

Null subjects and the EPP

Towards a unified account of *pro*-drop*

Christopher Johns (University of Durham)

Holmberg (2003) shows that the account of *pro*-drop in languages with rich agreement given in Rizzi (1986) is incompatible with current versions of checking theory (Chomsky, 2000, 2001), since uninterpretable features, specifically the ϕ -features manifested in verbal agreement, enter the derivation unspecified, and therefore cannot identify content of the null subject. If verbal ϕ -features were interpretable, as argued by Alexiadou & Anagnostopoulou (Alexiadou and Anagnostopoulou, 1998) among others, then they should be able to cooccur with an overt expletive. The Finnish example (1) shows that this is not possible, leading to the conclusion that the subject position is occupied by a fully specified pronoun which is either deleted at PF or just not spelled out.

- (1) (*Sitä) puhu-n englantia
EXPL speak-1SG English
'I speak English'

As it stands, this model provides no account of the relationship between the grammaticality of null subject constructions and rich agreement, since the valuing of the verbal ϕ -features and the deletion of the pronoun take place at two distinct stages in the derivation. Holmberg suggests that deletion of the pronoun is “restricted by a constraint on processing, namely ‘recoverability of deletion’” (Holmberg, 2003:20). Roberts (2003), on the other hand, maintains a syntactic representation of the relationship, by proposing deletion of the pronoun to be possible only if identity of features with the T-head obtains. By stipulation, any language with five or six distinct agreement morphemes has an uninterpretable D-feature on T and therefore allows null subjects. A language with fewer agreement forms does not have this feature, and deletion is therefore not possible.

The analysis proposed here follows Roberts in taking the relationship to be syntactic, but seeks to reduce the stipulative element to general, independently motivated properties of the grammar. Following Alexiadou & Anagnostopoulou (1998), verbal ϕ -features in rich agreement languages are taken to be lexically valued, but, at variance with their account, uninterpretable, as long as they remain attached to a head in the verbal domain. Crucial to this model is the possibility of feature movement¹, enabling syntactic and semantic features merged in positions in which they

* This research is supported by a grant from the AHRB (AR14458).

¹ Whether movement is a single operation or a combination of copying and deletion will not be discussed here, since the account offered is compatible with either version.

are not interpretable to move independently of their associated phonetic matrix to positions in which they are. This is shown to explain why the null pronoun *pro* may not be merged in the subject position of languages with no agreement. The obligatoriness of overt subjects in some languages with rich agreement (e.g. Icelandic) is shown to be due to an EPP-condition requiring the subject position to be filled at PF, and the possibility of null subjects in some but not all languages with no agreement is argued also to be due to differences between the EPP-conditions on the subject positions in these two groups of languages. This formalises the correlation between discourse *pro*-drop and the grammaticality of bare NP arguments observed by Tomioka (2003). By accounting for the traditionally awkward cases in this way, the grammaticality of null subjects is shown to fall out from an interaction of general principles and different formulations of the EPP, removing the need for a *pro*-drop parameter as such at all.

Feature interpretability as a derivative property

In Chomsky's (2000, 2001) model of agreement, features enter the derivation either with a lexically defined value or unvalued and hence uninterpretable. While being interpretable and therefore a legitimate LF-object implies that a feature is valued, the reverse is not true: in the course of the derivation, initially unvalued features may have a value assigned to them by a lexically determined feature, but this does not render them interpretable and they must still be deleted, typically under identity with the very feature that has valued them. Neither the existence of lexically unvalued features in the lexicon nor their valuing in the course of the derivation makes any positive contribution to LF-interpretability and their role in the syntax is motivated purely empirically. Clearly, a descriptively adequate theory that does not rely on such constructs is preferable to one that does.

A related question concerns how the set of uninterpretable features associated with a given head is learned. One possibility is that this is innately determined, but in its strongest form this is purely stipulative and fails to capture the generalisation that features which restrict the denotation of the head to which they are attached are interpretable, while those that do not are uninterpretable. It would furthermore run into difficulties explaining the wide range of agreement phenomena encountered cross-linguistically, since these appear to extend beyond the domain of the purely syntactic. While grammatical ϕ -features are realised on at least a subset of the nominal lexical entries (typically pronouns) and can thus be argued to be represented syntactically at some point in any nominal projection, it is difficult to make this claim for some cases of classifiers in the verbal domain, illustrated by the Waris sentence in (2).

