Domains and Parameters

Ian Roberts

Downing College, University of Cambridge

igr20@cam.ac.uk

Outline

Parameters as add-ons
Observ ations on cyclicity
Parameters and cycles
Rethinking word-order parameters
Skewed typological distributions

1. *Parameters as add-ons*

- An undoubtedly undesirable property of many approaches to parameters is that they add something to the theory simply in order to describe cross-linguistic variation:

1.1 In GB theory

(1) The bounding nodes for subjacency are NP and {S, S'}. (Rizzi 1982)

(2) a. *pro* (Chomsky 1982);

b. The licensing heads for *pro* are {I, V, P, …} (Rizzi 1986).

(3) X {precedes/follows} its complement in X' (Stowell 1981, etc.).

(4) Wh-movement is overt/covert (Huang 1982).

(5) V-to-I/I-to-V (Pollock 1989).

1.2 In Minimalism

(6) F is strong/weak.

(7) EPP/*/^ ….

(8) Operation Ω applies here and/or there (timing parameters, Müller 2005).

(9) Externalisation (PF as an add-on to the thought system: Berwick & Chomsky 2011, 2016).

(10) The Borer-Chomsky Conjecture (BCC):
Parameters reduce to formal features of (a subset of) functional heads.
(Limits the add-on domain).

"[W]e can ask the most difficult question of all, but one which the minimalist programme requires us to ask: why do we have parameters at all?" (Roberts & Holmberg 2010: 53).

- Clearly, none of (1-10) answers this question, and in fact all the approaches take us away from minimalist ideals by complicating the theory in a manner directly driven by empirical observation.

1.3 Emergent Parameters (ReCoS and beyond)

"An alternative is that parameters are just those grammatical options which are *not* specified by UG" (Roberts & Holmberg 2010: 55).

(11) The BCC as the underspecification domain: where UG doesn't mind.

(12)a. BCC_{bad}: functional head H has FFs {F, G, …};

b. BCC_{good}: UG makes available a set F of FFs, which are (all?) optionally assigned/scattered across functional heads.

(13) Parameters and the three factors of language design:

a. F1: UG as in (12b);

b. F2: the learner's interaction with PLD (see Crisma, Guardiano & Longobardi 2019);

c. F3: Feature Economy: avoid FFs.
Input Generalisation: maximise distribution of unavoidable FFs.
(Biberauer 2017 unites these as Maximise Minimal Means, MMM).

ReCoS idea: parametric variation (in the form of parameter hierarchies) is an emergent property of the interaction of F1, F2, F3 as specified in (13).

- Reduction of UG (as in (12b));
- some (much-needed) precision on F3;
- lots of results (see https://recos-dtal.mml.cam.ac.uk/, Roberts 2019, etc etc).

More radical ReCoS (Biberauer 2017): the FFs are not UG-given, as in (12b), but themselves emergent properties determined by (13).

(14) BCC_{better}: UG defines a possible FF (maybe as the pair {iF,uF}), which are optionally assigned/scattered across functional heads.

BUT: if there is one obvious property of natural language (in addition to Berwick & Chomsky's Basic Property, essentially Merge), it is variation.

What if variation is (virtually) conceptually necessary? Then even (14) doesn't capture this (as we could easily imagine a UG without it, or one which always defines a uniform set of FFs, uniformly scattered).

Perhaps we need to build variation in to the operations of the system itself. We probably can't do this with Merge (? see Hale 1983), so:

Proposal: parameters define and are defined by cyclic applications of operations in local domains. If cyclic application of operations (Merge, Agree) in local domains is (virtually) conceptually necessary, then so are parameters.

2. Observations on cyclicity

(15) *Dependency generalisation*

In a given local domain, if featural property P holds of H_1 where H_1 is asymmetrically c-commanded by H_2, then featural property P of H_2 is known.
(Roberts 2019:526)

(16) If a language has object agreement, then it has subject agreement.
(Moravcsik 1978:364, Corbett 2006:59)

(17) Dependent case Baker (2015:48-49):

a. If there are two distinct NPs in the same spell out domain such that NP1 c-commands NP2, then value the case feature of NP2 as accusative unless NP1 has already been marked for case.

b. If there are two distinct NPs in the same spell out domain such that NP1 c-commands NP2, then value the case feature of NP1 as ergative unless NP2 has already been marked for case.

- If heads (T and v, presumably) license case, then this is clearly a further instance of (15).

