Workshop: Perspektiven minimalistischer Syntax **27 October 2008**
Universität Leipzig

Eine minimalistische Perspektive auf parametrische Variation / A minimalist perspective on parametric variation

Marc Richards ● Universität Leipzig ● richards@uni-leipzig.de

1. INTRODUCTION: PARAMETRIC VARIATION WITHOUT UG

Chomsky 2006 on the biolinguistic perspective: "Development of language in the individual must involve three factors."

(1) *Three factors in language design*
- I. Genetic endowment (UG)
- II. Experience (external data, PLD, E-language)
- III. Principles not specific to the Faculty of Language (FL): principles of efficient computation, interface conditions (IC), "general properties of organic systems" (Chomsky 2004: 1)

- Two trends in the recent minimalist literature:
  - away from Factor I (UG must be as small, simple and empty as possible, on evolutionary grounds);
  - towards Factor III (taking us "beyond explanatory adequacy"; Chomsky 2004) – a new benchmark for what counts as a genuine explanation.
- Consequence: Not only must a maximally underspecified UG be devoid of language-specific principles (as far as possible), but also of parameters. *The range of variation thus cannot be part of the universal, genetic specification; variation (and the forms it takes) is no longer determined by the genes.*

Three recent relevant papers on this matter:
- Berwick & Chomsky 2008
- Boeckx 2008
- Holmberg & Roberts 2008

Problem (or challenge): How do we explain and accommodate language variation in a system based on third-factor explanation (i.e. minimalism)?

- Factors I and III are invariant;
  Factor II is the *trigger* for variation (different final states are acquired depending on the linguistic environment to which the child is exposed – the trigger experience); but it cannot be the *locus* of variation per se (since Factor II is language- and organism-external).

1.1 Uniformity of Factor I (UG)

(2) *Uniformity of Factor I* (Chomsky 2001: 2)
"[A]ssume languages to be uniform, with variety restricted to easily detectable properties of utterances."

- Under the Minimalist Program (MP), we abandon a direct parametrization of UG itself (Factor I) – i.e. parametrized principles – in favour of a uniform FL, with crosslinguistic variation restricted to the properties and features of lexical entries (i.e. that which must in any case be learned). In particular, under the "Borer-Chomsky Conjecture" (Baker 2008), it is restricted to the properties and features of functional categories:

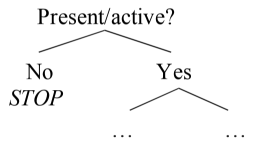
(3) *Borer-Chomsky Conjecture* (BCC; 'lexical parameters')
"Parametric variation is restricted to the lexicon, and insofar as syntactic computation is concerned, to a narrow category of morphological properties, primarily inflectional." (Chomsky 2001: 2)
"The availability of variation [is restricted] to the possibilities which are offered by one single component: the inflectional component." (Borer 1984: 3)

Example ('classical' minimalism): overt versus covert movement determined by strong versus weak categorical features on functional heads.

- Such parameters, affecting the properties of lexical items, are *microparameters* (largely independent of each other, applying to individual formal features and yielding small-scale variation with only limited clustering effects). The status of grand, typological *macro*parameters in MP is unclear (see section 2).

Acquisition and variation according to the BCC: interaction between Factors I and II:
- UG provides a universal inventory of formal features **F** from which each language (learner) makes a one-time selection of a subset [F]. This subset then defines the range of features that play an active role in the operations of the language in question; all other features are discarded. (See Chomsky 2000: 100-101.)
- From an acquisitional point of view, the universal inventory **F** defines the range of features that a child can postulate in constructing its lexicon. This is the starting point for theories of FC variation presented in, e.g., Thráinsson (1996), Bobaljik & Thráinsson (1998), and Koeneman & Neeleman (2001). ("Assume only those functional categories that you have evidence for", Thráinsson 1996:261; cf. Longobardi's (2001: 294) "Minimize feature content".)
- This yields variation in the actual features (and perhaps projections) present in a given language, and represents the first step on the parameter schemata familiar from such works as Uriagereka 1995, Roberts & Roussou 2003, Holmberg & Roberts 2008:

(4) *Partial parameter schema for feature F*



- If YES, then further options present themselves, e.g. is F morphologically expressed?, does F Agree?, does F trigger movement?, how much structure does F piedpipe? (cf. Biberauer & Richards 2006), etc. (See section 3.)

1.2 Uniformity of Factor III

Chomsky (2005, 2006): Minimalist enquiry may be defined as the pursuit of third-factor ("principled") explanations for linguistic phenomena and properties of the language faculty.

- Third-factor explanations in linguistic minimalism: general physical factors nonspecific to FL that constrain the growth, structure, form, evolution of FL as a biological organ.
- Goal of MP – to move descriptive technology from Factor I (genetic endowment, UG) to Factor III (i.e., show that that technology is dispensable and/or reducible to third-factor effects).
- **Strong Minimalist Thesis** (SMT): no aspect of FL is without a principled, third-factor explanation – specifically, we entertain the thesis that *all* properties of FL contribute to *efficient computation* satisfying *interface conditions* (IC), with FL in that sense a "perfect" solution to IC.
- It follows that the genetic endowment **UG** (Factor I) **should be maximally empty**, since this is where the unexplained, unprincipled residue resides:

(5) "Throughout the modern history of generative grammar, the problem of determining the character of FL has been approached 'from top down': How much must be attributed to UG to account for language acquisition? The MP seeks to approach the problem 'from bottom up': How little can be attributed to UG while accounting for the variety of I-languages attained, relying on third-factor principles?" (Chomsky 2006: 3)

➢ Thus we turn traditional linguistic inquiry on its head: MP is interested in what the study of language has traditionally abstracted away from, namely those properties of language which are *not* unique to this faculty.

- The question of FL's evolution is thereby simplified: the structure of FL would be "determined more by physics than by selection" if SMT holds (Saunders 1994 on Turing's work on morphogenesis [cf. "Turing's Thesis", Chomsky 2006]; see Ott 2006):

(6) "The task of accounting for the evolution of language would be correspondingly eased, for the same reasons that hold for enquiry into evolution generally: the less attributed to genetic information (in our case, the topic of UG) for determining the development of an organism, the more feasible the study of its evolution." (Chomsky 2006:2-3).

What about parameters in a Factor-III-driven approach? A maximally empty UG cannot contain FL-specific parameters.

- What must a maximally empty UG minimally comprise?
- Possible answer: just features (the inventory **F**, above).

Hypothesis: There are just two fundamental feature types provided by UG on which the Narrow Syntax operates: features for building structured expressions, and features for connecting those expressions with the external systems of sound and meaning.

- The former are Chomsky's (2005, 2006) Edge Features (EF); the latter are the uninterpretable features (uF), which trigger Transfer to the interfaces (Chomsky 2006).

- EF on lexical items yields *Merge* (recursion; cf. Chomsky, Hauser & Fitch 2002), *Move* (qua internal Merge).
- uFs yield Probe-Goal relations, *Agree*, *Transfer* (and thus phases – see Chomsky 2006).

➢ That is, just two basic types of feature suffice to give us the full range of operations postulated in minimalist grammars (and analyses): *Merge*, *Move*, *Agree*, *Transfer* (spell-out).

➢ Other minimalist principles, such as *Inclusiveness*, *No Tampering*, *Phase Impenetrability*, etc., are third-factor principles of efficient computation and thus not part of UG.

