COPiL CAMBRIDGE OCCASIONAL PAPERS in LINGUISTICS
Volume 6, Article 9: 267–294, 2012 | ISSN 2050-5949

TOWARDS A PARAMETER HIERARCHY FOR AUXILIARIES: DIACHRONIC CONSIDERATIONS *

THERESA BIBERAUER
Stellenbosch University
& University of Cambridge

IAN ROBERTS
Downing College, Cambridge

ABSTRACT This paper's objective is to reconsider the oft-discussed parameters determining verb- and auxiliary-movement in English in the context of the "parametric hierarchies" proposal of Roberts (2012) and its elaboration in Biberauer & Roberts (2012). In terms of the latter, parameters may be viewed as macro, meso, micro or nano, depending on the size of the class of elements whose behaviour they regulate (all elements of a given type, a featurally specifiable subset of the elements of a given type, the smallest definable sub-class of elements of a given type, and one or more individual lexical items, respectively). In relation to the history of English verb-movement, we show that English has "moved down" the parametric hierarchy, the loss of V-to-T movement having resulted in the rise of a number of micro- and nanoparametric properties. Strikingly, these smaller-scale parametric properties seem to be very unstable, constituting a domain in which syntactic change occurs very readily.

1 INTRODUCTION

The goal of this paper is to situate certain rather well-known parameters concerning verb- and auxiliary-movement in English in relation to the general concept of parametric hierarchies developed in Roberts (2012). In order to do

* The research reported here was funded by the European Research Council Advanced Grant No. 269752 "Rethinking Comparative Syntax" (ReCoS). We would like to thank our colleagues on the project Tim Bazalgette, Alison Biggs, Anders Holmberg, Michelle Sheehan and Jenneke van der Wal. In addition to DiGS 13, versions of this paper were presented at the Deutsche Gesellschaft für Sprachwissenschaft in Frankfurt in March 2012, and at the ReCoS Open Meeting in Cambridge. We would to thank the audiences for their questions and comments.

this, we begin by proposing a tentative classification of parameters in §2. In §3, we reconsider the data regarding the change in the value of the "V-to-T" parameter in Early Modern English. The goal of §4 is to situate the "V-to-T" parameter in relation to a parameter hierarchy for V-movement, whose nature will be explored. One aspect of this will involve looking at the nature of the auxiliary in contemporary colloquial varieties and some of the so-called *New Englishes*. Finally, §5 considers wider issues raised by this case study for the way in which the theory of parameters relates to syntactic change.

2 Parameters

We begin with the standard view of the nature of syntactic parameters as put forward in Chomsky (1995):

(1) Parametric variation is attributable to differences in the formal features of functional heads as specified in the lexicon.

Still following standard assumptions, we take formal features to include Case, person, number, gender (henceforth ϕ), categorial features, as well as movement-triggering features (EPP or Edge Features). These features are associated with functional heads such as C, D, T, v, etc. The idea behind (1) is simply that these features are subject to variation.

A central idea that we want to develop here is that the range of possible parametric variation characterised by (1) is not uniform, in that variation across certain sets of features may give rise to differing degrees of surface variation. More precisely, we suggest the following rough taxonomy (based on Roberts 2012, Biberauer & Branigan 2012):

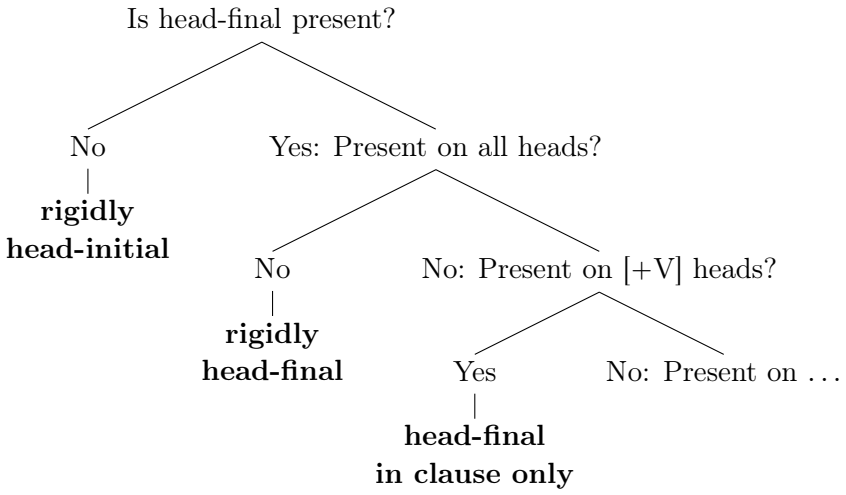
(2) For a given value v_i of a parametrically variant feature F:

a. **Macroparameters**: all heads of the relevant type share v_i;
b. **Mesoparameters**: all functional heads of a given category (e.g. all verbal heads, all nominal heads, all ϕ-bearing heads or all finite Cs) share v_i;
c. **Microparameters**: a small subclass of functional heads (e.g. auxiliaries, pronouns) share v_i;
d. **Nanoparameters**: one or more idiosyncratic lexical items are specified for v_i.

We will exemplify each of these types of parameter as we proceed. For now, the thing to notice is that the different kinds of parameters are hierarchically related to one another.

Roberts (2012) develops the idea of parametric hierarchies in more detail. A good example of such a hierarchy is that determining word order/linearization. The upper part of this hierarchy is presented in a preliminary way in (3):

(3)



Here the term *head-final* is a cover term for a more technical notion, whose precise nature need not concern us here. It can be reduced to a complement-movement feature, following the general approach in Kayne (1994) or to a PF Head Parameter as discussed by Richards (2004) and Sheehan (to appear).

A very important aspect of the approach put forward by Roberts and developed by Biberauer (2011), Branigan (2012) and Biberauer & Branigan (2012) is the idea that the hierarchies are not prespecified by UG, but instead are seen as emergent properties, arising from the interaction of markedness conditions (ultimately originating in "third-factor" principles optimising the acquisition process), a highly minimal underspecified UG and the Primary Linguistic Data (PLD). Two such conditions in particular are at work (although of course we do not exclude the possibility that there are others):

(4) a. Feature Economy (FE):
Given two structural representations R and R' for a substring of input text S, R is less marked than R' iff R contains fewer formal features than R'

b. Input Generalisation (IG):
If a functional head F sets parameter P_j to value v_i then there is a preference for similar functional heads to set P_j to value v_i.

The definition of FE given here is taken from Roberts & Roussou (2003:201), while IG originates in Roberts (2007:275).

Looking again at the word-order hierarchy in (3), and assuming either that head-final orders involve "roll-up" movement of complements triggered by a formal feature of a class of functional heads, or that the "final" value of the Head Parameter requires a special feature, consistently head-final order is one degree more marked than consistently head-initial order.[1] Importantly, all intermediate types are more marked than either of these.

The highest position in the hierarchy, then, conforms fully to both FE (since there is no feature) and IG (the absence of the feature is generalised). The next highest position is more marked in relation to FE, in that the feature is postulated, but still maximally unmarked in relation to IG in that the feature is fully generalised. In this sense, IG leads all the relevant functional heads to "point the same way." All lower positions in the hierarchy are relatively marked in relation to FE (as the feature is postulated) and in relation to IG (as the feature is not fully generalised).

As already mentioned, IG and FE are *not* grammatical principles, but rather acquisition strategies, which may be motivated by a general conservativity on the part of learner's use of cognitive devices. Mobbs (2008, in progress) suggests that this is a reflection of a non-language specific optimisation principle. It follows from this that the distinctions among the various types of parameters as presented in (2), derive from markedness, which in turn emerges from the conservativity of the learner.