(18) The Final Over Final Constraint (FOFC) (informal statement; Sheehan, Biberauer, Holmberg & Roberts 2017):

A head-final phrase βP cannot dominate a head-initial phrase αP where α and β are heads in the same Extended Projection.

(19) a. Dependent agreement (if the lower head agrees, the upper one does);
b. Dependent case (if the lower head licenses case K, the upper one licenses K');
c. Dependent linearization (if the lower head is initial, so is the upper one).

(20) Strict Cycle Condition (Chomsky 1973):
No rule R can apply to a domain dominated by a node A in such a way as to solely affect B, a proper subdomain of A.

(21) … $[_A$ … $[_B$ …] …] …

- If you apply agreement in B then you must in A (A is the agreement domain);
- If you license case K in B then you must license K' in A;
- If you don't roll up in B, then you can't in A (FOFC).

➔ Cyclic domains are more varied than we are used to thinking in the Phase Age: they include phases (for A'-movement etc), but also probably TP (= A for agreement and case-licensing in (21); maybe vP for erg-abs licensing) and the roll-up domain (any XP? See (28) below).
➔ NB how intervention locality can reduce to (20, 21): if there's a target in B for a B-level operation and one in A, then the B-target must be targeted (first). The B-target intervenes between the lower term of the operation and the A-target.
➔ E.g. the Head Movement Constraint (a classic case of intervention locality) can be seen as the head-movement pendant of FOFC.

3. Parameters and cycles

(22) Types of parameters (Biberauer & Roberts 2012, 2015; Roberts 2019:75-76):

For a given value v_i of a parametrically variant feature F:

a. Macroparameters: all heads of the relevant type, e.g. all probes, all phase heads, etc, share v_i;

b. Mesoparameters: all heads of a given natural class, e.g. [+V] or a core functional category, share v_i;

c. Microparameters: a small, lexically definable subclass of functional heads (e.g. modal auxiliaries, subject clitics) share v_i;

d. Nanoparameters: one or more individual lexical items is/are specified for v_i.

- Macroparameters define gross typological properties (polysynthesis, radical prodrop, harmonic word order, etc), are everywhere in the PLD, are (therefore) strongly conserved over time and (therefore) hold at genus (Dryer 1992) or higher taxonomic level;
- Mesoparameters correspond roughly to traditional GB parameters (consistent null subjects, verb-movement, etc), are well expressed in the PLD, are (therefore) fairly stable over time and (therefore) define genera up to (contact-induced) outliers, e.g. English in Germanic;
- Microparameters define often rather local axes of variation (clitics, determiners, auxiliaries, etc.) often within families, are reliant on subtle aspects of PLD, are (therefore) diachronically unstable and (therefore) define intra-family variation (e.g. Romance clitics, West Germanic verb clusters, etc);
- Nanoparameters are "irregular" lexically specific properties, usually diachronic residues which are analogised out of the system if not highly frequent/accessible in PLD (e.g. some aspects of English auxiliaries).

This taxonomy has proven descriptively fecund (Biberauer & Roberts 2012, 2015, 2017, van der Wal, to appear, Roberts 2019, Ledgeway, to appear, etc etc).

But:

- Although the taxonomy isn't (despite first appearances) subject to the "add-on" critique (it's the notion of emergent parameter plus the notion of set; Richie Kayne, p.c.), it does seem to be somewhat descriptively driven;
- It doesn't look very conceptually necessary;
- "what is lacking in the hierarchies as they are stated … is any clear formal aspect of the statement of a given option that allows us to immediately tell whether it is a macro-, meso-, or microparameter" (Roberts 2019: 616).

(23) A revision in terms of cycles:

a. Macroparameters: apply to all phasal domains;
b. Mesoparameters: apply across phasal domains;
c. Microparameters: apply within phasal domains;
d. Nanoparameters: apply across the minimal phasal domain, individual LIs (Marantz 1997).

- Meets the desideratum of 1.3: "parameters define and are defined by cyclic applications of operations in local domains."

- Involves basic quantification over cyclic domains (ALL, SOME) with SOME restricted by the privileged cyclic domain, the phase → close to conceptual necessity.
- Cycles define features/features define cycles; features define parameters/parameters define features; parameters define cycles/cycles define parameters: the three notions are interdefinable: the archi-thing that these notions define is the main other thing in the theory aside from Merge (see also Song 2019).