➢ As far as lexical, microparameters are concerned, i.e. those conforming to the BCC, it's a case of 'business as usual' (contra Boeckx 2008): uFs and EFs are provided by UG, and it is the job of the language learner to assemble these features into lexical items (see section 3 on feature associations); thus, the questions associated with parameter schemata like (4) are still formulable: "does F trigger Agree?" becomes "is F a uF?"; "does F trigger Move?" becomes "does F bear (undeletable) EF?" – see (35) below.

But what of macroparameters, in the sense of parametrized principles?
To the extent that UG is empty and such principles as those listed above are in fact third-factor principles, we would not expect variation here:

(7) *Uniformity of Factor III* – Boeckx 2008:5
"There is simply no way for principles of efficient computation to be parametrized. […] [I]t strikes me as implausible to entertain the possibility that a principle like 'Shortest Move' could be active in some languages, but not in others. Put differently, […] there can be no parameters within the statements of the general principles that shape natural language syntax. In other words, narrow syntax solves interface design specifications optimally in the same way in all languages." (Boeckx 2008:5)

➢ Boeckx calls this the "Strong Uniformity Thesis" (SUT) – essentially, not just Factor I is invariant, but also Factor III.

➢ In the following (section 2), I would like to explore and qualify this position somewhat. There is indeed room for 3rd-factor-based parametric variation in the way that ICs are satisfied on the sensorimotor ('PF') side. This provides a minimalist locus for macroparameters (cf. Richards 2004: Chapter 5).

Claims:

- Variation is itself a third-factor effect, the result of a minimally specified, maximally empty and underspecified UG (cf. Boeckx 2008:13, Holmberg & Roberts: 67).
- Such a UG is characterized by indeterminacies, points of optionality that arise where UG "doesn't mind" (Biberauer & Richards 2006) – e.g. in whether a particular is present in a language and how that feature is satisfied or lexicalised (thus yielding **micro**parameters – **section 3**), or in how IC is met (thus yielding **macro**parameters – **section 2**).

- These choice points are simply left open, as to specify or define them one way or another would be to *add* to UG, contra SMT (and SUT).
- Parameters, in the form of these choice points, thus emerge from SMT as a third-factor effect.

1.3 Interaction between Factors I, II and III

Example: a third-factor explanation of a syntactic universal – the *core functional categories*, C-T-v(-V)

- Each CFC defines a region of the clause that may be 'expanded' in certain languages (e.g. the articulated left periphery), in ways identified in 'cartographic' studies.
- But C-T-v are always minimally present (at the earliest stages of acquisition too – cf. Poeppel & Wexler 1993) – an apparent universal, and something which we might therefore want/need to attribute to UG (thus departing from SMT).

Can we provide a third-factor explanation of this universal property?

- A property is 'explained' insofar as it follows from SMT, i.e. contributes to computationally efficient satisfaction of IC.
- What are the ICs?
  - Perhaps just two: Full Interpretation (FI) holding at each interface (Chomsky 2006: 13). [Why FI? Because that's precisely what the SMT assumes – a perfect, nonredundant mapping to the interface, i.e. no redundant, uninterpreted symbols.]
- *Phase-cyclic computation* (Chomsky 2000, 2001, 2004, 2005, 2006) plays a crucial role in ensuring FI is met, thus contributing to efficient satisfaction of IC:
  - To satisfy FI efficiently, uFs must be deleted as soon as they are valued, cf. Epstein & Seely 2002. This is because interpretable and valued uninterpretable features are indistinguishable at the interface (Chomsky 2001, 2006). The other options for telling them apart, such as adding a delay to valuation or reconstructing the derivation (lookback), increase the computational burden – it is simpler just to transfer these features immediately.
  - Deletion occurs as part of Transfer (Chomsky 2006). Therefore, to ensure Value-Transfer simultaneity, Transfer must occur every time a uF is valued. That is, we have multiple Transfer – *phases*.
- However, if the entire phase (CP/*v*P) is transferred, phases always *violate* FI (due to uF within the phase, e.g. Case on the external argument, causing crash). Efficient satisfaction of IC (FI) is thus not yet achieved. To achieve it, some version of the *Phase Impenetrability Condition* (PIC) is required: only a subpart of the phase can be transferred (the complement of the phase head).

  ➢ Efficient satisfaction of IC (FI) requires both (A) Value-Transfer Simultaneity and (B) PIC.

- Unfortunately, (A) and (B) are incompatible as they stand (Richards 2007b). Valued uFs, located on the phase head, must be transferred immediately (by (A)), but the phase head (and its specifiers) cannot be transferred until the *next* phase level (by (B)).
  ➢ Valued uFs must therefore descend onto the complement domain, since that is the part which actually gets transferred (and deleted):

(9) uF must descend from edge to nonedge (i.e. from C to T, *v** to V, etc.).

- This is feature-inheritance, which now follows from 'good design'/SMT (it ensures FI is met, by enabling Agree-features to be valued and deleted as part of Transfer, at the phase level).

  ➢ We derive that efficient satisfaction of FI requires three components: Value-Transfer Simultaneity, PIC, and Feature-Inheritance.

- Further, it now follows from SMT that every phase head must be followed by a nonphase head (in order to receive its inherited features). However, any further nonphase heads beyond this do not follow from SMT – that is, Factor III provides only for sequences of functional heads of the kind in (10e), i.e. pairs of phase heads (P) and nonphase heads (N).

(10) a. * P – P – P – P ...
b. * P – P – N – P ...
c. * P – N – N – P ...
d. * N – N – N – N ...
e. √ P – N – P – N ...

  ➢ A phase consisting of two heads, P-N, is thus the minimum phase in conformance with SMT. The basic sequence provided for by Factor III (P-N-P-N) yields the sequence of Core Functional Categories C-T-v(-V), which we may thus take to be obligatorily present in every language, as a Factor-III effect.
  ➢ No positive evidence is thus required for children to acquire the basic sequence P-N-P-N (C-T-v-V): children minimally expect {P-N} pairs.

- Interaction with Factor II: Sequences such as (10c), with additional N heads, may still be acquired, as long as there is positive evidence for them in the input (i.e. Factor II; cf. Gallego 2008). Thus Factor III interacts with Factor II to yield richer functional sequences in accordance with the parameter schema in (4): this is a point of parametric variation compatible with the minimalist system.

2. MACROPARAMETERS AND EXTERNALIZATION

2.1 Macroparameters in MP

With parametrized principles unformulable under MP assumptions (SMT and SUT; section 1), the status of macroparameters is unclear and work on these has largely been abandoned (cf. Baker 1996, 2001).

- One possibility: Macroparameters arise through a conspiracy amongst microparameters to have identical or 'harmonic' settings (Boeckx 2008, Holmberg & Roberts 2008). As such, clustering becomes a third-factor effect due to general learning strategies that reflect the conservatism of the learner (bias towards simplicity):

Boeckx 2008's *Superset Bias*: "Strive for parametric value consistency among similar parameters" (thus economizing on memory);
Holmberg & Roberts 2008's *Generalization of the Input*: "If acquirers assign a marked value to H, they will assign the same value to all comparable heads" (a markedness preference)

- Another possibility: Macroparameters arise at the syntax-PF interface through the interface-imposed need to linearize a symmetrical syntax – Richards 2004.
- Chomsky 2005, 2006, Berwick & Chomsky 2008: There is a 'PF-LF' asymmetry in language design: SMT holds only of the mapping to the SEM-interface, with "externalization [mapping to PHON] a secondary process".

