

Emergent features: a minimalist perspective¹
Theresa Biberauer
University of Cambridge & Stellenbosch University
(mtb23@cam.ac.uk)

1. Introduction

- The “traditional” generative model of language acquisition:
 - (1) Universal Grammar (UG) + Primary Linguistic Data (PLD) → I-language grammar
 - where UG is **richly specified**, i.a. also in featural terms:
 - (2) a. ‘FL specifies the features F that are available to fix each particular language L...’
b. ‘We adopt the conventional assumption that L makes a one-time selection [FL] from F. These are the features that enter into L; others can be disregarded in use of L” (both Chomsky 2001:10)

→ a **universally given, pre-specified inventory of formal features**
- Since Chomsky (2005): a “third factor” should be added to the “traditional” two, with **UG** being assumed to be a **maximally impoverished (minimal) entity**.
- (3) UG + PLD + **3rd factors** → adult grammar
where 3rd factors are characterized as:
 - (4) a. principles of **data analysis** that might be used in language acquisition and other domains;
b. principles of **structural architecture and developmental constraints** that enter into canalization, organic form, and action over a wide range, including principles of **efficient computation**, which would be expected to be of particular significance for computational systems such as language.
(Chomsky 2005:6)
- My specific proposal:
 - (5) UG (F₁) + PLD (F₂) + **Maximise Minimal Means (MMM)** (F₃) → I-language grammar
 - where UG is **poor**, and does not contain the universal feature inventory in (2)
 - Most (all?) formal features are **emergent**, arising as a result of the interaction of the 3 factors in (5).
- Constraining feature postulation?
 - 2 elements:
 - (i) UG (=F₁) provides a basic feature template, while

¹ The research reported here is funded by the European Research Council Advanced Grant No. 269752

- (ii) MMM (=F3) drives acquirers to limit the number of [F]s they postulate (Feature Economy in (6)) while generalizing these to as many environments as are compatible with the systematic regularities in the input (Input Generalization in (7)).

- (6) **Feature Economy (FE):** postulate as few features as possible to account for the input (=intake) [adapted from Roberts & Roussou 2003]
- (7) **Input Generalisation (IG):** maximise use of postulated features [adapted from Roberts 2007]

The structure of this paper:

Section 2: The 3 factors model in more detail: emergent features how do they arise?

Section 3: New Predictions I: General formal properties of natural-language systems

Section 4: New Predictions II: Emergent features and Going Beyond the Input

Section 5: Conclusions

2. The 3 Factors model assumed here in more detail

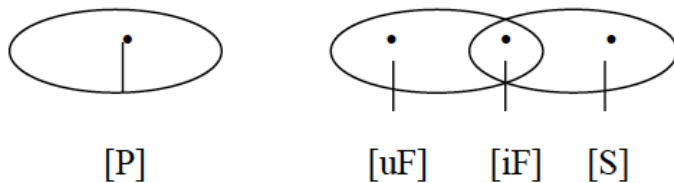
- In (5), the non-UG components take on much greater significance than in (1). So this section will consider UG only briefly (Section 2.1). The main focus will be on clarifying which aspects of the PLD serve as the basis for acquirers' postulation of the specific formal features ([F]s) that define their grammars (Section 2.2), and how MMM shapes feature- and hence parameter- and grammar-postulation (Section 2.3).

2.1. Minimal UG (the role of F₁)

- The point of departure: **Saussurean arbitrariness** → words as arbitrary sound-meaning mappings
- Chomsky (1995): the 3 feature-types defining human language = Phonological, Semantic and Formal (=grammatical) features (Chomsky 1995)
- Zeijlstra (2008): these 3 feature-types interact as in (8)

(8)

Phonological features Formal features Semantic features



(cf. Zeijlstra 2008)

- Saussurean arbitrariness viewed from this perspective: lexical items = lexically specified Phonological ([P]-) and Semantic ([S]) feature mappings

- BUT human language (uniquely?) goes beyond this: there is a **“higher” level of arbitrariness defined by Formal ([F]-) features**, which map onto the [P]- and [S]-feature pairs in systematic ways²
- (9) There are **degrees of arbitrariness** in human language:
- (i) (almost) entirely arbitrary **lexical** sound-meaning mappings: [P]- and [S]-features, and
 - (ii) less arbitrary **grammatically regulated** sound-meaning mappings: [P]-, [S]- and [F]-features
- Proposal:
- (10) a. [F]s piggy-back on [S]-features (cf. also Zeijlstra 2008)³
- b. A **UG-given [F]-template** – the precise nature of which is therefore a very important desideratum! – determines the format in which “higher-level” arbitrariness is encoded
- The more specific hypothesis I’m pursuing here: Many (perhaps all?) of the (substantive/non-diacritic⁴) [F]s defining I-languages are emergent (*pace* Chomsky 2001, Hegarty 2005, Gianollo et al 2008, Zeijlstra 2008, and pretty much everyone in generative syntax)

NB: this is a major respect in which the **current emergentist proposal** departs from **existing emergentist approaches** (cf. the work of Tomasello, Lieven, Ambridge and colleagues, Construction Grammar proposals more generally)

→ variation is **structured** on the basis of [F]s

So, **no UG-given inventory of [F]s**; but then where do they come from?

→ the interaction of the 3 factors given in (5) above.

2.2. The Primary Linguistic Data (PLD; the role of F₂)

- Even if one retains the notion that [F]s are UG-given, something along the lines of what is outlined below is necessary as part of a complete model of syntactic acquisition.
- Key assumption: **the PLD will “skew” acquirers’ postulation of [F]s**, depending on what is salient in a given language
 - We should not expect (many) universally formally identical categories (cf. also i.a. much work by Wiltschko and colleagues, Chung 2012, Ramchand & Svenonius 2014, Biberauer 2014, 2016a,b, Biberauer & Roberts 2015a,b, 2016)

² This proposal picks up on Tsimplici’s (2013:59ff) “instinct for syntactic categories”: acquirers are driven to systematize/categorise the form-meaning mappings they learn in such a way that syntactic categories arise. This renders much of the necessary detail – and thus, memory load – associated with individual lexical items superfluous when they are being manipulated by language users (Mobbs 2015).

³ It has recently been suggested that [F]s may piggyback on [P]-features (cf. Fujimori 2011 and work by Rose-Marie Déchaine, and also work on formalized “secondary meaning” by Norbert Corver). We leave this possibility – which is predicted in the context of an MMM system – aside here for expository reasons.

⁴ Though see Biberauer & Roberts (2015/16) for a proposal that does away with movement diacritics.

What the child is looking for: **systematic departures from Saussurean arbitrariness**, i.e. from one-to-one form:meaning mappings (cf. also Fasanella & Fortuny 2013 and Fasanella 2014 on the so-called Chunking Procedure).

•
Cues from the PLD that [F]s must be postulated

A. **Doubling/Agreement and expletives/dummy elements: 2/multiple forms and 1 meaning, or 1 form with no meaning**

(11) Hulle is **nie** laat **nie**. [Afrikaans]
they is NEG late NEG
'They are not late.'

(12) a. UJohn **u-** nik- a abantwana imali. [Zulu]
1a.John 1SM- give- FS 2.children 9.money
'John is giving the children money.'

b. UJohn **u-** ba- nik- a imali (abantwana).
1a.John 1SM -2OM -give-FS 9.money 2.children
'John is giving them money (the children).'

c. UJohn **u-** yi- nik- a abantwana (imali).
1a.John 1 SM -9OM - give- FS 2.children 9.money
'John is giving it to the children (the money).'

(Zeller 2011, via van der Wal 2015)

(13) "Dummy" cases: English *there, it* (subject expletives); Icelandic *það* (topic expletive); English *do* (polarity dummy); Basque *egin* and Korean *ha* (verb focalization, tense, etc. dummies, compensating for the nominal nature of focalized verbs), etc.

• Where doubling involves inflectional morphology, we know that children are very attuned to it.

B. **Systematic silence**, e.g. null exponence, null arguments, null complementisers, ellipsis, OCP effects (and thus also their failure), etc.: **no form with meaning**

(14) a. Chomsky **has** written a new paper (has/hasn't he?)
b. **Has** Chomsky written a new paper?
c. Chomsky **hasn't** written a new paper.