Two examples:

(24) A macroparameter: is [Person] present? N: radical prodrop (RNSL, e.g. Japanese); Y: non-RNSL (first option: agreement-licensed null arguments everywhere (Basque); second option: agreement-licensed null subjects (CNSL, e.g. Italian)).

- Obviously the negative answer affects all domains, and has numerous empirical consequences (argument ellipsis, lack of determiners, lack of agreement inflection, separation of Agree and Case-licensing; see in particular Saito 2016);
- The positive answer: agreement-licensed null arguments everywhere, e.g. Basque.
- First restriction: Person is present on T, D; null subjects (which, if they involve D-to-T incorporation – Barbosa 2019 – are clearly cross-phasal); presence of "rich" inflection; possibly verb-movement (see below); "rich" determiner system, etc.
- the CNSL property is mesoparametric, as it affects many but not all domains, and is cross-phasal: on the other hand full pronominal-argument systems like Basque contrast minimally with Japanese in having Person in all possible domains.

(25) A mesoparameter: does V raise into the TMA field (possibly in virtue of having uTMA features: Roberts 2019, Chapter 5)? Y: (most) Romance; N: English, Chinese, Haitian Creole, Niuean …

- This is basically the "Pollock parameter" (see (5)); it is cross-phasal in that it determines whether V can get out of the lower phase. Obvious consequences for word-order (including being a precondition for VSO/VOS alternations of the kind described Massam 2001), negation, answers to yes/no questions (Holmberg 2016), V-licensed VP-ellipsis, etc. Possible link to "rich inflection" on V (vs auxiliary/particle-based TMA systems); Tvica (2017).

(26) Microparameters of V-movement in Romance (Schifano 2018):

French/Romanian V > *probably* > (N)Italian V > *already* > EPort V > *always* > Sp V

- V targets different heads in the TMA field, determined by distribution of u/iAsp/T/M features on heads in that field and on V, correlating with TMA inflection on V (see Schifano 2018 for details, and Roberts 2019: 348-359 for a partial reinterpretation);
- This variation applies in the zone flanked by C and v, and hence is clearly intra-phasal;
- in non-English, non-Icelandic Germanic: probably (25) is negative (Holmberg & Platzack 1995, Biberauer 2003, Zwart 1997), the parameter(s) regulating V-movement into the left-periphery must be meso, but are distinct from (25).

Other consequences:

- erg-abs vs nom-acc alignment may be on the border of meso and micro: if the ERG parameter is "does v have ACC?" (Sheehan 2017) then, if positive, the case-licensing domain is cross-phasal (v and T); if negative, it is intra-phasal within vP (modulo "high ABS").
- Huang's wh-movement parameter is at least meso; micro wh-related parameters can't involve successive-cyclicity, which is cross-phasal by definition but can involve

(a) landing sites in the left periphery (Bošković 2002); (b) the DP-internal nature of wh-elements (Cheng 1991, Huang 2015).

- Subject clitics are always micro; object clitics may be meso (since they move at least to the vP-edge): this looks right given the differences in cross-Romance distribution of these.
- Clausal negation is arguably always in TP, and therefore the variation is micro/nano (a lot of it is a side-effect of verb-movement). See Roberts (2019: 604) on boring parameters.

4. Rethinking word-order parameters

Roll-up parameters in each main clausal domain (vP, TP, CP):

(27)a. **vP-internal rollup:**
Y/+: $[_{vP}\,[_{VP}$ O V] v (VP)] -- CTOV (NB rollup at both VP and vP)
N/-: $[_{vP}$ … v $[_{VP}$ V O]] -- CTVO

b. **TP-internal rollup:**
Y/+: $[_{TP}\,[_{vP}\,[_{VP}$ O V] v (VP)] T (vP)] -- COVT
N/-: $[_{TP}$ …T $[_{vP}$ .. v $[_{VP}$ V O]]]

c. **CP-internal rollup:**
Y/+: $[_{CP}\,[_{TP}\,[_{vP}\,[_{VP}$ O V] v (VP)] T (vP)] C (TP)] -- OVTC
N/-: $[_{CP}$ C $[_{TP}$ …T $[_{vP}$ .. v $[_{VP}$ V O]]]

Remember that FOFC (a case of the Strict Cycle in (20)) guarantees that positive (27c) implies positive (27b) implies positive (27a).
Definition of "across" phasal domains in (23b): movement which affects the computation of the PIC, i.e:

- Movement to a phase edge (standardly);
- Movement of an entire phase (since this affects the timing of PIC-computation)
- → all rollup is mesoparametric (this looks right from a typological/diachronic perspective).

