Parameters in minimalist theory: The case of Scandinavian

Anders Holmberg
Newcastle University

Abstract

The P&P theory of UG has come under heavy criticism, lately, from outside but also from inside generative grammar. The claim is that the search for ‘deep parameters’ underlying clusters of properties across languages has led nowhere, and should be given up. I have revisited a theory, now two decades old, which explained ten syntactic differences between Insular and Mainland Scandinavian as the effects of a single parametric difference (in a series of works by C. Platzack and A. Holmberg). The theory is shown to be fundamentally right, descriptively and theoretically. Later developments in generative theory only serve to sharpen the formulations, adding another layer of explanatory depth to the earlier theory. The conclusion is that there are parameters of the traditional kind. The problems encountered when the theory is tested on more distantly related languages is discussed on the basis of facts from Finnish. P&P theory is perfectly consistent with the minimalist approach to UG and variation when parameters are seen as points of underspecification in UG, and restrictions on variation are seen as, in part, third factor effects.

Keywords: parameters, P&P, $\phi$-features, null subject, agreement, third factor, stylistic fronting, Finnish

1. Introduction

The principles-and-parameters (P&P) theory of universal grammar (UG) proposed by Chomsky (1981) was by all accounts a very successful theory, which profoundly affected the course of research in linguistic theory. In particular it gave a huge boost to comparative

---

1 The research for this paper was in part carried out under the auspices of the projects ‘Null subjects and the structure of parametric theory’ and ‘Structure and linearization in disharmonic word orders’, both funded by the AHRC (UK). Thanks to Ian Roberts, Theresa Biberauer, and Michelle Sheehan, and Chris Johns for their contribution to this research. Thanks also to the organisers and audiences of the workshop ‘Holmberg & Platzack reloaded’ in Lund, Oct. 2008, and the ‘Student Conference on Theoretical and Applied Linguistics’ in Newcastle, March 2009. Special thanks to Fritz Newmeyer for his provocative views, to Hans-Martin Gärtner for encouraging me to write the paper, and to Halldor Sigurðsson for many discussions and much help with the data.
syntax, since, although comparative investigations were carried out before 1981, with the advent of P&P theory, cross-linguistic comparative investigation became the method of probing UG. It led to a flurry of papers, books, and dissertations on various aspects of the grammar of a wide variety of languages, including many hitherto poorly studied languages. It also gave a boost to research on language acquisition, both L1 and L2 acquisition, since the theory generated a range of testable predictions concerning acquisition.

For a long while, in the eighties and nineties P&P theory went virtually unchallenged within Chomskyan ‘mainstream’ generative linguistics. The position of P&P theory in current generative linguistic theory is less clear, though. It has been subjected to some severe critique lately, and not only from outside the theory, by functionalists (Croft 2001, Dryer 2009, Haspelmath 2008) and ‘usage-based’ linguists (e.g. Tomasello 2003), but also from inside, by scholars who are active supporters of ‘the generative enterprise’ (Bouchard 2003, Newmeyer (2004, 2006a), including people who are working within the minimalist framework (Boeckx, forthcoming). In particular, the notion of parameter has been called into question. Newmeyer (2004, 2006a) argues that two and a half decades of research within P&P theory has failed to substantiate the claim that linguistic variation is due to variation with regard to a small number of parameters; according to him, there are no parameters, only language particular rules. Haspelmath (2008) also argues that there are no ‘deep parameters’, that is parameters with a variety of surface effects. Boeckx (forthcoming) likewise does not believe in deep parameters, or the idea that variation is structured. Instead, all there is, according to him, is a range of isolated superficial parameters. He furthermore rejects the idea of parametrised principles on conceptual grounds, as being inconsistent with current minimalist theory.

This paper is a rebuttal of (some of) this critique. This will be done by way of inspecting a particular case of linguistic variation, and a parameter-based theory proposed to account for this variation. The case is that of the Scandinavian languages, and more specifically, variation between Insular Scandinavian and Mainland Scandinavian as regards properties to do with the subject. The theory is that of Platzack (1987), Platzack & Holmberg (1989), Holmberg & Platzack (1991, 1995). We isolated about ten constructions involving the subject, broadly speaking, which distinguished between the two groups of languages. We ascribed this variation to one parameter, to do with the features of I. What I will argue is that we were basically right, descriptively, in that most (though not all) of these differences are due to one parameter to do with the features of I. Later empirical findings and theoretical
developments do not justify abandonment of that theory, only a refinement of it. The conclusion is that there are ‘deep parameters’, and furthermore, I will argue that this is perfectly consistent with minimalist theorizing.

I will discuss to what extent the theory generalizes to other languages the way it should, if P&P theory is right. I will argue that it probably does, once the effects of other parameters are controlled for, though, in the absence of sufficiently detailed comparative work on many languages, the argument remains inconclusive.

I will begin by discussing parameters within minimalist theory. I will argue that the right way to think of parameters is in terms of underdetermination by UG. This notion of parameters is consistent with recent minimalist theorizing according to which UG, the specifically linguistic genetic endowment, is fairly simple, and where extralinguistic ‘factors of the third kind’ have a bigger impact on the form of language than used to be assumed. Furthermore, I will argue that it is, in fact, also consistent with parametric theory as implemented in the formative period of GB theory, despite the rhetoric that was often used, and does provide the required basis for a solution to the logical problem of language acquisition, as envisaged by Chomsky (1981).

2. The Principles and Parameters program

The purpose of the P&P program, famously, was to reconcile the following two facts: (a) A human language is an extremely complex system, which is nevertheless acquired in a very short time, on the basis of very scant evidence (the ‘logical problem of language acquisition’), which indicates that most of it is innate, and therefore does not need to be learnt. (b) There is a lot of cross-linguistic variation, which indicates that a lot must be learnt after all. Everything in a language which is not common to all languages must be learnt on the basis of exposure to the facts of the language in question, in L1 acquisition.

The leading idea of the P&P program is: If the learning path of language acquisition is narrowly circumscribed, in such a way that acquisition is a matter of making *a series of choices among a narrow range of options*, on the basis of the primary linguistic data, then this will help to solve the logical problem of language acquisition. And when these options are situated deeply enough in the network of interdependent properties of syntax and morphology, and perhaps phonology, then the effects of any one choice may be considerable, affecting different parts of the system, which helps explain why languages can look so
different, although, in this perspective they actually differ only with respect to a relatively small number of parameters.

Insofar as linguistic variation is due to variation with regard to parameters embedded in a complex network of interdependent properties, we should find clusters of surface effects of these deep-lying parameters in the languages of the world. Consequently, the discovery of clusters of properties, on the basis of comparison of sets of different languages and dialects, with ensuing explanation of the cluster in terms of a parameter of UG, became the favoured method of research in comparative syntax, in the period following Chomsky (1981).

One argument leveled against the P&P program, most eloquently by Newmeyer (2004, 2006a), is that the prediction regarding clustering has not been confirmed: Grammatical properties do not seem to cluster the way predicted by the various theories which adhere to the P&P idea. To take a classic example, one of the very first parameters to be proposed was the pro-drop, or null-subject parameter (actually two linked parameters) proposed by Rizzi (1982).

One argument leveled against the P&P program, most eloquently by Newmeyer (2004, 2006a), is that the prediction regarding clustering has not been confirmed: Grammatical properties do not seem to cluster the way predicted by the various theories which adhere to the P&P idea. To take a classic example, one of the very first parameters to be proposed was the pro-drop, or null-subject parameter (actually two linked parameters) proposed by Rizzi (1982).

(1) a. INFL can have the feature [+pronoun].

b. INFL which is [+pronoun] can have the feature [+referential].

Languages whose INFL is marked [+pronoun] can have a null NP as subject. If INFL is also [+referential], the null NP can be a definite pronoun, otherwise it can only be an expletive pronoun. Building on Perlmutter (1971), Rizzi claimed that languages with pronominal INFL have two other properties absent from languages without pronominal INFL, namely so called Free Inversion (the subject may occur in postverbal position) and violation of the COMP-t-filter (extraction of embedded subjects over an overt complementiser is allowed). The idea is, to put it very simply, that the pronominal INFL stands in for the subject, so that the subject pronoun which satisfies the EPP in spec,IP (in more recent terms) can be a featureless, null NP. The pronominal INFL also makes it possible to have the lexical subject placed in a lower position (hence Free Inversion), from which position it can be extracted without passing through spec,IP (hence violation of the COMP-t filter) (see Roberts & Holmberg (2009)).

The prediction is, then, that the three properties should cluster in this way in every language: Every language should either have all three properties or none of them, all else being equal. But this appears not to be the case. Gilligan (1987) did a comparative study of 100 languages from different corners of the world, and found that the properties do not show
any interesting correlations; see Newmeyer (2006a: 89ff.), but see Baker (2001: 83f.) and Roberts & Holmberg (2009) for discussion and part reappraisal of Gilligan’s findings. Note, however, the qualification ‘all else being equal’. The question is whether all else really is equal, in the languages Gilligan investigated, and in comparable cases where a hypothesis which posits a parameter to account for a cluster of properties in a set of languages is tested against a larger sample of languages. This is a question I will come back to below in section 5.

According to Newmeyer (2004, 2006a), the absence of any correlation between, say, the possibility of a referential null subject and Free Inversion implies that they are independent of each other. There is no parameter, just two syntactic properties which have to be acquired independently of each other, and indeed, perhaps independently of any other property of the language. Essentially the same point is made by Haspelmath (2008): The ‘deep parameters’, i.e. the ones with a range of surface effects, envisaged in early P&P theory, have not been substantiated by subsequent research, indicating that they do not exist.² The conclusion Newmeyer draws is linguistic variation, instead of being due to variation vis-à-vis a set of parameters, is due to language-particular rules. As he rightly points out (Newmeyer 2006: 87), language-particular rules may have a variety of surface effects as well, depending on their place in the grammatical system of the language. Furthermore, even though a rule is language-particular in the sense that it is not determined by UG, and is independently learnt, it may well be instantiated in many other languages, in which case the same surface effects may appear in many languages.