(15) a. Chomsky \emptyset wrote a new paper (did/didn't he?)
b. **Did** Chomsky write a new paper?
c. Chomsky **didn't** write a new paper.

C. **Multifunctionality: 1 form, multiple meanings** (cf. also Wiltschko 2014, Duffield 2013, 2014)

(16) a. Ông Quang **được** mua cái nhà. [Vietnamese]
PRN Quang CAN buy CL house
'Quang was allowed to buy a house.'
DEONTIC (PERMISSION)

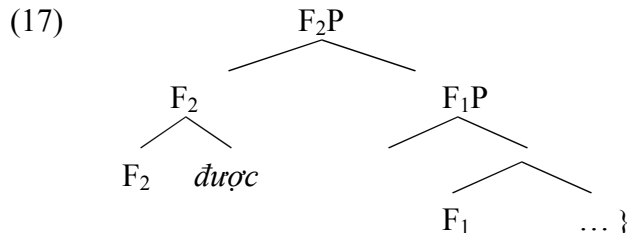
- b. Ông Quang mua **được** cái nhà.
 PRN Quang buy CAN CL house
 ‘Quang was able to buy a house.’ ASPECTUAL (ACCOMPLISHMENT)
- c. Ông Quang mua cái nhà **được**.
 PRN Quang buy CL house CAN
 ‘Quang may possibly buy a house/Quang is able to buy a house.’
 EPISTEMIC (ALETHIC)
 (Duffield 2007:101-102, 2013:2)

[As Duffield 2013, 2014 discusses in detail, the behaviour of *được* is not unique in Vietnamese:

- Two other modals, *phải* (cf. (ib) below) and also *nên*, exhibit the first 2 placement options; the final option is unique to *được*; see Biberauer in press
- many other lexical items (more precisely: units of language; Wiltschko 2014) are similarly multifunctional, creating the **impression of large-scale homophony.**]

In the context of the current emergentist proposal:

- acquirers encountering “large-scale homophony” will diagnose the presence of [F]s, associated with (null) functional heads.
- the units of language themselves will be analysed as **underspecified** elements, lacking [F]s and thus not being able to project, i.e. these elements are effectively like roots (on the view that these are [F]-less).
- their placement reflects the presence of different functional heads, to which these elements adjoin, as schematized in (17):



D. **Movement** – assuming Chomsky’s (2000) notion of **duality of semantics** (thematic + discourse/scopal meaning), movement results in “**extra**” meaning

- (18) a. [With no job] **would** she be happy. [Modern English]
 Basic order: She would be happy [with no job].
- b. [Never in my life] **did** I expect that to happen!
 Basic order: I [never in my life] expected that to happen!

- Also relevant here: the ‘higher-level’ **duality of patterning** deriving from “neutral” vs “marked” word orders.
 - “standard” **duality of patterning** (Hockett 1958): 2 levels of structuring
 - meaningless phonemes: /t/, /r/, /i:/, etc.
 - meaningful phoneme-combinations: *tree*, etc.
 - “higher-level” **duality of patterning** (cf. also Fortuny 2010)
 - meaningless “basic” word-order choice – e.g. OV vs VO (a Very Early parameter; cf. Tsimpli 2014)

- meaningless obligatory filling choices: e.g. V spellout position, Spec-TP, Spec-CP, etc.
- meaningful optional movements relative to the fixed higher-level conventions: e.g. T-to-C in English, the nature of the XP that raises to Spec-CP, etc.

i.e. “basic” ordering is simply a convention requiring fixing, just as phonotactic constraints require fixing, whereafter it can serve as the basis for further, potentially meaningful ordering patterns, which contrast with the “basic” one (e.g. V2, subject-auxiliary inversion, topicalization/focalization fronting, etc.)

- Having both levels of duality of patterning = **maximising the contribution of both the Lexical Items and (External and Internal) Merge (= minimal means)** – see section 3.3 below.
- Strikingly, there is now a significant body of evidence suggesting that basic word order (OV vs VO, etc.) is fixed very early indeed, apparently on the basis of prosodic cues, i.e. [P]-side features alone (cf. much work by Jacques Mehler, Judit Gervain and numerous colleagues).

E. **Recursion** (cf. much work by Tom Roeper and William Snyder, i.a. Roeper 2011, Roeper & Snyder 2011)

- Recursion differentiates “**exceptional**” domains from **truly productive grammar** (Roeper & Snyder 2011:158; cf. also Yang 2016)

(19) a. [frog man] [English]
b. [[frog man] team]

(20) a. [homme grenouille] [French]
b. *[équipe [homme grenouille]]

- ➔ noun-within-noun compounding is productive in English, not French (and we know that three-term compounds (e.g. [[Christmas tree] cookie] are robustly present in the input to English-acquiring children – Roeper, Snyder & Hiramatsu 2002)
- By contrast, V-serialising isn’t productive in English, though it is in other languages; it is simply an exceptional pattern in English (*Sue can come eat lunch now*, etc.)
 - Without some representation of NOUN vs VERB (i.e. [F]-defined categories), children wouldn’t be able to capture the systematic regularity difference in play here.
 - We also see that children are clearly sensitive to the extent to which a pattern is attested in their language – cf. Yang’s (2016) Tolerance Principle, in terms of which the number of elements conforming to a rule must be appropriately greater than the number of exceptions for it to be postulated by acquirers (change happens when a so-called Tipping Point is reached).

- Recursion also plays a second function, namely in driving acquirers to postulate **distinguishing [F]s** in cases where elements that appear to belong to “the same” category can be systematically stacked in a particular sequence.⁵

- (21) a. ... dass er singen₃ **können**₂ **muss**₁. [German]
 that he sing.INF can.INF must
 ‘... that he must be able to sing.’
- b. ... dat sy die boek **sal**₁ **moet**₂ koop₃. [Afrikaans]
 that she the book shall must buy
 ‘... that she will have to buy the book.’

- While the inflectional distinction in (21a) signals that the modals are not formally identical, this is not clear in (21b). Nevertheless, acquirers of Afrikaans still have a cue to postulate a distinguishing formal feature between *sal* and *moet* as there is a systematic regularity (the fixed ordering sequence) involved here, and we know that children are very sensitive to word-order phenomena, particularly the “meaningless” type highlighted below (18) (it’s a Very Early parameter and would therefore, in the context of the current emergentist approach, be expected to constitute a **category-defining feature** in the terms of Biberauer (2014, 2016, in press)⁶).

The overall picture:

- The **driving intuition behind A-E**: [F]s are postulated if they can be seen to regulate some form of **systematic contrast**, which cannot be explained by appealing only to semantic or phonological considerations (a higher level of Saussurean arbitrariness)
- The **morphosyntactic and morphosemantic contrasts** vary by language; hence the language-specific “content” of what it means to “be” categories of different types, and also what features are grammaticalised (i.e. [F]s) is expected to vary
- ➔ categories and features are **emergent**
- **BUT**: doesn’t this just predict rampant and unconstrained variation?

2.3. A closer look Maximise Minimal Means (MMM; F₃)

2.3.1. The basic idea

- The key idea here: a general (non-language-specific) learning bias:

(22) **Maximise minimal means**

which has 2 linguistic manifestations that I’m particularly interested in here:

⁵ **Distinctness** (in the sense of Richards 2010) is expected to be required in a system where distribution is governed by categorial “sameness”, i.e. formally identical elements compete for the same positions, meaning that co-occurring elements must be formally distinct in some way. Cf. also the OCP/haplology effects observed in various structural domains that arguably kick in at PF, and de Clercq & vanden Wyngaerd (2016) on a parallel effect at the LF interface relating to Double Negation. This property, which clearly surfaces across linguistic domains, may then also be understood as a consequence of our “language instinct” (see note 2).