(28) Word-order macroparameters? In:

a. $[_{\alpha}\ H_1\ H_2\ ]$ -- we can neither linearise ("bottom-pair" problem) nor label α, unless we assume one H, say H_2, has no categorial feature, i.e. it's a Root, and H_1 is a categoriser. Now we can label α as H_1, but we still can't linearise. So one of them must move. Move H_2:

b. $[_{\beta}\ H_2\ [_{H1}\ H_1\ (H_2)\ ]]$ – here α is successfully labelled H_1 and H_2 can be linearised as preceding H_1 (because H_2 asymmetrically c-commands H_1) and β must be labelled as H_1 as H_2 has no category. So H_1 is the head of β and H_2 is its complement → rollup of complement around head.

c. Move H_1 from (a):
$[_{\beta}\ H_1\ [_{H1}\ (H_1)\ H_2\ ]]$ – here H_1 will label β again but this time the moved H_1 asymmetrically c-commands H_2 and so is linearised as preceding it. This counts as head-movement.

- (28b) and (28c) are cases of rollup and head-movement as repairs of (28a). [NB we have already seen in Section 2 that head-movement and roll-up are subject to the same locality condition, giving rise to the HMC and FOFC respectively].

- Input Generalisation (of operations, not features) does the rest: (28b) will give rise to harmonic head-finality as rollup generalises; (28c) to harmonic head-initiality as head-movement generalises.
- So now the question becomes: what stops IG from creating maximal head-initiality/finality everywhere, i.e. when do the mesoparameters come in?
- **Proposal**: head-movement is required for labelling (defective goals do this; see Roberts 2010), and so is triggered by EM of some probing uF to α in (28a), in relation to which H_1 is a defective goal.
- XP rollup fails where X is a defective goal, so rollup parameters reduce to negative values of head-movement parameters.
- Possible typological consequence: head-movement systems have richer probes, able to license defective goals by head-movement → rich functional structure in head-movement, i.e. head-initial systems?

5. On skewed typological distributions

Analysis of basic word-order types:

(29) a. **SOV:** rollup, i.e. not head movement (no defective goals). Degree of rollup may vary (Japanese vs German vs Latin);
b. **SVO:** no rollup but head-movement (at the lowest structural level). Degree of head-movement varies, giving different amounts of (deep) analyticity (Huang 2015);
c. **VSO:** two types. Type A similar to SVO, but with slightly higher V-movement (McCloskey 2017). Type B is VOS + object shift (Massam 2001: EPP_V), O smuggled over S. Low S → ergative alignment; heavy restrictions on argument-extraction;
d. **VOS:** like Type B VSO without object shift.
e. **OVS:** like VOS but with low rollup. Other head-final properties (e.g. postpositions) < "harmonic" low rollup. Unclear how prevalent ergativity is;
f. **OSV**: extremely rare: *WALS* gives 4 languages out of 1,377. Derivable as OVS with (low) V-movement and fronting of remnant [O (V)].

Three questions:

(i) skewed frequencies (*WALS* Map/Feature/Chapter 81A; Dryer 2013).:

SOV	565 (47.6%);
SVO	488 (41.1%);
VSO	96 (8.1%);
VOS	25 (2.1%);
OVS	11 (0.9%);
OSV	4 (0.3%).

(NB also 189 "no dominant order" languages; percentages are calculated on the total leaving these out, i.e. 1,188).

Why?

(ii) diachronic factors: do the distributions change significantly over time? (See *WALS* Feature/Chapter 81, Figure 1: Order of Subject, Object, and Verb in "Ancient" Languages, Dryer 2013).

(iii) links to other variant properties: see (29) for some speculations, but this is mostly left aside here.