Essentially, we can think of the relation between syntax and the interfaces as 'SEM dictates, whereas PHON makes do'. That is, the mapping to PHON is imperfect – and thus open to variation:

(11) "Parametrization and diversity, then, would be mostly – possibly entirely – restricted to externalization. That is pretty much what we seem to find: a computational system efficiently generating expressions interpretable at the semantic/pragmatic interface, with diversity resulting from complex and highly varied modes of externalization, which, furthermore, are readily susceptible to historical change." (Berwick & Chomsky 2008:15)

Variation in a minimalist, third-factor-based system thus arises from (at least) two, related sources:

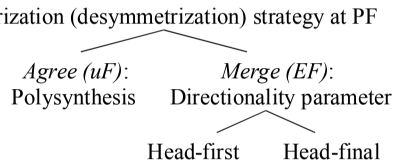
- imperfect externalization;
- underspecification of UG (leaving options open), yielding "variation in the externalized aspects of language" (Boeckx 2008).

A UG not specified for externalization leaves many options open in the morphophonological component, most notably, linearity/ordering (head-complement ordering, copy spell-out, etc).

- Merge is symmetrical (as part of a maximally empty UG, Merge is maximally unspecified, and thus specifies no particular ordering, contra Kayne 1994 – cf. Chomsky 2000, Boeckx 2008). The physics of speech demands that one or other logical order (precedence/subsequence) be imposed between heads and complements – if language were perfectly designed for PHON, then we might expect this requirement to be imposed on the syntax as a third-factor effect/IC, requiring syntax to yield asymmetric expressions to this interface. That this does not happen (i.e. Merge is symmetrical) thus supports Chomsky's position that "externalization is secondary", not part of optimal design.
- The symmetrical expressions transferred to PHON are thus illegible at that interface, i.e. they violate FI/IC at PHON. The sensorimotor interface has to 'make do' with what the syntax gives it, which means it must find its own solution to the legibility problem. Macroparameters might then naturally emerge as PF-repair strategies.
- Given the two basic operations provided by the minimal UG – Merge (from EFs) and Agree (from uFs) – we might expect two basic macroparameters, depending on whether linearization is determined by Merge or by Agree. The former gives us the traditional head-complement parameter (the intolerable optionality in head-complement ordering that results from an underspecified UG must be resolved one

way or the other, with two logical possibilities – head-first or head-final); potentially, the latter gives us something akin to Baker's (1996) polysynthesis parameter (linearization by Agree might involve arguments being expressed at the probe, in the form of agreement morphemes and/or incorporation – cf. Roberts 2006 for an analysis of head-movement and cliticization in these terms).
- Whether a language linearizes through Merge or Agree is then a super-macroparameter:

(12)



Let us consider how the Merge-based desymmetrization strategy might work.

2.2 VO/OV and Holmberg's Generalization

If the hallmark of a macroparameter is a clustering of grammatical properties around a single parametric setting, then we would expect a VO/OV linearization strategy located at the PF-interface to yield additional effects beyond directionality.

Richards 2004:

- Merge pairs (sisters, mutual c-command, the head-complement relation) are the basic unit of linearization (cf. Epstein et al 1998): Merge partners in the syntax become ordering partners at PF.

A good candidate for such a sisterhood-based linearization strategy is that offered in Epstein et al (1998: Chapter 5). They propose that it is simple c-command, rather than the arguably more complex notion of asymmetric c-command, that translates to precedence at PF.

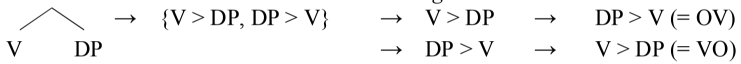
- Merge-pairs *over*determine linearization, rather than underdetermine it (cf. LCA under Bare Phrase Structure, which is unable to linearize the symmetrical base pair of every sub-tree): since head and complement mutually c-command each other, contradictory instructions are provided to PF by the sisterhood relation, such that each sister must precede the other.
- Linearization-by-sisterhood must therefore proceed via a strategy of deleting superfluous (contradictory, ambiguous, symmetrical) information at the interface.

To this end, Epstein et al propose a *Precedence Resolution Principle* (PRP), essentially (13).

(13) *PRP:* If two categories symmetrically c-command each other, ignore all c-command relations of one of the categories to the other. [Based on Epstein et al 1998:152]

From this interface deletion strategy, the effects of a head-complement directionality parameter are immediately implied:

(14)



- Richards 2004: (13)/(14) should be generalized to hold of *internal* as well as external Merge (thus imposing VO/OV 'shape' on Move and Merge alike). That is, the PRP must delete a *consistent* subset of c-command instructions in any given language.

Assume that the unordered set created by Merge(α,β) delivers contradictory instructions to PF (essentially, {<α,β>, <β,α>}), i.e. unordered sets translate to mutual precedence at PF. Then:

(15) *Parametrized desymmetrization*: given Merge(α,β) → {<α,β>, <β,α>}:
- a. Head-initial = Delete all Comp > Head [i.e. {<α,β>, <β,α>} → {<α,β>}]
- b. Head-final = Delete all Head > Comp[1] [i.e. {<α,β>, <β,α>} → {<β,α>}]

- (15) is a parameter operative at the syntax-PF interface, ensuring linearization of symmetrical syntax (a 'parametrized desymmetrization strategy'). It has the effect of a head parameter but maintains a uniform UG/syntax, in conformance with SMT/SUT.

Consequence: The linear order-preservation effect known as *Holmberg's Generalization* (HG) that constrains Germanic Object Shift now immediately follows from the 'head-initial' (VO) setting (15a) of this parameter.

- **Object Shift/Scrambling:** short leftward displacement of weak (destressed) objects in Germanic (Holmberg 1986, 1999, Vikner 1995, Collins & Thráinsson 1996, Thráinsson 2001, and many others) is constrained such that the shifted object cannot cross an in-situ (nonfinite) lexical verb (HG).
- HG makes little sense on familiar minimalist grounds, and the technology used to account for it in the literature (based on equidistance, Greed, government, phonological visibility, topic/focus-features, etc.) arguably exceeds SMT.
- HG is best (most simply) viewed as a verb-object order preservation constraint: *the derived order must reinstate the base order (VO)* – Müller 2000, 2001, 2006; Fox & Pesetsky 2003, 2005; Williams 2003; Richards 2004; Koeneman 2005.
- This constraint (HG) famously holds only of VO languages: compare VO (16) and OV (17). In (17), short scrambling of the object may occur irrespective of the finiteness of the main verb, yielding both OV (= (17b)) and, with V2-raising, VO orders (= (17a)) alike, i.e. 'anti-HG' effects.

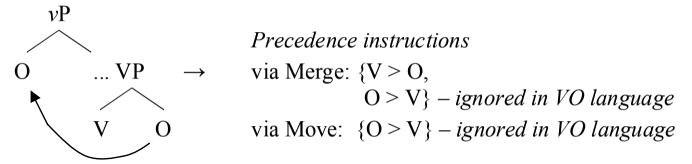
(16) VO (Icelandic)
- a. *Nemandinn las **(bókina)** ekki **(bókina)***
  The-student read (the-book) not (the-book)
  "The student didn't read the book."
- b. *Nemandinn hefur (***bókina**) ekki lesið **(bókina)***
  The-student has (the-book) not read (the-book)
  "The student hasn't read the book."

[1] The "all" in (15a,b) will be qualified immediately below.