There are other arguments levelled against P&P theory in Newmeyer (2006a), leading to the same conclusion that parameters ought to given up in favour of language-particular rules. Perhaps the most damning one is that the number of parameters that must be assumed, in the light of research carried out within this research program since the eighties, is so large that it threatens to undermine the argument for parameters from the logical problem of language acquisition. If the number of parameters that have to be set in the process of acquisition is not significantly smaller than the number of grammatical properties, then the logical problem of language acquisition remains essentially unsolved (see Newmeyer 2006: 83). This is related to the clusters issue. The purported absence of clusters of properties

² Haspelmath compares the history of ‘deep parameters’ in P&P theory with the history of ‘holistic types’ (e.g. analytic vs. synthetic) in early linguistic typology, and ‘deep implications’ in more recent linguistic typology, none of which, according to him, have been substantiated by subsequent research.
means that parameters, if they exist at all, are relatively superficial, with a limited number of effects. If so, the logical problem of language acquisition remains a problem.

This dilemma is resolved if parameters form hierarchies, since this may considerably reduce the number of choices that have to be made in language acquisition. In the ideal case (surely unrealistic) where parameters are all arranged in one big, symmetric, binary branching tree, and the setting of parameters in language acquisition proceeds from the top down, every choice will reduce the number of choices that remain to be made by half. Newmeyer (2006a) discusses, and refutes, extant theories of parameter hierarchies, particularly Baker’s (2001: 157-197). See Roberts and Holmberg (2009) for a defense of parameter hierarchies. I will return briefly to this issue in the conclusions.

3. Parameters in minimalist theory

3.1. The problem of parametrised principles

Consider the following quotations, from Boeckx (forthcoming): “/…/the idea that a GB-style Principles-and-Parameters architecture provides the right format for a solution to Plato’s Problem is, I think, seriously mistaken on both empirical and conceptual grounds.” and “For one thing, the traditional Principles and Parameters model is no longer compatible with the way minimalists think of Universal Grammar.”

The new view of UG that Boeckx is alluding to is found in Chomsky’s (2005, 2007) classification of the factors that determine the form of grammar:

(2) (a) The genetic endowment, i.e. UG (the first factor),
(b) the environment (the second factor), and
(c) extralinguistic factors (the third factor).

Factors of the third type include generic principles of ‘good design’, efficiency, and economy which are not specific to language but common to all cognitive systems, or all biological systems, or due to laws of nature. In his (2005) paper and subsequent works (Chomsky 2007, Berwick & Chomsky 2009), Chomsky argues that the role of third factor principles is more important than previously thought, and correspondingly, the role of first factor principles (i.e. UG) is less important. According to the traditional P&P view of the role of UG, it consists of (a) a set of universal categories and/or features (possibly a pool of categories/features from
which individual languages draw a subset), (b) a set of absolute universal principles, and (c) a set of parametrised principles, all of which together make up a rich and complex system which is specific to the language faculty, and part of the human genome. The system would contain principles concerning hierarchic structure and recursivity, X-bar theory, locality principles, movement, binding, selection, theta-roles, case, agreement, etc., some of which would be parametrised. The richer the system, the better it can explain how language acquisition is possible (see Chomsky 2007).

One problem with this theory is that it is implausible from the point of view of human evolution. If UG, as defined above, is genetically encoded, it must have evolved in the manner of other genetic endowments, essentially through Darwinian natural selection. It is unlikely that a complex cognitive system such as the one depicted above could have evolved in this way, though, particularly given the short history of modern humans. The language faculty as a whole has many obvious selectional advantages, but individual, purportedly universal properties of the system, such as, say, the Case Filter, the Left Branch Condition, the EPP, Principle B, the fixed order of tense over aspect (Julien 2002), or the fixed order of shape adjectives over colour adjectives (Cinque, forthcoming) do not. If so, they cannot be genetically encoded as such; see Boeckx & Hornstein (2009), Christiansen & Chater (2008).

A much more plausible scenario is that UG is fairly simple, encoding properties such as hierarchic, binary branching structure (which yields c-command as a crucial relation), recursivity, the lexical-functional distinction, the valued-unvalued feature distinction (Chomsky 2001), the minimal-maximal distinction, perhaps the verbal–nominal distinction, perhaps the LCA (Kayne 1994). All other universal properties would follow from the interplay of these fundamental linguistic properties with generic conditions on mental computation (including perhaps Relativised Minimality and other locality conditions) or to do with even more general efficiency conditions on computation, or storage, or access of information.

However, in this scenario there is no obvious place for parametrised principles in UG, if we exclude the possibility that fundamental properties such as hierarchic structure, or the lexical-functional distinction, or the LCA are parametrised. Besides, from the point of view of evolution, the notion that parameters would be part of the genetically encoded UG is implausible: What evolutionary advantage would they possibly have?

But is this the right way to think of parameters? The notion that the critics of the P&P program seem to have in mind is the ‘parametrised principle’, that is a principle of UG which
as part of its formulation specifies two or more options with regard to some property: a feature value, or a movement, or a linear order, etc. Viewed this way, a parameter is a principle *plus* something, namely a choice between two (or more) specified options. Effectively, it is two (or more) minimally different principles, of which the language learner has to choose one. It entails added complexity in UG, which is, indeed, inconsistent with minimalist assumptions. An alternative, certainly no less plausible or natural, is that a parameter is what we get when a principle of UG is underdetermined with respect to some property. It is a principle *minus* something, namely a specification of a feature value, or a movement, or a linear order, etc.

But above I stated, categorically, that the core idea of the P&P program is that language acquisition is a matter of making choices among a limited (and small) number of options. Can the number of options be limited, if they are not specified by UG? Yes they can. There is a variety of factors, linguistic and extralinguistic, which may have the effect of drastically restricting the number of options at points where a principle is underspecified. Consider the time-honoured head-complement parameter: UG specifies that a head can merge with a maximal category, its complement, but UG does not specify their linear order. There are two options: The head either precedes or follows the complement. This need not be specified by UG, though: The nature of our articulatory and perceptual apparatus makes it impossible to pronounce or perceive the head and the complement simultaneously, so they have to be pronounced one after the other. That is to say, the parameter is a third factor effect.\(^3\)\(^4\)

Or take the case of the wh-movement parameter, distinguishing wh-movement from wh-in-situ, another classical parameter proposed first by Huang (1982a,b). In terms of the theory assumed at the time, UG prescribes that wh moves to specCP (the principle), but does

---

\(^3\) The same point is made by Boeckx (forthcoming). Newmeyer (2006a: 74) proposes replacing the Head-Complement parameter with two language-particular rules, Head precedes Complement and Head follows Complement, of which languages select one. I take it that the fundamental difference between a parameter and a set of rules is that a parameter presupposes a choice among a limited number of alternatives, but rules do not. If so, the dichotomy between a binary parameter and two rules that Newmeyer sets up is spurious. There is no difference in this case, where the choice is restricted to two alternatives. The classical phrase structure rules (as in, say, Chomsky (1965)) on the other hand, are not parameters, since they are delimited only by the range of categories provided by UG (given the rewrite rule format). See the conclusions section.

\(^4\) It is sometimes intimated (e.g. Haspelmath 2008) that the head-complement parameter would be obsolete, because cross-categorial harmony is not prevalent enough. This is not the case. The format of the parameter is controversial (is it about base-generation or movement?), but cross-categorial harmony *is* universally valid, when properly defined and delimited; see Biberauer, Holmberg & Roberts (2008, 2009).
not specify at which point in the derivation the movement occurs (the parameter). This leaves open whether it applies before or after S-structure (or spell-out, in more recent terminology), that is to say, it leaves open the choice between overt and covert wh-movement, where the latter option yields wh-in-situ. The two options need not be specified as part of the principle, but are a consequence of the general architecture of the grammar, thus another third factor effect. This was not exactly how it was expressed at the time, because reducing the complexity of UG was not an issue at the time, but it could have been; see Roberts and Holmberg (2009).

The null subject parameter (1) looks formally different, though. It concerns the featural specification of a functional category. UG provides a pool of features, from which syntactic categories are constructed, but does not dictate the exact feature-composition of every category. I return to this issue in more detail below. Again, there is no parametrised, ‘overspecified’ principle, but underspecification of the featural content of a functional head. As will be discussed below, the options are restricted primarily by the type of nominal features that UG provides.

3.2. Parameter schemata

Several proposals exist from recent years addressing the conceptual and empirical problems that the idea of innate parameters poses. Gianollo, Guardiano & Longobardi (2008) propose that, instead of parameters, UG makes available a small set of parameter schemata, which, in interaction with the primary linguistic data create the parameters that determine the non-universal aspects of the grammatical system. Taking ‘the Borer-Chomsky conjecture’ that parameters are properties of lexical items, specifically functional heads, as their departure, they propose the following schemata:

(3)  a. Grammaticalisation: is F, a functional feature, grammaticalised?
    b. Checking: is F, a grammaticalised feature, checked by X, X a category?
    c. Spread: is F, a grammaticalised feature, spread on Y, Y a category?
    d. Strength: is F a grammaticalised feature checked by X, strong? (i.e. does it overtly attract X?)
    e. Is F, F a grammaticalised feature, checked by a category X°?
A case of a grammaticalized $F$ is when definiteness is obligatorily marked in an argument DP. (3b) asks (effectively) whether $F$ is unvalued/is a probe (in the sense of Chomsky 2001). (3c) asks whether $F$ has unvalued counterparts on other categories, as when definiteness is marked on adjectives in the DP as well as on the article. (3d) distinguishes the case where the probe-goal relation is accompanied by movement, and (3e) the case where head movement is triggered.

Roberts & Roussou (2003:213) propose a similar set of schemata of options concerning properties of a functional feature (Does $F$ enter an Agree relation?, Does $F$ attract? Does it attract a head or an XP? etc.). Yet another version of the same general idea is found in Boeckx (forthcoming), proposing that variation in the narrow syntax is reduced to the following:

(4) a. Features $F_1$ and $F_2$ may be expressed separately or as a bundle;
   b. $F$ may or may not exhibit a $uF$ variant;
   c. A given phase head may or may be strong, i.e. $uF$-bearing, or weak (defective).

(3b) and (4b) are two formulations of the same parameter schema. There are differences, too, though. In particular, Boeckx (forthcoming) rejects the idea that the inventory of grammatical features could vary across languages; they only differ with respect to their precise distribution (separately or bundled). The schemata define the properties that are allowed to vary at all. They are restricted to binary choices due to the nature of syntactic features: A feature either is or is not unvalued, does or does not trigger movement, etc.