As (2) implies, true macroparameters sit at the top of the hierarchy. As we move successively "downwards", the systems become more marked, the parameters become meso, then micro, then nano. Parameters in lower positions on the hierarchy have a longer description (the conjunction of all the higher "nodes"), and, in this sense, are intrinsically more complex. It is also plausible to think that these parameters are further along a learning path, as the least-marked values (the highest, macro ones) represent the acquirers' initial hypotheses, and assume the least amount of "knowledge" on the part of the acquirer; in fact, following Biberauer (2011), Branigan (2012) and Biberauer & Branigan (2012), we could assume that non- or less category-specific parametric choices (i.e. macro- and mesoparametric settings) are those automatically

1 The existence of the Final-over-Final Constraint (FOFC) supports the idea that head-final and head-initial orders are not equivalent. FOFC can be informally stated as follows:

(i) A head-final category cannot have a head-initial category as its immediate structural complement in a single Extended Projection, whilst a head-initial category may have either a head- final or a head-initial immediate complement.

For a more formal characterisation of FOFC, and an attempt to derive it from more general principles, see Biberauer, Holmberg & Roberts (2012).

"chosen" by the acquirer based on early "ignorance", which may or may not be subject to subsequent refinement.[2]

Most important for our present purposes, the hierarchies make predictions about diachrony. First, grammatical systems instantiating "lower" parameters are diachronically closer to one another than those in the higher reaches of the hierarchy. Second, if we think of language acquisition as involving the acquirer moving "down the hierarchy", driven from relatively unmarked hypotheses to more marked ones by the PLD which forces them to postulate ever more fine-grained parametric distinctions, and if we assume, following Lightfoot (1979, 1991), that language change is driven by reanalysis in language acquisition, then, it follows—all other things being equal—that languages will "drift upwards" in the hierarchies. Both of these predictions are relevant to the discussion that follows.

With this much background, we can now turn to the discussion of the history of English.

3 The "V-to-T" parameter in Early Modern English

It is well-known that lexical verbs could appear in higher clausal positions in earlier stages of English. Leaving aside Verb Second, which is usually thought to have died out in the 15th century (see in particular Fischer et al. 2000), verb-raising to T remained a productive option throughout the 16th century and into the 17th (see Roberts 1985, 1993 and Warner 1997 for discussion and details). This is shown by the position of the lexical verb in relation to *not* and light VP-adverbs such as *anon*, as illustrated in (5a,b), as well as the fact that main verbs could raise over the subject to C in main-clause interrogatives, as shown in (5c):

(5) a. *if I* ***gave not*** *this accompt to you*
'if I didn't give this account to you'
(1557: J. Cheke, Letter to Hoby; Görlach 1991:223, Roberts 1999:290)

b. *The Turkes .. **made anone redy** a grete ordonnaunce*
'The Turks ... soon prepared a great ordnance'
(c1482: Kaye, *The Delectable Newsse of the Glorious Victorye of the Rhodyans agaynest the Turkes*; Gray 1985:23, Roberts 1993:253)

2 Since certain core parameters (including word order) are set very early (see i.a. Wexler 1998), there may not always be direct acquisition evidence of the fundamental category-related "ignorance" postulated in relation to macroparameters; we do, however, expect initial overgeneralization errors in relation to more specified classes of elements.

c. *What* ***menythe this pryste****?*
'What does this priest mean?'
(1466-7: Anon., from J. Gairdner (ed), 1876, *The Historical Collections of a London Citizen*; Gray 1985:11; Roberts 1993:247)

In terms of the analysis of Modern French in Pollock (1989), we can think that Early Modern English (ENE) had the "French" value of the V-to-T parameter. It therefore follows that the value of this parameter changed at the earliest in the 17th century.

In fact, verb-movement to T has not been entirely lost in Modern English (NE). Following Emonds (1978), NE is generally thought to have an operation of "*have/be* raising", which moves the auxiliaries *have* and *be* to T. (6) illustrates this for perfect *have*:

(6) a. John **has often** kissed Mary.
b. John **has not** kissed Mary.
c. **Has** John kissed Mary?

Suppose that this involves v-to-T rather than V-to-T raising, with *have* and *be* first-merged in v. Modals and *do* were grammaticalised as functional elements in the early 16th century (see Roberts & Roussou 2003 and the references given there). Following Biberauer & Roberts (2010), we take it that these elements are also first-merged in v after this change. Then the characteristic NE split in main-verb vs auxiliary syntax follows from the fact that NE has retained v to T raising, but lost V-to-v raising. In relation to the characterisation of parameters given in (2), we can see that this is a case where a mesoparameter (all finite verbs raised to T) was replaced by a microparameter, whereby only finite auxiliaries raise to T. More precisely, the loss of V-to-v movement meant that v-to-T movement could no longer affect lexical verbs; the operation was retained, but was only able to affect auxiliaries first-merged in v. Thus the scope of the parameter changed from meso (affected all finite verbs) to micro (affecting only finite auxiliaries).

Furthermore Warner (1997:383) observed that there were specific lexical exceptions to the generalisation just made, in that certain lexical verbs, e.g. *know*, *doubt* and a few others, continue to raise to T until "some point in the eighteenth century" (Warner 1997:383). It seems then, that there was a period of nanoparametric variation affecting these few lexical verbs, raising them from V to v, in the 18th century. Thus the original productive option affecting all finite verbs split—presumably in the 17th century—into a microparameter

affecting auxiliaries and a nanoparameter affecting a handful of idiosyncratic lexical verbs.[3]

The question that now arises is how the non-UG-given parameter types that we have introduced in fact arise: how does a macro-, meso-, micro- or nanoparametric specification come to be part of an I-language? In general terms, we assume that acquirers will approach the input guided by UG-given knowledge, such as the fact that certain semantic features are formalised/grammaticalised, while others are not (the semantic vs formal feature distinction of Chomsky 1995) and that there is a UG-given movement diacritic (EPP-feature) which associates with substantive formal features in various ways. Gianollo, Guardiano & Longobardi (2008) present one proposal as to the kind of acquisition 'schema' that could lead acquirers to analyse the PLD in ways that would naturally lead to hierarchical parameter schemata of the type we are considering. We leave details of this complex question to future work. What is crucial for our purposes is that there is, owing to the fact that they require "less knowledge", a clear sense in which higher options in parametric hierarchies of the sort illustrated in (3) are simpler—and can thus be viewed as defaults—in relation to lower options. Defaults, then, are assumed to emerge from the emergent parameter hierarchies. To see how this is the case for V-to-T movement, we first need to say something more precise about the nature of head movement, however.

In this context, we follow Roberts (2010), who relates head-movement to the notion of defective goal. A defective goal is defined as follows:

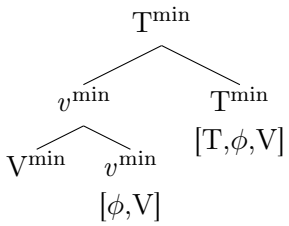
(7) A goal G is defective in relation to a probe P iff G's formal features are properly included in those of P.

