The Return of the Subset Principle: A Diachronic Perspective

Theresa Biberauer and Ian Roberts

Abstract

This paper reconsiders the role of the Subset Principle (Berwick 1985) in the context of lan-
guage acquisition and change. "Traditional" parameters (e.g. the null subject parameter) de-
tions of the kind entailed by the Subset Principle are to be defined. Here, we argue that, once
come possible to define grammars which generate languages that are in inclusion, rather than
chronic changes, the key idea being that the absence of sufficiently robust evidence for the
erasing a smaller language. We develop this idea in relation to two types of change attested in
obligatory standing, and the other what we call "restriction of function" in which an opera-
tion becomes restricted to a subset of the contexts in which it formerly applied.

1. Introduction and general preliminaries

In the context of generative theory, it has often been argued that aspects of language acquisi-
tion and/or constraints on learnability are a cause of change (see in particular Lightfoot 1979,
reanalysis and associated parametric changes. In this context, it is desirable to attempt to re-
gen programme for understanding language change in terms of acquisition and learnability
be advanced. The goal of this paper is to make a contribution in this direction.

The particular principle that we will be concerned with here is the Subset Principle
(henceforth: SP). Originally put forward by Berwick (1985), we can give the following in-
formal statement of this principle (cf. also Marzini and Wexler 1987):

(1) the learner must guess the smallest possible language compatible with the in-
put at each stage of the learning procedure (Clark and Roberts 1993: 304-5)

The conceptual interest of the SP lies in the fact that it arguably derives from a widely
recognised fact about language acquisition: that children do not seem to make use of (direct)
negative evidence. In other words, information that certain parts of the input text (sentences or
strings), are ill-formed is not available, or at least not made use of by acquirers. It follows
from this that acquirers run the risk of falling into "superset traps": if a grammar which gen-
does not contain this hypothesised system. Hence acquirers must always posit the grammar which
generating the "smaller language" consistent with the trigger experience; in this way positive
sense that evidence of this type may be sufficient to cause the acquirer to revise hypotheses,
that syntactic change is driven by acquisition, the SP may be relevant in accounting for
change (this point was also made by Clark and Roberts 1993: 305-307).

However, as has frequently been noted (cf. i.a. MacLaughlin 1995, LaFond 2001),
there is a problem with the SP as formulated along the lines in (1). This is that many, perhaps
all, of the parameters that have been independently proposed in the literature on comparative
syntact seem to define intersection relations rather than inclusion ones. Consider, for exam-
ple, the parameter which determines Object-Verb (OV) as opposed to Verb-Object (VO) order
in transitive VPs, which, for convenience and following tradition, we can call the Head Par-
parameter. Set to OV, it generates the grammatical strings in (2a,b) and not the ungrammatical
one in (2c), and set to VO it generates the strings in (3b,c) and not the ungrammatical one in
(3a):

(2) a. John Sue loves.
   b. John walks.
   c. *John loves Sue.

(3) a. *John Sue loves.
   b. John walks.
   c. John loves Sue.

Clearly, the OV grammar and the VO grammar are in an intersection relation, as
shown in (4):

(4) John Sue loves
    John walks
    John loves Sue

Even the classical Null Subject Parameter of Rizzi (1982), which Berwick (1985) took
as an example to illustrate the SP, defines an intersection relation if one takes into considera-
tion the fact that null-subject languages of the Italian type lack overt expletives while non-
null-subject languages lack null pronouns:

(5) a. He speaks.
   b. *Speaks.
   c. It seems that John speaks.

(6) a. Lui parla.
   b. Parla.
   c. *Ciò sembra che Gianni parla

Again, we have an intersection relation, as (7) shows:

(7) Null subjects
    Lexical subjects
    Expletive subjects

G1 (NULL)

G2 (non-Null)
For this reason, the SP has been thought to be of limited value. In fact, if all parameters turned out to define intersecting grammars, which appears likely on the basis of the above discussion, the SP would be of no real value at all. The point of this paper is to argue that this is not the case, once the role of true formal optionality (i.e. syntactic “free variation”) is fully taken into consideration. We will show that there are four parameter settings which give rise to grammars which generate languages which are in exclusion, rather than intersection, relations, and therefore the SP is relevant both for language acquisition and in language change.

Our central idea will be that, where the evidence for the grammar which generates the larger language is not sufficiently robust, acquirers “default” to a grammar generating a more evidence in language acquisition. So here we see a case where the conditions of language acquisition are more specifically, we will show that the SP is relevant to two types of change. First, where a system which at one stage allows a pied-piping option alongside a “standing” option at a later stage allows only the standing option (we will give a precise characterization of these notions of pied-piping and standing in section 2). Second, the SP is relevant in the case of formal support in the context in which it is applied.

The paper is structured as follows: section 2 presents the theoretical background underlying our assumptions about pied-piping, word-order variation and “true” optionality. Sections 3 and 4 then focus on changes that took place in the history of English that, we believe, illustrate the operation of the SP in the diachronic context. Firstly, section 3 considers the loss of pied-piping options that previously co-existed with “standing” structures, thereby bringing about word-order change in earlier English. Section 4 considers the two instances of the second case, restriction of function: firstly, the restriction of object movement during Middle English and secondly, the restriction of V-to-middle movement that led to the rise of the Modern English system of do-support. Section 5 concludes the paper.