(17) OV (German)
- a. *Der Student las **(das Buch)** nicht **(das Buch)***
  The student read (the book) not (the book)
  "The student didn't read the book."
- b. *Der Student hat **(das Buch)** nicht **(das Buch)** gelesen*
  The student has (the book) not read (the book)
  "The student hasn't read the book."

HG emerges from the linearization algorithm in (15), and thus conforms to SMT after all:

(18)



- ➢ Object displacement over V in a VO language is only orderable by (15a) where further movement of V over O takes place. That is, the displaced O must be the tail of a V > O chain, and not the head of an O > V chain.
- ➢ Thus HG is derived for exactly that subset of languages in which it holds (namely, those set to (15a), i.e. VO languages).
- ➢ OS is therefore only available in VO languages with independent verb movement (V-to-T, e.g. Icelandic).
- ➢ Scrambling is free in OV languages without any further V movement (i.e. anti-HG, cf. (17b)), as the resulting O > V instructions are uniformly legible at PF in languages set to (15b).

This yields order-preserving movement, as in (16)-(17b). What about order-distorting movement, cf. (19), (17a)? Clearly, HG does not hold of all movements, even in VO languages. Standard A/A-bar movement (passivization, topicalization, *wh*-movement, etc.) freely disrupts basic VO, resulting in surface OV:

(19)
- a. A man arrived (*a man*)
- b. John was rescued (*John*)
- c. John, I like (*John*)
- d. Which book did you read (*which book*)

  - ➢ The HG property of OS (i.e. head-complement order preservation) is unique to short-distance, A-type OS/Scrambling: it is perhaps its defining property.
  - ➢ That is, some types of movement conserve head-complement 'shape' whereas others do not.

QUESTION: Why? How can a single grammatical system (e.g. VO Icelandic) allow both shape-breaking (passive, *wh*-movement) and shape-preserving (HG/OS) movement operations?

Two reasonable claims/assumptions:

(20) All varieties of *v* are phase heads – i.e. both transitive *v** and passive/unaccusative 'defective' *v* (v_{def}); cf. Legate 2003, Richards 2004.

(21) OS/Scrambling (of the short, TP-internal, 'A-movement' type) targets spec-*v*P, as in (18); cf. Neeleman & Weerman 1999, Chomsky 2001, Kitahara 2002.

- Order-preserving object movement is thus 'shorter-distance' than order-disrupting object movement: the former targets spec-*v*P and is thus phase-internal; the latter targets spec-TP (A-movement / raising-to-subject) or spec-CP (A-bar movement), i.e. positions outside of the original phase.

We can now state a simple generalization covering all of (16)-(17), (19), as follows:

(22) Order-preserving movement is phase-internal;
Order-disrupting movement is cross-phasal.

- (22) is completely as expected given Chomsky's complexity-reducing view of phases as the unit of derivational memory – i.e. the idea that derivational information (including, plausibly, ordering information) is simply 'lost' ('forgotten') at the end of the phase. It follows that any constraints on linearization (such as order-preservation, i.e. the requirement of consistent precedence instructions) can only hold within a single phase (and not beyond, contra Fox & Pesetsky 2003/5).
- (22) thus draws a principled line between shape-breaking and shape-preserving movement:
  - Escaping a phase implies escaping shape (i.e. the order-preservation effects imposed by (15)).
  - This is due, simply, to the cyclic purging ('forgetting') of derivational information at the phase level: At the point where the higher phase (CP) is transferred and linearized, there is no longer any memory that the Object DP was originally merged with V in the lower phase (*v*P). Instead, the Object DP is 'relinearized' in the higher phase with its derived sister (a projection of T/C).

In sum:

- Head-complement order-preservation (HG) is a linearization effect confined to within the phase.
- VO/OV 'shape' is imposed at the phase level (perhaps by (15)), but beyond this level, the system no longer cares, hence the ordering freedom associated with longer-distance movement (again, an **underspecified UG** relying on third-factor principles yields **optionality**, here of ordering beyond the phase level).
- This result is exactly as expected if the relevant ordering information is lost at the end of every phase, i.e. if linearization is localized to the phase and thus strongly cyclic. Thus HG is a corollary of SMT-conforming phase-cyclic computation and third-factor (interface-imposed, syntax-external, non-UG-specified) linearization.
- Our macroparameter in (15) yields both OV/VO, +/− OS and +/−HG – a parametric cluster.

2.3 Optionality in phasal composition and crosslinguistic variation in CED effects

The directionality parameter of the previous section is located post-syntactically, in the morphophonological component. We now turn to a presyntactic indeterminacy in the

composition of phases, whose resolution also yields a macroparametric clustering of properties.

Empirical problem: Subextraction from subjects (CED effects)

- Much recent work has been undertaken in this area, inspired by Chomsky's (2005) phase-based analysis (see, e.g., Gallego & Uriagereka 2006a,b, Lohndal 2007a).
- DPs in derived positions are 'frozen'/'opaque' – they act as islands for subextraction:

(23) a. $[_{CP}$ [Which candidate]$_i$ were $[_{TP}$ there $[_{vP}$ [posters of t_i] all over town]]]?
b. * $[_{CP}$ [Which candidate]$_i$ were $[_{TP}$ [posters of $t_i]_j$ $[_{vP}$ t_j all over town]]]?

- *Spec-TP is a 'freezing' position*

- Certain base positions are also opaque for subextraction – specifically, the edges of (strong) phases (Chomsky 2005):

(24) a. *$[_{CP}$ [Of which car]$_i$ did $[_{TP}$ [the driver $t_i]_j$ $[_{v*P}$ t_j cause a scandal]]]?
b. $[_{CP}$ [Of which car]$_i$ was $[_{TP}$ [the driver $t_i]_j$ $[_{vP}$ awarded t_j a prize]]]?
c. $[_{CP}$ [Of which car]$_i$ was $[_{TP}$ [the driver $t_i]_j$ $[_{vP}$ arrested t_j]]]?

- *Spec-v*P, but not spec-vP (defective v), is a 'freezing' position*
  [hence (36a) fine too – *v*P there is not transitive *v**P so not a 'strong' phase]

- To capture the latter restriction, Chomsky (2005) proposes the following:

(25) *Edge Condition*
Syntactic Objects in *phase edges* become internally frozen.

- The internal structure of phase-edge material is inaccessible to higher probes.

- However, Subject Condition effects are subject to crosslinguistic variation – they do not hold of all languages, e.g. Spanish, where postverbal subjects *do* allow subextraction:

(26) *De qué conferenciantes*$_i$ *te parece que …*
of what speakers CL.to.you seem.3.SG that…
a. *me*$_j$ *van a impresionar*$_k$ $[_{v*P}$ $[_{DP}$ *las propuestas* $t_i]$ t_j t_k]?
CL.to.me go.3.SG to to.impress the proposals
b. *$[_{DP}$ *las propuestas* $t_i]_l$ *me*$_j$ *van a impresionar*$_k$ $[_{v*P}$ t_l t_j $t_k]$?
the proposals CL.to.me go.3.SG to to.impress
'Which speakers does it seem to you that the proposals by will impress me?'
(Uriagereka 1988: 118)

- (26a), where the subject stays in situ (spec-*v**P), nevertheless allows subextraction. Only the derived position (e.g. spec-TP in (26b)) blocks subextraction. Thus (25) does not appear to hold of Spanish.