For Roberts, head-movement can take place only where the features of the incorporee are properly included in those of the incorporation host, i.e. only where there is an Agree relation involving a defective goal (see Roberts 2010:54ff for an account of why this is so). In the context of v-to-T movement, the resulting structure is therefore as in (8):[4]

3 The nanoparametric property must be marked in the lexical entry of each individual verb. This kind of "syntactic irregularity" is thus like the morphological irregularity of certain verbs. So, for example, *sing* must simply be marked as having the past and participial forms *sang* and *sung*. But note that these, like the syntactic irregularity just observed, reflect a formerly productive option, in this case ablaut. Crucially, then, nanoparameters are predicted to involve item-specific instantiations of properties that are more systematically observed in other systems where the same property has a micro or meso or possibly even macro character. As such, they may, like their morphological counterparts, give us important clues to properties that earlier stages of a given system may have had.

4 The assumption here is that in general the number of formal features associated with each (non-defective) functional head increases monotonically as we move up the clausal hierarchy.

(8)



So we see that v is a defective goal in relation to T, and, as such, is able to incorporate with T. V, lacking formal features, is "inert" in relation to the probe-goal system and so cannot incorporate. V-to-v movement at earlier stages of English must then have depended on V's possession of formal features rendering it capable of participating in the probe-goal system and therefore capable of undergoing head-movement. Here we propose that lexical verbs in systematic verb-raising systems like earlier English differ from those in consistent non-raising systems (e.g. Niuean—see i.a. Massam 2005—and also some new Englishes—see §4) in that they enter the syntactic derivation in already-"verbalised" form, i.e. as heads specified [V]. This verbal specification is necessary to allow verbs to bear a formal feature for which higher heads can probe (see Biberauer & Roberts 2010) and also to allow the verb to bear unvalued V-related formal features (e.g. [Asp:___] and [T:___]) for which higher heads can probe; in more traditional terms, verbs must be specified [V] to enable them to be associated with the features that can be spelled out as verbal inflection. From the acquirer's perspective, this means that the presence of verbal inflection will signal that verbs in their language are specified [V].

Importantly, though, V-to-v movement does not automatically take place if lexical verbs are specified [V]; in terms of the system being sketched here, this depends on the featural specification of v and whether v's formal features represent a superset of V's. Following Biberauer & Roberts (2010), we assume NE auxiliaries to be v-elements which systematically undergo movement to T. More specifically, we assume the rise of this specific class of lexical items—in part, of course, connected to the loss of tense- and agreement-inflection—to have led to a change in terms of which the formal features previously associated with v came to be distributed in such a way that v retained, alongside its [V] specification, only its basic argument-related properties (e.g. capacity to introduce an external argument, capacity to probe an internal argument,

A lexical head may entirely lack formal features (on this point see text to follow), v has valued [V] and unvalued [ϕ], T has [T, V, ϕ], with [ϕ], in contrast to [V] and [T], once again unvalued, and so on. In practice, the actual distribution of formal features within an extended projection varies from language to language, with lower heads in some cases having more formal features than higher ones (see main text for illustration). This variation, we argue, is directly responsible for differences in the availability of V-movement in different systems.

where relevant); the verbal formal features previously associated with this head (e.g. [Asp], [Tense], [Polarity]—see Biberauer & Roberts 2010 for discussion of the types of features assumed to be encoded by NE auxiliaries) came to be lexicalised by the auxiliary elements.[5] As such, v crucially encodes the same formal feature as lexical V and no further verbal formal features; it does, of course, bear ϕ-features, which V lacks, but these do not probe V and, as such, will never form a specific syntactic link with V. We assume this last point to be crucial in determining the possibility of V-to-v movement: where V shares v's [V] specification, it will only incorporate (i.e. undergo V-to-v raising) if v bears one or more additional formal features creating a v-V Agree relation. In general terms, this analysis predicts that V-raising will be dependent on a monotonic rise in verbal features along the clausal extended projection (see Grimshaw 1991, 2005 and also Note 5) to the point where V-raising stops, with any change to the feature make-up of intermediate heads (e.g. that brought about by the rise of a class of analytic forms, see Note 6) necessarily interfering with V-movement possibilities.

Returning to the question of default parameter values, then: if options requiring the postulation of fewer formal features constitute defaults and both inflection and verb-raising are reflexes of the presence of formal features (by virtue of the fact that both require Agree), it is clear that systems exhibiting these properties cannot instantiate initial defaults in a parametric hierarchy. In the following section, we will consider in more detail how such systems can be incorporated into a parametric hierarchy. For the moment, we conclude this section by highlighting the importance of verbal inflection in the context of our proposal: by hypothesis, its presence unambiguously signals the presence of formal features which have to be acquired and about which acquirers may then "ask" the types of hierarchy-defining questions illustrated in (3) above. Like earlier proposals (i.a. Kroch 1989, Roberts 1993, Lightfoot 1997, Warner 1997 and Biberauer & Roberts 2010), then, we propose that the morphological trigger, extant until the 16$^{\text{th}}$ century, was central in supporting Earlier English-type verb movement.[6] In this connection, the much-debated question of a

5 These auxiliaries are v-elements in the sense that they lexicalise features associated with the vP-domain (e.g. voice, aspect, etc.), but they lack specifically argument-related formal features. As such, we assume them to be merged immediately above argument-related ϕ-bearing v. In other words, we assume that the introduction of a special class of auxiliaries (analyticization) resulted in the necessary projection of a more articulated vP-domain.

6 We do not, however, exclude the possibility, also mentioned by Warner in relation to earlier English, of "secondary" syntactic triggers such as movement. That is, in addition to relying on inflectional cues, acquirers can also establish the presence of formal features on the basis of structures that must be movement-derived (cf. also Gianollo et al. 2008, Zeijlstra 2008 and Biberauer 2011 on this). Movement is clearly available in morphologically impoverished languages—consider the Mainland Scandinavian languages and Afrikaans, all of which are V2 languages, despite famously having lost their verbal agreement morphology. See main

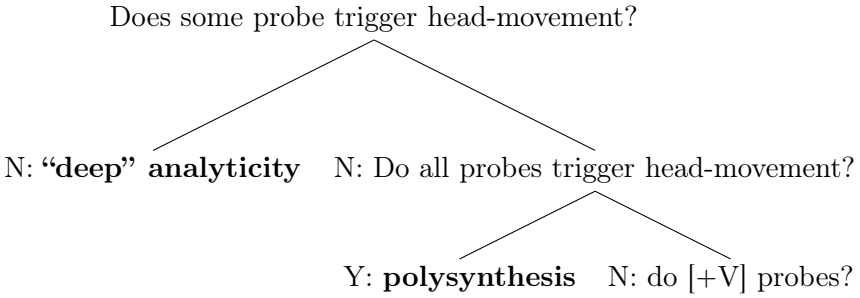
possible morphological trigger for V-movement becomes clearer.

Now it is time to consider in detail how the verb-movement parameters fit into a parametric hierarchy.

4 A parameter hierarchy for V-movement

Roberts (2012) proposes a parametric hierarchy determining word structure, which has the following form at the highest levels:

(9)



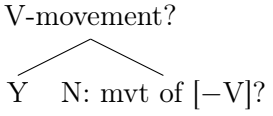
Assuming that an extra feature is needed to trigger head-movement (which follows from the approach to head-movement sketched in the previous section), the highest option is maximally unmarked in relation to both FE and IG (cf. (4) above): for head-movement to occur in the system outlined in § 3, a higher head must (i) locate a lower head on its extended projection bearing at least one probable (i.e. valued) formal feature and (ii) itself bear a superset of the formal features located on the goal. If formal features of the relevant type (V-related for a V-oriented probe, D-related for a D-oriented probe, etc.) are entirely absent, we predict a head-movement-less system. Thus the default option in relation to a word-structure parameter of the type Roberts (2012) envisages arguably gives rise to "deep" analyticity of the kind described for Mandarin and other Chinese varieties by Huang (2007). The next option features maximal head-movement to all available targets; this gives rise to a polysynthetic system of the type described in Baker (1996). Such a system is more marked than a fully analytic one in terms of FE (formal features need to be postulated), but still maximally unmarked in terms of IG, since the features are fully generalised: every formal feature-bearing head (probe and goal) behaves in the same way. Fusional languages are inherently more marked, involving feature specifications that vary across heads and thus violate both FE and IG. This results in meso- or smaller variation and here we can see that the familiar V-movement parameters come into the picture.[7]

text for references considering the differences between V-to-T and V-to-C triggers.