2. Theoretical preliminaries

We assume the Probe/Goal/Agree system of Chomsky (2000 et seq). In terms of this system, heads are syntactically active as long as they bear one or more unvalued feature(s). To see how it works, consider (8) below:

(8) a. \( \ldots X_{\text{PROBE}} \ldots \text{YP} \ldots \text{Z}_{\text{GOAL}} \ldots \ldots \)

b. \[ \text{X} \]

Here, a head \( X \) bearing an unvalued feature (or features; for expository convenience, of a valued counterpart of its unvalued feature) acts as a probe in search of its c-command domain and which, in turn, bears an unvalued feature of its own, which renders it active and therefore visible to the probe. Agree takes place between the probe and goal, with the result that the unvalued features on both heads are valued. Feature valuation, which corresponds to the feature checking of earlier versions of the theory (cf. Chomsky 1995), is therefore achieved without movement. For movement to take place, a probe must bear an additional movement-triggering feature, conventionally referred to as an EPP-feature (although see Chomsky 2005 et seq. on so-called Edge Features). Following Richards and Biberauer (2005) and Biberauer and Richards (2006), we assume that Universal Grammar (UG) offers various parametric options in respect of the “size” of the constituent that undergoes movement wherever X probes Z and also bears an EPP-feature (i.e. movement diacritic). These are given in (9):

(9) a. Z-movement only: \[ \text{XP} \rightarrow \text{Z} \rightarrow \text{Y} \]

b. ZP-movement only: \[ \text{XP} \rightarrow \text{ZP} \rightarrow \text{Y} \]

c. Obligatory pied-piping: \[ \text{XP} \rightarrow \text{ZP} \rightarrow \text{Y} \rightarrow \text{Z} \]

d. Optional pied-piping: \[ \text{XP} \rightarrow \text{Y} \rightarrow \text{Z} \]

In the case where the probe \( X = T \), the goal \( Z = A \)-feature (or, more probably, the phi-features associated with an active DP) and \( Y = v \), we therefore expect the following range of movement possibilities:

(10) a. D-movement only: \[ \text{XP} \rightarrow \text{D} \rightarrow \text{Y} \rightarrow \text{Z} \]

b. DP-movement only: \[ \text{XP} \rightarrow \text{DP} \rightarrow \text{Y} \rightarrow \text{Z} \]

c. Obligatory pied-piping: \[ \text{XP} \rightarrow \text{DP} \rightarrow \text{Y} \rightarrow \text{Z} \]

d. Optional pied-piping: \[ \text{XP} \rightarrow \text{DP} \rightarrow \text{Y} \rightarrow \text{Z} \]

(10a) involves adjunction of the D-head of the DP-subject to the probing T-head, a movement operation that has been argued to take place in various VSO languages (cf. i.e. Baker and Hale 1990, Chomsky 1992 and Travis 2006 for discussion). As this mode of EPP-satisfaction does not play a role in the specific contexts with which we will be concerned, we will leave it aside here. (10b – d) are, however, directly relevant to our concerns in this paper.

(10b) instantiates the mode of EPP-satisfaction standardly assumed for D-seeking T in Modern English (and many other languages), i.e. subject DP-raising to [Spec,TP]. (10c) represents a mode of EPP-satisfaction that Richards and Biberauer (2005) and Biberauer and Richards (2006) argue to differ minimally from those more standardly assumed: instead of a D-head (which could also be associated with finite V; cf. Alexiadou and Anagnostopoulou 1998) or DP undergoing raising to [Spec,TP], it is also possible that either of these D-bearing categories may constitute the goal of Agree, while a larger category properly containing this goal is targeted for movement. In other words, the authors exploit the distinction between two heads which may have in the context of Chomsky’s Probe/Goal/Agree system, namely that of being a probe on the one hand and that of being associated with one or more EPP-features on the other: in terms of this system, it does not follow that the category that a given head probes has to correspond to the category that ultimately undergoes movement under the influence of the probe’s EPP-feature; as long as the moving category contains the goal, the computational system will “not mind” (cf. also Baker and Roberts 2005, 2006a)

We will present our discussion in terms of single unvalued features (or features; for expository convenience, of a valued counterpart of its unvalued feature) acts as a probe in search of its c-command domain and which, in turn, bears an unvalued feature of its own, which renders it active and therefore visible to the probe. Agree takes place between the probe and goal,
erauer (2005) and Biberauer and Richards (2006) propose that inflectionally rich German represents a language in which T agrees with D-on-V (cf. Alexiadou and Anagnostopoulou's (1998) proposal for Greek) and then moves the entire vP to [Spec, TP]. Inflexionally more impoverished Dutch and Afrikaans, by contrast, are argued to be languages which target the subject DP, but then move the whole containing vP (cf. also Biberauer 2003; Joutteau 2005 proposes that Breton is another language that employs vP-raising as a means of satisfying T’s EPP-requirements, and cf. also Chomsky 2001: 38).

The important point about the pied-piping modes of EPP-satisfaction in the present context is that they, under currently rather poorly understood circumstances, appear to permit formal optionality which is not available to languages employing non-pied-piping modes of EPP-satisfaction (i.e. D- or DP-raising). In particular, there appear to be synchronically attested languages that allow EPP-satisfaction via vP-raising (i.e. pied-piping) to alternate with satisfaction via DP-raising (i.e. stranding), modern spoken Afrikaans and various dialects of German and Dutch being cases in point (cf. Biberauer 2003, Biberauer and Roberts 2005).

The crucial consideration for our purposes is that a grammar which permits this optionality would correspond to (10d) above, i.e. to a grammar which generates both strings generated by (10b) and those generated by (10c). (10d) therefore represents a superset language in relation to both (10b) and (10c). As such, we assume that it must be very robustly favoured by the SP (cf. discussion in section 1). Biberauer and Roberts (2005) therefore propose that earlier stages of English permitted pied-piping modes of EPP-satisfaction that are no longer available in Modern English (NE). According to B&R, various non-syntactic factors interacted with existing syntactic properties to bring about the loss of these options in fact be understood as following from the operation of the SP, understood as a guiding principle in language acquisition.

3. Case Study I: word-order change in Middle English

As noted above, B&R propose that Old and Middle English (henceforth: OE and ME) represent stages of English that feature pied-piping modes of EPP-satisfaction in domains in which this is no longer synchronically possible. More specifically, B&R propose that OE differed only, as in NE, act as probes for D-features (those associated with the subject and the direct object respectively), but they also brought about movement of categories containing these D-bearing goals.