Boeckx (forthcoming) makes the point that (4) is the only variation allowed in the narrow syntax. All other variation, such as the linear order between head and complement, which copy in a chain is pronounced, or whether a head wants its specifier position to be filled by overt material, is a matter for the post-syntactic morpho-phonological component. Which variation is allowed in the narrow syntax as opposed to morpho-phonology is obviously an interesting question. According to the most radical, minimalist view narrow syntax is completely uniform, with all variation relegated to the morpho-phonological component, i.e. to externalization of syntactic structure (see Berwick & Chomsky (forthcoming), Burton-Roberts & Poole (2006a,b), Hinzen (2009), Sigurðsson (2004b)). This makes no difference to the status of parameters in the theory, though. Wherever the variation
is located, it must be constrained so as to allow only a limited number of options, or else we have not addressed the logical problem of language acquisition.\(^5\)

4. A theory of subject-related variation among the Scandinavian languages

I have argued that the parametric theory of linguistic variation is perfectly compatible with a minimalist approach to UG, once parameters are taken to be points of underdetermination in UG. I have also argued that this does not entail a break with the tradition; the classical parameters of the formative period of P&P theory can be viewed in this light, with virtually no change in their formulation. The reason why this was typically not done, is that, in the period in question, enrichment of UG was not seen as a problem, but in fact, a desideratum; see Chomsky (2007), Berwick & Chomsky (forthcoming), Boeckx & Hornstein (2009).

In this section I will address the critique levelled against the favoured method of probing UG within the P&P program, that of comparing a set of languages, isolating a cluster of properties present in a subset of the languages, but absent in the complement set, and formulating a parameter to account for the clustering. According to critics such as Newmeyer (2004, 2006a), Haspelmath (2008) and Boeckx (forthcoming) the clusters invariably fail to hold up when a wider range of languages are taken into account, hence no ‘deep parameters’ have ever been proposed that would have stood the test of time, and the research program has essentially led nowhere.

What I will do in the following is scrutinise a particular theory which adopts the P&P approach, and is based on a clustering of properties across a group of closely related languages, that is the theory of Scandinavian sentence structure articulated in work that Christer Platzack and I did in the late eighties and early nineties: Platzack (1987), Platzack & Holmberg (1989), Holmberg & Platzack (1991, 1995), henceforth collectively called P&H. The reason for picking this theory is that it is a good example of a theory where a large number of differences among two sets of languages are explained as effects of one parameter.

I will first establish to what extent the conclusions we reached then are still valid, empirically as well as theoretically, and whether more recent developments in generative

\(^5\) One consequence of this way of looking at things is that the Borer-Chomsky Conjecture loses some of its force. It does not matter, for the purposes of solving the logical problem of language acquisition, where the variation is located. All that matters is that it is restricted to a few alternatives.
linguistic theory invalidate the conclusions we reached in P&H. I will argue that they do not, but, on contrary, strengthen the conclusions, once the empirical scope of the theory is reduced somewhat.

4.1 Differences between Insular Scandinavian (ISc) and Mainland Scandinavian (MSc)

‘Insular Scandinavian’ includes Icelandic and Faroese, while ‘Mainland Scandinavian’ includes Danish, Norwegian, and Swedish. In terms of syntax and morphology this is clearly the main typological distinction among the Scandinavian languages. Old Norse also falls clearly in the ISc camp. More controversially, this is also the case for the Swedish Oevdalian dialect (see Rosenkvist 2008). In order to simplify the presentation, I will disregard Faroese and Old Norse throughout this paper, so that ISc here actually only includes Icelandic.6

The following is a list of the differences between ISc and MSc dependent on a parameter to do with AGR, according to P&H.

(5) AGR-related differences

<table>
<thead>
<tr>
<th></th>
<th>ISc</th>
<th>MSc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Rich subject-verb agreement</td>
<td>+</td>
</tr>
<tr>
<td>2.</td>
<td>Embedded V-to-I</td>
<td>+</td>
</tr>
<tr>
<td>3.</td>
<td>Oblique subjects</td>
<td>+</td>
</tr>
<tr>
<td>4.</td>
<td>Stylistic Fronting</td>
<td>+</td>
</tr>
<tr>
<td>5.</td>
<td>Null expletives</td>
<td>+</td>
</tr>
<tr>
<td>6.</td>
<td>Null generic subject pronoun</td>
<td>+</td>
</tr>
<tr>
<td>7.</td>
<td>Transitive expletives</td>
<td>+</td>
</tr>
<tr>
<td>8.</td>
<td>Heavy subject postposing</td>
<td>+</td>
</tr>
<tr>
<td>9.</td>
<td>Indirect subj questions without res. element</td>
<td>+</td>
</tr>
<tr>
<td>10.</td>
<td>VP fronting</td>
<td>–</td>
</tr>
</tbody>
</table>


6 On Old Norse, see Faarlund (2006: esp. 217ff., 236f.). Faroese is an interesting case in this connection, since it is undergoing changes that seem to crucially involve the parameter discussed in the text below. Thus with respect to the properties in (5) and (67) below there is optionality and variation across dialects and generations; see Thráinsson (2001), Thráinsson & al. (2004), Jonas (1996), Platzack (1987), Heycock, Sorace & Svabo Hansen (2009). As regards Oevdalian, from Rosenkvist (2006) and Garbacz (2006) I get the impression that, with regard to the parameter discussed in this paper, Oevdalian is a MSc language but with some archaic features (case morphology, vestiges of verb agreement, verb raising ) and some new features (pro-drop of 1PL and 2PL pronouns).
I will reduce the list, as follows:

(6) *The reduced list of AGR-related differences:*

<table>
<thead>
<tr>
<th></th>
<th>ISc</th>
<th>MSc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Rich subject-verb agreement (7)</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>2. Oblique subjects, (9)</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>3. Stylistic Fronting, (10)</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4. Null expletives, (11)</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>5. Null generic subject pronoun, (12)</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>6. Transitive expletives, (13)</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>7. Heavy subject postposing (14)</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

Richness of subject-verb agreement is illustrated in (7), with the present and past indicative of the verb ‘call’ in Icelandic and Swedish. As shown, Icelandic has 2/3 syncretism in the present indicative singular (of most verbs), and 1/3 syncretism in the past indicative singular (and also in the subjunctive singular), but all other forms are distinct. MSc, here standard Swedish, does not have any subject-verb agreement on the finite verb, not even on auxiliaries or the copula ‘be’.

(7) **Present indicative** of ‘call’

<table>
<thead>
<tr>
<th>Icelandic</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG kalla</td>
<td>kallar</td>
</tr>
<tr>
<td>2SG kallar</td>
<td>kallar</td>
</tr>
<tr>
<td>3SG kallar</td>
<td>kallar</td>
</tr>
<tr>
<td>1PL köllum</td>
<td>kallar</td>
</tr>
<tr>
<td>2PL kallið</td>
<td>kallar</td>
</tr>
<tr>
<td>3PL kalla</td>
<td>kallar</td>
</tr>
</tbody>
</table>

**Past indicative** of ‘call’

<table>
<thead>
<tr>
<th>Icelandic</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG kallaði</td>
<td>kallade</td>
</tr>
<tr>
<td>2SG kallaðir</td>
<td>kallade</td>
</tr>
<tr>
<td>3SG kallaði</td>
<td>kallade</td>
</tr>
<tr>
<td>1PL kölluðum</td>
<td>kallade</td>
</tr>
<tr>
<td>2PL kölluðuð</td>
<td>kallade</td>
</tr>
<tr>
<td>3PL kölluðu</td>
<td>kallade</td>
</tr>
</tbody>
</table>
The other properties are exemplified in (8)-(13):

**Oblique subjects**

(8) a. **Mér voru gefnir peningar.** [Icelandic]
me-DAT were given money-PL
‘I was given money.’
b. *Mej blev givet/givna pengar. [Swedish]
me was given-SG/given-PL money

**Stylistic Fronting**

(9) a. **[Þeir sem í Osló hafa búið segja að það sé fínn bær.** [Icelandic]
those that in Oslo have-3PL lived say that it is nice town
‘Those that have lived in Oslo say that it’s a nice town.
b. *[De som i Oslo har bott] säger att det är en fin stad. [Swedish]

**Null expletives**

(10) a. **Nú rignir (*það).** [Icelandic]
now rains it
‘Now it’s raining.’
b. **Nu regnar *(det).** [Swedish]

**Null generic subject**

(11) a. **Hér má ekki dansa.** [Icelandic: Sigurdsson & Egerland 2009]
here may not dance
‘One must not dance here.’
b. *(Man) má ikke danse her.** [Norwegian]
one may not dance here

**Transitive Expletive Construction (TEC)**

(12) a. **Það hefur einhver köttur étið mýsnar.** [Icelandic: Vangsnes 2002]
there has some cat eaten the-mice
there has a cat eaten mice-the
Heavy subject postposing

(13) a. Það lásu hana því ekki margir stúdentar fyrir prófið. [Icelandic]
    there read it thus not many students for exam-the
    ‘Consequently not many students read it for the exam.’

b. *Det läste den således inte många studenter inför provet. [Swedish]
    there read it thus not many students for exam-the

In P&H, the null generic pronoun is not listed as a separate case, although it could have been. The reasons for removing V-to-I from the list are summarised in a separate section below. VP-fronting plausibly has complementary distribution with V-to-I, and is therefore removed as well. Indirect subject questions without a resumptive element are not included because this point of variation is not straightforwardly predicted by the theory articulated below.

4.2. V-to-I

The hypothesis that there is a crucial relation between rich agreement and V-to-I, called the Rich Agreement Hypothesis (RAH) in Bobaljik (2002b), has been intensely investigated ever since the eighties, in relation to Scandinavian (in addition to H&P, see Kosmeijer (1986), Falk (1993a,b), Rohrbacher (1999), Vikner (1995, 1997), and in relation to stages in the evolution of English, see Roberts (1994), Pintzuk (1991). There are strong indications, though, that rich agreement is neither a necessary nor a sufficient condition for V-to-I. Bobaljik (2002b) discusses cases of V-to-I in the absence of rich agreement, concluding that the strong version (14a) of the RAH cannot be right, but the weaker version (14b) can.

(14) a. Rich agreement is the cause of V-to-I.
    b. If a language has rich agreement, then it has V-to-I.