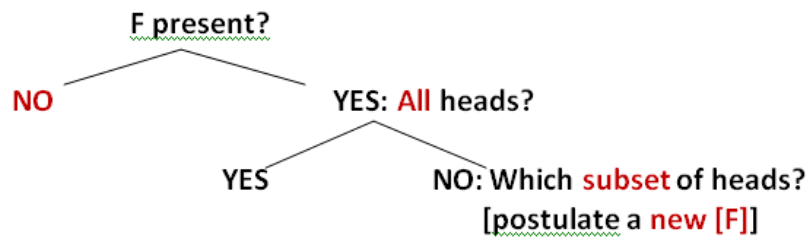
⁶ Biberauer (*op. cit.*), for example, proposes that head-finality is part of “what it means to be verbal” in languages with a head-final clausal domain.

(23) **Feature Economy (FE):** postulate as few features as possible to account for the input (=intake) [adapted from Roberts & Roussou 2003; = (6) above]

(24) **Input Generalisation (IG):** maximise use of postulated features [adapted from Roberts 2007; = (7) above]

- And recall:
 - [F]s are only postulated where the child detects departures from Saussurean arbitrariness
 - [F]s “piggy-back” on semantic features [S]
 - together, FE and IG form a **minimax search/optimisation algorithm** (FE: minimise features; IG: maximise detected features).
- FE and IG also naturally result in a **learning pattern/path** (hierarchy) with the following general “shape” (cf. Biberauer & Roberts in press)

(25)



- **NO > ALL > SOME** acquisition sequence
 - **NO = default** as the acquirer doesn’t receive (systematic) input pointing to need for [F] → **FE & IG respected**
 - **ALL:** plausibly follows from the acquirer’s initial “ignorance” (Biberauer 2011, Branigan 2012) → **IG respected & FE minimally violated (“Make maximal use of minimal means”)**
 - **SOME:** both IG & FE violated, but FE will dictate that the violation should be minimal, picking up on the nature of the input cues
- Importantly, the **NO** option does not involve an actual question triggered by the input data; this option emerges as a natural **default**, a setting that arises where the input an acquirer experiences does not motivate the postulation of a feature to account for some regularity in the system. There is thus no need to stipulate defaults on the system proposed here.
- “Recovery” from **superset traps** (Berwick 1985, Manzini & Wexler 1987, Hale & Reiss 2008) possible as the acquirer must postulate appropriately specified heads to capture the regularities in the data (s)he is exposed to, with the [uF]/[iF] template in combination with IG and FE gradually reducing the acquirer’s “ignorance”.
- And we do seem to see these “recovery” phenomena in language acquisition (see following discussion, and also work in Developmental Psychology, such as that by Bornstein & Arterberry (2010), which shows that object classification also seems to develop on the basis of ‘hierarchical inclusiveness’, with superordinate/more

inclusive/less specified categories being acquired before subordinate/less inclusive/more specified categories.)

- How do acquirers know where to **start** in constructing [F]-systems, though? Which [F]s are postulated first and then either generalized or further articulated?
- Proposal:
 - (i) the earliest properties (e.g. headedness) are **prosodically mediated** (cf. much work by Mehler, Nespors, Gervain, Guasti and colleagues; cf. also Tsimpli 2014)
 - (ii) subsequent properties are signalled by A-E phenomena, particularly in certain high-frequency, relatively simple, but strikingly syntax-rich structures, notably **questions** and **imperatives** (Biberauer 2015a, 2016, and see also below)

i.e. to get the categorial system off the ground, the child need only be sensitive to prosodic properties, which we know they are. Once an appropriate initial division has been made, sound:meaning-based regularities can serve as the basis for the postulation of substantive [F]s.

- Importantly, the system being outlined here, in terms of which F_3 interacts with F_2 and F_1 as outlined above in order to “create” syntactic categories and features foregrounds – in what I hope will ultimately become a suitably explicit way – what Tsimpli (2013:67) calls ‘the critical role of **structure** for shaping language units in the lexicon (my emphasis – TB)’.
- And this, in turn, allows us to understand the “language instinct” as “a cognitive ability which exclusively combines hierarchical representations with units whose substance can only be fully conceived within the system that creates them” (Tsimpli 2013:67).

2.3.2. MMM in different domains

- If MMM is genuinely a 3rd factor, we would expect to see its operation outside syntax, and also outside the formal language system.
- The following sections consider the operation of MMM in Phonology and the domain of Concept Formation.
- Probably most stunningly, Jaspers (2012) describes the same pattern in the development of colour perception.

2.3.2.1. *NO>ALL>SOME in Phonology*

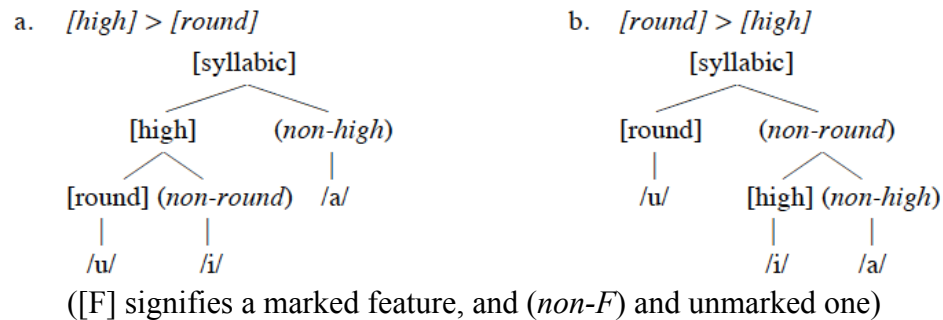
- the **Successive Division Algorithm** (Dresher 2009:16):
 - (26) a. Begin with *no* feature specifications: assume all sounds are allophones of a single undifferentiated phoneme.
 - b. If the set is found to consist of more than one contrasting member, select a feature and divide the set into as many subsets as the feature allows for.
 - c. Repeat step (b) in each subset: keep dividing up the inventory into sets, applying successive features in turn, until every set has only one member.
- Only features that **contrast** play a role in structuring the phoneme inventory (see also Avery & Rice 1989):

(27) The **Contrastivist Hypothesis** (Hall 2007)

The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another.

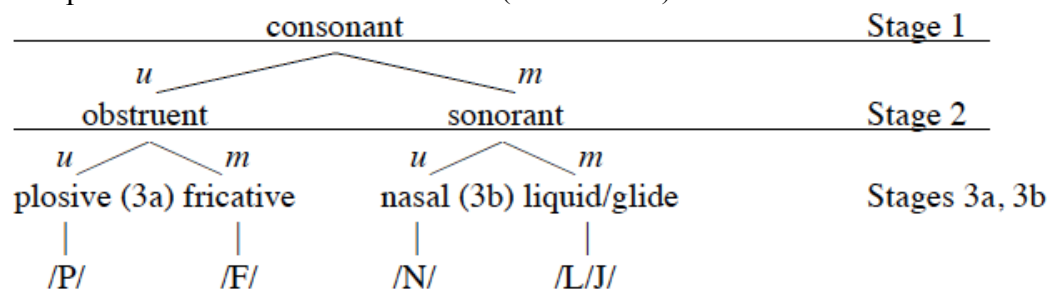
- Dresher (2014:177) describes (27) as the “**formal organizing principle of the phonology**”, contrasting it with the “functional phonetic explanations for substantive phonological patterns” appealed to by both “substance-free” and “substance-based” theorists (cf. also Hall 2012, 2014 for discussion)
- The syntactic model presented here similarly assumes that “higher-level” Saussurean contrasts are the organizing principle in syntax (Clark 1987; Cowper & Hall 2014), i.e. syntax and phonology are **not** that different (*pace* Bromberger & Halle 1987)
- As phonological features are not pre-given, their hierarchical organization may vary across languages:

(28)



- There appears to be **acquisition evidence** for this approach (cf. i.a. Fikkert 1994, and also Dresher 2014, Mobbs 2015 and Bazalgette 2015)

(29) Development of Dutch onset consonants (Fikkert 1994)