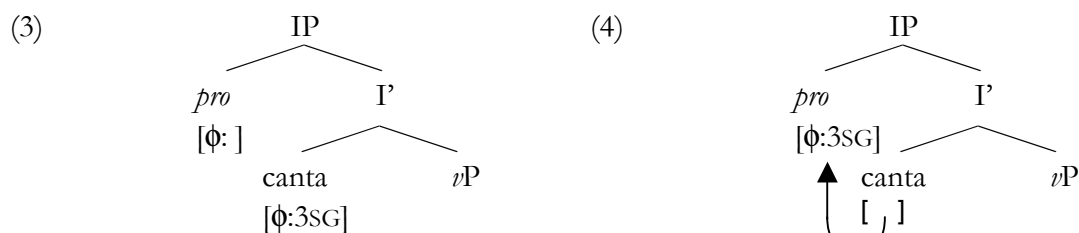
- (2) Nelus ka-m ninge-ra-ho-o
greens 1SG-DAT CLASS_[leaf-wrap]-get-BEN-IMP

‘Give me some greens (cooked in their leaf wrapper)’ (Brown (1981:95ff), quoted in Benedicto (2004))

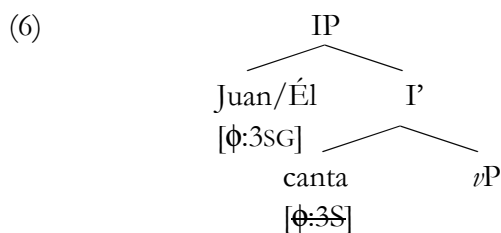
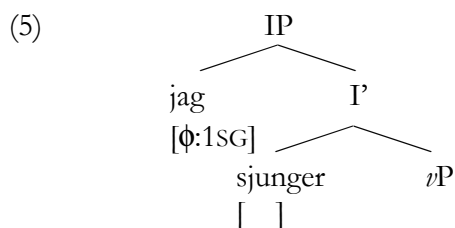
The classifier *ninge*, although realised on the verb, restricts the denotation of the noun *nelus*, but to claim that it duplicates some part of the featural composition of the noun would entail positing a series of null morphemes on the noun, each corresponding to a different way of cooking greens! A generic classifier feature capable of deleting any classifier in the verbal domain does not solve the problem either, since the semantic information associated with the classifier would then be lost before the derivation reaches the LF-interface. The solution to be adopted here involves movement of such features from the position in which they are not interpretable to one in which they are, and the hypothesis is explored that this holds of agreement generally. Apart from the empirical benefits this hypothesis has for the treatment of null subjects, it has the advantage for the learnability problem of enabling the featural composition of lexical items to be learned at face value by the child: for the following discussion this means that verbs in languages with rich agreement may be assumed simply to have valued ϕ -features rather than an unvalued [$\mu\phi$:] in the lexicon.

Pro-drop in rich agreement languages

A consequence of this approach is that the ϕ -features of the phonetically empty pronoun *pro* cannot be lexically valued, since they may take on a range of values (typically a subset, proper or otherwise, of the range of possible combinations of person, number and gender features available in the language). It is clear, however, that these features restrict the denotation of the null pronoun and that their presence is in some sense obligatory if it is to receive an interpretation. This can be formalised by taking the ϕ -features of *pro* to be lexically specified as [ϕ :]. The Spanish sentence *canta* ‘He/She sings’ is thus taken to have the structure shown in (3).



As it stands, this structure will crash at LF, since the ϕ -features of *pro* are unvalued and those of *canta* uninterpretable in their current position. However, if the features [3SG] move to the subject position (4), *pro* has the ϕ -values it requires, no uninterpretable features are left on the verb and the derivation can converge. In a language with no agreement and obligatorily overt subjects (such as Swedish), such movement is not necessary, since the ϕ -features are already in a position in which they can be interpreted (5).



It is less clear how languages with obligatorily overt subjects and rich agreement or indeed sentences with overt subjects in *pro*-drop languages (6) should be dealt with in this model, since the ϕ -features of the subject position are already valued and there is nowhere for those of the verb to move to. Roberts' (2003) definition of Agree (7) and condition of deletion under feature identity (8) provide the starting point for one possible solution.

(7) α Agrees with β iff α c-commands β and α and β are non-distinct in features (and there is no category γ equivalent to α which c-commands β but not α);

(8) α deletes under identity of features with β only if β Agrees with α .

This definition facilitates the deletion of whole lexical items (pronouns) just in case they have the same features as a head (agreement inflection) which they locally c-command. Rich agreement languages are assumed to have a D-feature on I, ensuring the identity of features between the verb and its subject which facilitates the deletion of the subject. This kind of deletion is crucially different from that which takes place in the normal course of a derivation, in that it involves the deletion of a bundle of interpretable features under identity with uninterpretable ones rather than the reverse, and as such renders the account less appealing than one in which only uninterpretable features are deleted. The analysis proposed above, apart from relating the grammaticality of null subjects directly to the ϕ -featural composition of the verb rather than by means of an additional D-feature, allows just this. In (6), the identity of the interpretable ϕ -features of the subject and the uninterpretable ϕ -features of the verb enables the deletion of the latter, ensuring LF-convergence only in the event that the subject and verb agree. At the stage of the derivation at which the subject is merged, a language with rich agreement such as Italian therefore has two possibilities: either an overt subject with ϕ -features matching those of the verb, or the phonetically empty pronoun *pro* must be merged. Merger of *pro* is not an option in a language lacking agreement, since there is no means of fully specifying its ϕ -features.