7 Traditional agglutination, to the extent that it is exclusively suffixal or exclusively prefixal, may in fact be viewed as a case of analyticity. This would be the case if this type of

The next part of this hierarchy, below the deepest right branch in (9), must involve the most basic categorial distinction, which we take to be V vs N. Hence it may look as follows:

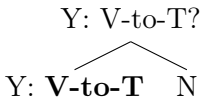
(10)



The right-branch option gives a subhierarchy relating to N-movement, of the kind studied by i.a. Longobardi (1994), Bernstein (2001) and Julien (2005). The left branch gives the V-movement subhierarchy. The highest option on this subhierarchy must give the most general (i.e. default) option concerning V-movement. We take this to be movement to T. This is not because we consider C to be outside the extended projection of V; *pace* Grimshaw (1991, 2005) we take C to be in this extended projection (see Biberauer, Holmberg & Roberts 2012 for empirical justification of this). Instead, as already mentioned in the previous section, we take V-movement to v and T to be bound up with the inflectional system, with V-movement to C being driven by a different class of features (see i.a. Holmberg & Platzack 1995, Roberts 2010).

In (11), we see the next part of the verbal subhierarchy, dominated by the "Y" node in (10):

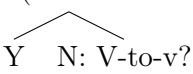
(11)



The V-to-T option here gives a grammar with the properties of French or Early Modern English (ENE), as described in §3. The cue for this setting, as stated in §3, is "relative" richness of verbal inflection, which signals the presence of formal features, all of which behave in the same way, respecting IG (cf. Biberauer & Roberts 2010 for further discussion of the types of 'richness' involved here).

At the "N" node in (11), the choice concerns the availability of more restricted V-movement to T, namely auxiliary-movement:

(12)



agglutination can be shown to involve heads consistently bearing the same formal features, i.e. probes with movement-feature triggering comp-to-spec movement in the former case (cf. Julien 2002) and, potentially, probes lacking a movement-feature and also not bearing a superset of the goal's formal features in the latter.

Here we see that a negative setting of the V-to-T parameter does not rule out the possibility of verbal elements appearing in T since auxiliaries (which we take to be first-merged in v) may constitute a class of verbal elements that move there. This, of course, is the Modern English (NE) setting. Nor does the negative setting of V-to-T entirely rule out movement of the lexical verb: it may still move to v. According to Biberauer & Roberts (2005 *et seq.*), this is the setting in the OV West Germanic, languages, which, in combination with VP-movement to Spec,vP (an option regulated by a distinct, but intersecting parameter hierarchy) gives surface verb-final order in subordinate clauses.

We can now see that the loss of V-to-T movement for lexical verbs, with the retention of auxiliary (i.e. v-) movement to T makes NE a more marked system than ENE. This seems correct as it has often been observed that the split syntax of auxiliaries vs lexical verbs in NE is cross-linguistically rare and late acquired (cf. i.a. Davis 1987, Stromswold 1990, Rowland & Theakston 2009 and Theakston & Rowland 2009 for varied discussion).

The availability of movement to T, whether of v or of both v and V, of course sets up the possibility of further movement to C. Taking up an idea first proposed in Holmberg & Platzack (1995), we assume that V-to-T is not required to feed Verb Second, however, a state of affairs plausibly instantiated in the modern Mainland Scandinavian and OV Germanic languages (see Roberts 2010, chapter 4 and 5 for discussion and for motivation of the non-adoption of the Head Movement Constraint that this view entails). As noted above, we assume that movement to C is not triggered in the same way as the movements within the core verbal domains (for which we have employed the cover-terms *vP* and *TP*). As such, it is could feasibly be viewed as an option regulated by a distinct hierarchy, namely that concerned more generally with the grammatical options (both head- and XP-movement, and agreement) associated with the features triggering V-to-C. As before, we remain noncommital as to the precise nature of these features, venturing only that these will be the sorts of features associated with the extended left periphery in the sense of Rizzi (1997), i.e. broadly discourse-oriented features. More specifically, consideration of the range of V2 patterns found in modern Germanic suggests that features relating to polarity (veridicality/affirmation—hence the matrix bias and embedded V2 and possibly even Quotative Inversion effects, and conceivably also focus- and topic-relevance; non-veridicality—hence conditionals, interrogatives and fronted negation and maybe even the exclamative V2 discussed in Biberauer (2010) in NE) may be central. Clearly, the specifications associated with languages opting for a generalised V2 pattern will be simpler than those opting for a more restricted pattern. In this sense, then, it is easy to see that an NE V2 system, requiring a restriction to a subclass of nonveridical

matrix Cs, is more marked than a full V2 West Germanic-style V2 system (see § 4 for further discussion).[8] Strikingly, the range of inversion options taken in Standard NE do not feature in all modern varieties. Conditional, Negative and Quantifier Inversion are, for example, restricted to high-register varieties, while many studies of Indian and Singaporean English have additionally registered the (frequent) non-occurrence of subject-auxiliary inversion in matrix interrogatives (cf. i.a. Bhatt 2004 and Gupta 1994). Similarly, In English-based creoles, subject-auxiliary inversion appears to be systematically ruled out. The following examples illustrate:

(13) Colloquial Singaporean English

a. How much it will be?

b. What the cruise is like? (Gupta 1994:8)

c. The tea very hot, **is it**? (Wee 2008:595)

d. You ever work in other jobs **ah**? (Ziegeler 2011:5)

(14) a. Jamaican Creole
Im no love dem ting?
'Doesn't she like those things?' (Patrick 2004:11)

b. Kamtok
Wuna go bay moto smol taim, **no bi so**?
'You will soon buy a car, won't you?' (Ayafor 2004:926)

As shown above, non-inversion may correlate with the presence of an interrogative particle, arguably another case where the introduction of new lexical items produces parametrically relevant knock-on effects in the syntax (cf. the discussion of the introduction of English auxiliaries above).

Non-standard modern varieties also exhibit a number of differences from standard American or British NE in relation to the nature of their verbal and auxiliary systems. Thus some English-based creoles show serial-verb constructions (SVCs). Consider (15):

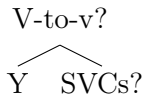
(15) a. Uh **run go** home.
'I run home'

b. Uh **tuhn look up** fuh heh.
'I turned [and] looked up for her' (Mufwene 2001:307)

8 It is less clear that (standard) NE V2 is more complex than that of OE and early ME with its in part idiosyncratic lexical triggers (*þa, þonne*, etc.), a matter we leave aside here.