Let us firstly consider OE. At this stage, both v and T were associated with EPP-features which could be satisfied either by movement of the goal alone (i.e. DP-movement) or movement of the category immediately dominating the goal-VP (i.e. VP- or vP-movement). In respect of the satisfaction of the EPP-features on v and T, therefore, OE was an optional pied-piping language of the kind schematised in (9d) above, i.e. a superset language in which movement was not relative to one permitting only one or none of these modes of EPP-satisfaction.

In (13a), the goal probed by v is the direct object has bac. In moving to [Spec, VP], under the influence of v’s EPP-feature, this goal pied-pipes the additional material contained in VP (i.e. to Englishe sparace (V), with the result that the lexical verb ultimately surfaces as string-finally. In (13c), by contrast, the same goal (has bac) moves independently of the rest.

In (13a), a grammatical analysis proposed in B&R, in terms of which V consistently raises to v (cf. section 4.2 for further discussion), it is clear that orders like (13c) must involve movement not only of the probed object-VP, but also of the particle; B&R propose that these elements raise together as another case of VP pied-piping (cf. Elfenbaas 2007 for further discussion). Particle-verb structures, alongside rigidly verb-final structures such as that illustrated in (13a), therefore constitute a key piece of data indicating the availability of pied-piping as a means of satisfying v’s EPP-feature. (13d), in turn, represents further evidence (alongside (13c)-type structures) that VP-standing is also an option: here the reflexively used personal pronoun that constitutes the goal probed by v, with only this element undergoing raising under the influence of v’s EPP-feature and the PP being stranded when the object undergoes movement.
The T-related modes of EPP-satisfaction available in OE are illustrated in (14):

(14) Satisfaction of T's D(subject)-oriented EPP-feature:

a. VP pied-piping:

just as Gregorium to papan geocran [hayde (V)]

"...that the people Gregory to pope elected had

(ALT 1987, 10/4; van Kemenade 1987: 34)

b. VP stranding:

just as [he] made (a) generian

that he could

"so that he could save his property"

(Oxford, 48, 18; van Kemenade 1987: 59)

(14a) illustrates the structure that results when the subject-goal of the T-probe pied-pipes the VP within which it is contained. Note that this typically West Germanic SOV/Ad ordering also requires pied-piping satisfaction of v's EPP-feature, i.e. although Gregorium is the goal to the merger of the subject-DP. In (14b), only the subject-goal, he, undergoes movement to

his feorh (depending on one's analysis of the structure; cf. Biberauer and Roberts (2006a) for further discussion) surfaces in post-T position.

B&R propose that VP pied-piping as in (11a) and (13a,b) was lost as an option along-side VP-stranding (i.e. sole movement of the goal) in early ME. Biberauer and Roberts (2006b) attribute this to a number of factors, including:

1. The relative low incidence following the Norman invasion of French borrowings which i.e. replaced particle verbs and thereby removed from the input O-Part-of the types illustrated in (13b) (cf. Spasov 1966, cited in Kisch and Taylor 2000: 146, who notes that particle verbs are vanishingly rare during the 12th and 13th centuries; see also Fischer 1992: 368 and the discussion in Biberauer and Roberts 2006a)

2. The relatively low incidence of compound tense-containing structures in OE and early ME (cf. Traugott 1972). This consideration is relevant because the majority of OE and early ME main clauses were V2 structures in which the finite verb surfaced in clause-second position (cf. Fischer et al. 2000: 118f for discussion of structures featuring personal pronouns). As such, these clauses, which are usually assumed to play a crucial role in the acquisition context (cf. i.e. Lightfoot's 1991 'degree zero learnability' proposal), would have been ambiguous in respect of the nature of the operation that has taken place to satisfy v's EPP-feature; once the lexical verb has raised to clause-second position, it can no longer serve as 'signpost' to the size of the category that has undergone raising to [Spec, V] = c. (15) where the XP-VP-Subj-Obj-PP output string does not allow one to detect whether the string has in fact been derived via VP pied-piping as in (15a) or VP-stranding as in (15b):

(15) a. VP pied-piping in an auxiliary-less V2 clause:

[c xp v+v+T+c [v subj (v+v+T) [u subj v obj pp (V) (v+v)]]](VP)

b. VP-stranding in an auxiliary-less V2 clause:

[cp xj v+v+T+c [v subj (v+v+T) [u subj v obj pp (V) (v+v)]]](VP)

3. the loss of dative case (cf. Allen 1995: 441, Table 10.1) and the concomitant rise in indirect-object-PPs. This resulted in a rise in the number of argumental (as opposed to ad-junct) PPs surfacing in "lacking" configurations of the kind illustrated in (13a) above at the expense of the TO-D0-V orders which had formerly triggered the VP pied-piping option in (11a). Thus structures like that in (16) depicting a dative-marked indirect object were replaced by structures like (17) in which the indirect object is realised as a PP:

(16) if the priest cannot say a homily to the lady people (Ecclesiastes's Homilies II, 41.306-66; Koopman and van der Wurff 2000: 259)

(17) He shew his brother Amon that such dishonesty and untruth had done to his sister (Caxton Knight of the Tower 87.15; Fischer et al. 2000: 169)

Biberauer and Roberts (2006b) suggest that the lexical and morphological considerations in (i–iii) conspired to affect the PLD to which early ME children were exposed in such a way that the VP pied-piping operation in (11a) was insufficiently robustly triggered. In other words, they propose that independent and contingent non-syntactic factors were behind the syntactic change that took place in early ME. Since the SP inherently favours a grammar-generating fewer strings (cf. Note 1), the drop in sufficiently unambiguous triggering evidence for VP pied-piping led to the loss of (11a) as a means of satisfying v's EPP-feature. In respect of this parametric setting, early ME therefore became a subset language of the (9)-type relative to OE which was a (9)-type language.