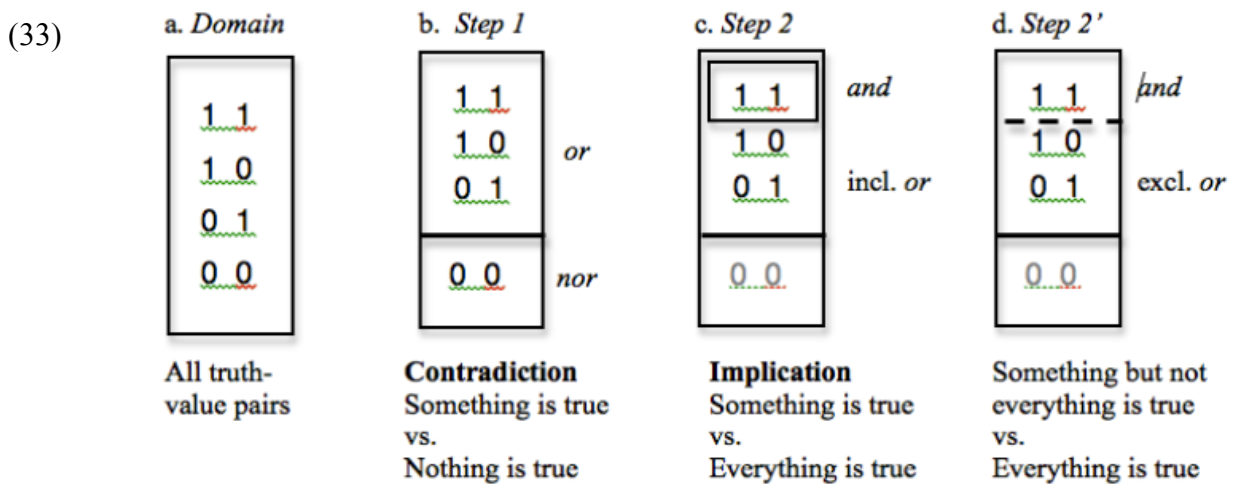
- Stage 1: no contrasts; the value of the consonant defaults to the **least marked onset**, i.e. an **obstruent plosive**.
- Stage 2: **first contrast** made between **obstruents** and **sonorants**, with the former remaining unmarked (*u*) the unmarked option. The **marked option** (*m*), **sonorant**, defaults to **nasal**. Then, 2 possibilities:
- Stage 3a: expansion of the obstruent branch, introducing **marked fricatives** to contrast with plosives
- Stage 3b: expansion of the sonorant branch, introducing **marked sonorants**, either **liquids** or **glides**

2.3.2.2. *NO>ALL>SOME in the domain of Concept Formation*

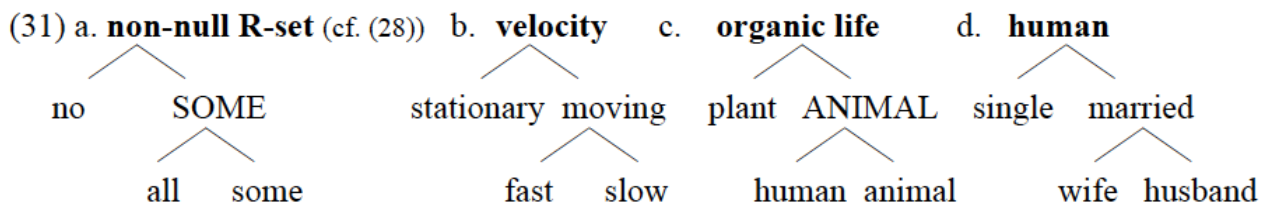
(30) Jaspers’ (2013) **Concept Formation Constraint** (cf. also Seuren & Jaspers 2014)
 ‘a set of four natural operators is generated by making subtractions from a fixed domain space of values via a series of **two successive binary divisions**. There is an initial exhaustive division between the contradictories NOR and OR ... within the remaining non-NOR space of values, we can either carve out the subset AND, leaving inclusive OR as superset space ... or we can divide the inclusive OR space exclusively into AND and exclusive OR’.

(31) Predicate calculus oppositions: $[[\text{Some}_2 \text{ All}; \text{Some}_1] \text{ None}]$

(32) Propositional calculus oppositions: $[[\text{Or}_2 \text{ And}; \text{Or}_1] \text{ Nor}]$, as illustrated in (33):



- And consider the “logico-cognitive” structure of the Lexicon more generally (Seuren & Jaspers 2014; cf. also Bazalgette 2015 for discussion)



THUS: There appears to be good evidence for **NONE>ALL>SOME** as an organizational pattern found both in language and in other cognitive domains, just as we might expect if it’s something that arises under the (partial) influence of a general third-factor force.

3. Predictions I: General formal properties of natural-language systems

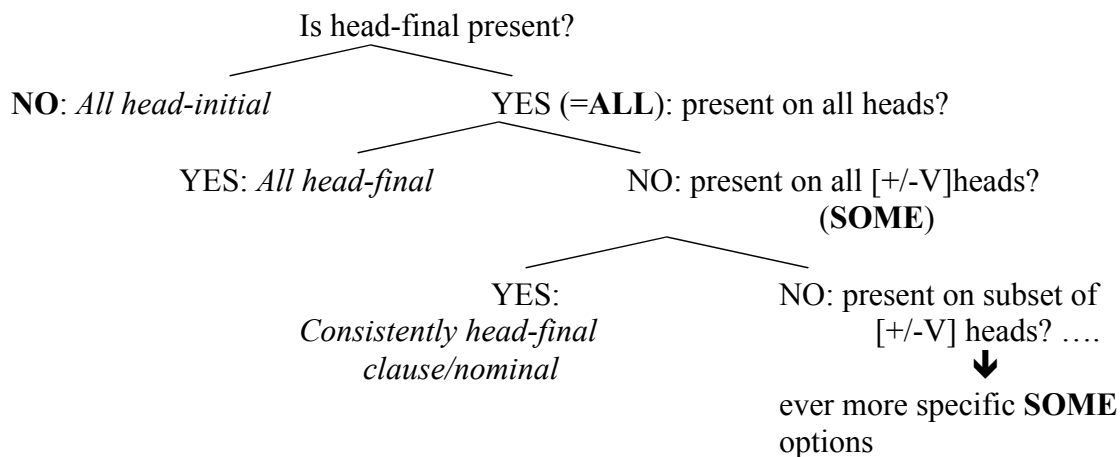
3.1. The shape of syntactic (=parametric) variation

- On the approach outlined above, we expect “the same” phenomenon to surface in languages in **different sized versions**

Consider head-initiality/finality

- Assume head-finality to be formally marked by means of the diacritic \wedge and thus to be the property a child will need to postulate a formal marking for, head-initiality not requiring any specific linearization diacritic (cf. Biberauer, Holmberg & Roberts 2014, and Biberauer, Roberts & Sheehan 2014 for discussion)

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



- Types of head-finality (cf. also Biberauer & Sheehan 2013):

- (36) a. “rigid” head-finality: Japanese, Malayalam, etc.
 b. clausal head-finality, nominal head-initiality, and vice versa
 c. “leaking” OV of different kinds, e.g. West Germanic
 d. OVX, where O is the direct object (Baker 2005, Hawkins 2009)
 e. O_[F]VX, where O_[F] is a restricted object-type (e.g. Neg, Focused, Specific, etc.)
 f. more lexically restricted OV (e.g. only with certain predicates) ...

Thus: macro, meso, micro, nano “sizes” (Biberauer & Roberts 2012, 2014, 2015a,b, 2016a,b), where these terms must be understood in **relative** not **absolute terms**

- The SOME options are of particular interest as the approach outlined here makes it possible to think of **distinct formal features** [F]s fulfilling **parallel roles** in **structuring different grammars** (cf. what was illustrated in (28) for phoneme inventories).
 - In the case of word order, for example, the clausal vs nominal (36b) options (i.e. completely head-final clause vs completely head-final nominal) are “the same” in the sense that both involve a single [F] – [+V] and [-V] respectively – which determines the distribution of head-finality. [+V] and [-V] are here fulfilling the same function in determining the distribution of head-finality, producing systems that are, at the level of “grain” being considered here – clausal vs nominal headedness, i.e. quite a gross-grained level – **typologically equivalent**.
 - Cf. also, following Wiltschko (2014), the choice to formalize clausal anchoring (INFL) could be [tense], [person] or [location], making these features formally parallel in this domain, but leaving open properties such as (i) the extent to which they play a role in other domains (see

Section 4.2 below), (ii) the ways in which the INFL domain is further articulated (e.g. to distinguish between finite and non-finite forms, different moods, etc.), etc.