SVCs arguably involve lexicalisation of v without V-movement, with the "light", serialising verb instantiating v. To the extent that they systematically lack inflection and also do not affect the shape of the lexical verbs they co-occur with, they plausibly instantiate components of a system lacking verbal formal features, i.e. the maximally unmarked system falling out from the "N" option defined by the top of the V-movement hierarchy given in (10). In systems where they do exhibit and/or trigger inflectional variation, we expect them to feed the negative option under the right branch of (12), viz.:[9]

(16)



Also characteristic of creoles, both English-lexifier and others, is the presence of immobile, invariant tense-mood-aspect (TMA) particles. These are, for example, found in Jamaican Creole:

(17) a. *Jan* ***did*** ***a*** *nyam unu bami.*
John PAST PROG eat your bammy
'John was eating your bammy.' (Durrleman-Tame 2008:33)

b. *Im* ***did*** ***jos*** ***a*** ***go*** *dw i'.*
s/he PAST RETRO PROG PROSP do it
'S/he was just about to do it.' (Durrleman-Tame 2008:34)

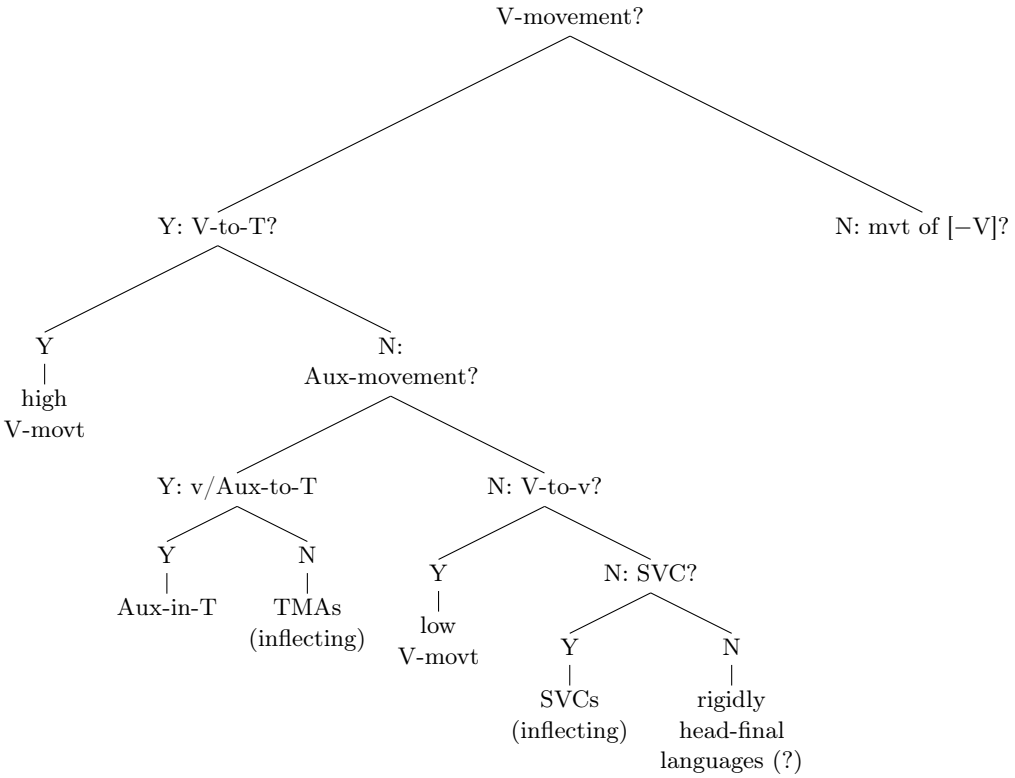
c. *Im* ***wi*** ***mos*** ***(h)afi*** *tek dat.*
S/he will must have-to take that
'S/he will be obliged to take that.' (Durrleman-Tame 2008:30)

Since these are non-inflecting elements which also do not affect the form of the lexical verbs they co-occur with, the acquirer receives no evidence signalling the presence of verbal formal features, either on the lexical verb or on the TMA particles. If this is correct, Jamaican Creole-type systems will necessarily instantiate the "N" option at the top of the V-movement hierarchy in (10), i.e. a maximally unmarked system.

9 In this case, movement-triggering formal features are absent, but Agree operations may nevertheless occur, giving rise to the observed inflectional variation. Worth bearing in mind, of course, is the possibility that the observed inflectional variation may simply be a PF-effect. If both types of formal variation exist, one would, however, expect Agree (i.e. syntactically-driven formal variation) to exhibit systematic differences from its PF-driven counterpart. This is a large topic which we leave aside here, but which is clearly centrally relevant to our parametric proposals.

Having outlined various V-movement-related parametric options, it remains to show how they may be incorporated into a hierarchical structure. Consider (18), which combines (10-12) and (16), in this connection:

(18)



Combining (16) with (9), we start to get a picture of what the head-movement hierarchy might look like. As noted in the main text, V2 appears to be an option (or, more accurately, a range of options) relating to a hierarchy defined by distinct features. As likewise noted, certain phenomena—e.g. TMA markers and SVCs—may fall out from different parts of the hierarchy, depending on the nature of their interaction with the verbal systems—specifically, whether they trigger form-change or not. Thus the TMAs and SVCs illustrated in (18) are not the invariant ones found in analytic creole systems—these belong to systems lacking V-movement and additionally offering acquirers no other (e.g. inflectional) evidence of the presence of verbal formal features; instead, they represent TMAs and SVCs in systems where either the TMAs/SVCs or associated verbs or both inflect. We leave to future research whether the predictions regarding the V-movement properties of such TMA and SVC systems match those predicted in (18) and also questions such as the diagnostics that would

allow acquirers to distinguish, for example, the "Y" and "N" options under v/Aux-to-T.

5 Nanoparameters in the English auxiliary system

In terms of the definitions given in (2), V-to-T movement-regulating (11) is part of the mesoparametric part of the hierarchy (as would be the parameters governing the V-movement part of V2), while the parts of the hierarchy concerning auxiliaries are microparameters. Additionally, the contemporary English auxiliary system also has many examples of variation restricted to individual lexical items, i.e. nanoparameters. One well-known case is possessive/modal *have*, which differs across the main varieties of English as regards its behaviour under negation and inversion:

(19) a. I don't have any money. (US, "advanced" UK)

b. I haven't got any money.
(most UK; earliest attestation mid-18thC)

c. I haven't any money. (conservative N. England, Scots)

It is noteworthy that perfect *have* never shows *do*-support and causative *have* always does, across all these varieties. So we need to distinguish three types of *have*: the modal/possessive type, which has the variants in (19a) and (19c), i.e. it shows "auxiliary syntax" in some varieties but acts as a main verb in others; the perfect *have*, which always shows auxiliary syntax in all varieties we are aware of, and finally the causative *have*, which is always a main verb. In most varieties of British English, we find (19b): here there is no possessive *have*; instead possession is expressed by what is historically the perfect tense of *get* in this variety, so *have* always has auxiliary syntax since this is in fact perfect *have*. There are, however, varieties in the Southern United States where this form of *got* has been regularised so that it appears without *have* (as in *I got rhythm*) and shows *do*-support in the relevant contexts (*I don't got money*).

There is also variation concerning *do*. One observation is that British and American English differ regarding the availability of *do* in VP-ellipsis contexts, as in (20) (cf. Kayne 2005):

(20) A: Will John come to the party?

B: He might $*^{\text{US}}$(do).

British English tolerates non-finite *do* as a VP pro-form here, while American English does not. Note that this is not the *do* of *do*-support, with which it can co-occur in some varieties, giving forms like *he didn't do*.