5.1 Some general points

When we observe any kind of unexpectedly uneven distribution of properties we know to be variable, there are in principle four types of account we can give:

a. **It's not true**: data like the above is misleading. Two versions of this: (i) the typologists' answer: go and find more languages and you'll see it's not the case.
(ii) the generativists' answer: the analytical categories you're dealing with are at the wrong level of abstraction. Get your analysis right and the skewing will disappear.
[NB (i) and (ii) are not exclusive].

b. **Diachrony**: (i) deep history: processes of change cause oscillations in distributions – we happen to observe one particular skewing right now but that's a product of our historical moment, rather like observing a continent at the South Pole: it doesn't have to be there and won't be if we wait long enough (interesting questions here about the Uniformitarian Principle: see Roberts 2017 for discussion).
(ii) surface history: what if the Hixkaryana (who speak an OVS language) had colonised Europe (or Africa) and extirpated the local languages?

c. **Psycholinguistics**: maybe UG sees all word orders as equal but some are easier to process than others: see Hawkins (1994, 2004, 2014). NB, combining this with (b), we could think that some orders are easier to acquire, and therefore less subject to change and more easily spread through contact than others, giving rise over time to skewed distributions. The dispreferred orders are winnowed out of the "gene pool" (i.e. parameter pool) over time. NB the implication for deep history.

d. **Universal Grammar**: parameters come with built-in markedness preferences favouring certain types over others.

- NB the possibility of combined explanation: maybe (d) is right, but UG-preferred systems are easier to acquire and process (c) and so spread more readily giving rise to a certain skewing that oscillates over deep and surface history but remains broadly constant (b). Nichols (1992) can be read this way (although she does not advocate UG).
- It does now seem that large language families of the Indo-European/Afro-Asiatic/Austronesian type may be recent phenomena connected to technological advances (horses, agriculture, boats; Ringe 2013).
- But first we need to tackle (a).

5.2 Observing the typological variation

- How much credence should we attach to statistics like those for word-order types above?
"to what extent can a sample of the languages that happen to be spoken now (and happen to be investigated) ever tell us anything about human language in general? The problem is that the current range of languages may not (and is indeed unlikely to) instantiate everything that is humanly possible because languages arise and die for reasons that have nothing to do with their structures" (Bickel 2017:20).

(30) General uniformity of relative frequencies of basic types (from Peeters 1991):

Word Order	Ruhlen		Mallinson and Blake		Hawkins		Tomlin	
	No.	%	No.	%	No.	%	No.	%
SOV	222	51	41	41	174	52	180	44.78
SVO	155	35	35	35	109	32	168	41.79
VSO	47	11	9	9	45	13	37	9.20
VOS	8	2	2	2	8	2	12	2.99
OVS	2	0.5	1	1	—	—	5	1.24
OSV	1	0.25	1	1	—	—	—	—

- all of these surveys have two major things in common:
  (i) the order of frequency: SOV > SVO > VSO > VOS > OVS > OSV;
  (ii) SOV just beating SVO, a big gap to VSO, then the OS orders as minority patterns and O-initial extremely rare.
- So the basic picture (at least at our historical moment, with the empirical resources available) is clear.

So: the typologists should of course continue to report more languages (especially with so many endangered) but it doesn't look likely that the overall distribution will change.
[NB the question of patterning by genera: see Dryer 1992].

But what about the generativists' point in (a, ii) above? We know for sure that SOV, SVO and VSO, at least, are non-uniform "types".

5.3 A simple system of (meso-)parameters designed to capture the basic word orders

- Six orders, so three binary options ($2^3 = 8$) with two excluded:

	±OV (head par.)	**±SO (EPP?)**	**±VS (V(P)-mvt)**
SOV	+	+	-/+
SVO	-	+	-
VSO	-	+	+?
VOS	-	-	+?
OVS	+	-	+
OSV	+	-	+
SOV again	+	+	+
** (in-situ)*	-	-	-

- Too simple: doesn't capture preferences/frequencies, doesn't capture known differences within types (VSO, SVO, etc.); the "VS" parameter is equivocal between V and VP-movement.
- We need something more:

(23b) **Definition of mesoparameter:**
Variant property holding across some phasal domains.

(31) Roll-up parameters in each main clausal domain (vP, TP, CP), repeated from (27):

a. **P1: vP-internal rollup**
Y/+: $[_{vP}\,[_{VP}$ O V] v (VP)] -- CTOV (NB rollup at both VP and vP)
N/-: $[_{vP}$ … v $[_{VP}$ V O]] -- CTVO

b. **P2: TP-internal rollup**
Y/+: $[_{TP}\,[_{vP}\,[_{VP}$ O V] v (VP)] T (vP)] -- COVT
N/-: $[_{TP}$ …T $[_{vP}$ .. v $[_{VP}$ V O]]] (or = P1+)
[NB FOFC creates a dependency here: P2+ only if P1+/*P1- & P2+]

c. **P3: CP-internal rollup**
Y/+: $[_{CP}\,[_{TP}\,[_{vP}\,[_{VP}$ O V] v (VP)] T (vP)] C (TP)] -- OVTC
N/-: $[_{CP}$ C $[_{TP}$ …T $[_{vP}$ .. v $[_{VP}$ V O]]] (or = P1+, or P1+ & P2+)
[NB only allowed if P1+ & P2+, FOFC again]