The cases which Bobaljik mentions include a stage in the evolution of Danish and English, where V-to-I occurred for a relatively long period (100-200 years, in the case of Danish)

7 Thanks to Halldór Sigurðsson for the example.

8 The following is an example of the construction in question.

(i) Finnur spyr hvað sé í pokanum. [Icelandic]
    Finn asks what is in bag-the

In the corresponding Mainland Scandinavian example there would be a resumptive element in specTP, either *der (Danish) or som (Swedish and Norwegian); see Taraldsen (1991)
after subject-verb agreement had become poor. This situation is also found in certain current MSc dialects (see Bentzen 2005). Siewierska & Bakker (1996), investigating the correlation of agreement and sentential word order in a large sample of languages, have many examples of VSO languages without agreement. If VSO order is derived by V-to-I, these languages constitute counterexamples to the strong RAH. Bobaljik (2002b) also notes that there is no credible explanation for the strong RAH in the literature. He constructs an explanation of the weak version in terms of Distributed-Morphology-style late insertion and the hypothesis that rich agreement means having separate AGR and T heads (Bobaljik & Thrainsson 1998).

The weak version of the RAH appears not to be right, either, though. In Icelandic, a notoriously V-to-I-moving, rich agreement language, V-to-I is, in fact, optional in adverbial clauses, relative clauses, and embedded questions; see Sigurðsson (1986, 1989/1992: 44-45), Angantysson (2007), Wiklund & al. (2007). That is to say, V-movement is not a prerequisite for licensing an agreement affix on the verb, even in a rich agreement language such as Icelandic – which means that Bobaljik’s (2002b) explanation of the weak RAH is flawed.

Not implausibly, VP-fronting, as in (20), is bled by V-to-I, which would explain why MSc has VP-fronting but Icelandic does not (Holmberg & Platzack 1995: 223f., Wiklund & al. 2007). There is no other explanation of this variation in terms of the theory articulated below. For this reason VP-fronting is removed from the list of properties dependent on the AGR-parameter.

4.3 Reasons for thinking that these difference are due to variation with respect to one parameter

The reasons given in H&P for thinking that the differences in (10) are due to a single parameter is that certain other languages exhibit the same cluster, at least in part. The languages listed in Platzack & Holmberg (1989) are Middle English, Old French, and Yiddish. They show that the three languages have, in addition to rich agreement, oblique

---

9 There are more ways than one to derive VSO order, though. One way would be by VP-fronting (Massam 2005), which yields VSO if it is preceded by object shift out of VP. The RAH makes no predictions regarding such languages.

10 Angantýsson (2007) rejects the idea that this is a question of verb movement, in favour of an analysis where it is a matter of variation in adverb placement. Thráinsson (in press) argues that it is a case of variation in adverb type. It is highly suggestive, though, that the very same pattern is found in Kashmiri: This basic-SOV language has V2 order in main clauses and complement clauses, but SOV in adverbial clauses and relative clauses; see Bhatt (1999), Holmberg (forthcoming). Here the SVO-SOV contrast excludes an analysis where the difference is a matter of adverb placement or type.
subjects, Stylistic Fronting, null expletives, and the TEC, though without any clear example of oblique subjects or the TEC in Old French, and with a clear example of Stylistic Fronting only in Middle English. Thus the only language of the three which incontestably has five of the seven properties in (7-13) is Middle English. The examples are given in (15) (from Platzack & Holmberg 1989):

(15) Middle English:
   a. Eet this when ðe hungrep. [oblique subject]
      eat this when you-DAT hunger
   b. that ladyes...might se Who that beste were of dede [Stylistic Fronting]
      that ladies might see who that best was of deed
   c. Now es arly, now es late. [null expletive]
      now is early now is late
   d. there woulde some Jewes reprove this his doing [TEC]

Like MSc are Modern English and Modern French, exhibiting none of the properties in (7-13), except that agreement is clearly less poor in Modern French than in Modern English and MSc (see Roberts 2009b).

A second reason is that the properties (7-13), all characteristic of older stages of Scandinavian, disappeared around the same time in the history of MSc.; Platzack (1988), Falk (1993).\(^{11}\)

4.4 The parameter according to P&H

On the most general, informal level, the intuition that P&H tried to formalise, is that the crucial difference between ISc and MSc, from which everything else follows, is that ISc, but not MSc, has rich subject-verb agreement. This idea was attractive because subject-verb agreement provides a particularly natural cue for the setting of the parameter in L1 acquisition, being sufficiently frequent and salient in the primary linguistic data. This intuition is, in fact, captured best in Platzack (1987), while the more sophisticated theory in Platzack & Holmberg (1989) and Holmberg & Platzack (1995) could not formalise this

\(^{11}\) Falk’s (1993) detailed study shows that verb agreement, Stylistic fronting, and the expletive pronoun (and, as it looks irrelevantly, V-to-I) underwent a change in Swedish from an ISc system to a MSc system in the same period, roughly the 16\(^{th}\) century.
intuition directly. According to Platzack (1987), the variation is an effect of Rizzi’s (1982) Null Subject parameter (1) (see above section 2). In ISc, INFL (or rather C, according to Platzack) encodes the feature [+pronoun], in MSc it does not. Only languages with (rich) agreement can have a pronominal INFL (or C), so the fact that MSc languages have no agreement means that INFL cannot be [+pronoun]. The pronominal feature in INFL (or C) absorbs nominative case, and this makes it possible for non-nominative categories, including oblique subjects, SF-moved categories, and null expletive pronouns, to occupy spec,IP.

Holmberg & Platzack (1995), adopting a lexicalist approach, proposed that the crucial difference is whether AGR, i.e. the $\phi$-features of I, has inherent nominative case or not. In ISc it does. The inherent nominative case explains why I can only agree with a nominative argument. Furthermore, assuming a universal requirement that finite C must check nominative case (the formulation is ‘Finite C must govern nominative’), AGR in I can satisfy this requirement in ISc. This allows for movement of non-nominative categories to specIP, to satisfy the EPP. MSc has no AGR in I (hence no agreement). Consequently, a nominative DP must be placed in specIP, to satisfy C, ruling out oblique subjects, SF, or null expletives (which, by assumption, are caseless). Languages like English and French have AGR in I, hence agreement, but AGR is not nominative. Therefore, in these languages, too, a nominative-marked DP must move to specIP, ruling out oblique subjects, SF, or null expletives.

4.5. Adapting P&H into a Chomskyan ‘probe-goal framework’

This section will present an account of the findings of P&H within the theory of agreement and movement articulated in Chomsky (2000, 2001), as developed in Roberts (in press, 2009a) and Holmberg (2009a,b). The resulting theory will be closer to Platzack (1987) than to the later works by P&H. The idea that I has inherent nominative case, as in P&H’s later works, does not sit well in current minimalist theory of case and agreement, but the idea that I has nominal features spelled out as agreement, is, of course, quite standard. The distinction between interpretable and uninterpretable features was not part of the theory in 1987. It is clear, though, that Platzack (1987) intended the feature [+pronoun] to be interpretable, actually substituting for a pronoun. It is a ‘clitic’, which absorbs nominative case, and is even

---

12 One reason for the modifications in P&H’s later works is that, unlike Platzack (1987), we wanted to include variation with regard to verb movement among the effects of the parameter. But this was probably a mistake, as discussed in section 4.2.
assigned a theta-role in languages with referential null subjects (p. 384f.). But INFL (or C) is not actually a pronoun, in for example Icelandic, as it can co-occur with a pronoun.

(16) þeir hafa+INFL komið. [Icelandic]
they have+PRS.3PL come
‘They have come.’

If INFL were a pronoun, this sentence ought to violate the theta-criterion and Principle B. The alternative is that the feature is uninterpretable. In this view, the difference between ISc and MSc is that ISc has some uninterpretable nominal features as part of the make-up of finite T, lacking in MSc.

In Holmberg (2009a), I argue that Rizzi’s (1982) formulation of the null subject parameter is very close to the right theory (and superior to later formulations). The following is a proposal how to embed this parameter in a more explicit theory of UG.

Let us say that UG requires a dependency-relation between (finite) T and an argument. Let us say, furthermore, that the mechanism that UG provides for such dependency relations is feature-valuation: A functional head can enter a dependency-relation with a lexical category if it has the unvalued counterpart uF of an inherently valued feature F of the lexical category. So to enter the required relation with a nominal argument, T needs at least one unvalued nominal feature. Relevant nominal features include definiteness, number, gender/class, and (for pronouns) person. Which of these features are selected is not dictated by UG, but varies across languages. The ISc languages have unvalued number [uNr] and person [uPn] in T. The MSc languages have neither. Languages which have definite null subjects (Italian, Spanish, Greek, Turkish, etc.) have unvalued definiteness [uD] as well, in T (this corresponds to Rizzi’s (1982) distinction between languages that have just [+pronoun] and languages that have [+referential] as well, in T). The choice of features will have other effects, including the possibility of different types of null subjects; see Holmberg (2009a).

Roberts (in press, 2009a) articulates the idea that Agree in Chomsky’s (2000, 2001) sense between a functional head (a probe) and a pronoun (the goal), depending on the feature

---

13 See, however, Platzack (2004), who develops the idea that AGR, in languages like Icelandic, is an anaphor, in the sense of Chomsky (1981), and thus can, and must, have a local antecedent.
14 The reasons for this requirement remain elusive. According to Miyagawa (in press), it is in order to make movement of the argument possible, thereby increasing the expressivity of language. See also Alexiadou & Anagnostopoulou (2007).
make-up of the probe, may lead to formation of a chain, that is when the features of the goal, as a result of Agree, are a subset of the features of the probe. This will be the case when T has unvalued features for number and person, and the subject is a pronoun which is ‘weak’ or ‘deficient’, in terms of the typology of Cardinaletti and Starke (1999), consisting of nothing but valued person and number features and an unvalued case feature.

Consider first the case of Agree between T and a lexical subject (i.e. an NP with a lexical root, e.g. hund, nominative hundur,’dog’), in for example, Icelandic (D = definite, PST = past tense, SG = singular). The derivation of the relevant part of (17) is shown in (18).

(17) Hundurinn át fisk
dog,NOM.D ate fish
’The dog ate fish.’