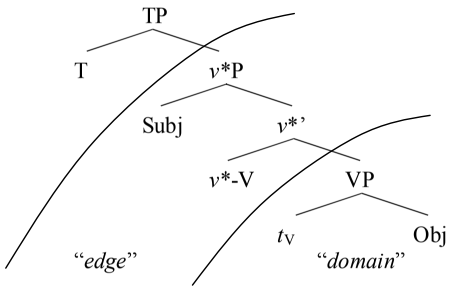
- Gallego's (2005) solution/proposal:
  A phase-based implementation of the old idea that T (S/INFL) has a special status in Romance (cf. Kayne 1989 on L-marking of VP by INFL through V-to-I movement; Rizzi 1978, 1982 on Italian bounding nodes; Jaeggli 1982, Uribe-Etxebarria 1992 on Romance spec-T as an A-bar position, etc.):

  - 'Modern translation': TP behaves like a phase in Spanish; *v*P behaves like it isn't one.

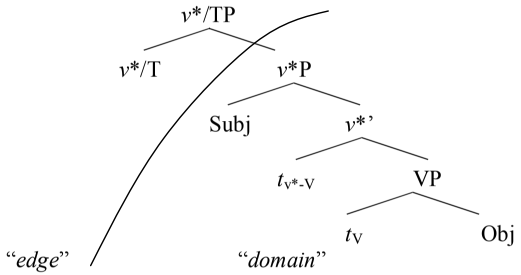
- Does this mean phases themselves are parametrizable? This would be problematic from the point of view of an invariant computational system (one that conforms to SMT and SUT), with phases motivated by considerations of complexity reduction. Thus, we would not want to associate uFs with T instead of C/v in Spanish.
- Instead, the proposal is that phase effects on T are simply derivative, a side-effect of *v**-to-T movement (i.e. *v** remains the phase *per se*; it simply extends its domain; cf. also den Dikken 2006):

(27) ***Phase sliding***:
Movement of a phase head effects an "upwards percolation" of phase properties.

(28) a. <u>*English*: no *v**-to-T</u>



b. <u>*Spanish*: *v**-to-T</u>



- As a result of (28b), the in-situ subject in spec-*v**P in (26a) is no longer in the edge of the phase but rather in the complement domain, which explains why it is not frozen: (25) does not apply to this position.
- Further, the edge of the extended (slidden) phase is now T (the "amalgamated" *v**-T head) and its specifiers. The opacity of spec-TP for subextraction in (26b) thus now reduces to (25) (that is, the 'freezing' property of the derived position now emerges as a phase-edge effect).
- As well as deriving the anti-CED effect in (26a), the status of T as phase should also be responsible for the rest of the cluster of properties associated with T/I's special status in Spanish/Italian/Romance (cf. Rizzi 1982), including lack of *that*-trace effects and the availability of free inversion (/extra subject position).

<u>Problem I</u> (see also Lohndal 2007b):

- Phase Sliding cannot hold of all *v*-to-T languages, even within Romance, where French is a clear counterexample (postverbal subjects are not subextractable despite verb raising to T):

(29) a. *[*De quelle équipe*]$_i$ *est-ce que John a dit* [$_{CP}$ *que* [$_{TP}$ [*beaucoup de* of which team is-it that John has said that many of *joueurs* t_i]$_j$ *ont accusé* t_j *l'arbitre*]]?
players have accused the.referee
'Of what team has John said that many players have accused the referee?'

b. *[*De quelle équipe*]$_i$ *est-ce que John a dit* [$_{CP}$ *qu'ont accusé*
of which team is-it that John has said that.have accused
[$_{v*P}$ *l'arbitre*$_j$ [*beaucoup de joueurs* t_i] t_j]]?
the.referee many of players

(Gallego 2005: 86)

[To address this point, Gallego 2005 (p.29ff.) proposes additional parameters to ensure that the phase-sliding effect of head movement only obtains in rich-agreement null-subject languages.]

<u>Problem II</u>:

- Phase-sliding is incompatible with feature-inheritance (section 1.3). If the property that defines phases is uF (Chomsky 2005, 2006), then movement of v-to-T will fail to carry 'phasehood' up to T, since upon completion of vP, v's uF is inherited onto VP and transferred.

<u>Alternative to Phase-Sliding</u>: variable phase composition (Richards 2008b)

- Chomsky's (2000) original conception of phases as lexical subarrays (LAs): subparts of the numeration placed in active memory, each containing a single instance of a phase head (C or v) – derives anti-Merge-Over-Move effects.
- Richards 2008b: We should retain this conception of phases, since it allows all the otherwise stipulated properties of phases encoded in the PIC to be eliminated.
- In particular, LAs define the units actually sent to spell-out.

To see this, consider the sequence of Core Functional Categories (phase heads and nonphase heads) derived from SMT in section 1.3:

(30) P – N – P – N

If every LA must contain one instance of a phase head, then only two logical possibilities for forming LAs arise:

(31) a. {P – N} – {P – N} *or*
b. … P} – {N – P} – {N …

Substituting the CFCs, C-T-v:

(32) a. {C – T} – {v – V} *or*
b. … C} – {T – v} – {V …

- In principle, either is possible – another indeterminacy, and thus another point of parametric variation.

If (32b) is correct, at least for English, then we immediately derive various mysterious properties of phases, such as what gets spelled out (the complement domain instead of the whole phase}, why v and C act as phase heads and not T, and why v acts as an edge for VP-internal material:

(33) Spell out phase LA_n at LA_{n+2} (ensures the provision of an edge; cf. section 1.3):

{... C}	{T, *v*}	{V, DP}
LA_{n+2}	LA_{n+1}	LA_n
↑	↑	↑
'spellout trigger'/ 'phase head'	'edge'	'spelt-out unit'/ 'complement domain'

- VP gets spelled out (and not vP) since it is the phase defined by the LA {V, DP} that is transferred.
- C and v act as phase heads (Transfer triggers) since they are the first heads to be introduced in a new LA, thus marking the completion of a previous phase and triggering its Transfer.

- Under Chomsky's (2005, 2006) LA-less characterization of phases as the locus of uFs, in order for T to become a 'phase head', uF would have to be associated with T instead of C/v – a mere lexical change perhaps, but one that would come at the expense of weakening the computational rationale for phases (contra SMT).

- However, the approach to phases in (31)-(32) is readily parametrizable, due to the inherent indeterminacy of (31a)/(32a) versus (31b)/(32b): For T to act as 'phase head', a language just has to choose (32a) instead of (32b).

- Should the cluster of properties associated with anti-CED effects in Spanish indeed be adequately capturable in terms of T being the phase in this language (as per Gallego 2005, 2008), then (31) provides just the parameter we need, allowing for exactly this dimension of variation, which again results from a uniform, SMT- and SUT-conforming UG:

(34) a. Spanish = {C – T} – {v – V} (32a)
b. English = C} – {T – v} – {V (32b)

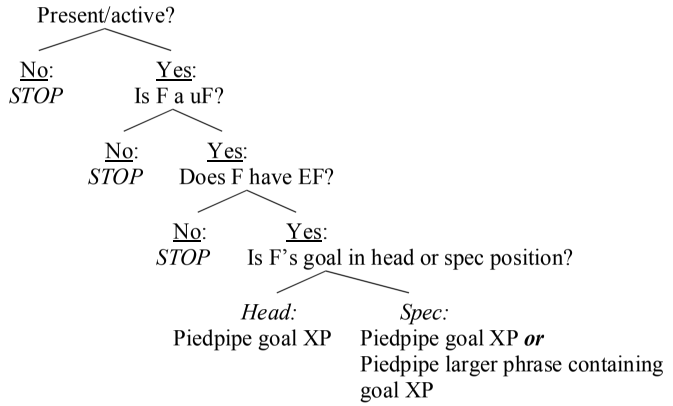
3. MICROPARAMETERS AND SYNTACTIC OPTIONALITY

Roberts 2007:

- All parameters must be fixed one way or another – thus for all features F, (4) must be set to either 'Yes' or 'No'.
- Further, there can be no optional associations of, e.g., F with EF to yield optional movements. Either F is associated with EF (the movement trigger) in a given language or it is not; to allow for optional EF-associations is tantamount to allowing for optional, competing grammars (cf. Kroch 1989), since it is this very feature (EF, its presence or absence) which defines a parameter.