Let us now consider the changes that affected T during ME. Following the SP's mediated reanalysis of the ways in which v's EPP-feature might be satisfied (i.e. the loss of VP-internal. In the context of the phase-based Probe/Coaf/Agre system of Chomsky (2000 et seq.) and, in particular, the strict version of the Phase Impenetrability Condition (PIC), material located within the VP-domain is subject to Spellout upon completion of the VP phrase. Thus an object which has not undergone raising to [Spec, V] is sent to Spellout.
upon completion of the vP phase and will be spelled out in its VP-internal position (along
with any other VP-material). In the context of the ME structures that we are considering here,
this proposal has the specific consequence that unmoved objects will no longer be part of the
vP which is available for raising to [Spec,TP]. In other words, the loss of generalised object
movement in later ME had the consequence that structures in which T’s EPP-requirements
were satisfied by vP-raising became harder to distinguish from those in which DP-raising sat-
sisfied these requirements. In a variety of structures, including matrix and embedded clauses
lacking an auxiliary, it would in fact have become impossible to determine which raising op-
eration had taken place as the output string in both cases will be S-V-O. Consider the embed-
ded clause case illustrated in (18) (material sent to SpellOut indicated in outline font):

\[
\begin{align*}
(18) & \quad [\text{w} \{ \text{Subj} \} \quad T \quad \phi \{ \text{Subj} \} \quad V^{++} \} \quad \text{vP} \{ V \} \quad \text{Obj}] \quad \rightarrow \quad \text{DP-raising} \\
& \quad [\phi \{ \text{Subj} \} \quad V^{++} \} \quad T \quad \phi \{ \text{Subj} \} \quad V^{++} \} \quad \text{vP} \{ V \} \quad \text{Obj}] \quad \rightarrow \quad \text{vP-raising}
\end{align*}
\]

In (a), DP-raising results in an S-V-O string in which the finite verb is located in v within the
unraised vP, while vP-raising in (b) produces a surface-identical string in which the finite verb
is once again located in v, but vP is located within the TP-domain. V2 matrix clauses likewise
constitute ambiguous input for reasons similar to those outlined for (15) above. Clearly, then,
the changes affecting the manner in which v’s EPP-feature could be satisfied had a knock-on
effect in the TP-domain. This can be schematised as follows:

\[
\begin{align*}
(19) & \quad \text{a. Changes in the satisfaction of v’s EPP-feature} \\
& \quad \text{Reanalysis I:} \\
& \quad \text{Early ME – loss of the VP pied-piping mode of satisfaction} \\
& \quad \phi \{ \text{w} \} \{ V \ \text{O} \} \quad V^{++} \quad \{ \text{vP} \} \quad \rightarrow \quad \phi \{ \text{w} \} \{ V \ \text{O} \} \\
& \quad \text{Later ME – restriction of the presence of EPP on v} \\
& \quad \phi \{ \text{w} \} \{ V^{+} \quad \{ \text{vP} \} \quad \{ \text{vP} \} \quad \rightarrow \quad \phi \{ \text{w} \} \{ V^{+} \quad \{ \text{vP} \} \quad \{ \text{vP} \} \} \\
& \quad \text{Thus, for non-Neg O:} \quad \phi \{ V^{+} \quad \{ \text{vP} \} \quad \{ \text{vP} \} \} \\
& \quad \text{b. Changes in the satisfaction of T’s EPP-feature} \\
& \quad \text{Reanalysis II:} \\
& \quad \text{Late ME, around 1450 – loss of the vP pied-piping mode of satisfaction} \\
& \quad \phi \{ \text{w} \} \{ S \quad V^{+} \} \quad T \quad \{ \text{w} \} \{ \text{vP} \} \{ \text{vP} \} \quad \{ \text{vP} \} \{ \text{vP} \} \quad \rightarrow \quad \phi \{ \text{w} \} \{ S \quad V^{+} \} \quad T \quad \{ \text{w} \} \{ \text{vP} \} \{ \text{vP} \} \}
\end{align*}
\]

Reanalysis II effectively created a “canonical subject position” [in SpecTP] (cf. Biberauer
2006), with consequences for expletive distribution, raising-to-subject in passive and unaccu-
sative contexts and so-called “Stylistic Fronting” structures (cf. Biberauer and Biberauer
2006a,b for further discussion). In the context of our present discussion, the point worth
noting is that the T-related reanalysis, like its v-related counterpart, involved the loss of a
pied-piping option that had, for independent reasons, become difficult to discern on the basis
of the PLD. In both cases, therefore, the operation of the SP resulted in a grammar permitting
both pied-piping and stranding being replaced by an innovative one generating a smaller lan-

4. Case Study II: "Restriction of Function"

By “restriction of function” we mean the case where a change operates in such a way as to
limit the set of contexts in which a movement operation applies. More specifically, restriction
of function denotes those cases where, at an earlier stage, a movement operation affects a rela-
tively large class E of elements and, at a later stage, only a subset D ⊆ E. The SP is relevant to
this kind of change at the stage where the movement process applied optionally to E – D (the
part of E that is not in D). A grammar where a movement operation applies optionally in E –
D is a superstep both of the one requiring movement in the larger set of contexts E and of the
requiring it only in the smaller set of environments D. Again, optionality gives rise to the su-
perstep grammar.²

We consider two cases here: first, object movement in ME (which follows on directly
from the discussion of word-order change in the previous section) and the more complex case
of the development of auxiliary do in Early Modern English (ENE).