(32) V-movement parameters in each clausal domain:

a. **P4: V-movement in vP**
Y/+: (S) … {TMA/Neg} … V+v O
N/-: (S) … {TMA/Neg} … v V O (Huang-style "analyticity" effects)

b. **P5: V-movement in TP**
Y/+: (S) … {TMA,Neg,V} … O (Pollock/Schifano)
N/-: see P4 [NB HMC entails *P4- & P5+][1]

c. **P6: V-movement in CP**
Y/+: V2
N/-: P4- or P5-
[NB other factors needed for XP-fronting/root-embedded distinction, although neither of these varies, we have assumed].

(33) Subject-placement parameters in each clausal domain:

a. **P7: Subject in vP**
Y/+: … $[_{vP}$ S ….] no $EPP_D \rightarrow EPP_V$ (*_in-situ_*)
N/+: $[_{TP}$ S ….] EPP_D (link to accusative vs ergative alignment?)

b. **P8: Subject in TP**
Y/+: $[_{TP}$ S ….]
N/-: … $[_{vP}$ S …]
[NB P7 and P8 are just inverses, so almost a single $EPP_{V/D}$ parameter:
*P7- & P8- could follow from LA, but would be allowed if P1+; the result of LA is *P1- & P7- & P8-, hence P7 and P8 are separate parameters after all. LA may also rule out P7+ & P8+].

[1] If FOFC and the HMC are both cyclicity effects, then the implicational relations among P4/P5 (HMC) and those among P1/P2/P3 derive from the same thing: locality/cyclicity defined as in (20).

c. **P9**: **Subject in CP**
Y/+: subject in left-periphery (obligatory topic or focus)
N/-: P7+ or P8+ or P1+ (see (b)).

(34) Object-shift (question of OS to T- or C-domains left open here):

P10: Object shift to vP edge
Y/+: $[_{vP}$ O .. v .. $[_{VP}$ V (O)]]
N/-: $[_{vP}$ .. v .. $[_{VP}$ V O]]

NB: in general we can leave aside P6 and P9, as these concern the left-periphery and "marked" word orders. The parameters which determine basic word order are thus P1-P5, P7, P8 and P10.

➔ 8 parameters, so 2^8 = 256 possible combinations in principle, BUT we know the following, at least:

(35)a. *P1- & P2+ (FOFC)
b. *P2- & P3+ (FOFC)
c. *P1- & P3+ (FOFC)
d. *P1- & P7- & P8- (LA)
e. *P4- & P5+ (HMC)
f. *P7+ & P8+ (See above)
g. *(P1+ & P4+) (28)
h. *(P2+ & P5+) (28)

5.4 Addressing the skewed distribution in meso-parametric terms

I. Which p-settings are required for a given order to be possible (are p-expressed by a given order; see Clark & Roberts 1993, Roberts 2019: 65, 76, 89, Longobardi & Roberts, in progress)?
II. Which p-settings are excluded as a consequences of (I)?
III. Which p-settings are allowed and what are the consequences?

		P1	P2	P4	P5	P7	P8	P10	
(40)	**SOV:**	+	+	-	-	-	±	±	(4 derivations)
		+	-	-	-	-	+	±	(2 derivations)
		-	-	±	-	-	-	+	(2 derivations)

■ 8 derivations for SOV

(41)	**SVO:**	**-**	**-**	**-**	**+**	**±**	**-**	**+**	**±**	(4 derivations)
(42)	**VSO:**	-	-	-	+	++	-	+	±	(2 derivations)
		-	-	-	-	-	+	-	+	(1 derivation)
(43)	**VOS:**	-	-	-	-	-	+	-	+	(1 derivation)
(44)	**OVS:**	+	-	-	-	-	+	-	-	(1 derivation)
(45)	**OSV:**	+	-	-	+	±	+	-	-	(2 derivations!!)

