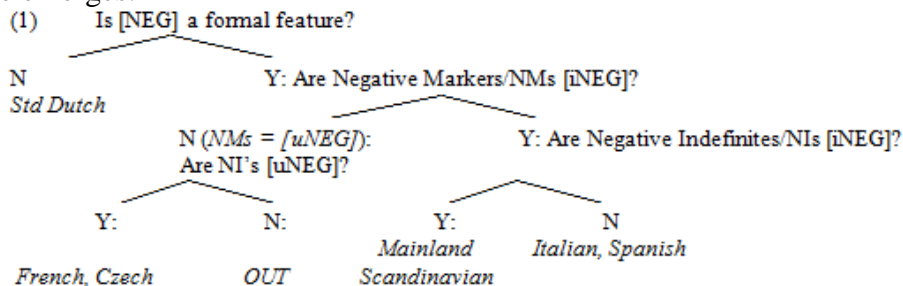
Take, for example, Kayne's (1994) Linear Correspondence Axiom (LCA). While the numerous left-right asymmetries of natural language (Greenberg's Universal 20, the ban on rightwards *wh*-movement, the Final-over-Final Constraint; see Cinque 2007, Kayne 2012) are manifest, the correct explanation for them remains controversial. One oft raised objection to LCA-based explanations is that there is no deep reason why asymmetric c-command should map to precedence rather than subsequence. On the Mafioso approach, however, the LCA is simply a linearization parameter, with the subsequence/precedence option requiring setting during acquisition. That precedence always emerges as the selected option is the consequence of the processing-shaped PLD, with processing pressures of the type discussed by Neeleman & van de Koot (2002) and others, notably filler-gap relations, being crucial here. As such, the LCA can be considered an emergent property of language: linearization must rely on independently attested syntactic relations of the relevant (asymmetric) kind (Kayne 1994), and this asymmetric relation is mapped to precedence for syntax-external reasons. Moreover, the precedence setting implies that all movement is leftward, including cases where processing cannot explain the leftward preference (e.g. VP-remnant topicalisation in German, where leftward movement arguably *introduces* processing challenges; Den Besten & Webelhuth 1989), and also that first-merged specifiers will be leftward, which does not follow directly from the processing account. Clearly, then, structural precedence phenomena cannot just be reduced to the effects of processing.

This account of the LCA implicitly assumes that all languages have filler-gap relations, i.e. movement. While the basic combinatorial operation (Merge) makes internal merge available in all languages, its actual application in a given context, we assume, results from the presence of a UG-given movement diacritic ^ which may be variously associated with a given feature/head. As the presence of ^ is in principle optional, it remains unclear why all languages should have to employ it. We propose that this too is a Mafioso Effect. Consider for example Alexiadou & Anagnostopoulou's (2001) observation that either the external or the internal argument must vacate vP. Chomsky's (2013) account of this effect is that it is forced by the interface-driven need for labels: X in $[_X EA\ [v\ VP]]$ requires a label; as, by assumption, discontinuous elements cannot supply a label, EA-movement will eliminate EA as a potential label for X, with IA- and, we argue, "VOS"-style VP-movement (Massam 2001), similarly facilitating labelling. The precise location in the higher phase of ^ is, as noted above, a parametric option, conventionally fixed via exposure to the PLD, but the need for at least one ^ is a universal property of linguistic systems which is not UG-specified, but rather forced by system-internal pressures.

We also see more sophisticated Mafioso Effects of this kind in Case/alignment contexts. According to Aldridge (2004, 2008), syntactic ergativity results where a v assigning theta-related ERG Case to its specifier also bears ^ triggering object movement past the subject, thus ruling out A-bar

extraction of DP_{ERG}. In morphologically ergative languages like Basque, unergative v also assigns ERG, yielding a morphologically ergative split-S (or "stative-active") system. What appears to be ruled out, though, is a *syntactically* ergative split-S system (Deal 2012). In our terms, this too is a Mafioso Effect. In cases where unergative v assigns ERG and all ERG-assigning heads are associated with ^, the result is a derivation which can never converge (there is no XP which can raise to satisfy v's ^).

Following Gianollo et al. (2008), we assume acquisition to entail i.a. the determination of which features are grammaticalised (participate in Probe/Goal relations) in a given language, and how these formal features interact with ^. The 'sequence' in which these facts are established is guided by restricted UG-specified elements (the availability of a [uF]/[iF] distinction, ^, the operations Merge and Agree) and 3rd factor-imposed acquisition strategies, including a version of Feature Economy/FE and Input Generalization/IG (Roberts & Roussou 2003, Roberts 2007). By the former, acquirers posit as few formal features as possible; by the latter, they assume the minimum number of distinct elements/operations compatible with the PLD, maximally generalising input patterns. The nature of the PLD, though, excludes certain potential parametric options. Consider the case of negation. In terms of the system in Biberauer & Zeijlstra (2012), the child must establish whether negation is grammaticalised, which classes of negative elements are specified [iNEG] and [uNEG], and whether an abstract [iNEG]-encoding negative operator is required (Ladusaw 1992). Assuming the child to follow the "learning path" given by the emergent parametric hierarchy in (1), this being determined by the interaction of the minimally specified UG proposed above and what is independently known about the salience of different types of negation elements (Klima & Bellugi 1966 *et seq.*), a further Mafioso choice emerges:

(1)   Is [NEG] a formal feature?

N
Std Dutch
               Y: Are Negative Markers/NMs [iNEG]?

N (*NMs = [uNEG]*):
Are NI's [uNEG]?
               Y: Are Negative Indefinites/NIs [iNEG]?

Y: N: Y: N
                            Mainland   Italian, Spanish
French, Czech   *OUT*   Scandinavian

Here languages with [uNEG] NMs, but [iNEG] NIs are ruled out as there is no unambiguous input leading to the postulation of this system-type (Double Negation structures only unambiguously signal [iNEG] in all-[iNEG] systems of the Mainland Scandinavian type; DeSwart & Sag 2002), and credible 3rd factor motivations (FE, IG, and the general biases discussed by Pearl (2012)) also work against it: everything, then, pushes the acquirer towards extending the previously established [uNEG] analysis of NMs to NIs, leading to the seemingly correct prediction that mixed negation systems with [uNEG] NMs and [iNEG] NIs cannot exist.

We also discuss cases where two emergent choices produce superficially identical outputs that cannot be distinguished, with implications for the synchrony and diachrony of verb-movement, and, more generally, the nature of choices located at the "bottom" of the learning path-defining parametric hierarchies resulting from the interplay of the minimal UG we assume, the processing-shaped PLD, and the partially 3rd factor regulated acquisition biases of learners. What emerges from the discussion as a whole is that "emergent" parameter hierarchies are restricted by a range of 1st, 2nd and acquisitional and computational 3rd factor considerations. In short, there will be many parametric "offers that cannot be refused", a state of affairs that enhances the explanatory power of a model of the proposed type, while minimising, but crucially not eliminating, the role of UG.