Word-order variation and change in systems that maximize minimal means

This paper revisits the question of “basic” vs. “marked” word-order in the context of the so-called emergentist generativist approach to syntactic variation and change (Biberauer 2011 et seq.).

Our point of departure is Chomsky’s (2005) Three Factors model (1). Here UG is minimally specified, crucially lacking pre-given parametric options like those traditionally assumed during the classic parametric era (Biberauer 2008, Biberauer & Roberts 2016, 2017). Neither the classic Head Parameter nor Kaynean or other proposed alternatives to it are hard-wired, therefore. Similarly, UG is not assumed to entail a universal inventory of formal features from which “one-time selections” are made (pace Chomsky 2001:10). Instead, acquirers are assumed to tap into systematic departures from Saussurean arbitrariness to establish the identity of the formal features defining their target grammar. More specifically, we harness Chomsky’s (1995) semantic ([S]), phonological ([P]) and formal ([F]) features to distinguish “basic” vs “higher” degrees of arbitrariness in human language: (i) lexically stored, idiosyncratic conventionalized sound-meaning mappings involving [P]- and [S]-features, and (ii) grammatically regulated and thus more systematic conventionalized sound-meaning mappings requiring the postulation of [P]- and/or [S]- and [F]-features. Systematic departures in the PLD from one-to-one form-meaning mappings like those in (2) signal the presence of [F]s, i.e. grammaticalised [S] and/or [P]-features (cf. also Zeijlstra 2008, Fujimori 2011, Déchaine 2015). Crucially, the postulation of [F]s is constrained by the general cognitive bias to Maximise Minimal Means (MMM). In the domain of language, MMM underlies the acquisition biases in (3-4).

This has consequences for our understanding of word-order acquisition, variation and change. First, “basic” word order is acquired very early (Wexler 1996, Tsimpli 2014), and much work points to prosodic (i.e. exclusively [P]-type) properties, to which neonates are highly attuned, being key (see i.a. work by Jacques Mehler and colleagues). Minimal means thus provide the stepping-stone into grammar. Since the linearization convention underlying “basic” word order instantiates a systematic idiosynchrony, it must be captured by the postulation of a suitable [F]: we employ I[initial] and L[ast] for expository convenience. Once present, MMM predicts that [I/L] will be the basis for further grammar structuring. Given (3-4), we expect ‘categorial harmony’ effects (Hawkins 1983), and, more generally, monotonicity effects within Extended Projections (e.g. all-[I/L] clausal or nominal projections, or all hierarchically contiguous heads within a given domain sharing [I/L]), with predictable opportunities for ‘disharmony’ (Dryer 1992, Biberauer, Holnberg & Roberts 2014, Biberauer 2017): monotonicity effects follow where properties are distributed across natural classes (4), with larger classes requiring the postulation of fewer [F]s (3); see (5-6). As “larger” (more macro) choices require fewer [F]s, we expect greater stability, which seems true: rigid head-finality seems very stable, for instance, while West-Germanic-style OV is less so (Biberauer & Roberts 2012). Change in the direction of “smaller” (more micro) choices is not ruled out, however; the expectation is that such change will reference [F]s that are already present in the system to account for other grammatical regularities. OV-loss/-restriction, for example, frequently seems to involve states where OV is limited to [negative]-, [topic]- and/or [focus]-marked objects or, in systems with prominent aspect- or tense-based structuring, to [aspect]-/tense-defined domains.

MMM further leads us to expect that the fixing of “basic” word order (e.g. a specific type of VO (V-in-situ, V-to-T, etc.) or OV (“rigid” OV, West Germanic-type OV, OVX-type OV)) will provide a reference point in relation to which other ordering patterns (e.g. (V)-to-T-to-C) may be interpreted. This is essentially “higher-level” duality of patterning (Hockett 1960), in terms of which meaningless “basic” word-order choices (OV vs VO), which are fixed early, pave the way for other meaningless obligatory-filling choices (e.g. V spellout position, Spec-TP, Spec-CP, etc.), and, crucially, also meaningful optional movements (e.g. T-to-C in English, the nature of the XP that raises to Spec-CP, etc.). Since (i) “basic” OV/VO (head-directionality) effectively “comes for free” via prosody in systems of the kind investigated by Mehler and colleagues, and (ii) higher-level duality of patterning permits a grammar to maximize the contribution of both the “virtually conceptually necessary” components of syntactic structure-building – Lexical Items and (External and Internal) Merge – it is difficult to sustain the idea of a grammar with truly “mixed”, “basic” order-eschewing OV/VO of the sort quite commonly assumed during the classic parametric era (and beyond) for systems like earlier English. Instead, we expect “mixed” OV/VO systems to exhibit [F]-regulated OV/VO patterning, with one of these being “basic”. Drawing on both historical and contemporary examples, I show that this seems to be correct. Looking beyond these, arguably, not maximally “mixed” systems, I also briefly show for representative ‘free word-order’ systems that the MMM prediction that “basic” OV/VO/head-directionality should be fixed seems to be borne out. The potential for variation in the domain of “basic” and “marked” word order thus seems to be considerable, but nevertheless constrained.
(1) UG (=Factor 1) + PLD (=Factor 2) + Factor 3 \(\Rightarrow\) Steady-state Grammar/I-language

(2) a. **Doubling**: “extra” form (2/multiple forms and 1 meaning, 1 form with no meaning, etc.) – e.g. agreement, concord, and expletive phenomena.
   b. **Silence**: “missing” form (no form, but meaning) – e.g. null exponence, null arguments, null complementisers, ellipsis, OCP effects (and thus also their failure).
   c. **Multifunctionality**: 1 form, multiple meanings – e.g. systematic homophony (cf. also Wiltschko 2014, Duffield 2013, 2014).

(3) **Feature Economy** (adapted from Roberts & Roussou 2003): Postulate as few features as possible to account for the regularities in the PLD.

(4) **Input Generalization** (adapted from Roberts 2007): maximise use of postulated features.

(5) FE- and IG-regulated formal-feature learning hierarchy: NO>ALL>SOME

Is [F] present?

- NO: YES: ALL heads?
  - YES: NO: Which subset of heads? (SOME) (postulate new [F])

(6) NO>ALL>SOME applied in the domain of word order

Is head-final present?

- NO: All head-initial
  - YES: (=ALL): present on all heads?
  - YES: All head-final
    - NO: present on all [+/-V]heads? (SOME)
      - YES: Consistently head-final clause/nominal
      - NO: present on subset of [+/-V] heads? ….

\[\downarrow\] ever more specific SOME options