4.1. Object Movement in Middle English

Recall that in Section 3 we stated that the following reanalysis took place in Early ME:

\[
\begin{align*}
(20) & \quad \text{Reanalysis I} \\
& \quad \phi \{ \text{w} \} \{ V \ \text{O} \} \quad V^{++} \quad \{ \text{vP} \} \quad \rightarrow \quad \phi \{ \text{w} \} \{ V^{++} \quad \{ \text{vP} \} \quad \{ \text{vP} \} \}
\end{align*}
\]

This change affected different types of objects differently. In fact, according to Kroch and
objects of different types underwent movement under slightly different conditions in OE too;
specifically, it is proposed that three distinct operations targeting quantified, negative and
non-quantified, non-quantified objects respectively were responsible for the OV orderings in
OE. B&R propose that there are in fact only two distinct movement triggers at issue here: an
classified function which resulted in a moved object being interpreted as defocused (i.e. as
an optional EPP-feature which resulted in a moved object being interpreted as defocused
in the manner of Germanic scrambling) and an obligatory EPP-

old information, in the manner familiar from Germanic scrambling) and an obligatory EPP-

A specific interpretation of the quantified expression was required (cf. i.a. Diesing 1990).
Assuming this to be correct, the idea is that (20) represents a change after which the various
types of objects surfaced prevalently independently of each VP content, whereas they previously
did so alongside the material in VP. After this change, movement of negative objects

remain typically ordered, as we shall see below. Optional movement of the other object types also
remained obligatory, as we shall see (Kroch and Taylor 2000). In

In 1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In

1400 (van der Wurff 1997, 1999 and Foster and van der Wurff 1997). In
negative, non-quantified direct objects was lost, while at least negatively quantified DPs continued to move.

Here are some examples of pre-1400 object shift of non-negative, non-quantified direct objects:

(21) a. 
> hast e rule he forsaken
that I will-not you forsake
(St. Juliana (Bod) 278; Koopman and van der Wurff 2000: 269)

b. 
ðat we mote then swa his sorte liif her laden
that we moy thus this short life here lead
(Fices and Virtues 21.23; Koopman and van der Wurff 2000: 269)

c. 
ðat he no schel pene stude
neaster mare changing bute for need ane
that she NEG shall the abode never more change but for need alone
(CMANCRW, I.46.52; Kroch and Taylor 2000: 148)

Examples of this type are only found with defocusing after 1400 (see B&R), and are not found at all after approximately 1500. As mentioned above, in later ME and ENE, i.e. in the 1999; Kroch and Taylor 2000; Moerenhout and van der Wurff 2000; Pintzuk 2002; Ingham 2000:

(22) a. 
I may no rest have a-mongys ow
I may no rest have among you
(MKempe A 122, 19-20; Roberts 1997: 425)

b. 
Þei shuld no meythir have
they should no mayor have
(Capgrave Chronicles 62.23; Koopman and van der Wurff 2000: 271)

As Moerenhout and van der Wurff (2000: 527) comment “... the pattern with an auxiliary is predominant” (cf. also Note 8). See Biberauer and Roberts (2006a) for a speculation as to why object shift occurred in precisely these contexts at this period.

We can schematise the changes affecting object shift during ME as follows:

(23) Pre-1400 ME:
a. 
... O V ...
b. 
... V O ...

(24) Post-1400 ME:
a. 
... O [O] V ...
b. 
... V O ...

In (23), “O” is intended to designate objects in general, although this glosses over the fact that different types of objects are in fact subject to distinct movement triggers, with movement of non-negative objects being optional, i.e. the consequence of an optional EPP-feature, as just pointed out, while negative objects moved obligatorily. In the later grammar (24), however, only negatively-quantified objects could move. The instance of restriction of function in ques-

The Return of the Subset Principle

(25) EPP-D  $\rightarrow$ EPP-D[+neg]

It is clear that the set of negative D-elements is a subset of the set of D-elements; hence we have a paradigm case of restriction of function, as characterised at the beginning of this section. In other words, we can take the set of direct objects to correspond to E in the above discussion, and the set of negative direct objects to correspond to the subset D C E. Now, the earlier grammar, that of (23), was one where non-negative direct objects underwent optional movement. Thus we have an optional rule applying to E D; as pointed out above, a grammar movement of this type is a subset both of the one requiring movement in the larger set of contexts E and of that requiring it only in the smaller set of environments D. Again, optionality therefore gives rise to the superset grammar. So we see that the grammar in (24) would, in the absence of robust input data signalling the contrary, be favoured by the SP.

In the ME case, optional movement of non-negative objects for defocusing was disfavoured since it was hard to detect in certain rather common types of constructions. For example, in a V2 clause where the verb is in a simple tense (i.e. no auxiliary is present, with the result that the verb is second), it is impossible to tell what structural position the object occupies, since, moved or not, it is final in the string. Thus, a sentence like (26a) could have either the structure in (26b) or that in (26c).

(26) a. 
In pus many maners touches þe ymage of dremes men
In so many ways touches the image of dreams men
(CMROLLEP, 93.499; Trips 2002: 251)

b. 
 [...] XP V+v+T+C [TP S (V+V) [¥ (V) (O)]]

c. 
 [...] XP V+V+T+C [TP S (V+V) [¥ (V) (O)]]

Verb Projection Raising is another context where the surface order S-Aux V-O is structurally ambiguous, as (27) shows (cf. also the discussion in Note 8):

(27) a. 
pat e mahen ane pine me here
that you may alone torture me here
(St. Juliana; Fischer et al. 2000: 161)

b. 
VR [TP PRO V+V+T [¥ (V) (O)]]

c. 
VR [TP PRO V+V+T [¥ (V) (O)]]

The ambiguity of cases like (26) and (27) meant that the cue for the optional EPP-feature was often obscured. We take it that it was therefore insufficiently robust to support the postulation of the superset grammar, since this is inherently disfavoured by the SP.