South-Western dialects of England also allow auxiliary *do* in positive declarative sentences, with what appears to be a habitual reading:

(21) a. Then he **did cut up** the various joints what you **wanted**.
b. We **did** come back then and we **did** have a glass or two of cider, and then we **did** go and have a bit of breakfast, come out again and then we **did** have another drink before we **did** start off. (Ihalainen 1991:154)

Furthermore, there are varieties (in the same general region, but slightly more geographically widespread), where *do* appears in the protasis of conditionals:

(22) a. If the boss **did** see that you was a bit pushing, .. he would ...
b. If you **did** buy up a load of peat in them days, it used to cost you ten shillings. (Ihalainen 1991:156)

But the most intricate and interesting variation arguably concerns modals. Here we see some fairly clear cases of recent and ongoing change. First, there are archaic forms which have recently gone out of normal British usage:

(23) a. **Mayn't** we go out to play now? No you mayn't.
b. **Shall** you miss your train? (1862, see Denison 1998:168)
c. They **dare** not be confused (1995(!), Denison 1998:167)
b. But I dare say he **might** [would be able to] come if he **would** [wished] (1816, Jane Austen; Denison 1998:168)

Regarding (23a), Denison (1998), citing Palmer (1990) says that contracted negation on *may*, giving *mayn't*, is no longer found in any variety. However, it is readily attested in 18th- and 19th-century English. (23b) is a fairly clear example of *shall* being used as a simple future auxiliary, with no sense of obligation; this, as far as we are aware, is no longer possible in most varieties of English today. (23c) is an example of a "raising" interpretation of *dare*; the antecedent of *they* is *ideas*, an inanimate DP, and therefore incapable of bearing the thematic role assigned by root *dare* to the external argument, which is instead here borne by the implicit argument of the passive (the content of the role is something like "possess courage"), so the intended interpretation is "One dare not confuse these ideas". Finally, (23d), a slightly older example, illustrates a conditional-ability sense of *might* and conditional-volition sense of *would*, both entirely unavailable in all varieties of contemporary English, as far as we are aware. So here we see several cases of recent, fairly subtle,

item-specific change affecting modals, all of which have taken place in the past two hundred years.

A second class of cases involves uses which are unremarkable in British English, but either not accepted or considered marked or marginal (or British-sounding) in other varieties, such as American English. These include inverted and negated *might* and *must*:

(24) a. **Might** it rain tomorrow?
b. It **mightn't/may not** rain tomorrow.
c. **Must** we? We **mustn't.**

Third, the "conditional inversion" construction no longer accepts any auxiliary other than *should* or *had* (it is also incompatible with contracted negation; these facts are discussed in Pesetsky 1989 and Roberts 2010)

(25) a. **Had** I been rich, everything would have been great.
b. **Should** he do that, everything will be great.
c. *__Did__ I do that, everything would be great.
d. ?__Were__ I/he to do that, everything would be great.
e. *__Were__ I rich, everything would be great.
a. If I **were** rich, everything would be great.

Note in particular that "modal" *be* is marginally possible in this construction, as seen in (25d), but not predicative *be* as in (25e). (25f) shows that there is nothing semantically wrong with this kind of conditional; it is the syntactic operation of inversion which appears to lead to ungrammaticality here.

In his very interesting and detailed discussion of this construction in the recent history of English, Denison (1998:298f.) shows that, up to the mid-19th century, this construction was productive with all auxiliaries. The following are attested examples of what are now ungrammatical sentences:

(26) a. And **were** she a little less giddy than she is ...
(Dickens, 1843-4; Denison 1998:298)
b. And **could** I read yours [face], I'm sure I should see.
(1864: Denison 1998:298)
c. My dear friend, **did** I want your aid I would accept it.
(1840: Denison 1998:299)

In fact, Denison's Table 3.11 (1998:300) indicates that the productive option died out between 1850 and 1900, leaving just the residual possibilities with *should*, *had* and, at that stage, *were*. It is natural to think of this inversion operation, like others in NE, as involving an irrealis C which attracts T (hence the fact that lexical verbs aren't attracted in NE, but can be in German, for example). Irrealis is, of course, a sub-type of non-veridical polarity, alongside interrogativity and negation, the categories which we noted in the previous section as apparently being of particular relevance in relation to NE V2 structures. In observing the loss of productive fronting in conditionals, then, we see that a grammatical operation that previously applied to a larger grammatically defined class (i.e. an option that we might think of as microparametric) lost its unity, becoming in part lexically (i.e. nanoparametrically) specified. As was the case with the instances of loss of V-to-T movement and loss of interrogative inversion discussed in the previous section, we see alternative lexicalization (here, by conditional *if*) playing a role in the loss of a formerly productive option, with different exponents of the originally unified class retaining varying degrees of productivity.

Finally, we can observe variation and recent change involving *be*. One emerging possibility, attested even in literary texts, is *do*-support with "main-verb" *be*:

(27) "Why don't you be happy as well?" she said.
(1921, D.H. Lawrence; Denison 1998:203)

The appearance of *be* as the perfect auxiliary with unaccusatives is attested until at least Jane Austen's time (early 19$^{\text{th}}$ century):[10]

(28) Jenny & James are walked to Charmouth this afternoon.
(1804, Jane Austen; Denison 1998:136)

A further striking change in the relatively recent history of English is the emergence of the passive progressive involving two occurrences of *be*. In Jane Austen, and throughout the 17th and 18th centuries, we find the older construction in which the progressive could have a passive interpretation, as in (29):

(29) Our Garden is putting in order
(1807, Jane Austen; Denison 1998:148)

10 See McFadden & Alexiadou (2010) for discussion of the status of BE+perfective structures throughout the history of English.

See Warner (1995), Denison (1998) and the references given there for more discussion of this striking change. Here we do not attempt an analysis, but merely note the existence of another relatively recent change, one which appears to affect a single lexical item, non-finite passive progressive *be*, the rise of which undermined the availability of passive interpretations for progressive participles.

The cases mentioned above illustrate that there are a number of clear cases of nanoparametric change in recent and contemporary English. At this point, it is appropriate to consider why we refer to this change as "nanoparametric", given that it involves specific lexical items rather than independently definable syntactic classes (can *parameters* reasonably be said to be involved?) and also given Kayne's well-known work on lexically based *micro*parametric syntax (is the reference to *nano-* justified?[11]). Firstly, we make the general observation that the variation we describe as *nanoparametric* fits exactly into the basic characterisation of parametric variation given in §2 as involving movement and lexicalisation properties of functional heads. Our intention in distinguishing micro- and nanoparameters in the way we do (see (2)), then, is precisely to highlight the fact that languages may exhibit one or more individual lexical items whose behaviour (movement, agreement, etc.) replicates that applying to *classes* of elements of varying size (micro or meso or macro) in other languages. To the extent that the same general, abstractly defined, feature-related operations can be shown to occur time and again across languages, affecting differing groups of elements, including single items, we believe there is a strong case for pursuing a parametric-hierarchy approach of the type proposed in Roberts (2012) and for asking why the observed recurring patterns are as they are. From a diachronic perspective, "relic syntax" of the sort associated with nano- and microparametric properties may, of course, give us a window on patterns that may have been more productive earlier in the language's history, many of the verbal raising-patterns discussed in this paper being a case in point. In the case of well-documented languages like English, this may not seem such an important consideration; in languages where documentation of earlier stages is sparse or non-existent, however, knowledge of the ways in which "larger" parametric options may be instantiated by one or more small classes of elements or even by individual elements may prove rather valuable.

11 It will be clear that our use of *nano-* is distinct from that employed within the framework of Nanosyntax (see http://nanosyntax.auf.net/output.html for representative papers).

6 Conclusion

At the beginning of this paper, we presented a four-way taxonomy of parameters; see (2), which we proceeded to discuss on the basis of the history of English and focusing in particular on changes in verbal-movement parameters of different kinds. To conclude, we would like to briefly consider wider issues raised by this case study for the way in which the proposed theory of parameters relates to syntactic change.