(18) 1. \([TP \{PST, uNr, uPn, NOM, EPP\}] \rightarrow [vP [SG, uCase, D, hund]]\)
2. \([TP \{PST, SG, 3d, NOM, EPP\}] \rightarrow [vP [SG, NOM, D, hund]]\)
3. \([TP [SG, NOM, D, hund]] \rightarrow [TP \{PST, SG, 3d, NOM, EPP\}] \rightarrow [vP [SG, NOM, D, hund]]\)

In Step 1 a tensed T probes for a goal to value its unvalued \(\phi\)-features. The closest goal is the subject DP in spec,vP. In Step 2 Agree has applied: T has copied the number value of the goal. The person feature has got a default 3rd person value, called 3d (on the assumption that lexical DPs do not have a person feature). The subject has had its unvalued Case feature valued NOM by T; by assumption tensed T has a NOM feature. In Step 3, a copy of the subject has been remerged with TP, triggered by the EPP feature of T which is deleted as a result. Eventually, given that the verb is étta ‘eat’, and following copy deletion and other spell-out operations, the derived structure is spelled out as (17).

Consider the case when the subject is a weak pronoun, made up solely of the features [SG, 3, uCase], lacking a D-feature.

(19) 1. \([TP \{PST, uNr, uPn, NOM, EPP\}] \rightarrow [vP [SG, 3, uCase]]\)
2. \([TP \{PST, SG, 3, NOM, EPP\}] \rightarrow [vP [SG, 3, NOM]]\)

\(^{15}\) An alternative idea, also compatible with the theory sketched here, is that the uCase feature is valued by T’s tense feature. NOM would be tense, when copied by a nominal category; Pesetsky & Torrego (2001).
In this case Agree between T and the subject has the result that the features of the subject are (properly) included among the features of T. This is what Roberts (in press, 2009a) refers to as incorporation of the pronoun in T. The distribution of features is the same as if the subject had undergone head-movement, incorporating in T, but no movement has taken place, only Agree. The result is, however, that the subject is formally a copy of T. Still following Roberts, this means that T and the subject form an argument chain, headed by T. This, in turn, has two interesting consequences. First, as a chain, it is subject to ‘chain reduction’: only one copy is spelled out. The principles of chain reduction are controversial: see Nunes (1995, 2004), Bobaljik (2002a), Landau (2006), Trinh (forthcoming). Typically, though, only one copy is spelled out, and typically, though by no means always, the copy spelled out is the highest copy (the one which c-commands the other copies). There are two reasons, in the present case, why T is the copy spelled out: One is that T is the highest copy. The other, more compelling, reason is that T is specified for tense, in addition to valued \( \phi \)-features, so that deletion of T/not spelling T out would violate recoverability. Consequently, the subject copy in spec,vP is not spelled out, deriving a null subject.

Second, since the subject is now the non-head member of a chain headed by T, it cannot satisfy the EPP of T: A constituent cannot both be a part of T and a specifier of T. This means that T, in such a language, must either not have an EPP feature, or the EPP of T can be satisfied by some other category than the nominative subject.

The latter is, of course, what we find in ISc: In the oblique subject construction a non-nominative argument satisfies the EPP. Stylistic Fronting, on the other hand, is when some non-argument category, an adverbial, or a participle, or an adjective, etc. satisfies the EPP. Before going into details, consider the situation in languages of the MSc type, here Swedish.

MSc has no overt subject-verb agreement. I assume that this is a reflex of the lack of unvalued number and person features in T. I will assume, nevertheless, that T has a generalised unvalued feature \([uN]\), which makes T a probe, in Chomsky’s (2000) sense, as required by UG.\(^{16}\) In the case where the subject is a lexical DP, the derivation works as in Icelandic. In Step 1, T probes for a nominal category and finds the subject DP in spec,vP (the

\(^{16}\) On a more general level, the claim is that there is a mutual dependency relation between (finite) T and a nominal constituent in the predicate. T needs an NP, but every NP needs case, so the subject NP needs (finite) T. An alternative formal account of T’s dependency on NP is that finite T has a NOM case which must be discharged (by the ‘inverse case-filter’; see Boskovic (1997)). In this version of the theory, the \([uN]\) feature is not needed.
closest nominal category which does not already have a case value). In Step 2 the uN-feature of T is valued, and the subject’s case-feature is assigned NOM.

(20) 1. [TP T[PST, uN, NOM, EPP] ...[vP [N, SG, uCase, D, hund] \(\rightarrow\)
   2. [TP T[PST, N, NOM, EPP] ...[vP [N, SG, NOM, D, hund] \(\rightarrow\)
   3. [TP [N, SG, NOM, D, hund] [TP T[PST, N, NOM, EPP] ...[vP [N, SG, NOM, D, hund]

In Step 3, the EPP-feature of T has triggered remerge of a copy of the subject with TP. Given appropriate lexical values for the verb and the object, the sentence can be spelled out as

(21) Hunden åt fisk.
    dog.D ate fish

Consider the situation when the subject is a defective pronoun made up of the features [SG, 3, uCase], and, we assume, a feature N, common to all nominal categories.

(22) 1. [TP T[PST, uN, NOM, EPP] ...[vP [N, SG, 3, uCase] \(\rightarrow\)
   2. [TP T[PST, N, NOM, EPP] ...[vP [N, SG, 3, NOM]

Agree applies, T’s uN-feature and the pronoun’s uCase-feature are valued. However, unlike the situation in Icelandic, the features of the subject are not thereby a subset of the features of T, as the inherently valued number and person features of the pronoun do not have counterparts in T. Consequently T and the pronoun do not form a chain in this case anymore than in the case of a lexical subject. The EPP-feature is therefore free to trigger remerge of a copy of the subject pronoun with TP.

The reason why we only ever find movement of the nominative subject, pronoun or lexical DP, to specTP in MSc is, then, the fact that the subject can never be incorporated in T and is therefore always available for the EPP. This follows given that (a) Agree will always establish a relation between T and the closest nominal argument (the subject) anyway, (b) the grammar will always go for the most economical option, (c) few dependency relations is more economical than many, (d) Agree is one dependency relation, EPP is another. It follows that the EPP will always trigger movement/remerge of the subject, ruling out not just Stylistic Fronting, but any construction where a non-nominative category is in spec,TP.
4.6. The ISc/MSc facts in the light of the theory of incorporation by Agree

4.6.1 The null generic subject pronoun

Consider first the null generic subject construction.

(23) a. Hér má ekki dansa. [Icelandic: Sigurdsson & Egerland 2009]
    here may not dance

b. Man må ikke danse her. [Norwegian]
    one may not dance here

‘You mustn’t dance here.’/ ‘Dancing is not allowed here.’

Consider first Icelandic. This is precisely the situation where the subject is a 3rd singular weak, D-less (indefinite) pronoun (Holmberg 2009a,b). The underlying structure is (24).

(24) T
    TP
      T
        PRS  NEG
          uNr
            uPn
              ekki V vP
                NOM EPP
                  mega
                    SG 3 uCase dansa+v dance
                      v'    VP
                        dansa hér here

The subject pronoun is assigned a role by v (or dansa+v). Agree applies, valuing the unvalued uϕ-features of T and the case feature of the subject pronoun. The auxiliary verb eventually moves to T. The negation is a particle (inherently maximal) rather than a head, allowing head movement (see Holmberg & Platzack 1995: 16f.). The locative hér is analysed as a VP-internal constituent (along the lines of Larson (1988)). Alternatively it is an adjunct to VP, or even an adjunct to vP; this will not affect the argument here. The distribution of feature values following Agree will be as in (25):
Here the subject’s features are properly included among the features of T, so the subject is a copy of T, a non-head member of a chain headed by T, and as such is not spelled out. This makes it unavailable for the EPP. The result is that the EPP instead attracts the locative proform hér. Following V-movement of the auxiliary to T, the spelled-out form is (23a). The subject is a null copy of the features which are spelled out as an affix (or as part of a suppletive form) of the auxiliary, in spec,vP. The interpretation of the pronoun is not definite 3SG (not ‘he’, ‘she’, or ‘it’), but indefinite-generic 3SG, corresponding to English one, Norwegian man, French on, etc. This is due to the absence of a D-feature, required for definite interpretation; I return to this point below.

An alternative way to satisfy the EPP is to merge an expletive with TP:

(26) það má ekki dansa hér.

EXP may not dance here

‘One mustn’t dance here.’/’Dancing is not allowed here.’

Now consider Norwegian (23b), the structure of which is (27).
The only unvalued $\phi$-features in T is [uN], valued by the N-feature of the subject pronoun. Thus, even though T assigns NOM to the subject, the pronoun is not a copy of T. Consequently it can be, and in fact, must be attracted by the EPP. The ‘must’ is plausibly a result of the fact that there is already a feature-valueing relation between T and the subject, hence, to minimise structural relations, the EPP of T attracts the subject, not some other category. The subject pronoun remerges with TP. Since there is no D-feature associated with it, it is interpreted as an indefinite-generic pronoun, and is spelled out (in the higher position) as *man*. The result is the string (23b).

The absence of a null generic pronoun in Norwegian and the other MSc languages is a direct consequence of the absence of u$\phi$-features in T which would make possible establishing a chain (T, subject pronoun). Note that complete absence of $\phi$-features from T is not required, to rule out a null generic subject. Given that the pronoun has both number and person, unless T has both number and person features, the pronoun will not be a copy of T, and thus will be accessible to the EPP, and will be spelled out (as the highest member of an A-chain). English is a language which only has unvalued number in T (following Kayne 1989). Thus English, predictably, does not have a null generic pronoun either. The generalisation is: Whenever T has fewer $\phi$-features than a subject pronoun, the pronoun will be overt (although, to complicate matters, it can still be deleted by other mechanisms).\(^{17}\)

No ISc language has definite null subjects, i.e. null ‘I’, ‘you’, ‘he’, ‘we’, etc., except as a result of topic-drop in restricted circumstances (for instance diary-drop; Haegeman (2000)), and, in Old Icelandic/Old Norse, in cases where the null subject is controlled from a higher

---

\(^{17}\) Some languages which do not allow copy-deletion of a referential pronoun in spec,vP, allow control of a null pronoun in spec,TP of a finite clause; see Holmberg (2005, 2009a), Holmberg, Nayudu & Sheehan (2009), Holmberg & Sheehan (2009).
clause (Sigurðsson (1993); footnote 16). Following Holmberg (2009a), I assume this is because these languages do not have a uD-feature in T. In the absence of a uD-feature in T, a definite subject pronoun will always have at least one feature distinct from the features of T, consequently must be spelled out, and is attracted by the EPP. So called consistent null subject languages, where definite subject pronouns are null except when they are focused or when they are shifted topics, are thus characterised by having not only unvalued person and number features in T, but also an unvalued definiteness-feature.