4.2. Do-support

The modern do-support system arose in the 17th century from the 16th century system in which do was a freely inserted auxiliary. Auxiliary do itself arose from the reanalysis of an earlier
verb which had an ECM, causative and raising interpretation with an infinitival complement (see Denison 1985, 1993; Roberts 1993): quite possibly this reanalysis is exactly the one 1983, 1993, 1997; Roberts 1985, 1993). In the second half of the 16th century, do was an aux-
positive declarative clauses without needing to bear emphatic stress. This use of do was re-
ferred to by Jespersen (1909-1949) as “exuberant do”. The following quotation from Palsgrave’s (1530) manual of French grammar for English speakers illustrates this aspect of the distribution of do at around this time:

(28) I do is a verbe moche commonly used in our tounge to be before other verbes, as, it is all
one to say [Visser’s emphasis] I do speake and such like, and I speake ...
(Palsgrave (1530) Esclaireissement de la langue francoyse 84, 380, 523; in Visser 1962-73: §1419; Roberts 1993: 294)

Also, examples of the following kind show that do in positive declarative clauses did not have to be stressed, as here the line of verse does not scan correctly if do is stressed, as the accents are intended to indicate:

(29) Rough winds do shake the dairling birds of May

(Shakespeare, Sonnet 18)

(30) [cp DP [c V] [r τ [r tv] [s τ [s tv] VP ]]  
    → [r τ DP [r V] [s τ [s tv] VP ]]  

(See also Adams 1987 and Roberts 1993 on Old French, and Willis 1998 on Middle Welsh.)

V undergoes “French-style” raising to T, and possibly beyond (see Roberts 1985; Polлок 1989):

The Return of the Subset Principle

(31) a. My wyfe rose not (Roberts 1985: 23)
   b. Se ye not how his herte is enduret...?
   (Anon. The Examination of Master William Thorpe, 44; Roberts 1993: 239)
   c. Looks it not like the king?
   d. And gife he be nought so, then...  
      (Hamlet I, I, 43; Roberts 1993: 293)
   and if he be not so, then...
      (1420s: James I, King’s Quair, 62; Gray 1985: 73; Roberts 1993: 323)

(32) Romance:
   French: parle (present indicative/subjunctive), parlera (future), parlerais (conditional), parlais (imperfect), parla (perfect), parlos (past subjunctive);
   Italian: parlo (present), parlerò (future), parlerai (conditional), parlavo (imperfect), parli (present subjunctive), parlasti (past subjunctive)

(33) Germanic:
   German: spreche (present indicative/subjunctive), sprach (past), spräche (past subjunctive),
   English: speak (present), spoke (past)

Biberauer and Roberts thus postulate that there are two quite distinct types of “richness” of verbal inflection agreement inflection and tense inflection. “Rich” agreement inflection has many of the properties which are standardly attributed to it: it triggers movement of a D-bearing category (perhaps V in a null-subject language, but crucially not in a non-null-subject language); and it licenses null subjects. “Poor” agreement inflection, on the other hand, determines the presence of overt expletives, perhaps for the reasons given in Richards and Biberauer (2005). On the other hand, “rich” tense inflection triggers V-movement and is irrelevant to subject-licensing. Biberauer and Roberts’s proposals entail a new typology of the relation between tense and agreement marking and various syntactic properties. The typology is given in (33):

(33) a. Rich agreement and rich tense inflection: hence V-to-T and null subjects, e.g. Italian, Greek, Spanish, etc.
b. Poor agreement but rich tense: hence V-to-T, but no null subjects, e.g. French (and see below on NE).
c. Poor tense and poor agreement: hence no V-to-T and no null subjects, e.g. Modern English, Mainland Scandinavian.
d. Rich agreement and poor tense: null subjects, but no V-to-T; no clear example.

With this background, we can look again at the status of verb movement to T in the history of English. Until the 15th century, when V2 creoded significantly (see Fischer et al 2000: 128-129) on this, English was essentially like West Germanic in having V2, but no independent verb-movement to T (and, of course, no null subjects). With the loss of V2 and the reanalysis of subject-initial clauses shown in (29), English developed verb movement to T. However, there is a sense in which (late) 15th-century English had or developed a rich tense inflection; tense-marking at this period was more or less what it is in NE. Given the analysis of the relation between "rich" tense-marking and verb movement to T in Biberauer and Roberts (where further technical details and motivation are provided), a Romance-style V/T-Agree system could not be supported by the feeble tense-inflection of English. B&R suggest that this contributed to the reanalysis of modals and do as auxiliaries in early 16th century.

As already mentioned, it is well-known that "Pollockian" verb-movement to T was lost around 1600 (ca. 1575 according to Kroch 1989, but see Warner 1997). In terms of the clause structure in Chomsky (1995, 2001), this must be treated as the loss of V-to-T movement. We propose that, after the loss of V-to-T movement, V-to-T movement remained (i.e. T retained a V-oriented EPP feature). In other words what was lost was V-to-v movement. Two innovations in the English of this period, both of which also hold in PDE, support this idea. The first is that (non-heavy) direct objects and finite main verbs are required to be rigidly adjacent from this period on. This observation has also been made by Zwart (2005) (who interprets it rather differently). Orders of the following kind, where material intervenes between the finite main verb and the direct object, are not found after approximately 1600 (intervening adverb in bold).

(34) Here men vadoonston ofte by this nyght the nyght of synne

(Wycliffe Sermons 1,477,605: Han and Kroch 2000: 2)

The second observation has to do with inflectional morphology. The loss of almost all agreement marking (except 2sg- -st and 3sg present -s/-d) by the early 16th century and the loss of infinitival -en, combined with the loss of subjunctive inflection and the general absence of a dedicated present-tense marker had the consequence that verbs frequently appeared as bare roots. From the 16th-century onwards, verbs in the present tense, both indicative and subjunctive, the infinitive and the imperative all appear as bare verbal roots. This contrasts markedly with the situation in the Romance languages and in most other Germanic languages (in fact, Afrikaans is the only other Germanic language which systematically allows bare roots to appear; although the Mainland Scandinavian languages lack agreement marking, they have both an infinitive ending and a present-tense marker). If we think of V-to-V movement as creating a verbal category from an acategorial root, then it is natural to think that a reflex of this would be the marking of verbs as such with specifically verbal inflection; the "theme vowels" of the Romance conjugation classes are arguably realisations of features of V (see, for example Kayne 1994: 42-46). Conversely, the fact that words which are interpreted as verbs regularly appear as roots can be interpreted as indicating the lack of v-to-V movement. In this connection, a further innovation of 16th-century English is worth mentioning here: the development of zero-conversion of verbs to nouns and nouns to verbs as the productive way in which to derive new elements in each category (see, for example Görlach 1991: 178). The development of zero-conversion as a productive process is a natural consequence of the absence of verb-marking inflectional morphology, which in turn, we suggest, is linked to the loss of V-to-V movement at this period.