This is shown in (35), a more articulated and specified version of the parameter schema in (4), based on Holmberg & Roberts 2008, slightly modified in light of section 1.2, and enriched in line with the theory of (head- vs. specifier-)piedpiping in Biberauer & Richards 2006.

(35) *Partial parameter schema for feature F*



- The indeterminacy left open by UG (i.e., whether a feature F is present in a language or not, and whether it is associated with EF (etc.) or not) yields a parameter that must be fixed one way or another (e.g., for (35), a feature F may be set to {Yes-Yes-Yes-Spec}).
- However, certain settings may yield systems with further indeterminacies left open by the underspecified, maximally empty UG. The {Yes-Yes-Yes-Spec} setting of (35) yields such a scenario. Biberauer & Richards 2006 discuss numerous empirical cases. One is given below, by way of illustration.

3.1 True optionality in Russian *wh*-movement (Biberauer & Richards 2006)

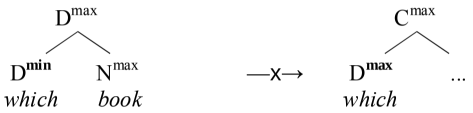
(36) a. *Whose book did you read?*
b. **Whose did you read book?*

(37) a. *Č'ju knigu ty čital* [Russian]
Whose book you read
b. *Č'ju ty čital knigu*
Whose you read book

- Movement of the *wh*-element to C[+*wh*] requires the whole DP to be piedpiped along in English – the Left-Branch Constraint (LBC).
- As is well-known, the LBC does not appear to be operative in all languages. Thus, as (37b) shows, the non-piedpiped counterpart of (37a) is grammatical in Russian. However, piedpiping is not barred in Russian, but simply optional (cf. (37a)). This is a case of semantically vacuous, true optionality (i.e. optionality which cannot be captured in terms of the Chomsky-Reinhart-Fox rationale for optional operations as having an interpretive effect on outcome).

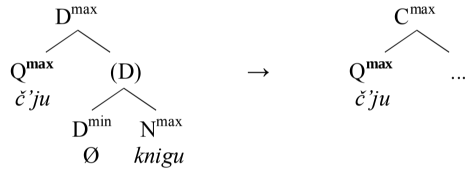
  ➢ Why this optionality in Russian?

(38)



- Piedpiping is forced in English if we assume that *wh*-elements such as *what* and *which* are *wh*-determiners and thus the head of their respective *wh*-DPs. If, for some reason (e.g. Chain Uniformity, Chomsky 1995), heads cannot undergo *wh*-movement, these elements cannot raise independently to spec-CP. Piedpiping of the whole *wh*-DPs is thus forced in English (Biberauer & Richards refer to this as *head-piedpiping*).

(39)



- Assuming Russian to lack (overt exponents of) the category D, Russian *wh*-elements such as *kakoj* ('what', 'which'), *kotoryj* ('which', 'what'), *čej* 'whose' must be quantificational phrases (QP) occupying the specifier of the relevant DPs. As such, they have the status of maximal projections and are thus able to raise independently to spec-CP without incurring a uniformity violation; hence the lack of LBC effects illustrated in (37).
- Alternatively, the entire DP (containing the goal QP in its specifier) may be piedpiped, since this too conforms to uniformity and allows satisfaction of the relevant EF on C. Thus as long as the C probe's EF is satisfied, the grammar doesn't care how. Such true optionality is characteristic of *spec-piedpiping*.

Essentially, system-internal, formal optionality results from the following configuration (adapted from Roberts 2007: 308):

(40) uF_{EF} … [$_{XP}$ … YP_{goal} …]

➢ To satisfy EF on uF, either YP may raise or the containing XP may do so – the grammar doesn't mind.

Biberauer & Roberts (2005, 2007) and Roberts (2007) apply this idea to the domain of *language change* in the history of English and elsewhere.

- As Roberts (2007:309, 332-3) points out, such instances of formal optionality provide a possible solution to the logical problem of language change in an approach based on imperfect acquisition (parameter resetting/missetting/abductive change; cf. Lightfoot 1979, 1991, 1999): namely, how can Generation 1 produce a trigger experience for Generation 2 sufficient for Generation 2 to set a different parameter value from Generation 1? In order to produce that PLD in the first place, it would seem that Generation 1 would need to have the alternative, innovative setting too.
- As long as doublets (as in (37)) are generable by a single grammar (/a single generation), there is variation inherent in the PLD. All that is then required is for one of the doublets to become more frequent than the other – e.g. by taking on sociolinguistic, pragmatic prominence – and it will oust the other, in so far as a retreat to the non-doublet-generating grammar is possible.
- The latter depends on linguistic factors: is a grammar that just generates the prominent alternant possible? Thus a grammar that just generates option (37b) and not (37a) is impossible in this system, whereas a grammar just generating option (37a) is fine – cf. English.
- Such a retreat also depends on psycholinguistic factors (learning strategies implicated in acquisition). Thus, as Biberauer & Roberts (2007) note, the two grammars involved in cases of true optionality stand in a subset-superset relation – in the case of Russian and English *wh*-movement, English is a subset of Russian. Should the non-piedpiping doublet (37b) lose sociolinguistic value and thus disappear sufficiently from the PLD, then a retreat to the subset grammar, generating only (37a) as in English, is predicted (syntactically, this would involve a categorial reanalysis of Russian *wh*-words from QPs to D heads), since there would be no positive evidence for the superset grammar. Exactly such a change has arguably taken place in the history of Greek (see Biberauer & Richards 2006 for details).
- Thus an acquisitional strategy, the Subset Principle, may well be implicated in language change in precisely these scenarios – where true optionality involving subset-superset relations arises from an underspecified UG. A convergence between language acquisition and language change thus becomes possible and detectable here.

3.2 Optional EPP effects

The phase heads (C and v) allow optional movements to their edges (specifier region), associated with extra interpretive effects ('edge semantics'), in line with the Chomsky/Fox/Reinhart rationale in (41):

(41) *Interface economy* (Chomsky 2001: 34)
An optional rule/operation can apply only when needed to yield a new outcome

Optional movements to C: topicalization, focus, etc.: (42).
Optional movements to v: Object Shift and scrambling: (43).

(42) a. John, I like (but not Paul).
b. ...dat [zo'n foto van haarzelf$_i$]$_j$ zelfs Jan deze actrice$_i$ niet graag t$_j$ toont
that such picture of herself even Jan this actress not gladly shows
(Dutch focus scrambling; Neeleman 1994:399)

(43) a. *Er hat oft ein Buch gelesen* (German)
he has often a book read
'He often read a (non-specific) book.'
b. *Er hat ein Buch oft gelesen*
he has a book often read
'There's a book that he often read.'

- In line with (41), these optional movements are associated with such interpretive effects as topic, focus, old information, specificity, etc.