4.6.2. The null expletive

(28) a. Nú rignir. [Icelandic]
   now rains

   b. Nu regnar *(det). [Swedish]
      now rains it
      ‘Now it’s raining.’

The underlying structure of (28a) is the same as (24), without the negation and the auxiliary, and, more importantly, with no theta-role assigned by rignir+v (or, at most a ‘quasi-role’ typical of weather-predicates). After Agree, the subject will, again, be a copy of T, hence a non-head member of a chain headed by T, will therefore not be attracted by the EPP, and will not be spelled-out. The EPP therefore has to either attract another category (the time-adverbial in (28a), or an overt expletive has to be merged.

(29) Það rignir.
    it rains

As in the case of (26), the overt expletive is not ‘the subject’; the subject is a null pronoun in spec,vP.

(30) is not well formed as a declarative, because it violates the EPP.

---

As mentioned in section 3.1., gender/class is another nominal feature. The Scandinavian languages all have nominal gender (a two or three way system), but even ISc does not exhibit subject–verb agreement for gender. In order to maintain the incorporation-by-Agree hypothesis of the null generic pronoun, it must either be the case that T has an unvalued gender feature, even though it is not reflected in morphology, or that deficient, D-less pronouns lack a specified gender feature – since otherwise the the pronoun would always have one more feature than T, preventing copy deletion. See Holmberg (2009b) for discussion.
(30)  *Rignir.
    rains

For reasons which are not entirely clear to me, it is well formed as a yes-no question (‘Is it raining?’). Assuming (a) that T moves to C in yes/no-questions, and (b) that there is a covert question operator in spec,CP, it seems that this null question-operator satisfies the EPP of T.

Consider Swedish (28b): The structure is as in (27) (minus the negation, the auxiliary, and the theta-role assigned by v). Agree does not make the pronoun a copy of T, so it is attracted by the EPP, and gets spelled out in spec,TP. This is the word order seen most clearly in embedded clauses such as (31):

(31) Jag undrar [CP om [TP det T [vP det regnar]]]
    I wonder if it rains

In a main clause, T moves to C (the verb-second rule), and, in declaratives, some phrasal constituent moves to spec,CP. In (28b) the time-adverbial nu is in spec,CP.\(^{19}\) Alternatively, the expletive subject (which really is the subject, in MSc) moves from spec,TP to spec,CP, which yields (32), with det in spec,CP, regnar in C.

(32) Det regnar.
    it rains

4.6.3. The TEC and Heavy Subject Postposing

The analysis sketched in the previous section entails that the expletive is a different sort of element in Icelandic (ISc) and Swedish/Norwegian (MSc): In MSc it is ‘the subject’, a nominal category merged in spec,vP, assigned NOM by finite T, and moved to satisfy the EPP of T. According to Holmberg (2002) the expletive in MSc is merged in spec,vP even in cases when no theta-role is assigned to that position in order to satisfy the EPP of v. Another

---

\(^{19}\) In Icelandic, the corresponding word order is not well formed:

(i)  *Nú rignir það.
    now rains it

Icelandic is a V2 language, too, with T-to-C in main clauses and movement of a constituent to spec,CP. So how come the EPP cannot be satisfied by merge of the expletive in (i), and move of the adverbial directly to spec,CP? Plausibly the reason is that merging of the expletive is a last resort, in Icelandic, applying in the phonological component. If the EPP is satisfied in the syntax by a moved adverbial, then merge of the expletive is not called for. In Swedish the expletive is moved to spec,TP, a narrow-syntactic operation, which may be followed by another narrow-syntactic operation, move of an adverbial to spec,CP.
reason is that T needs to have its uN-feature valued, and needs to discharge the NOM feature (the inverse Case Filter idea). In Icelandic the expletive has no other function than satisfying the EPP of T (and the EPP of C, in the V2 case). Case, agreement, and (insofar as Icelandic has it) EPP of v is taken care of by the subject pronoun in spec,vP, not pronounced because it is the non-head member of a chain headed by T.20

This is why ISc can, but MSc cannot, have the TEC.

(33) a. Það hefur einhver köttur etið mýsnar. [Icelandic: Vangsnes 2002]
   there has some cat eaten the-mice

   there has a cat eaten mice-the

In MSc (here Norwegian) the subject competes with the expletive for case. In Icelandic, and ISc generally, it does not.

It is well established that the position of the lexical subject in the TEC is not spec,vP, but rather some position between vP and T; see Bobaljik & Jonas (1996), Vangsnes (2002). Bobaljik & Jonas (1996) proposed a parameter, the ‘TP-parameter’, which allowed some languages but not others to have the lexical subject appear in this intermediate position. We do not need to resort to such a parameter; the independently established differences in the feature-make up and derivational history of the expletive is sufficient to rule out the TEC in MSc, while ruling it in in ISc.21

For the same reason ISc can, but MSc cannot, have Heavy Subject Postposing.

(34) a. Það lásu hana því ekki margir stúdentar fyrir prófið. [Icelandic]
   there read it thus not many students for exam.the
   ‘Consequently not many students read it for the exam.’

20 Another argument that the expletive in MSc (and also in English), unlike the ISc expletive, is assigned case, is that it can occur in ECM contexts; see Holmberg (2002).

21 Bobaljik & Jonas’s TP-parameter is dubious anyway as an explanation of the contrast in (33) given that the subject in MSc, particularly Norwegian and Swedish, can occur in a position lower than the ‘canonical position’ (here taken to be spec,TP), as discussed by Holmberg (1993). In (i) an indefinite subject occurs to the right of the negation, which is standardly taken to be merged with vP.

(i) Om inte en katt har ätit mössena, vem har då gjort det? [Swedish]
   if not a cat has eaten mice.the who has then done it
   ‘If a cat hasn’t eaten the mice, who has?’
b. *Det läste den således inte många studenter inför provet.\footnote{Swedish} there read it thus not many students for exam.the

In MSc the expletive and the lexical subject compete for nominative case, in ISc they do not. This is sufficient to rule it out in MSc.\footnote{MSc does have presentational sentences with unaccusative or passives verbs, with an expletive and a VP-internal argument. (i) Det har inte kommit många studenter. [Swedish] there have not come many students ‘Not many students have come.’}

4.6.4. Oblique subjects

Next, consider the case of oblique or ‘quirky’ subjects, common in ISc, non-existent in MSc.

\begin{enumerate}
\item a. Jenni líka hestar. [Icelandic] her-DAT like-PL horses-NOM ‘She likes horses.’
\item b. Hun liker hester. [Norwegian] she likes horses
\end{enumerate}

There is no good reason to assume that incorporation of a nominative NP in T is involved in (35a), or other oblique subject constructions. Instead, if there is an object, T typically agrees with the object. A straightforward account of (35a) is that T agrees with the object because T can only agree with a nominal which does not already have a case, and that the EPP attracts the dative NP, assigned its case by the verb along with the theta-role, because it is closer to T than the nominative object.\footnote{This means that there is a ranking of the economy conditions such that locality wins over minimize relations. EPP-triggered movement of the nominal argument which values T’s features would minimize relations but violate locality. An alternative, explored by Boeckx (2000) and Sigurðsson (2006a, 2006b) is that there is a probe-goal relation between T and the oblique subject, such that the oblique subject values the unvalued \textit{person} feature of T, while the nominative object values the \textit{number} feature of T. The overt effect of this is that the nominative object in the quirky subject construction can only ever be 3rd person, which, by hypothesis, means no person. See, however, Sigurðsson and Holmberg (2008), who argue that this is a ‘defective intervention’ effect: The dative, higher argument blocks person agreement but allows number agreement between T and the lower argument, but doesn’t actually value the person feature of T.}
As for MSc, a straightforward reason why MSc does not have oblique, quirky subjects is that it does not have quirky case. However, arguably this does not explain why (36a) is not well formed, in the manner of its Icelandic counterpart (37b) (NEU = neutre, COM = common gender).

(36) a. *Henne blev givet/givna presenter. [Swedish] 
    her was given-SG.NEU/given-PL presents 

    b. Henni voru gefnar gjafir. [Icelandic] 
    her were given-PL presents 

    ‘She was given presents.’

(36a) falls under the generalisation that the EPP in MSc only ever attracts a nominative NP. The two grammatical counterparts of (36a) are (37a,b).

(37) a. Hon blev given presenter. [Swedish] 
    she was given-SG.COM presents 

    b. Det blev givet henne presenter. 
    EXP was given-SG.NEU her presents 

In Holmberg (2009a) I proposed a variant of the theory articulated here which postulates two parameters distinguishing ISc and MSc, accounting for the list of syntactic differences in (6):

(38) a. T has/does not have [uPn] and [uNr]; 
    b. EPP is or is not ‘ϕ-dependent’, where ϕ-dependent EPP can only ever attract the goal probed by the relevant head’s ϕ-features.

In MSc T’s EPP feature is ϕ-dependent, in ISc it is not. This will rule out oblique subjects in MSc, and Stylistic Fronting, as well (discussed below). See Ott (2009) for the same idea, in the context of a theory of Stylistic Fronting.

Note, however, the link between (38a and b): If a head H has [uPn] and [uNr], and an EPP-feature, and the language also has defective (D-less) pronouns which can be merged in the c-command domain of H, then the EPP-feature cannot be ϕ-dependent. It cannot, because
incorporation of the defective pronoun in T cannot be excluded, in which case the EPP has to be satisfied in some other way. Now assume (39).

(39) $\phi$-dependent EPP is the unmarked case.

Given (39), we can maintain that the presence/absence of oblique subjects is the result of variation with regard to the same parameter as in the case of the null subject pronouns, namely (38a). In MSc, where pronoun incorporation is never an option, the unmarked case holds: EPP is $\phi$-dependent. This will rule out any non-nominative subjects, including (36a). The $\phi$-independence of the EPP of T in ISc is not the effect of a separate parameter, but a consequence of other properties of the grammar: (a) T has [uPn] and [uNr], (b) T has an EPP feature, and (c) the lexicon contains defective pronouns. 24

4.6.5. Stylistic Fronting

Stylistic Fronting (SF) is an operation which applies in ISc in clauses where the subject is, for some reason, not available for ‘filling the subject position’. What SF does, is move some other sentential constituent to the vacant position, between C and the finite verb. The constituent may be an adverb, the negation, a participle, a predicative adjective, or a verb complement. According to Holmberg (2000, 2005) the landing site of SF is spec,TP, and the trigger of the movement is the EPP of T; see also Ott (2009). One case where SF is particularly common is subject relatives. 25

(40) [Þeir sem í Osló hafa búið segja að það sé finn bær.]  [Icelandic]  
those that in Oslo have-3PL lived say that it is nice town

‘Those that have lived in Oslo say that it’s a nice town.