Concerning auxiliaries, we propose, still following Biberauer and Roberts, that at this period auxiliaries were merged in v and raised to T. Moreover, this operation consistently applied in all finite contexts, and so, where there was no overtly realised auxiliary, "null" v (i.e. minimally specified v which was not phonologically realised) raised to T. This accounts for a further property of ENE that is distinct from PDE: the fact that do was not required in negative clauses, even in the absence of main-verb movement. We thus find examples with a superficially "Scandinavian" negation pattern such as those in (35):

(35) a. Or if there were, it not belongs to you.
   (1600: Shakespeare 2 Henry IV, IV, i, 98; Battistella and Lobeck 1988: 33)

b. Safe on this ground we not fear today to tempt your laughter by our rustic play.
   (1637: Jonson Sad Shepherd, Prologue 7; Kroch 1989: 33)

The difference between examples like these and comparable examples with "do-support" is simply that in the latter type, do (or rather, the features ultimately lexicalised as do—see Note 15) is merged in v and raised to T over not; this operation was not yet obligatory in the English of this period, with the only obligatory operation at the time being the raising of the feature-bundle located in v to T, regardless of whether this would ultimately be spelled out or not. A crucial step in the development of the modern system of do-support, where do is of course obligatory in negative clauses like (35), from the earlier optional system had to do with the development of negative auxiliaries. In NE, negative auxiliaries such as won't (can't, won't, didn't, and, crucially, don't) are lexically derived by suffusion of n't to the auxiliary, rather than any syntactic operation of cleftisation or the like. One reason to think this comes from the fact that syntactic operation of cleftisation or the like. One reason to think this comes from the fact that syntactic operation of cleftisation or the like. One reason to think this comes from the fact that syntactic operation of cleftisation or the like. One reason to think this comes from the fact that syntactic operation of cleftisation or the like. One reason to think this comes from the fact that
A consequence of the development of negative auxiliaries (i.e. elements first-merged in T) is that, from the first half of the 17th century, T was readily able to be lexicalised independently of the presence of an EPP-feature. It is also possible that some modals which were formerly merged in v were reanalysed as T-elements at this stage, thus removing further cases of v->T movement. So these changes meant that T no longer had to have an obligatory EPP-feature in order to be lexicalised; instead, T’s EPP-feature could become optional, with the result that attraction of the remaining v-elements began to have a discourse effect (see again Chomsky’s 2001 proposal that optional EPP-features must have an effect on output). Under these conditions, moved v clearly has to be overt. Minimally specified v, which lacks phonological realisation can therefore no longer move to T, with the result that it consistently remains in situ and do becomes the minimally specified T-element. In contrast to the earlier system, then, v->T movement in NE always and only applies when v is realised as an overt auxiliary. The discourse effects associated with merger of do in v include interrogativity and emphasis, as well as the licensing of VP-fronting and VP ellipsis, both of which are also associated with emphatic readings (given our proposals, it is worth noting that the VP-constructions just mentioned are constructions which only emerged from ca. 1600 on; cf. Plank 1984, Roberts 1993):

(36) a. Did he smoke?
   b. John does (so too) smoke
   c. He threatened to bring some sandwiches and [bring some sandwiches] he did --
   d. I hoped he would bring some sandwiches and he did [bring some sandwiches]

It is easy to see that the replacement of general v->T movement by v->T only under specific conditions represents another case of restriction of function as defined at the beginning of this section. The SP is relevant here since the earlier grammar permitted minimally specified v to be realised as either do or zero whereas the later grammar requires it to have a null realisation when it does not move and to be realised as do when it does. The earlier grammar, then, allows a superset of the strings allowed by the later grammar because it allows strings of the type [v (do)]->T X Vp where X is an adverb, floated quantifier or negation and VP contains the main verb and relevant complements. The later grammar requires the realisation of do wherever T has a discourse effect or is negative and disallows it otherwise; in other words, where minimally specified v moves to T in PDE, this is always realised as do + T, i.e. the resulting string is necessarily [v do]+T X Vp. So we see that do is never optional.

This instance of the SP does not involve optional movement, as we have seen, but instead involves a change in the optional realisation of the minimally specified v to obligatory realisation as do whenever movement takes place. So we see how the SP can play a role in eliminating optional realisation of formatives just as it can play a role in eliminating optional movements. In other words, it influences the diachronic incidence of both movement and realisation by influencing how acquirers set the parameters relating to these options as they apply to functional heads.

5. Conclusion

This paper has tried to show that, contrary to what is often thought given the interective nature of many parameters, the SP does appear to be relevant as a causal factor in syntactic change. We have seen two general types of effect of the SP resulting in the elimination of formal options: in one case, this directly affected movement; in the other case, realisation of a formative. It is entirely possible that the SP has other effects in diachrony. For example, one issue that we have touched on here without developing is the characterisation of the "robustness" of the trigger experience necessary both to preserve and to cause a system to move to a superset grammar. Notice that the issue of preservation of superset systems relates to the Inert- Tia Principle of Longobardi (2001) and Keenan (2002). In fact, there may be a tension between the simplest understanding of the notion of inertia and the SP; until the question of robustness is fully clarified, we cannot really tell. The question of how systems can move from a relatively unmarked to a relatively marked state (i.e. from a subset to a superset grammar in the present case) is much more difficult; it is clear, though, that this must be driven above all by the nature of the PLD.

The conclusion that the SP is after all relevant to syntactic change is positive, since, as we saw in the introduction, the SP is based on a recognized fact about language acquisition: the fact that language acquirers do not have access to negative evidence. More generally, in showing how the SP plays a causal role in syntactic change, we see one small way in which the study of syntactic change and language acquisition may begin to converge.