Firstly, given that options higher up the hierarchy involve "less knowledge" and also affect larger numbers of elements than those lower down, we might expect syntactic changes to proceed "upwards", all things being equal: in terms of Roberts & Roussou's (2003) approach to syntactic change, this typically involves feature loss, with acquirers striving to postulate the simplest featural analyses compatible with the PLD. The TMA-related developments discussed in relation to creoles arguably instantiate a case where this prediction is borne out: while NE auxiliaries clearly bear formal features allowing them to be spelled out in various inflected forms and also to undergo movement, creole TMA particles lack these properties; accordingly, they may lack formal features as proposed for analytic systems more generally, meaning that the creole TMA system is higher up the parametric hierarchy than the NE auxiliary system.[12] As mentioned in Note 9, the reconfiguration of the V2 component of earlier English may also constitute a simplification. In this case, however, it is worth noting that the creation of a new grammatically defined class of V2-triggering Cs does not obviously reflect a direct continuation of the earlier system, a problem we leave aside here. Unlike those just mentioned, the majority of the case studies we have considered in this paper, however, show that things very often are *not* equal: thus the mesoparametric V-to-T setting has, for example, given rise to a range of micro- and nanoparametric settings in NE. While microparametric settings do require the specification of more formal features than mesoparametric settings, it is worth noting that the same is not obviously true for all (or maybe even any) nanoparametric settings: here a given property is tied to one or more specific lexical items which represent(s) an outlier in relation to others of the same type (e.g. the case

12 Combinatorial considerations (e.g. sequencing of TMA markers, combinability with VPs and adverbs of different types, etc.) would then be regulated on a purely semantic basis. To the extent that the distribution of creole TMA markers is governed by considerations of purely semantic compositionality, the proposal that no formal features are involved is plausible. A question that we leave aside here is whether the complete absence of formal features in (a part of) a system necessarily leads to a simpler system: while fewer formal features suggest a simpler *syntactic* system, their complete absence may greatly complicate the overall system. See i.a. Mobbs (in progress) for discussion.

of *had* and *should* in relation to other modals in the context of conditionals). Where such items are frequent in the PLD, it seems plausible that the number of formal features encoded by the lexical item concerned is less relevant than its prominence in the PLD; hence the oft-observed irregularity-frequency effect—featural make-up simply isn't at issue, in the same way as it fails to be relevant in the context of Elsewhere scenarios more generally. Evidently, then, the basic upwards directionality-of-change prediction that falls out on the basis of independent consideration of a parametric hierarchy can readily be overridden at the "bottom" of hierarchies. Whether this is also possible in relation to non-nanoparametric properties and, if so, what considerations determine "downward" change in these cases is an open question.

Another prediction that the hierarchical approach to parametric variation outlined here would seem to make is that we might expect differences in the diachronic stability of the parametric options in (2). (30) represents one proposal as to the types of differences one might expect:

(30) a. Macroparameters are stable over millennia.
b. Mesoparameters are somewhat stable, often characterising genera.
c. Microparameters are somewhat unstable.
d. Nanoparameters are highly unstable.

Biberauer & Roberts (2012) discuss these predictions in detail, giving examples of macro-, meso-, micro- and nanoparametric properties that appear to support the idea that there is a parametric "size"-stability correlation. This paper has presented various examples of micro- and nano-change. The behaviour of the English modals, for example, neatly illustrates how microparameters tend to be relatively short-lived: the system of English modals may have existed as a genuine system only from when it emerged through the grammaticalisation of various lexical verbs in the mid-16th century until the end of the 18th century, a period of just over two centuries. In relation to nanoparameters, we have seen that the nanoparametric options in English verbal system behave in many respects like irregular verbs in that they have item-specific specifications which override the defaults of the system; as a result, we expect them to "regularize" diachronically or to disappear, unless they are sufficiently frequent. We see examples of this in the class of English modals: *dare* is almost entirely regularized as a main verb in nearly all varieties of contemporary English (leaving the formulaic expression *I dare say* as a relic of auxiliary syntax); while various former modals have disappeared, *shall* having all but disappeared in contemporary varieties, as already noted.

We contend, then, that our taxonomy of parameters, related to the hierarchies, can give us a good descriptive approach to many different kinds of change while at the same time retaining a robust and conceptually clear concept of what a parameter of UG is.

References

Ayafor, Miriam (2004). 'Cameroon Pidgin English: morphology and syntax', in B. Kortmann, E. Schneider, C. Upton, R. Mesthrie & K. Burridge (eds), *A Handbook of Varieties of English. Vol 2: Morphology and Syntax,* Berlin: Mouton de Gruyter, 909-28.

Baker, Mark (1996). *Polysynthesis Parameter.* New York: OUP.

Bernstein, Judy (2001).'The DP Hypothesis: identifying clausal properties in the nominal domain', in M. Baltin & C. Collins (eds), *The Handbook of Contemporary Syntactic Theory.* Malden, Mass.: Blackwell Publishers, 536-61.

Bhatt, Rakesh (2004). 'Indian English: syntax', in B. Kortmann & E. Schneider, C. Upton. K. Burridge & R. Mesthrie (eds), *A Handbook of Varieties of English.* Berlin: Mouton de Gruyter, 1016-30.

Biberauer, Theresa (2010). 'MAN is that weird!: Exclamative V2 in Modern English and beyond', paper presented at the LAGB Annual Meeting (Leeds).

Biberauer, Theresa (2011). 'In defence of lexico-centric parametric variation: two 3rd factor-constrained case studies'. Paper presented at the Workshop on Formal Grammar and Syntactic Variation: Rethinking Parameters (Madrid).

Biberauer, Theresa & Phil Branigan (2012). 'Microparametric expression of a macroparameter: Afrikaans verb clusters and Algonquian grammars'. Unpublished ms: University of Cambridge and Memorial University, Newfoundland.

Biberauer, Theresa, Anders Holmberg & Ian Roberts (2012). 'A syntactic universal and its consequences'. Unpublished Ms: Universities of Cambridge and Newcastle.

Biberauer, Theresa & Ian Roberts (2005). 'Changing EPP-parameters in the history of English: accounting for variation and change', *English Language and Linguistics* 9(1):5-46.

Biberauer, Theresa & Ian Roberts (2010). 'Subjects, Tense and Verb-Movement', in T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan, *Parametric variation: null subjects in minimalist theory.* Cambridge: CUP, 263-303.

Biberauer, Theresa & Ian Roberts (2012). 'The significance of what hasn't happened', paper presented at DiGS 14 (Lisbon).

Branigan, Phil (2012). 'Macroparameter learnability: an Algonquian case study'. Ms. Memorial University of Newfoundland.

Chomsky, Noam (1995). *The Minimalist Program.* Cambridge, MA: MIT Press.

Davis, Henry (1987). *The acquisition of the English auxiliary system and its relation to linguistic theory.* Ph.D. dissertation: University of British Columbia.

Denison, David (1998). Syntax, in Suzanne Romaine (ed), *The Cambridge Grammar of the English Language: Volume IV, 1776-1997*, Cambridge : CUP, 292-329.

Dresher, Elan (1999). 'Charting the learning path: cues to parameter setting', *Linguistic Inquiry* 30:27-68.

Durrleman-Tame, Stephanie (2008). *The Syntax of Jamaican Creole: A Cartographic Perspective.* Amsterdam: Benjamins.