3.2. Recycling

- FE would lead us to expect postulated [F]s to be maximally exploited, and there seems to be copious evidence of this.
 - Multi-domain use of **case**:
 - It's typically (in the Western European case) connected with grammatical functions (subject, object, etc.)
 - But, it can also be harnessed for discourse-related reasons (“prominence marking” in the sense of Biberauer 2011 et. seq.; cf. also Naess 2011 on the difference between so-called *speaker-* vs *referent-determined salience*; and i.a. Aikhenvald 2008, Richards 2013, Levin 2013, Pesetsky 2013 on so-called *case stacking*):
- (37) a. John-hant^{hey} Mary-ka mwusewe. [Korean]
 John-DAT Mary-NOM be.afraid
- b. John-hant^{hey-ka} Mary-ka mwusewe.
 John- DAT- **NOM** Mary-NOM be.afraid
 ‘JOHN is afraid of Mary.’ (via Pesetsky 2014)
- And in the thematic domain (i.a. also to contribute to the realization of Aktionsart, etc.)
- And similarly, the multi-domain use of **tense**, not only in clauses (cf. Ritter & Wiltschko 2009, 2014, Wiltschko 2014), but also in nominals (cf. Sadler Nordlinger 2001, Nordlinger & Sadler 2004) and on adverbials (as in Malagasy “oblique marking”, discussed by Matt Pearson in various places)
 - And then a short (and very partial) list of further examples:

Agreement

 - multi-domain/“extreme” use of agreement (e.g. Archi – Bond, Corbett, Chumakina & Brown 2016, Chamorro, etc.)
 - “monstrous” agreement in Tamil and other languages (Sundaresan 2012)
 - allocutive agreement as in Basque, and confirmational particle agreement as in Austrian German (Wiltschko 2016)

Lexically based recycling

 - “doubling” pronouns (as in clitic-doubling, resumption, and also in focus contexts, as i.a. discussed in Leffel, Simik & Wierzba 2013)
 - “double life” co-ordination markers in Japanese, Malayalam, etc. (Jayaseelan 2014, Mitrović 2015)
 - “multi-life” focus/question/polarity, etc. particles crosslinguistically (Bailey 2012, Biberauer 2016)
 - **grammaticalization** phenomena generally (including also expletives, “fake forms” (Wiltschko 2014), “multi-grammaticalisations”, e.g. *go*, which functions in all clausal domains in modern English: lexical verb, linking verb [*I want you to go (and) write your essay*],

future marker, speaker-oriented [*He told everyone he was so worried about the exams, but then he (only) went and got a First*]

Phonological recycling

- “specialised” use of C(onsonant) and V(owel), stress, and basic linearization in acquiring lexicon and morphosyntactic regularities (cf. i.a. Mehler and colleagues; Fujimori 2011; Déchaine 2015; Bazalgette 2015)
- “specialized” use of schwa, C lengthening, duplication, repetition and other phonological mechanisms, and also “exaptation” of “fossil” morphosyntax to convey affective meaning (Corver 2015, Biberauer 2016)

“Base category” recycling

- **verbalisation** and **nominalisation** effects
- the existence of **extended projections** (Grimshaw 1991 *et seq.*): typically thought to be defined by lexical categorial features (e.g. V, N, P, etc.)
- **Relativised Minimality** (Rizzi 2013, Starke 2001), which defines **local domains**, constraining syntactic relations (e.g. Agree and movement operations). This wouldn’t be formulable if it weren’t the case that grammars employ **small**, and thus **formally comparable feature sets** which **facilitate the computation of RM-defined intervention**.

3.3. Contiguity effects

- If grammars make maximal use of minimal means, we expect natural classes to be constructed on the basis of “**nested**” **featural specifications** (cf. (25)), creating **featurally defined natural classes** similar to those that have long been assumed in phonology.
- More specifically, we expect the acquirer’s keenness to generalize over as large a domain as possible to create **formally defined domains** sharing a particular property, with **no “on/off” patterns > contiguity**.
- And, again, there seem to be many natural-language phenomena that exhibit this character.
- The Final-over-Final Constraint (FOFC; Biberauer, Holmberg & Roberts 2014, Sheehan 2013, Biberauer in press), for one.

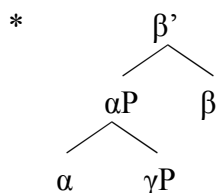
(38) **The Final-over-Final Constraint (FOFC)**

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

(cf. Biberauer, Holmberg & Roberts/BHR 2008 *et seq.*, notably BHR 2014)

- (38) rules out structures like (39):

(39) * $[\beta_P [\alpha_P \alpha \gamma_P] \beta]$



where α P is the complement of β and γ P is the complement of α , and α and β are part of the same projection line (cf. Grimshaw 1991 *et seq.*)

- What FOFC requires is that **head-finality start at the bottom of an Extended Projection** (see Grimshaw 1991 *et seq.*); and once a head-final sequence has “stopped”, it cannot restart within the same EP.

(40) A very basic FOFC-respecting pattern:

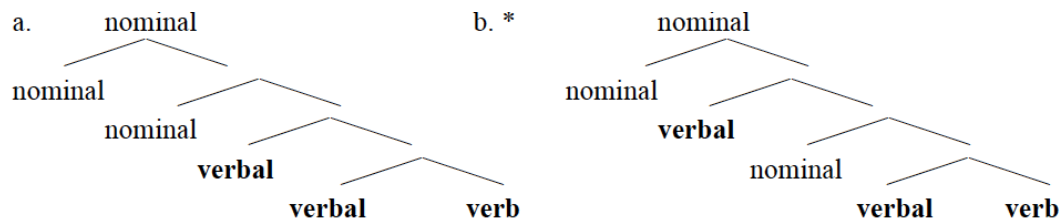
[_{CP} C[^] [_{TP} T[^] [_{VP} V[^]]]]
 > **contiguity**: structurally adjacent heads consistently bear [^]

(41) Three basic FOFC-violating patterns:

- *[_{CP} C[^] [_{TP} T [_{VP} V[^]]]]
 - *[_{CP} C[^] [_{TP} T [_{VP} V]]]
 - *[_{CP} C[^] [_{TP} T[^] [_{VP} V]]]
- > non-contiguity: structurally adjacent heads vary in their [^]-specification; an “on-off” pattern

- Phrasal Coherence is another case in point (cf. Panagiotidis 2014)

(42)



And similarly:

- (Non)agreement “cut-off” effects (see Pesetsky’s 2013 LSA slides for the details)

(43) Russian

- An optional feminizing null morpheme \mathcal{K} may be merged at any point above a **certain structural threshold** within the noun phrase. Low adjectives fall below this threshold.
- Once \mathcal{K} merges, **the nominal counts as feminine for agreement purposes from then on** (from that point up the tree), i.e. both higher nominal categories and verbal agreement will be feminine.

(42) Lebanese Arabic:

- An optional null morpheme # that creates semantic pluralities (Borer 2005) may be merged at any point above a certain structural threshold within the noun phrase. Low adjectives fall below this threshold.
- Once # merges, the nominal counts as plural for agreement purposes from then on (from that point up the tree).

[Puškar (2016) shows that the same kind of consideration holds in the determination of grammatical vs semantic agreement in SerBoCroatian]

- Animacy and Case Hierarchy (Caha 2009) effects
- Extended Auxiliary Hierarchy Effects (Sorace 2004)
- Syncretism patterns: *ABA (Nanosyntax; Caha 2009), and so on.

In all cases, there is generalization over as large a domain as possible to create **formally defined** domains sharing a particular property; there are no “on/off” patterns, but variation emerges on the basis of the interaction between IG and FE (which, in some cases, means IG could go either way).

4. New Predictions II: Emergent features and Going Beyond the Input

- There are circumstances where it must be uncontroversial that children “go beyond the input” > **extreme poverty of the stimulus** scenarios.
- On the 2-factors model, the best generativists could do was point to UG as a source for observed innovations.
- On the 3-factors model outlined here, we have a further possibility: MMM.