[NB here the left-periphery-related parameters P3, P6 and P9 have been left out for simplicity]

4. Conclusions

- If each derivational possibility is a type (a cluster of mesoparametric settings), then we have the right order of frequencies (more languages < more types < more derivational possibilities): SOV > SVO > VSO > VOS > OVS (except for OSV).
- But the ratio of languages to types, in terms of raw numbers, is badly off:
  8 SOV derivations for 565 languages: 70 languages per type;
  4 SVO ones for 488 languages: almost 120 languages per type;
  3 VSO ones for languages: 32 per type;
  1 for VOS: 25 per type.

Still too many SOV and SVO languages, but one more binary independent parameter for SOV and two more for SVO would bring us to 30-35 languages per type;
NB the rarer V-initial types are well-behaved: the limited areal and typological distribution of these languages may be giving us a clearer picture of type-distribution than the very widespread SOV/SVO ones.
BUT what's so special about 30 ± 5 languages per type?
Now we look for diachronic/contact/markedness explanations.

6. Conclusions

- (23) is an improvement on (22);
- Parameters, domains and features drawn more tightly together;
- Uniform statements of parameters across domains;
- Potential for an approximately uniform notion of "word-order type."

References

Alexiadou, Artemis, & Elena Anagnostopoulou. 2001. The Subject-in-Situ Generalization and the role of case in driving computations. *Linguistic Inquiry* 32: 193‑231.

Baker, Mark C. 2015. *Case.* Cambridge: Cambridge University Press.

Berwick, Robert, & Noam Chomsky. 2011. The biolinguistic program: the current state of its development. In Anna-Maria di Sciullo & Cedric Boeckx (eds), *The Biolinguistic Enterprise*, 19–41. Oxford: Oxford University Press.

Berwick, Robert, & Noam Chomsky. 2016. *Why Only Us: Language and Evolution*. Cambridge, Mass.: MIT Press.

Biberauer, Theresa. 2003. *Verb second (V2) in Afrikaans: a Minimalist investigation of word-order variation*. PhD dissertation, University of Cambridge.

Biberauer, Theresa. 2017. Factors 2 and 3: a principled approach. *Cambridge Occasional Papers in Linguistics* 10: 38–65.

Biberauer, Theresa, & Ian Roberts. 2012. The significance of what hasn't happened. Paper presented at Diachronic Generative Syntax Conference XIV, Lisbon.

Biberauer, Theresa, & Ian Roberts. 2015. The clausal hierarchy, features and parameters. In Ur Shlonsky (ed.), *Beyond Functional Sequence*, 295‑313. Oxford: Oxford University Press.

Biberauer, Theresa, & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway and Ian Roberts (eds), *The Cambridge Handbook of Historical Syntax*, 134‑62. Cambridge: Cambridge University Press.

Bickel, Bathazar. 2017. Areas and universals. In R Hickey (ed) *The Cambridge handbook of areal linguistics*. Cambridge: Cambridge University Press, 40-55.

Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33: 351–83.

Cheng, Lisa. 1991/2007. On the typology of wh-questions. PhD dissertation, MIT.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson and Paul Kiparsky (eds), *A Festschrift for Morris Halle*, 232–86. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding.* Cambridge, MA: MIT Press.

Clark, Robin, & Ian Roberts. 1993. A computational approach to language learnability and language change. *Linguistic Inquiry* 24: 299–345.

Corbett, Greville. 2006. *Agreement.* Cambridge: Cambridge University Press.

Crisma, Paola, Cristina Guradiano & Giuseppe Longobardi. 2019. Syntactic parameters and language learnability. Ms Universities of Trieste, Modena/Reggio Emilia and York.

Den Besten, Hans. 1983. On the interaction of root transformations and lexical delete rules. In Werner Abraham (ed.), *On the Formal Syntax of the Westgermania*, 47–131. Amsterdam: Benjamins.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68: 81–138.

Dryer, Matthew S. 2013. Order of Subject, Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/81, Accessed on 2019-07-06.)

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 58–90. Cambridge, Mass.: MIT Press.

Hale, Kenneth. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1: 5–47.

Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Hawkins, John A. 2014. *Cross-Linguistic Variation and Efficiency.* Oxford: Oxford University Press.

Holmberg, Anders. 2016. *The Syntax of* Yes *and* No*: Studies in Rethinking Comparative Syntax*, vol. 1. Oxford: Oxford University Press.

Holmberg, Anders, and Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. New Oxford: Oxford University Press.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. PhD dissertation, MIT.