---

24 I have suggested above that the EPP feature in T is parametrised, but this is controversial; see Biberauer & Roberts (2006), Massam (2005). It seems plausible enough that the existence of deficient pronouns in the lexicon is parametrised, not dictated by UG, but this is an empirical matter.

25 In this case the PP complement of the verb has undergone SF. An alternative to (41) is (i), where the nonfinite verb has undergone SF.

(i) [Þeir sem búið hafa í Osló] ...
those that lived have in Oslo

It is characteristic of SF that it can move what looks like a head, as well as moving constituents that are clearly phrasal. According to Holmberg (2000) this is due to the ‘phonological’ nature of the movement. See, however, Ott (2009) for a theory where the moved category is always phrasal, but can be a remnant phrase with only a head remaining.
In this case the subject is, by standard assumptions, a null operator, which apparently does not satisfy the EPP, which triggers SF.

Another case is that of impersonal passives, as in (41):

(41) Ef *gengið* er eftir Laugaveginum, ...

if walked is along Laugavegur.the

‘If you walk along the Laugavegur, …’

In this case the past participle has moved to specTP, apparently thus satisfying the EPP.

An alternative is using the overt expletive pronoun það.

(42) Ef það er gengið eftir Laugaveginum, ...

MSc does not have SF in either of these constructions.

(43) a. *De som i Oslo har bott ...

those that in Oslo have lived

b. *Om gått är längs Sveavägen...

if walked is along Sveaväg.the

Consider first the case of the null pronoun in (42), a special case of the generic pronoun. This is straightforwardly a case of incorporation of the subject pronoun in T, as discussed in sections 4.5 and 4.6.1.

(44) \[ T_{[\text{PRS} \equiv \text{uNf}, \text{uPn}, \text{NOM}, \text{EPP}]} \left[ \text{vP} \ 	ext{vera} \left[ \text{vP} \ gengið+ \text{v} \left[ \text{vP} \ [\text{S3}, \text{SG}, \text{uCase}] \ [\text{vP} \ ganga eftir L \ ] \ ] \ ] \right] \left[ \text{be} \ \text{walked} \ \text{walk} \ \text{along L.} \right] \]  

This is the structure (roughly) and the distribution of features before Agree, assuming a VP-shell analysis of the unaccusative VP (the construction involves passivisation of an unaccusative verb, unusual in the context of Germanic languages, but this is unimportant for the present discussion). The subject pronoun, I assume, is in spec,VP. The copies which are
not pronounced at PF are marked as struck out. After Agree and V-movement to T, it is as in (45):

\[(45) \text{er}^+T_{[\text{PRS}, \text{SG}, \text{3}, \text{NOM}, \text{EPP}]} [\text{vp} \text{vera} [\text{vp} \text{gengið}^+ \text{v} [\text{gl}^3, \text{SG}, \text{NOM}] [\text{vp} \text{ganga} \text{eftir L.}]]]\]

The subject’s features are included in T’s features; consequently T and the subject form a chain, and the subject, as the lower (and also less specified) copy, is not spelled out, and is not available for attraction by the EPP. What happens in Icelandic (and ISc generally), is that the EPP attracts the next category down the tree which has a phonological matrix, that is the participle, in this case. The result is the string seen in (41). Alternatively an expletive can satisfy the EPP of T, with the string in (42) as result.

Now consider the case of SF in the subject relative. The gist of the analysis I want to articulate is that the null operator subject of the relative clause in (40) does not satisfy the EPP of T, but instead moves directly to the C-domain (Rizzi (1982), Rizzi & Shlonsky (2007), Holmberg & Hróarsdóttir (2004), Chomsky (2005, 2008). The EPP of T therefore attracts the next available category down the tree, in this case, the PP. As in connection with oblique subjects, this is possible in ISc because the EPP of T must be allowed to attract other categories than the subject NP; see also Ott (2009). In MSc, on the other hand, there is no pressure to deviate from the default, which is to have the EPP strictly attract the NP probed by T, that is the subject. Consequently SF is not found in MSc subject relatives.

Following Chomsky (2005, 2008), we may assume that the operator subject in a subject relative without SF undergoes two independent movements, one A-bar movement to spec,CP, one A-movement to spec,TP, the latter to satisfy the EPP of T. The copy in spec,TP is not spelled-out because it is c-commanded by the copy in spec,CP. The copy in spec,CP is not spelled out either (in Icelandic), presumably because it is c-commanded by the head of the relative.\(^{26}\) Ott (2009) notes that the optionality of SF in subject relatives (see (46)), always a problem for extant theories of SF, can then be viewed as optionality of A-movement by the operator subject.

\(^{26}\) This is if we assume that the category that moves is a pronominal operator, which is linked with the head of the relative by an additional operation. Alternatively the category that moves is the head of the relative (Kayne 1994). Either theory is compatible with the analysis of SF proposed here, and below.
(46) [Þeir sem hafa búið í Osló] segja að það sé finn þær.

those that have lived in Oslo say that it is nice town

If A-movement applies, no SF is called for (or even possible), if A-movement does not apply, SF must apply, to satisfy the EPP.

If this is true, then the role of the \( \phi \)-features of T in relation to SF in subject relatives is indirect, just as in the case of oblique subjects: The richness of the \( \phi \)-features in T means that the EPP cannot be restricted to attracting the probed subject, making SF possible. The following is an attempt to link SF in subject relatives more directly to the richness of \( \phi \)-features in T.

Assume that the null operator has the structure (47): a tree made up of an OP feature and a bundle of \( \phi \)-features, including case.

(47) \[
\text{[ OP [\( \phi \)Nr, Pn, uCase]]}
\]

The structure of the relative clause in (40), repeated here as (48a), is roughly (48b), prior to Agree

(48) a. Þeir sem í Osló hafa búið

those that in Oslo have lived

b. \[
[CP C [TP T[PRS,uNr, uPn, NOM, EPP] [VP hafa [VP [ OP [\( \phi \)Nr, Pn, uCase]] buið í Oslo]]]]
\]

To be interpretable, OP must move to C. Assume that C has the required features [uOP, EPP] attracting OP, and, crucially, that OP can be detached from the \( \phi \)-feature bundle, moving without pied-piping the \( \phi \)-feature bundle. The \( \phi \)-features of the subject, on the other hand, enter into Agree with T. As they are included in T’s features, a chain is formed. The EPP, therefore cannot attract the \( \phi \)-feature bundle. And given that OP moves directly to the C-domain, the EPP of T is left unsatisfied. The solution is SF: The EPP triggers movement of the next available category down the tree.

We can now assume that (46) is the result when OP in Icelandic is not detached from the \( \phi \)-feature bundle. In that case, the subject cannot be incorporated in T, and thus it is available for movement and remerge with TP, satisfying the EPP.
In this scenario, there are two preconditions for SF in subject relatives: OP must be able to detach from the \( \phi \)-feature bundle, and (b) T must have sufficiently articulated \( u\phi \)-features to incorporate a subject pronoun. We know that MSc does not meet (b). Since this is sufficient to rule out SF in this construction, whether it meets precondition (a) or not will not make a difference.

4.7. Summary
The seven differences between ISc and MSc I started out with, a reduced version of P&H’s list, have been explained as effects of variation with regard to \( \phi \)-features in T: ISc has \([uPn]\) and \([uNr]\), MSc has neither. This is morphologically reflected in the agreement on the finite verb. It causes incorporation, in the sense of Roberts (2009a, in press) of a defective, D-less pronoun. A direct result of this is the null generic pronoun and the null expletive. An indirect result is the pure expletive, merged directly with TP, which, in turn, makes the TEC and Heavy Subject Postposing possible. Another indirect result is the possibility of oblique subjects and SF.

The parameter is the choice of unvalued nominal features in T, the same parameter which determines (certain forms of) the null subject property (as intimated by Platzack (1987)). This is a ‘deep parameter’, in that there is a range of surface effects associated with the choice, thus vindicating the theory in P&H, even though the range of effects is reduced somewhat, compared to that theory. The parameter conforms broadly to the parameter schemata of Gianollo, Guardiano & Longobardi (2008) and Boeckx (forthcoming): UG provides for the possibility, but does not dictate, that a given grammatical feature has an unvalued counterpart, and also does not dictate the distribution of the unvalued feature, say, whether it is employed for the T-NP relation in the clause, or the Adjective-NP relation in the noun phrase.

A difference between the theory articulated here and P&H (which can be ascribed to the minimalist program) is that another layer of explanation has been added. It is not sufficient to postulate a feature \([+\text{pronoun}]\) in T; The question is why? The answer (at least the beginning of an answer) is that dependency-relations are established by feature-valueing. In the case at hand, the dependency involves NP, so nominal unvalued features need to be enlisted. The derivation of pro-drop has also been sharpened: The relation between the ‘licensing head’ and the null argument is a special case of Agree, that is when the features of the argument are a
subset of the features of the probing head (only the case of indefinite and non-referential null subjects are discussed in any detail in this paper, though).

5. Effects of the \( u_\phi \)-feature parameter in other languages

If the cluster of properties (6) is the effect of the choice of unvalued features in \( T \), then other languages which have made the same choice as ISc ought to exhibit the same cluster of properties, all else being equal. As discussed in section 2, given the large number of parameters we must assume, it is unlikely that all else will ever be equal, except in the case of very closely related languages. This, unfortunately, can make it difficult to either support or falsify a proposed parameter, by means of comparison with more distantly related languages; see Roberts & Holmberg (2009). Apart from agreement morphology, the properties of the cluster (6) are not the most easily observed. It is not, for example, possible to check the predicted correlations using the *World Atlas of Linguistic Structure* (Haspelmath & al. 2005), in its present form. Whether a language does or does not have a null generic pronoun is not the sort of information which is regularly included in descriptive grammars. And although the occurrence of oblique cases is easily observed, it is notoriously difficult to determine whether an initial oblique experiencer NP is a subject in specTP or a topicalised object, and more generally, it can be difficult to determine the syntactic status of fronted categories; see Sigurðsson (2002, 2004a), Holmberg & Nikanne (2002). It thus requires detailed investigation of a language to determine whether it does or does not exhibit the cluster (6).