The question: Do we see **featureally defined NONE>ALL>SOME** in acquisition/learning contexts where there is variable input, and where acquirers appear to be building on early/robustly attested [F]s? (see Biberauer 2016 for more detailed discussion and further case studies)

4.1. Case studies showing how the model deals with variable input

4.1.1. Artificial Language Learning Experiments

- Hudson Kam & Newport (2005) and subsequent work: experimental work has shown **regularization of variable input** in action.
- Hudson Kam & Newport (2005):
 - under experimental conditions, children learning SillySpeak regularize variable input in various ways, while adults do not
 - “children learn unpredictable variation differently than adults. They have a stronger tendency to impose **systematicity** on inconsistent input ...” (Hudson Kam & Newport 2005:184; cf. also Hudson Kam & Newport 2009, Hudson Kam 2015 Reali & Griffiths 2009, Smith & Wonnacott 2010, and, outside language, the non-veridical learning work starting with Gardner 1957; see Mobbs 2015 for overview discussion)

- (45) The types of regularization that children impose on the input:
- minimization**: use the variable form none of the time (NONE)
 - maximization**: use the variable form all the time⁷ (ALL)
 - linguistically governed selection**: use the variable form in a grammatically defined subset of contexts (e.g. only with transitive Vs) (SOME)

⁷ Cf. Bitterman (1965) on maximization as strategy that is also found elsewhere in nature, and also the speculation that different species may have different learning preferences.

4.1.2. *Real-life situations: illustrations of children “going beyond the input”*

4.1.2.1. *Inflection-related innovations*

- Inflection is something children are known to be very sensitive to.
- A “traditional” generalization: rich agreement > movement (e.g. V-to-T)
 - But this is too simplistic (cf. Schifano 2015).
 - And it doesn’t follow at all on the present model:
 - Agreement is only **one** potential source of evidence for the presence of [F]s, which are, in turn, a necessary, but not sufficient condition on movement (by hypothesis, movement triggers must associate with [F]s; an [F]-bearing element can, however, remain Agree in situ)
 - Needing to account for “basic” placement (the “meaningless” type of movement in D in Section 2.3 above) is sufficient motivation for the postulation of an [F].
 - There is therefore no necessary correlation between movement and [F]-signalling inflection (cf. also Guasti 2013, who, however, suggests on the basis of both acquisition and typological evidence that [F]-specifications requiring both Agree and Move operations are in some sense more robust and thus acquisitionally preferred than [F]-specifications requiring only an Agree operation. This preference is expected on the MMM perspective: it makes maximal use of the relevant [F], allowing it to serve as the Narrow Syntax-internal “reference point” for both of the operation types given by UG (Agree and Merge).
- Even though children are good at acquiring inflection, inflectional patterns can nevertheless become obscure/optional for independent reasons ... and then we sometimes see new patterns emerging:

Example 1: Various varieties of non-standard English

- Willis (2015) shows that acquirers deal with the steady loss of agreement (person and number)-marking in the English verbal system over time in different ways:⁸

(46) Present-tense number marking in modern English

- Standard English: present tense number marking on 1st and 3rd person for *be* (*am/are, is/are*), and consistently for 3rd person singular on lexical verbs and other auxiliaries
- Vernacular varieties: different patterns, including:
 - (i) generalization throughout the paradigm, either to *s*-forms throughout (*she sings, they sings*) – **ALL** – or to *s*-less forms (*she sing, they sing*) throughout – **NONE**.
 - (ii) Use with specific sub-types of subjects, as in the Northern Subject Rule, which takes a number of different forms – **SOME**.
(cf. Isaac 2003, Klemola 2000, McCafferty 2003, Pietsch 2005, Childs 2013)

⁸ As noted in Miller (2007), Miller & Schmitt (2009, 2010, 2012a,b), verbal number marking is acquired later than nominal number marking, with English children only mastering the comprehension of this marking by age 5, with the comprehension of nominal plural marking being successfully in place by age 3. In our terms – and as one might expect in the context of the type of emergentist generative model being outlined here – [Number] thus has to be generalised from the nominal domain, where it is initially acquired, leaving scope for reanalysis if the cues become sufficiently restricted.

(47) Past-tense number marking in modern English

- Standard English: past tense marking again only on *be* (*was/were*).
- Vernacular varieties: again different patterns, including:
 - (i) generalization throughout the paradigm, either to all *was* or all *were* (**ALL/NONE**)
 - (ii) specialization relative to **polarity**: *were* (as *weren't*) = negative clauses, regardless of person and number, while *was* – affirmative clauses, regardless of number. This is i.a. seen in Reading and York (Tagliamonte 1998), Fenland varieties (Britain 2002), and outer East London (Cheshire & Fox 2009). **SOME**

- (48) a. They **was** writing a lot of tests that time.
 b. They **weren't** doing much else.

- The grammatically defined **SOME** choices that emerge in the past tense centre on [polarity]
 - Why?
 - Proposal: children get very clear evidence from interrogative structures that auxiliaries are fundamentally concerned with polarity:

- (49) a. They **were** all picnicking in the sunshine.
 b. **Were** they all picnicking in the sunshine?
 c. They **ate** a lot of cake.
 d. **Did** they all eat a lot of cake?

➔ This very fundamental **contrast** between declarative and interrogative auxiliary positioning and realization shows that auxiliaries in English have a [polarity] feature.

Thus: what the British English speakers who analyse *was* as the basic affirmative auxiliary and *were* as the basic negative auxiliary are doing is (i) discarding the agreement (ϕ) features, as it is not clear to them that [person] and [number] are relevant, and (ii) giving these elements **fixed and distinct polarity specifications**. Taking Zwicky & Pullum's (1983) arguments in favour of English's *n't* auxiliaries as distinct lexical items from their positive counterparts seriously (see also Biberauer & Roberts 2010), the change can be represented as follows:

(50) Standard English:

- a. *was* [Tense: Past; Person: 3; Number: Sg; Pol: Aff]
- b. *were* [Tense: Past; Person: 3; Number: Pl; Pol: Aff]
- c. *wasn't* [Tense: Past; Person: 3; Number: Sg; Pol: Neg]
- d. *weren't* [Tense: Past; Person: 3; Number: Pl; Pol: Neg]

(51) Dialectal British English:

- a. *was* [Tense: Past; Pol: Aff]
- b. *weren't* [Tense: Past; Pol: Neg]

- **Basic questions and negative structures** – initially lacking any form of auxiliary – are produced early, i.e. [**pol**], **which unites these two** (cf. Holmberg 2016 and references cited therein) is **an early-acquired [F]** (see also Biberauer & Roberts 2015a,b on C being an early-acquired category, one

which precedes T and other lower clausal functional heads). In the context of the present system, this means that [pol] is an [F] which is **in place early** and thus **available as a target for IG**.

- For acquirers of the relevant varieties of Dialectal British English, then, the conclusion that *was* and *were* are specified as in (51) is an IG-based one centring on an already available [F], i.e. these acquirers are **building on what is already in the system** to make sense of input that isn't clear to them – **MMM** (see also the following example, which again builds on the presence of [polarity] in a system)
- Even for acquirers of English who do ultimately acquire a system featuring the auxiliaries in (50), we know that [number] is fully acquired much later than [pol]; cf. i.a. Johnson, de Villiers & Seymour (2005), and de Villiers & Johnson (2007), who show that initial production of –s does not entail comprehension of this inflection.
- The (re)analysis in (51) evidently involves simplification – fewer [F]s are required to specify the relevant lexical items⁹ – and there are also fewer lexical items to be stored.

4.1.2.2. Lexically based innovations

Example 2: Afrikaans negation (or: Why Standard Afrikaans negation does not constitute a naturally acquirable grammatical system! See Biberauer 2012, 2015b, Biberauer & Zeijlstra 2012a,b)

- (52) a. Hulle is **nie**₁ laat **nie**₂. > **DOUBLING** (cf. (11))
 they is not late POL
 'They are not late.'
- b. **Niemand** is laat **nie**₂.
 no-one is late POL
 'No-one is late.'
- c. Hulle sien **niks** **nie**₂.
 they see nothing POL
 'They don't see anything.' > all **Negative Concord**

BUT:

- (53) a. **Niemand** sien **niks** **nie**₂.
 no-one see nothing POL
 'No-one sees nothing.', i.e. everyone sees something

⁹ Importantly, what is being said here is not that [person] and [number] are absent from the grammars of the relevant speakers; just that these features do not play a role in the characterization of the relevant lexical items, whose specifications are fixed early, before [person] and [number] are stably available as candidate [F]s for the specification of auxiliary elements. [person] and [number] **are** available in the nominal domain, for which they are acquired earlier than in the clausal domain (cf. much work by i.a. Jill de Villiers and colleagues, Cristina Schmitt, Karen Miller, and others). In systems where agreement marking is more transparent than in English, these [F]s are will ultimately be harnessed in the verbal domain too, but there is evidence even from richly inflecting Romance languages that this cross-domain phi-harnessing is not as simple as one might imagine – a question that clearly deserves closer attention.

