EMERGENT PARAMETERS AND SYNTACTIC COMPLEXITY: NEW PERSPECTIVES

Theresa Biberauer[1,2] & Ian Roberts[1]
University of Cambridge & Stellenbosch University
mtb23@cam.ac.uk & igr20@cam.ac.uk

Background: Here we develop ideas in Biberauer, Holmberg, Roberts & Sheehan (2014, BHRS) concerning the formal complexity of linguistic systems viewed from the perspective of an emergentist approach to parametric theory, and its implications for language acquisition and change. Our central idea is that, rather than postulating a rich UG-specified parametric endowment, parameters are emergent properties falling out of the interaction of Chomsky (2005)'s 3 factors: a minimally specified UG (F1), the PLD (F2), and non-language-specific cognitive optimization strategies (F3). Crucially, under F3 we assume a general cognitive economy principle, Maximise Minimal Means, which has 2 major language-oriented reflexes: Feature Economy (FE: postulate as few formal features as possible) and Input Generalization (IG: generalize features as much as possible). Together FE and IG constitute a minimax search/optimization procedure. The goal of this paper is to consider this approach's consequences for how we understand the complexity of grammatical systems in relation to language acquisition and change, focusing in particular on a subset of seemingly privileged formal features.

The proposal: Key to the proposal is the idea that FE and IG combine to create a learning path of the general type in (1): postulate NO features (satisfies FE and IG); if a feature F is detected, posit it in ALL (relevant) domains (satisfies IG but not FE); if F is absent in expected parts of the PLD (given the previous step), posit it only in SOME relevant domains.

(1) Is F present?
- **NO**
- Yes: Is F present on **ALL** heads?
  - Yes
  - No: F and not-F are present (**SOME**)

The first NO is a default: F is only postulated if PLD points to its existence. The last step creates a distinction between domains where F is present and where it is absent, thereby effectively creating a new feature distinction (cf. also Dresher 2009, 2013, where essentially the same idea is applied in phonology, and Jaspers 2013, Seuren & Jaspers 2014 for application in the domain of concept formation). After the last step, the NO>ALL>SOME procedure is repeated for the restricted version of F, and for not-F (i.e. G). Equating parameters with (a subset of) formal features which are open to cross-linguistic variation (Chomsky 1995), this produces parameter hierarchies of a highly constrained and hence comparable kind. We will demonstrate how (1)-type hierarchies work for 4 central features which seem to regulate many properties, i.e. they are pleiotropic, in the sense familiar from Genetics: **Person**, **Tense**, **Case** and **Order**. Call these **Pleiotropic Formal Features** (PFFs).

PFFs have "strong" and "weak" variants: a strong PFF controls more formal features (FFs) and acts in more formal domains than a weak one. If **Person** is strong, it controls the properties of FFs like Number, Gender, etc.) and also plays a role in multiple domains (potentially, all phasal (sub)domains, across all categories); if it is weak, Person is simply instantiated with its standard values (1st, 2nd, possibly 3rd), and other φ-features are not grammaticalised (thus not participating in Agree relations, i.e. they are present only as semantic features; Wiltschko 2014). Similarly, **Tense** can be weak or strong. If strong, verb-movement into the higher inflectional field is found, along with restricted VP-ellipsis and few or no auxiliaries; Tense will also function in numerous domains (cf. Ritter & Wiltschko 2014 on CP-Tense, Pearson 2001 on vP-Tense, and Nordlinger & Sadler 2004 on nominal Tense). Moreover, strong Tense controls further FFs (e.g. future, modal and aspectual features). The position of the verb may thus be relativized to the nature and realisation of other FFs in an intricate way (cf. Schifano 2015 on Romance). If Tense is weak, we see no verb movement to "high" clausal positions, a relatively rich auxiliary system and more liberal VP-ellipsis, with modal and aspectual Fs either not grammaticalised or functioning independently of Tense, which simply has the values Past/Non-Past. **Case** regulates positional argument "licensing" (Vergnaud 1997/2008). Strong Case is associated with the presence of lexical, inherent and/or quirky Case, a rich inventory of Case features, and a range of domains in which it is active (e.g. CP and vP besides the usual nominal and TP-domains). Weak Case entails either an undifferentiated F which merely functions to make arguments active for Agree (as in Chomsky 2001), or a minimally distinct Nominative-Accusative clausal opposition (and possibly Genitive in DP). Strong Case is associated with verbal semantics (argument structure, aspect) and can determine DP semantics (various kinds of partitive case, genitive of negation, specificity, focus, etc.); Weak Case simply licenses arguments in given positions, thereby restricting the range of argument positions (as observed by Vergnaud; cf. GB Case theory). Weak-Case systems are likely to have a richer array of adpositions, especially semantically empty "linker" elements like English *of*, whose sole role is argument-licensing. Finally, **Order** contributes to interpretation in that basic word order identifies unmarked interpretations, thereby serving as a reference point for non-neutral, discourse-marked structures (extended Duality of Patterning). Weak Order means that constituents are linearised without the need to postulate an additional F, e.g. ^ in Biberauer, Holmberg & Roberts (2014), i.e. head-initially. Strong Order requires the presence of movement-triggering Fs, giving rise to fully harmonic head-final order and, in keeping with Maximise Minimal Means, scrambling. In (1), "strength" thus corresponds to the ALL option, and "weakness", a relative and gradient property, either to NONE (Order in fully head-initial systems) or to various sub-options under SOME.

Evaluating complexity: As we move downwards along (1)-type hierarchically defined routes, parameters become more "micro", behaving in a non-uniform, differentiated fashion which is inherently more complex and governed by more Fs than higher systems. The higher options are inherently preferred by the acquirer because FE and IG favour them, absenting PLD forcing more articulated options. Both strong and (default) NO-type weak PFFs can therefore produce non-complex grammatical (sub)systems and might thus be expected to be acquired readily (Tsimpli 2014). Where the systematic nature of more restricted options becomes obscure, we assume a tipping point (Yang 2013), beyond which acquirers no longer postulate an F-defined parametric domain, but begin to learn exceptions. Where minority options of this type rely on low-frequency components of the PLD, we predict IG-conditioned overgeneralisation, which may, in turn, trigger the loss of such options, producing a less complex system. This perspective highlights the need to distinguish not only differing types of parametric complexity (different types of strong vs weak systems, as above), but also to consider the distinction between lexical-item- and feature-based complexity: in parametric terms, a system with lexical exceptions is as simple as one lacking this complication and simpler than one which requires an F-specific rule (the pre-tipping point system). The present approach makes various predictions regarding the relative complexity of systems of different types, which we will attempt to verify. One question is the extent to which all 4 features may pattern together. Since strong PFFs must have robust morphosyntactic exponence to be acquired, we can understand the cross-linguistic tendency for morphologically rich (especially agglutinating) languages to be head-final, while analytic languages tend to be head-initial: this follows from "harmonic" setting of the 4 PFFs to strong and weak, respectively. More generally, following the method introduced in BHRS, we take each successive hierarchical step to be half as probable as the preceding one, and thereby quantify the morphosyntactic complexity of a system in relation to the application of (1)-type hierarchies to the PFFs. We consider the implications of this exercise for both language acquisition and language change.