Analysis (Chomsky 2000): as phase heads, C and v can have EPP-features optionally added to them, thus allowing phase-internal material to be displaced to their edge and move into higher phases.

In terms of the Edge Feature of Chomsky 2005, 2006, we have pure EF-movement (i.e. movement without Agree): see Nevins 2004, Richards 2007a, Yang 2007.

PROBLEM: Optional EPP on nonphase heads, e.g. T, as in Icelandic (44), versus obligatory EPP on T, as in English (45):

(44) Icelandic 'optional EPP' effects (from Bobaljik & Thráinsson 1998)
a. Í fyrra luku **þrír stúdentar** [$_{VP}$ víst öllum prófunum]
last year finished three students apparently all exams-the
'Three [=specific] students apparently finished all the exams last year.'
b. Í fyrra luku [$_{VP}$ víst **þrír stúdentar** öllum prófunum]
last year finished apparently three students all exams-the
'Last year, there were three students [=existential] who finished all the exams.'

(45) a. There appeared **a face** at the window
b. **A face** appeared at the window
c. * Appeared **a face** at the window

- Whilst the obligatory EPP property on T in a language like English can be easily captured in terms of the parameter schema in (35) – the head T (or its Agree-feature, uF) is associated with EF, i.e. 'Yes-Yes' – optional EPP (as in Icelandic) should not be possible, since (i) T is not a phase head (and so cannot have Agree-less EF/EPP-features added to it), and (ii) there can be no optional associations of EFs with heads (cf. Roberts 2007, above): either a language is set to 'Yes' or to 'No'.

- Thus, for movement to spec-TP to be possible at all in languages like Icelandic, the EF-property must be set to 'Yes'. The lack of movement in (19b) is then a problem – EF/EPP on T goes unsatisfied (this is esesntially just the old problem of optional movements in Minimalism).

Proposal: Underspecified UG leaves open the possibility of associating the obligatory EF/EPP-feature with a *subset* of a head's uFs.

- Assuming the uFs on T to comprise at least two probes, Person and Number, there are at least two possibilities for associating EF + uF:

(46) a. EF + [uPerson, uNumber]
b. EF + [uPerson] (or EF + [uNumber])

Claim: Optional EPP effects arise where EF is associated with Person alone:

(47) $EPP_{optional}$ = EF + [uPerson]

(48) PREDICTION: Lack of Person-checking/valuation should correlate with lack of movement (cf. Boeckx 2006).

(49) OBSERVATION: As has been widely remarked, nonreferential arguments (nonspecific indefinites) are less conducive to raising than referential (index-bearing, specific, definite) arguments – e.g. Icelandic subjects (44), Mandarin Chinese (50), etc.

(50) a. *Kèren lái-le* (Mandarin Chinese)
guest come-$_{PFV}$
'(The) guests came.'
b. *Lái-le kèren*
come-$_{PFV}$ guest
'There came (some) guests.'

(48) and (49) are transparently linked if nonreferential arguments lack a Person specification (Richards 2008a,c).

- *Indefinites* (nonreferential, non-D-linked, new information, etc.) are often claimed to be structurally different from definites (referential, index-bearing, D-linked DPs) – they lack D, or Case, or phasehood, etc.
- *Claim*: indefinites and bare nouns lack Person, i.e., they are defective in the agreement system (Richards 2004). Person is a lexically specified property only on definites (perhaps contributed by the D head), as here it is not independently predictable.
- This claim seems semantically plausible: bare nouns and indefinites are invariably (interpreted as) third-person, thus a third-person specification (feature/value) on indefinite DPs or bare nouns would seem redundant.
- Thus, in languages with optional EPP effects (on T, *v*, etc.), the Person-associated EF will trigger movement only for definites, i.e. those arguments with which the EF-associated Person probe enters (Match and) Agree.
- There is no need to add 'EPP-features' optionally to the T head: their EPP property is always present, in the form of Person-associated EF.
- We thus have a further point of syntactic indeterminacy left open by UG, yielding a parameter of EPP-types:

(51) a. EF + [uPerson, uNumber] = obligatory EPP (e.g. English, French)
b. EF + [uPerson] = optional EPP (e.g. Icelandic, German)

4. SUMMARY

- Parameters – both micro and macro – arise in an SMT-conforming FL (one with a uniform and maximally empty UG), as points of indeterminacy/optionality, principally at the PHON-interface (externalization): none of these parametric choices affects meaning, lending support to Chomsky's (2005, 2006) and Berwick & Chomsky's (2008) position that language is optimally designed to meet IC only at SEM (SEM is thus uniform, with a single design solution and no room for variability).
- The points of indeterminacy left open by the underspecified UG may be lexical (Factor I), involving feature distributions (sections 1.3 and 3), such as the presence versus absence of a feature and/or the way in which a feature is associated with other features (e.g. EF-uF associations, section 3.2).
- Additionally, there may be nonlexical points of indeterminacy, such as those arising through the symmetry of Merge or the basic composition of lexical subarrays. These have their basis in third-factor principles (linearization, phase-cyclic computation).

REFERENCES

Baker, M. (1996) *The Polysynthesis Parameter*. Oxford/New York: OUP.
Baker, M. (2001) *The Atoms of Language*. New York: Basic Books.
Baker, M. (2008) *The Syntax of Agreement and Concord*. Cambridge: CUP.
Biberauer, T. & M. Richards (2006). True Optionality: When the grammar doesn't mind. In C. Boeckx (ed.) *Minimalist Essays*. Amsterdam: Benjamins, pp. 35-67.
Berwick, R. & N. Chomsky (2008). The Biolinguistic Program: The Current State of its Evolution and Development. Ms., MIT. Forthcoming in *Biolinguistic Investigations*, di Sciulo & Aguero (eds.), MIT Press.
Biberauer, T. & I. Roberts (2005). Changing EPP-parameters in the history of English: accounting for variation and change. *English Language and Linguistics* 9: 5-46.
Biberauer, T. & I. Roberts (2007). The Return of the Subset Principle. Ms., University of Cambridge.
Bobaljik, J. & H. Thráinsson (1998). Two heads aren't always better than one. *Syntax* 1: 37-71.
Boeckx, C. (2006). The Syntax of Argument Dependencies. Paper presented at Linguistics Colloquium, University of Leipzig, November 2006.
Boeckx, C. (2008). Approaching Parameters from Below. Ms., Harvard.
Borer, H. (1984) Parametric Syntax: Case Studies in Semitic and Romance Languages. Dordrecht: Foris.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass: MIT Press.
Chomsky, N. (2000). Minimalist Inquiries: the Framework. In *Step by step*, R. Martin, D. Michaels, & J. Uriagereka (eds), 89-156. Cambridge, Mass.: MIT Press.
Chomsky, N. (2001). Derivation by Phase. In *Ken Hale: A Life in Language*, M. Kenstowicz (ed.), 1-52. Cambridge, Mass.: MIT Press.