- b. Ek gee **niemand niks nie**₂.
 I give no-one nothing POL
 ‘I give nothing to no-one.’, i.e. I give something to everyone
- c. Ek gee **nooit niks nie**₂.
 I give never nothing POL
 ‘I never give nothing.’, i.e. I always give something > **Double Negation**

And another BUT: in spoken Afrikaans, the most natural interpretations are Negative Concord ones

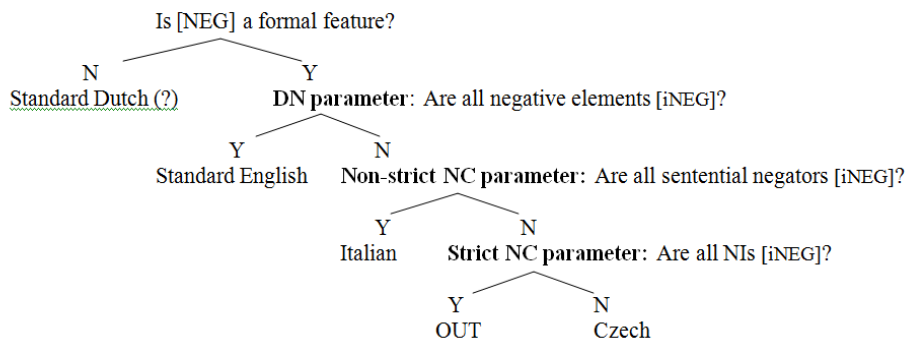
- (53’) a. **Niemand sien niks nie**₂.
 no-one see nothing POL
 ‘No-one sees anything.’
- b. Ek gee **niemand niks nie**₂.
 I give no-one nothing POL
 ‘I give nothing to anyone.’
- c. Ek gee **nooit niks nie**₂.
 I give never nothing POL
 ‘I never give anything.’ > **Negative Concord**
- (54) a. Ons is nog **nooit nêrens betrap nie**₂.
 us is still never nowhere trapped POL
 ‘We have never been trapped anywhere’ (*Volksblad* 2002-01-19)
- b. En bowe-al futiel omdat dit **niemand nêrens bring nie**₂.
 and above-all futile because it nobody nowhere bring POL
 ‘And above-all futile, because it gets nobody anywhere.’
 (*Beeld* 1997-02-07)
 [Examples from Huddleston 2010]
- (55) a. Hy wil (**g’n niks** luister **nie**₂.
 he want no nothing listen POL
 ‘He won’t listen at all.’
- b. Jy gaan met hierdie ding (**g’n nêrens** kom **nie**₂.
 you go with this thing no nowhere come POL
 ‘You won’t get anywhere at all with this thing.’ (Biberauer 2009)

i.e. an innovated emphatic structure that is also Negative Concord

Thus: the overall sense is that Afrikaans **actually** wants to be fully Negative Concord, an ALL rather than a SOME system (compare many colloquial Englishes too)

➔ WHY?

- (56) A Negation hierarchy (Biberauer 2011) constructed on the basis of the NONE>ALL>SOME schema



- Once children have determined on the basis of the doubling input that Afrikaans definitely grammaticalises negation ([negation]) and, moreover, that it is Negative Concord (NC) – i.e. that every negative-marked element doesn't actually **mean** “negative” – they are biased by IG to analyse **all** negative elements as [uNEG].
- The Strict NC parameter, which is set once the child has decided that the sentential negation marker is [uNEG] = a **no-choice** parameter (Biberauer, Roberts & Sheehan 2014): there is no PLD that could lead the acquirer to retreat from the assumption that, if the sentential negator is [uNEG], the negative quantifiers could be [iNEG]
- Input that would show this would involve Double Negation *Nobody found nothing*-type structures, but these are (i) not that common in the input, and, more problematically, (ii) they are also available in Strict NC languages. In other words, there is **no disambiguating data** that could show the child who has decided that sentential-negation elements are [uNEG] that this feature should not also be extended to negative indefinites
- The Standard Afrikaans negation system, which combines [uNEG] sentential negation markers (the medial and final *nies*) with [iNEG] negative quantifiers (the OUT system in (56)) exists because it is taught in schools, etc.
- Colloquial Afrikaans thus features real “two grammars”-type optionality (Kroch 1989 *et seq.*): both the “natural” Strict NC one also instantiated by Czech-type systems and the OUT system are represented.¹⁰
- Strikingly, though, **children naturally opt for the “natural” Strict NC system: they get mixed input, but go beyond it to postulate a system that fits into the most naturally NONE>ALL>SOME-defined acquisition path.**

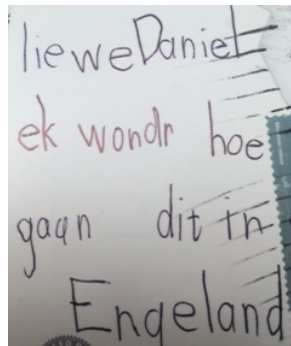
¹⁰ **Types of optionality:** The type of optionality evident in the Afrikaans negation context is crucially different to the piedpiping vs-stranding type instantiated, for example, in the Afrikaans embedded *wh*-interrogative case to be discussed here as Example 7. The latter springs from an **option-permitting grammar** (essentially, the size of the moving element may vary) which is acquired during the course of L1 acquisition, while the former is the consequence of **two distinct (sub-)grammars, acquired in different ways** – one as, effectively, an L2 (see Meisel, Elsig & Bonnesen 2011) – **but co-existing within individual speakers.**

4.1.2.3. *Word order-related innovations*

Example 3: Afrikaans V2 in embedded wh-interrogatives

- This case involves both lexical and word-order-related innovation.
- Word order is again a property that children are very sensitive to.
- No Germanic language permits V2 in genuinely embedded *wh*-interrogatives (not even the relevant colloquial varieties of English; McCloskey 2006, Woods 2016a,b); but Afrikaans allow V2 to alternate freely with V-final structures in embedded *wh*-interrogatives (Biberauer 2014b, 2015b)

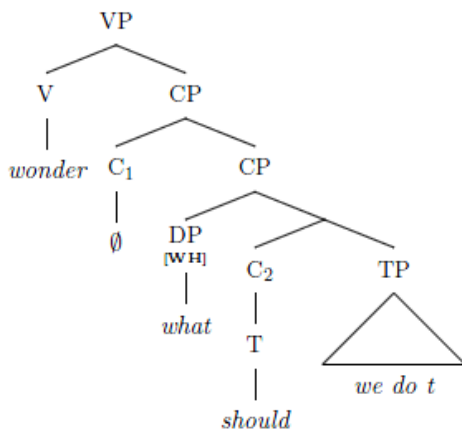
- (57) a. Ek wonder [wat eet hulle saans (eet).]
 I wonder what eat they evenings eat
 ‘I wonder what they eat in the evenings.’
- b. Ek sal uitvind [hoe kom ons by die gebou in (kom).]
 I shall out.find how come us by the building in come
 ‘I will find out how we (can) get into the building.’
- c. Ek sal kyk hoe lyk die weer as ek aankom.
 I shall see hoe look the weather when I arrive
 ‘I’ll look to see what the weather looks like when I arrive.’
 (Source: email from 40 year-old Afrikaans native-speaker, 23 November 2015)
- d. Source: native-speaker of Afrikaans, aged 7 (November 2015)
 [literally: “Dear Daniel. I wonder how goes it in England.”]