A language which has all the ISc values for the cluster (6), except, arguably, SF, is Finnish.

5.1. ISc-like properties of Finnish

5.1.1. Agreement morphology

The verb paradigm morphologically distinguishes 3 persons and two numbers in all finite tenses and moods, in Standard Finnish and some spoken varieties. Other spoken varieties exhibit syncretism between 3PL and 3SG.
5.1.2. Oblique subjects
Finnish has oblique subjects in several constructions. One of them is the so called neessive
construction, exemplified in (50) (GEN = genitive, PX = possessive suffix); see Laitinen &

(49) Sinun ei pidä ottaa itseäsi liian vakavasti
you-GEN not should take self-PX.2SG too seriously
‘You shouldn’t take yourself too seriously.’

5.1.3. Stylistic Fronting
Although Finnish does not have an operation with the precise properties of SF in ISc, it
allows movement of non-nominative categories to specTP (the ‘EPP-position’) in active
finite clauses (see Holmberg & Nikanne 2002). In (51), the position to the right of the verb
bearing the question particle excludes an analysis where the fronted object is in the C-
domain. The verb form is active.

(50) Onko tämän kirjan kirjoittanut Graham Greene?
has-Q this book-ACC written Graham Greene-NOM
‘Has Graham Greene written this book?’/’Is this book written by Graham Greene?’

SF in ISc can front a surprisingly wide variety of categories, subject only to locality:
adverbs, the negation, predicative adjectives, non-finite verbs, verb particles, PPs, some NPs,
but not VPs or clauses (Holmberg 2000, 2005). As discussed by Holmberg &Nikanne (2002),
Finnish allows fronting of any category which can function as topic: thus NPs and PPs, but
not adverbs, predicative adjectives, or non-finite verbs. Why this difference? In the case of
Finnish, another parameter enters the picture: Finnish is a ‘discourse-configurational
language’ (Vilkuna 1995, Miyagawa, in press), that is to say, it expresses information-
structural functions primarily by movement in the syntax. This may provide a principled
reason why Finnish does not attract the same range of non-nominative categories as ISc, to
spec,TP.
5.1.4. Null expletive and null generic pronoun

Finnish has a null expletive with weather predicates, and a null generic pronoun; see Holmberg & Nikanne (2002), Holmberg (2009b).

(51) Nyt sataa.
    now rains
    ‘It’s raining now.’

(52) Täällä ei saa polttaa.
    here not may smoke
    ‘One mustn’t smoke here.’

5.1.5. The TEC: Finnish also has an overt ‘pure’ i.e. case and $\emptyset$-featureless, expletive sitä, which can be combined with a transitive predicate, as in (53), which yields a construction very similar to the ISc TEC (Holmberg & Nikanne 2002).

(53) Nyt sitä ovat nämäkin lapset oppineet uimaan. (Holmberg & Nikanne 2002)
    now EXP have these-too children learnt swim-INF
    ‘Now these children, too, have learnt to swim.’

The placement of the lexical subject in construction with the expletive is rather free, more so if it is heavy, again not unlike Icelandic.

(54) Sitä ei ole oppinut uimaan vielä kovin monta lasta.
    EXP not have learnt swim-INF yet that many children
    ‘Not many children have learnt to swim yet.’

As mentioned, P&H supported their theory by data from Middle English, Old French, and Yiddish, all of which exhibited at least parts of the criterial cluster of properties characterising ISc. However, Middle English and Yiddish are Germanic languages, and thus close relatives of Scandinavian, while Old French was in close contact with Germanic languages. This means that the ISc-like properties observed in these languages may just be accidental and unrelated remnants of an earlier proto-Germanic grammar, or, in the case of
Old French, accidental and unrelated loans induced by language contact (a point made by Newmeyer 2006b).

As a Finno-Ugric language, Finnish is unrelated to Germanic. It has been in close contact with Germanic, though, ever since the days of Proto-Germanic (as evidenced by a number of loan words; Koivulehto (1984)). The historical relations among the languages of Northern Europe (Germanic, Baltic, Slavic, Finno-Ugric) are complex, and controversial (see Wiik 1997, 2000). We cannot, therefore, exclude the possibility that some syntactic similarities between ISc and Finnish are the result of ancient language contact (ancient, because in the last 800 years or so, Finnish has been in close contact only with Swedish, among the Germanic languages, which in this period has evolved into a modern MSc language). However, closer analysis of the constructions (49)-(53) does not offer much support of the hypothesis that they are independently borrowed from Germanic/Scandinavian (or vice versa, that the ISc constructions (6) are independently borrowed from Finnish). For example, the necessive construction is found in Finnish, Estonian, and the Baltic language Latvian (see Hedlund 1990), but is unknown in Germanic, so at least that construction has not been taken over as such from Germanic. If it is an effect of ancient language contact, the effect must be of a more profound nature, affecting parameters, not just constructions.

As for the expletive sitä, there is no reason to think that it would be an ancient loan from Germanic, since ancient Germanic probably did not have an expletive. In ISc it is a recent phenomenon (Rögnvaldsson 2002). 27

Even if we accept that the cluster (50)-(54) in Finnish is an independent development, the occurrence of a particular cluster in one unrelated language is obviously not enough to convince a skeptic that deep parameters are real. Unfortunately we are unlikely to find many other languages that would have the complete cluster. Consider the properties that Finnish happens to share with ISc, in addition to rich agreement in T, which are crucial for this cluster:

(a) morphological case, a condition for oblique subjects;
(b) partial pro-drop, a condition for a null generic pronoun (Holmberg (2009a) and section 4.6.1.);
(c) an EPP-feature in T; it is not a V-initial language (Holmberg & Nikanne 2002);
(d) SVO order; SF is difficult or impossible to detect in an SOV language;

27 There is also little evidence that it would be a more recent loan from Swedish. In particular, the regional distribution of the expletive does not support such an origin (Urpo Nikanne, p.c.).
(e) a pure expletive; a precondition for the TEC (Holmberg & Nikanne 2002).

This underscores the observation that we cannot expect even moderately complex clusters such as (6) to reappear in toto in languages which are not closely related; there are quite simply too many interfering parameters. 

6. Conclusion

It may be true that the complexity of the task of characterising the system of parameters underlying syntactic variation was sometimes underestimated, in the early days of the P&P program. I hope to have demonstrated, nonetheless, that the arguably modest success of the theory, relative to the most inflated expectations, does not warrant the conclusion that there are no parameters, only language-particular rules (Newmeyer (2004, 2006a), or only unconnected superficial parameters (Boeckx (forthcoming), Haspelmath (2008)). I hope to have demonstrated, also, that some of the results achieved in the ‘GB-era’ based on P&P theory, have stood the test of time, testifying to the soundness of the method of identifying a cluster of properties by comparison of a set of languages, preferably closely related, in order to control for irrelevant variables, and deriving the properties of the cluster from a parameter setting.

The parameter has a certain depth, in the sense that the choice of $u\phi$-features for T has a range of surface effects. In the big scheme of things it is more of a microparameter than a macroparameter, though. In the theory outlined in Roberts & Holmberg (2009), the parameter would be embedded in a hierarchy of parameters to do with $u\phi$-features, with the most general parameter (a macroparameter) at the root, and progressing down to increasingly specialised parameters. The most general parameter is the choice between having or not having $u\phi$-features in the grammar, distinguishing between languages without any agreement (Japanese, Korean, Thai, etc.) and other languages; see Miyagawa (in press). The next parameter distinguishes between languages that have a full set of $u\phi$-features on all probes (consistently head-marking languages), and other languages. At the next level down the

---

28 In Holmberg, Nayudu & Sheehan (2009) two other partial pro-drop languages, Marathi and Brazilian Portuguese, are compared with Finnish. They each exhibit part of the cluster (6). Both have a null expletive and a null generic pronoun. Marathi has ergative and dative subjects. Brazilian Portuguese does not have any clear cases of oblique (pronominal) subjects, but does have PP in specTP; see Avelar & Cyrino (2008), Holmberg, Nayudu & Sheehan (2009: p. 74) (thanks to Sonia Cyrino, p.c.). Neither of them have an expletive pronoun, though, so we do not expect to see a TEC or heavy subject postposing.
hierarchy we find parameters like the one discussed in this paper, specifying a subset of $\phi$-features, encoded on specific probes.29

As discussed by Chomsky (2007), Berwick and Chomsky (forthcoming), Boeckx (forthcoming), Boeckx & Hornstein (2009), the minimalist conception of UG is in part different from the traditional P&P conception. In the minimalist program, in part as a consequence of taking phylogenetic evolution into account (Hauser, Chomsky & Fitch 2000), a central concern is to reduce the complexity of UG, while in traditional P&P theory complexity of UG was not seen as a problem, as such. It is notable, however, that this change of perspective does not, in itself, invalidate the aims, or the methods, or the results achieved in the GB-era (a point also made by Boeckx & Hornstein (2009)), nor is it inconsistent with P&P theory, once parameters are seen as points of underspecification, where, crucially, the alternative specifications are always strictly limited in number. And the question how syntactic variation is restricted is as relevant as ever; the logical problem of language acquisition has not gone away.

Finally, I would like to give an indication of what would count as an empirically based counterargument against the P&P theory of linguistic variation. We have seen that MSc, ISc, and Finnish differ with regard to which categories can satisfy EPP of T. It is conceivable that investigation of a wider range of languages will show that the combinations of categories that can satisfy the EPP across languages are basically random: In some languages AP and PP, in other languages the negation and NP, in yet other languages participles, the negation and AP, and so on. If so, this would be a strong counterargument against P&P theory. The number of alternative systems would still logically be restricted by the range of syntactic categories provided by UG, but the number of combinations would be so large as to render a parametric account pointless. Then we would be justified in saying that the variation is a matter of language-particular rules.

References


---

29 In defence of macroparameters, see also Baker (2008).
Language? Chomsky's minimalism and the view from syntax and semantics (pp. 31-60). Berlin: Mouton de Gruyter.


NEWMEYER, F. (2006b) A rejoinder to ‘On the role of parameters in Universal Grammar: a reply to Newmanr’ by Ian Roberts and Anders Holmberg. people.pwf.cam.ac.uk/mtb23/NSP/Rejoinder%20to%20R&H.pdf


47