Chomsky, N. (2004). Beyond Explanatory Adequacy. In *Structures and Beyond. The Cartography of Syntactic Structures* (volume 3), A. Belletti (ed.). Oxford: Oxford University Press.
Chomsky, N. (2005). On Phases. Ms. MIT.
Chomsky, N. (2006a). Approaching UG from below. Ms. MIT.
Chomsky, N. (2006b). Turing's Thesis. Talk given at *InterPhases*, Cyprus, May 2006.
Collins, C. & H. Thráinsson (1996). VP-internal Structure and Object Shift in Icelandic. *Linguistic Inquiry* 24: 391-444.
den Dikken, M. (2006). Phase Extension: Contours of a Theory of the Role of Head Movement in Phrasal Extraction. In *Theoretical Linguistics* (2007).
Epstein, S. D., E. Groat, R. Kawashima & H. Kitahara (1998). *A Derivational Approach to Syntactic Relations*. New York: OUP.
Epstein, S. D. & T. D. Seely. (2002). Rule Applications as Cycles in a Level-free Syntax. In *Derivation and Explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 65-89. Oxford: Blackwell.
Fox, D. & D. Pesetsky (2003). Cyclic Linearization and the Typology of Movement. Handout. Paper presented at GLOW 2003 (Lund).
Fox, D. & D. Pesetsky (2005). Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31: 1-46.
Gallego, Á. (2005). Phase sliding. Ms., Universitat Autònoma de Barcelona.
Gallego, Á. (2008) The Second Factor and Phase Theory. Ms., Universitat Autònoma de Barcelona.
Gallego, Á. & J. Uriagereka. (2006a). Conditions on Sub-extraction. Ms., Universitat Autònoma de Barcelona & University of Maryland, College Park.
Gallego, Á. & J. Uriagereka. (2006b). Sub-extraction from subjects: A Phase Theory Account. Ms., Universitat Autònoma de Barcelona & University of Maryland, College Park
Hauser, M., N. Chomsky & W. T. Fitch (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298: 1569-79.
Holmberg, A. (1986). *Word Order and Syntactic Features in the Scandinavian Languages and English*. Doctoral dissertation. University of Stockholm.
Holmberg , A. (1999). Remarks on Holmberg's Generalization. *Studia Linguistica* 53: 1-39.
Holmberg, A. & I. Roberts (2008). Introduction. Ms., Universities of Cambridge and Newcastle.
Jaeggli, O. (1982). *Topics in Romance Syntax*. Dordrecht: Foris.
Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
Kitahara, H. (2002). Scrambling, Case, and Interpretability. In S. D. Epstein & T. D. Seely (eds.) *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell, pp. 167-83.
Koeneman, O. (2005). Shape conservation, Holmberg's generalization, and predication. Ms., Amsterdam.
Koeneman, O. & A. Neeleman (2001). Predication, Verb Movement and the Distribution of Expletives. *Lingua* 111: 189-233.
Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199-244.
Lasnik, H. & M. Saito (1991). On the subject of infinitives. In *Papers from the 27th regional meeting of the Chicago Linguistic Society*, 324-43. Chicago Linguistic Society, University of Chicago, Chicago, Ill. Reprinted in Howard Lasnik. 1999. *Minimalist Analysis*. Oxford: Blackwell.

Legate, Julie Anne (2003). Some Interface Properties of the Phase. *Linguistic Inquiry* 34: 506-16.
Lightfoot, D. (1979) *Principles of Diachronic Syntax*. Cambridge: CUP.
Lightfoot, D. (1991) How To Set Parameters: Arguments from Language Change. Cambridge, MA: MIT Press.
Lightfoot, D. (1999) The Development of Language. Oxford: Blackwell.
Lohndal, T. (2007a). Sub-extraction and the Freezing Effect: A Scandinavian Case Study. Ms., University of Oslo.
Lohndal, T. (2007b). V-T-C languages and sub-extraction: Against phase extension. Ms., University of Oslo.
Longobardi, G. (2001) Formal Syntax, Diachronic Minimalism, and Etymology: The History of French *Chez*. *Linguistic Inquiry* 32: 275-302.
Müller, G. (2000). Shape Conservation and Remnant Movement. Ms. University of Tübingen.
Müller, G. (2001). Order Preservation, Parallel Movement, and the Emergence of the Unmarked. In G. Legendre, J. Grimshaw & S. Vikner (eds.) *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press, pp. 279-313.
Müller, G. (2006). Towards a Relativized Concept of Cyclic Linearization. Ms., University of Leipzig.
Neeleman, A. (1994). Scrambling as a D-structure phenomenon. In N. Corver & H. van Riemsdijk (eds.) *Studies in Scrambling: Movement and Non-Movement Approaches to Free Word-Order Phenomena*. Berlin: de Gruyter, pp. 387-429.
Neeleman, A. & F. Weerman (1999). *Flexible Syntax: A Theory of Case and Arguments*, Dordrecht: Kluwer.
Nevins, A. (2004). Derivations without the Activity Condition. In M. McGinnis & N. Richards (eds.) *Proceedings of the EPP/Phase Workshop* (MITWPL).
Ott, D. (2006). Phases, Interfaces, and the Logic of Linguistic Minimalism: Some Conceptual Remarks. Ms., Harvard.
Poeppel, D. & K. Wexler (1993). The Full Competence Hypothesis of Clause Structure in Early German. *Language* 69: 1-33.
Postal, P. (1974). *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.
Richards, M. (2004). *Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax*. PhD dissertation. University of Cambridge.
Richards, M. (2007a). Interpretive Asymmetries at the Edge. Paper presented at GLOW XXX (Tromsø).
Richards, M. (2007b). On feature-inheritance: An argument from the Phase Impenetrability Condition. In *Linguistic Inquiry* 38(3): 563-572.
Richards, M. (2008a) Defective Agree and Case Alternations. Ms., University of Leipzig.
Richards, M. (2008b) Deriving The Edge: What's In A Phase? Ms., University of Leipzig.
Richards, M. (2008c) Quirky Expletives. In Roberta d'Alessandro, Gunnar H. Hrafnbjargarson, and Susann Fischer (eds) *Agreement Restrictions*. Berlin: Mouton de Gruyter, pp. 181-213.
Rizzi, L. (1978). Violations of the Wh-Island Constraint in Italian and the Subjacency Condition. *Montreal Working Papers in Linguistics* 11: 155-190.
Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
Roberts, I. (2006). Clitics, Head Movement, and Defective Goals. Ms., University of Cambridge.
Roberts, I. (2007). *Diachronic Syntax*. Oxford: OUP.
Roberts, I. & A. Roussou (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: CUP.

Saunders, P. (1994). Evolution without Natural Selection: Further implications of the Daisyworld Parable. *Journal of Theoretical Biology* 166: 365-73.
Thráinsson, H. (1996) On the (Non-)Universality of Functional Categories. In W. Abraham, S. D. Epstein, H. Thráinsson & J-W. Zwart (eds.) *Minimal Ideas: Syntactic Studies in the Minimalist Framework*. Amsterdam: Benjamins, pp. 253-81.
Thráinsson, H. (2001) Object Shift and Scrambling. In M. Baltin & C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 148-202.
Uriagereka, J. (1988). On Government. Ph.D. dissertation, University of Connecticut.
Uriagereka, J. (1995) An F Position in Western Romance. In K. Kiss (ed.), *Discourse Configurational languages*, 153-175. Oxford (New York): Oxford University Press.
Uribe-Etxebarria, Myriam (1992). On the Structural Positions of the Subject in Spanish, their Nature and their Consequences for Quantification. In J. Lakarra & J. Ortiz de Urbina (eds.) *Syntactic Theory and Basque Syntax*. San Sebastian: ASJU, 447-491.
Vikner, S. (1995). Verb Movement and Expletive Subjects in the Germanic Languages. New York: Oxford University Press.
Williams, E. (2003). *Representation Theory*. Cambridge, MA: MIT Press.
Yang, D-W. (2007). On Edge Feature Movement. Ms, MIT.