- e. Sien hoe sy haar kop skuins draai om te hoor hoe skinder
 see how she her head sideways turn INF-C to hear how gossip
- ‘n trossie voëls in ‘n tak langs hulle.
 a cluster birds in a branch beside them
 ‘Observe how she inclines her head to listen how a flock of birds gossips on a branch beside them.’
 (Biberauer 2003:191)
- Why should Afrikaans permit this option?
 - McCloskey (2006): embedded V2 is possible in the complements of

verbs that select a “large” clausal complement, thus circumventing the ban on selected Cs containing lexical verbs (his **KRR effect**). *Wonder*-predicates c-select larger complements than *discover*-predicates:

- (58) a. * $[_{CP^{matrix}} \dots \text{discover } [_{CP} \text{ what should-C } [_{TP} \text{ we should } [_{VP} \text{ do what}]]]]]$
 b. $[_{CP^{matrix}} \dots \text{wonder } [_{CP} [_{CP} \text{ what should-C } [_{TP} \text{ we should } [_{VP} \text{ do what}]]]]]]]$
 (a form of **CP-recursion**)
 c.



- Biberauer (2015b): Afrikaans clauses are consistently bigger than just a single CP owing to an innovation in the negation system: the establishment of clause-final *nie* in negative clauses – cf. (11) above and the previous example. This *nie* was integrated into the clausal domain as a left-peripheral (outermost) Pol-head (cf. Laka 1990, 1994 and much subsequent discussion of the variation associated with the placement of Pol).

- (59) a. $[_{CP^{matrix}} \dots \text{wonder } [_{CP} [_{CP} \text{ what should-C } [_{TP} \text{ we should } [_{VP} \text{ do what}]]]]]]]$ (= (46c))
 b. $[_{PolP} [_{CP^{matrix}} \dots \text{wonder/discover } [_{PolP} [_{CP} \text{ what should-C } [_{TP} \text{ we should } [_{VP} \text{ do what}]]]]]]]]]$

- Although it would initially only have been [Pol:Neg] that required postulation of CP-peripheral PolP, it is clear that speakers have generalized this Pol-projection to (i) **affirmative** structures [evidence from Verum Focus-type predicate-doubling and expanded complementiser-less V2 (the null complementiser = a spellout of [Pol:Aff]), and (ii) **interrogative** structures [the above embedded *wh*-V2 data], and (iii) **emphatic negation** in sub-clausal structures (cf. Biberauer 2009).

Thus: IG has led to the generalization of a feature/head beyond what it was initially only required for negative structures.

- Significantly, we also see **lexically conditioned** IG, which interacts with the above change:

- (60) a. Ek wonder **of** daar vir ons werk is. [standard Afrikaans]
 I wonder if there for us work is
 ‘I wonder if there is work for us.’
- b. Ek wonder **of is** daar werk vir ons. [dialectal Afrikaans]
- For (60b)-speakers, the interrogative-marking clausal C-head *of* has **lexically merged** with the already available, acategorial disjunction marker *of* (‘or’, ‘either ... or’), which is plausibly a spellout of acategorial Pol (cf. Jayaseelan 2001, 2008, 2014, and also the frequent attestation crosslinguistically of Q-particles deriving/lexically indistinguishable from the disjunction marker – Biberauer 2016).
- (61) a. Koffie **of** tee?
 coffee or tea
 ‘Coffee or tea?’
- b. **Óf** die heining **óf** die posbus moet jy hierdie naweek regmaak!
 or the fence or the postbox must you this weekend right.make
 ‘Either the fence or the postbox you have to fix this weekend!’
- (62) Conservative *of*:
- a. Ek wonder **of** daar vir ons werk is.
 I wonder if there for us work is
 ‘I wonder if there is work for us.’
- b. ... *wonder* [_{PolP} Pol [_{CP} **of-C** [_{TP} *daar vir ons werk is*]]]
 i.e. interrogative-specific *of* in C
- (63) Innovative *of*:
- a. Ek wonder **of is** daar vir ons werk.
- b. ... *wonder* [_{PolP} **of-Pol** [_{CP} *is-C* [_{TP} *daar vir ons werk is*]]]
 i.e. underspecified (disjunctive) *of* in Pol
- (64) Upwards reanalysis:
- a. *of* = C-head: [_{PolP} Pol [_{CP} **of-C** [_{TP} T ...
- b. *of* = Pol-head: [_{PolP} **of-Pol** [_{CP} C [_{TP} ...
- **More conservative speakers** resist (60b/63b/64b) because they **generalise the input** they get for **overt complementizers**, treating all overt complementizers as the spellout of C.
 - **More innovative speakers** entertain (60b/63b/64b) because they **generalise the input** they get in relation to *of*.
 - Crucially, we see that **Input Generalisation has affected all speakers of modern-day Afrikaans**, with the **difference** between them being **the pattern that serves as the basis for their extended generalizations**: a class of phonologically distinct C-elements (overt complementisers) or a set of homophonous C-elements (*of*), i.e. syncretism.
 - Differences of this type more generally, which we might expect to arise wherever the input underdetermines or radically lacks an unambiguous indication of the formal specification acquirers should postulate, can be

expected to lead to **variability between speakers**, depending on the significance they attach to other (non-ambiguous) properties in their system. ... and then they influence each other, with the result that we might also expect intra-speaker variation in domains of this type.

- An interesting question: modern colloquial varieties of English have also extended their “residual” V2 inventories in various ways (e.g. via Exclamative V2 – *Man is that annoying!* Biberauer 2012) suggesting expansion on the basis of a coherent, featurally defined class of elements (focused quantified elements trigger inversion) ... though, presumably, by adults, rather than children (?); see Sailor (2015, 2016) for discussion of an even further elaborated innovative system.

5. Conclusions

- Making room for a suitably specific third-factor may facilitate the advancement of the types of analyses generativists have come to feel offer satisfying accounts of crosslinguistic variation (typology), acquisition and change, even without a richly specified UG.
- On the approach advocated here, the formal features [F]s structuring natural-language systems have little to do with Universal Grammar, deriving in largest part from specific components of the linguistic input – e.g. (A-E) in Section 2.2, with their manifestations in questions and imperatives being particularly significant – and, no less importantly, the third-factor-regulated ways in which children go “beyond the input”.
- There is, then, no reason to let go of the idea that variation is structured on the basis of formal features.
- On an MMM-type 3 factors model, we expect these formal features to be organized in particular ways, creating patterns of variation that have a parametric character in that they can be formulated in familiar terms (e.g. the nature and location of [F]s on functional heads – the Borer-Chomsky Conjecture (Baker 2008); how featural distributions and dependencies are spelled out – PF/spellout parameters relating (Berwick & Chomsky 2011); and how formally present features are interpreted at LF – LF parameters (Ramchand & Svenonius 2008, Wiltschko 2008, etc.)), and also in that it’s possible to see how their effects might go beyond “rules”/lexically based stipulations in facilitating into the grammar more generally (cf. i.a. Newmeyer’s criticisms).
- The model crucially leads us to expect variation to exhibit particular formal properties, some of which are already well established in the literature, and some of which are only emerging more clearly now.
- It also, no less crucially, seems to facilitate insight into the long-standing puzzle of how acquirers can “go beyond the input”, and how this ability might relate to our more general cognitive abilities.
- In the context of the model presented here, we expect that quite distinct situations in which the **input is compromised in some way** – e.g. via **irregularity, incompleteness** (e.g. colloquial structures that fall “beyond” the prescriptive radar), etc. – will deliver feature-based NONE, ALL and/or SOME-type generalisations.
- The fact that this seems to be correct suggests that the acquisition biases we have identified here are of the kind that can productively be investigated as potential factors in understanding **why acquirers are not stymied by partial or irregular input or, indeed, by variation between native-speakers.**

- In combination with our assumptions about **aspects of the input that are particularly significant to acquirers**, we can also make progress in understanding **why certain seemingly “complex” or “redundant” properties should prove stable** (Nichols 1992 *et seq.*). **Inflectional morphology**, for example, will always trigger the postulation of one or more [F]s, and IG means that ALL-type systems, in which the relevant [F]s are invariantly associated with their associated category – as in noun-/verb-class marking, agglutinating morphology, etc. – will be acquisitionally favoured, and thus expected to be stable, which is correct.

References available from me