Emonds, Joseph (1978). 'The verbal complex V-V' in French', *Linguistic Inquiry* 9:151-75.

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff (2000). *The Syntax of Early English.* Cambridge: CUP.

Gianollo, Chiara, Cristina Guardiano & Giuseppe Longobardi (2008). 'Three fundamental issues in parametric linguistics', in T. Biberauer (ed.). *The Limits of Syntactic Variation.* Amsterdam: Benjamins, 109-42.

Görlach, Manfred (1991). *Introduction to Early Modern English.* Cambridge: CUP.

Gray, Donald (1985). *The Oxford Book of Late Medieval Prose and Verse.* Oxford: OUP.

Grimshaw, Jane (1991). *Extended Projection.* Unpublished ms: Brandeis.

Grimshaw, Jane (2005). *Words and Structure.* Stanford: CSLI Publications.

Gupta, Anthea (1994). *The Step-Tongue: Children's English in Singapore.* Clevedon: Multilingual Matters.

Holmberg, Anders & Christer Platzack (1995). *The Role of Inflection in Scandinavian Syntax.* Oxford: OUP.

Huang, C-T. James (2007). 'The macro-history of Chinese syntax and the theory of language change', unpublished handout: University of Chicago.

Ihalainen, Ossi (1991). 'Periphrastic *do* in affirmative sentences in the dialect

of East Somerset', in P. Trudgill & Chambers (eds), *Dialects of English: Studies in Grammatical Variation*. London: Longman, 148-160.

Julien, Marit (2002). *Syntactic Heads and Word Formation*. New York: OUP.

Julien, Marit (2005). *Nominal Phrases from a Scandinavian Perspective.* Amsterdam: Benjamins.

Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, Richard (2005). 'Some notes on comparative syntax, with special reference to English and French', in G. Cinque & R. Kayne (eds), *The Oxford Handbook of Comparative Syntax*. Oxford: OUP, 3-69.

Kroch, Anthony (1989). 'Reflexes of Grammar in Patterns of Language Change', *Language Variation and Change* 1:199-244.

Lightfoot, David (1979). *Principles of Diachronic Syntax*. Cambridge: CUP.

Lightfoot, David (1991). *How to set parameters: arguments from language change.* Cambridge, MA: MIT Press.

Lightfoot, David (1997). 'Shifting Triggers and Diachronic Reanalyses', in A. van Kemenade & N. Vincent (eds), *Parameters of Morphosyntactic Change*, Cambridge: CUP, 253-72.

Longobardi, Giuseppe (1994). 'Reference and proper names: A theory of N movement in syntax and logical form'. *Linguistic Inquiry* 25(4):609-65.

Massam, Dianne (2005). 'Predicate fronting and lexical category in Niuean', in A. Carnie, H. Harley & S. Dooley-Colburg (eds), *Verb First: Studies in Predicate Initial Languages*. Amsterdam: Benjamins, 227-42.

McFadden, Thomas & Artemis Alexiadou (2010). 'Perfects, resultatives and auxiliaries in Earlier English', *Linguistic Inquiry* 41(3):389-425.

Mobbs, Iain (2008). ''Functionalism', the design of the language faculty, and (disharmonic) typology'. M.Phil. dissertation: University of Cambridge.

Mobbs, Iain (in progress). 'Minimalism and the Design of the Language Faculty'. Ph.D. dissertation: University of Cambridge.

Mufwene, Salikoko (2001). African American English. In J. Algeo (ed), *The Cambridge History of the English Language Volume VI: English in North America*. Cambridge: CUP, 291-324.

Palmer, Frank (1990). *Modality and the English Modals*. London: Longman.

Patrick, Peter (2004). 'Jamaican Creole: morphology and syntax', in B. Ko-

rtmann, E. Schneider, C. Upton, R. Mesthrie & K. Burridge (eds), *A Handbook of Varieties of English. Vol 2: Morphology and Syntax,* Berlin: Mouton de Gruyter, 407-38.

Pesetksy, David (1989). 'Language-particular processes and the Earliness Principle'. Ms: MIT.

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

Richards, Marc (2004) *Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax.* Ph.D. thesis: Department of Linguistics, University of Cambridge.

Rizzi, Luigi (1997). 'The fine structure of the left periphery', in L. Haegeman (ed), *Elements of Grammar: Handbook of Generative Syntax.* Dordrecht: Kluwer, 281-337.

Roberts, Ian (1985). 'Agreement parameters and the development of English modal auxiliaries'. *Natural Language and Linguistic Theory* 3:21-58.

Roberts, Ian (1993). *Verbs and diachronic syntax: a comparative history of English and French.* Dordrecht: Kluwer.

Roberts, Ian (1999). 'Verb movement and markedness', in Michel deGraff (ed.), *Language creation and change.* Cambridge, Mass.: MIT Press, 287-328.

Roberts, Ian (2007). *Diachronic Syntax.* Oxford: OUP.

Roberts, Ian (2010). *Clitics, Head Movement and Incorporation.* Cambridge, MA: MIT Press.

Roberts, Ian (2012) 'On the nature of syntactic parameters: a programme for research', in C. Galves, S. Cyrino, R. Lopez & J. Avelar (eds), *Parameter Theory and Linguistic Change.* Oxford: OUP, 319-334.

Roberts, Ian & Anna Roussou (2003). *Syntactic Change: a Minimalist Approach to Grammaticalization.* Cambridge: CUP.

Rowland, Caroline & Anna Theakston (2009). The acquisition of auxiliary syntax: a longitudinal elicitation study. Part 2: the modals and DO. *Journal of Speech Language and Hearing Research*, 52(6):1471-92.

Sheehan, Michelle (to appear) 'Some implications of a copy theory of labeling'. *Syntax.*

Stromswold, Karen (1990). *Learnability and the acquisition of auxiliaries.* Ph.D. dissertation: MIT.

Theakston, Anna & Caroline Rowland (2009). The acquisition of auxiliary syntax: a longitudinal elicitation study. Part 1: Auxiliary BE. *Journal of Speech Language and Hearing Research* 52(6):1149-70.

Warner, Anthony (1995). 'Predicting the progressive passive: parametric change within a lexicalist framework', *Language* 73:533-57.

Warner, Anthony (1997). 'The structure of parametric change, and V movement in the history of English', in A. van Kemenade & N. Vincent (eds), *Parameters of Morphosyntactic Change*, Cambridge: CUP, 380-93.

Wee, Lionel (2008). 'Singapore English: morphology and syntax', in R. Mesthrie (ed), *Varieties of English 4. Africa, South and Southeast Asia*. Berlin: Mouton de Gruyter, 593-609.

Wexler, Ken (1998). 'Very Early Parameter Setting and the Unique Checking Constraint: a new explanation of the Optional Infinitive stage', *Lingua* 106:23-79.

Zeijlstra, Hedde (2008). 'On the syntactic flexibility of formal features', in T. Biberauer (ed), *The limits of syntactic variation*. Amsterdam: Benjamins, 143-73.

Ziegeler, Debra (2011). 'Experiential aspect in Singapore English: the depolarisation of *ever*', *Groninger Arbeiten zur germanistischen Linguistik*53(2):139-162.

Theresa Biberauer
Department of Theoretical &
Applied Linguistics
Faculty of Modern and Medieval Languages
Sidgwick Avenue
Cambridge, CB3 9DA

mtb23@cam.ac.uk

Ian G. Roberts
Department of Theoretical &
Applied Linguistics
Faculty of Modern and Medieval Languages
Sidgwick Avenue
Cambridge, CB3 9DA

igr20@cam.ac.uk