Huang, C.-T. James. 2015. On syntactic analyticity and parametric theory. In Y.-H. Audrey Li, Andrew Simpson, and W.-T. Dylan Tsai (eds), *Chinese Syntax in a Cross-linguistic Perspective*, 1–47. Oxford: Oxford University Press.

Jouitteau, Mélanie. To appear. Rethinking V2 typology from expletives: EPP inquiries on stylistic fronting, excorporation and verb-doubling. In Sam Wolfe, & Rebecca Woods (eds), *Rethinking Verb Second*. Oxford: Oxford University Press.

Ledgeway, Adam. In press. From Latin to Romance syntax: the great leap. In Paola Crisma & Giuseppe Longobardi (eds), *The Oxford Handbook of Diachronic and Historical Linguistics*. Oxford: Oxford University Press.

Longobardi, G. & I. Roberts. In progress. The Restricted List: expressing and explaining grammatical diversity. Ms Universities of Cambridge and York.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language & Linguistic Theory* 19: 153–97.

Moravcsik, Edith A. 1978. Reduplicative constructions. In Joseph H. Greenberg et al. (eds), *Universals of Human Language*, vol. 3: *Word Structure*, 297–334. Stanford, Calif.: Stanford University Press.

Nichols, Joanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

Panagiotidis, Phoevos. 2015. *Categorial Features: A Generative Theory of Word Class Categories*. Cambridge: Cambridge University Press.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365‑424.

Ringe, Don. 2013. The linguistic diversity of Aboriginal Europe. In S.-F. Chen & B. Slade (eds) *Grammatica et verba, glamour and verve: A Festschrift for Hans Heinrich Hock.* Beechstave Press, pp. 202-12.

Rizzi, Luigi. 1982. *Issues in Italian Syntax.* Dordrecht: Foris.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17: 501–57.

Rizzi, Luigi, and Ian Roberts. 1989. Complex inversion in French. *Probus* 1: 1–30.

Roberts, Ian. 2004. The C-system in Brythonic Celtic languages, V2, and the EPP. In Luigi Rizzi (ed.), *The Structure of IP and CP*, 297–328. Oxford: Oxford University Press.

Roberts, Ian. 2010. *Agreement and Head Movement: Clitics and Defective Goals*. Cambridge, Mass.: MIT Press.

Roberts, I. 2017. Uniformitarianism. In A. Ledgeway & I. Roberts (eds) *The Cambridge Handbook of Historical Syntax.* Cambridge: Cambridge University Press, 338-359.

Roberts, Ian. 2019. *Parameter Hierarchies and Universal Grammar*. *Studies in Rethinking Comparative Syntax, Volume IV.* Oxford: Oxford University Press.

Roberts, Ian, & Anders Holmberg. 2010. Introduction: parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts, & Michelle Sheehan. *Parametric Variation: Null Subjects in Minimalist Theory.* Cambridge: Cambridge University Press, 1–57.

Roberts, Ian, & Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization.* Cambridge: Cambridge University Press.

Saito, Mamoru. 2016. (A) case for labeling: labeling in languages without phi-feature agreement. *The Linguistic Review* 33: 129–75 [special issue, *On Labels*, ed. Željko Bošković].

Schifano, Norma. 2018. *Verb Movement in Romance: A Comparative Study.* Oxford: Oxford University Press.

Sheehan, Michelle. 2017. Parametrizing ergativity: an inherent case approach. In Jessica Coon, Diane Massam, & Lisa Travis (eds), *The Oxford Handbook of Ergativity*, 59–85. Oxford: Oxford University Press.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts, & Anders Holmberg. 2017. *The Final-Over-Final Condition: A Syntactic Universal*. Cambridge, Mass.: MIT Press

Stowell, Timothy. 1981. The origins of phrase structure. PhD dissertation, MIT.

Tvica, Seid. 2017. *Agreement and Verb Movement: The Rich Agreement Hypothesis from a Typological Perspective.* Utrecht: LOT.

Van der Wal, Jenneke. To appear. From microparameters to macroparameters: a Bantu case study. András Bárány, Theresa Biberauer, Jamie Douglas, & Sten Vikner (eds), *Nominal Architecture and Its Consequences: Synchronic and Diachronic Perspectives.* Berlin: Language Science Press.

Zwart, Jan-Wouter. 1997. *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch.* Dordrecht: Kluwer.