Acknowledgements

We would like to thank the audiences at CGSW21 (Santa Cruz), SHE55 (York), DGS9 (Trieste) and the LAGB Annual Meeting (Newcastle) for their questions, comments and suggestions at various presentations of this material during 2006, in particular: Helen Goodluck, Caroline Heycock, Richie Kayne, Ed Keenan and Peter Svenonius. Our sincere thanks are also due to two anonymous reviewers and to the editors of this volume for the various incivilities they picked up in the originally submitted version of the paper. Usual disclaimers apply.

Finally, we would like to acknowledge the financial support of the AHRC (project AR14458 – "Null subjects and the structure of parametric theory").

Notes

1. A grammar G generates a language L which is smaller than another language L' generated by G' iff G' generates at least one grammatical string S which is not generated by G and G generates no grammatical string S' which is not generated by G'. Taking grammars to be defined as sets of values of parameters of UG, we can then say that a given parameter-value v of parameter P, is a subset value of P, where P defines a grammar of type G as just defined and value v defines a grammar of type G', assuming (crucially) all other parameters are set to the same values in both G and G'.

   Regarding the determination of subset and superset parameter values, clearly this cannot be done by inspection of the sets of strings generated by the different grammars defined by the different parameter values, since these will presumably always be infinite. This is an aspect of what we call the "implementation problem" associated with the SP – see Note 2.

2. The intersective nature of many parameters is an empirical problem stemming from comparative linguistics, and represents in a way the main focus of this paper. There is also a conceptual problem with the SP, which may have a psycho-linguistic dimension: the "implementation problem" alluded to in Note 1. The issue is how the learning device (whether a child or an algorithm) "knows" which are the subset and superset values of a given parameter. Since this cannot be directly determined by inspection of the trigger experience, we must, as it were, "build it in" to the learning process or the structure of parameters. One possibility is to assume that the learning device is equipped with an inherent system of default values for pa-
Theresa Biberauer and Ian Roberts

parameters which correspond to the subset values (as proposed by Manzini & Wexler 1987). Alternatively, the relevant notion of markedness may be more abstract, being a property of systems of parameters rather than of individual parameter settings: in this case, too, the learning device could be structured so as to inherently favour certain combinations of parameter values over others, and these could be combinations yielding subset grammars. In this connection, the discussion of the SP in relation to “shifted” parametric systems in Clark (1992) is relevant. Deciding amongst these and other imaginable options is well beyond the scope of this paper; what seems clear, however, is that some further specification of the nature of UG and/or of the learning device is required. On this, see Clark & Roberts (1993); Hale & Reis (2003); Fedor & Sakas (2005).

3. It is worth noting, as pointed out to us by an anonymous reviewer, that the illustrations depicted in (4) and (7) fall outside the scope of the parameters concerned. The questions that, however, remain are what status these illustrations have for acquirers and, related to this, how they are to be viewed when linguists calculate sub- and superset grammars.

4. Biberauer & Roberts (2006a) speculate on the factors that may potentially determine the nature of the category that counts as the goal for a given probe, suggesting that the distribution of the features being probed, as “signaled” by morpho-lexical cues, may play a crucial role. Thus, if a probe seeks a complete set of phi-features, as is usually assumed for finite T, and these features are consistently present on a D-head in T’s c-command domain, D may become the designated goal for phi-seeking T in that language. Similarly, if these features are spread across various DP-internal heads, the entire DP may count as the goal.

5. It is worth noting that a system consistently requiring a particular movement operation (e.g. a C vs. V(TP)-type system) and one consistently lacking this movement operation are not in a superset-subset relationship. A movement-requiring grammar will not generate more structures than a movement-lacking one; the two grammars will simply consistently deliver different output strings wherever the movement operation applies in the former, but not the latter. From the perspective of the SP, then, movement per se is not associated with any “cost”.

6. On the approach advocated in B&R, the “head-final” orderings in OE and ME therefore fall out from a consistently head-initial underlying structure (cf. Biberauer & Roberts 2006a,b for more discussion).

7. Roberts & Roussou (2003:162) use the term ‘restriction of function’ in a slightly different way, referring for example to the change in distribution of Ancient Greek indefinite, which in Modern Greek can only function as wh-words. The two notions are related if we think of the “function” in question as the environments in which (first or second) Merge is possible.

8. An anonymous reviewer points out that B&R’s proposals would not seem to make the correct predictions for negative objects in either OE or ME as negative objects are able to appear both preverbally and postverbally throughout, i.e. negative-object movement does not ever appear to have been obligatory. It is, however, worth noting that the examples cited in order to support the idea that negative-object movement was only ever optional all involve modal in verb-raising (VR) and verb-projection raising (VPR) structures (cf. Evers 1975 and Hagehman and van Riemsdijk 1986). As discussed in Biberauer and Roberts (2006a), there appears to be quite good evidence that the V(PR) structures in which modal appears throughout OE and ME are biclusal. According to Biberauer and Roberts (ibid.), these structures are best analysed as involving a PRO subject and an infinitival T which attracts V to it. If this analysis is correct, negative (and other) objects can in fact undergo floating onto the VP-domain in the manner outlined in (20), but still surface in a postverbal position: if the VP into which the object has raised subsequently undergoes raising to [Spec, TP], i.e. if the pied-piping mode of EPP-satisfaction is employed to satisfy infinitive T’s EPP-requirements, the object will ap-
hard to be sure what the situation was. It has been suggested that do had a perfective interpretation; extrapolating from the facts of contemporary South-Western dialects of British English which allow positive declarative do with a habitual interpretation (Chalcraft 2007; Trudgill & Chambers 1991), we might ascribe an imperfective feature of some kind to ENE do. Given that auxiliary do arose from the reanalysis of an ECM, causative and raising verb, as mentioned above, there may have been various homophonous auxiliaries do in ENE. The important thing is that these elements all had some extra feature in addition to those borne by "null" v.

References