Obligatorily overt subjects with agreement morphology

In common with any analysis proposing a direct relationship between agreement and the grammaticality of null subjects, the account proposed here runs into difficulty when faced with languages with rich agreement requiring overt subjects (such as Icelandic) or which allow null subjects despite the complete absence of verbal morphology (so-called discourse *pro*-drop

languages). Given the fact that the Icelandic paradigm is by any standards as rich as that of the Romance *pro*-drop languages, this problem cannot be solved by reference to agreement alone and interference from an EPP-requirement of such languages must be invoked. The fact that the combination of obligatory overt argument and morphological agreement appears only ever to occur in case of subjects (cf. Gilligan, 1987:195 & 210) lends credence to this hypothesis, since EPP-requirements are typically formulated as conditions on the content of subject positions.

An EPP-requirement must also be assumed to account for the obligatoriness of null subjects with the inflected forms of degenerate paradigms (such as first and second person plural in French), distinguishing such languages from those with similarly defective paradigms that allow null subjects only in the presence of unambiguous agreement. It might be objected that this introduces a degree of redundancy into the system, since the presence of an overt pronoun in the subject position of a language such as French is rendered obligatory by both the EPP and the lack of overt morphology in the case of verbal forms without agreement (singular and third person plural). However, it is precisely this overlap which predicts that the subject position should be a locus of interaction between the null subject parameter and the EPP, a hypothesis borne out in the loss of null subjects in the history of French. According to the model presented in (3) to (6), loss of distinct verbal forms should lead to the obligatory use of overt pronouns in contexts where the morphology is ambiguous. The concomitant increase in overtly filled subject positions might subsequently be reanalysed in the course of language acquisition as being due to an EPP-requirement absent from the grammars of earlier generations, leading in turn to overt subjects in all contexts, regardless of verbal morphology. In the case of Icelandic, the continued presence of overt pronouns in the primary linguistic data available to the child ensures that an EPP-condition continues to be acquired, regardless of the presence of rich agreement. It does predict, however, that if subject pronouns should ever become optional in Icelandic, they would be unlikely to become obligatory again, without concomitant loss of agreement morphology, which perhaps provides insight into why languages of this type are comparatively uncommon.

Discourse *pro*-drop

It is possible to adapt the proposed analysis to accommodate languages which allow null subjects despite a complete lack of agreement, although doing so raises problems of a different kind. It has hitherto been assumed that sentences such as (3) are uninterpretable at LF on two counts: the subject *pro* does not have the ϕ -features it requires in order to be interpretable, while the verb does have ϕ -features which are uninterpretable as long as they are not attached to a nominal head. However, relaxing the first of these requirements to allow *pro* to host ϕ -features without stipulating that it must is still sufficient to derive the grammaticality of the Spanish examples and

furthermore enables discourse *pro*-drop to be assimilated into the same theory: since Mandarin verbs do not agree with the subject, there are no ϕ -features needing to escape from the verbal head to ensure LF-convergence.

This hypothesis appears to be wrong for a number of reasons. Firstly, it requires an alternative account of why *pro*-drop is not possible in other languages with no agreement, such as Mainland Scandinavian. This problem does not at first appear insuperable, since it has been argued above that an EPP-requirement is only indirectly dependent on lack of verbal ϕ -features, and the fact that Mainland Scandinavian has obligatory overt expletives is incontrovertable evidence for an EPP in these languages, but the analysis falls down when faced with defective paradigms which require overt subjects only in those parts of the paradigm where there is no agreement. Assuming the EPP-requirement to be a condition on the properties of a structural position, it cannot be taken to apply selectively according to the ϕ -features of another head, except by stipulation. Secondly, completely divorcing the possibility of null subjects from the presence of rich agreement in this way precludes the analysis offered for the history of French, which depends crucially on the loss of agreement morphology leading to an increase in the use of overt pronouns. Finally, relaxing the requirement that the featural composition of *pro* include ϕ -values at LF admits the possibility of a language with degenerate agreement allowing null subjects throughout the paradigm, the verbal ϕ -features (where present) moving to *pro* to become interpretable, and no movement taking place otherwise. Such a paradigm is not possible if *pro* is required to have ϕ -values. Whether or not such languages exist is a question that will not be addressed here.

Discourse *pro*-drop thus continues to pose a problem, since there is no means of valuing the ϕ -values of the *pro* in the subject position. Various solutions to this are conceivable: the *pro* in such cases may be different from that found in agreement *pro*-drop languages in not requiring ϕ -values or allowing its ϕ -values to be anaphorically or discourse bound, or there may simply be no *pro* present. Whatever the correct formulation, it is clear that it must be formulated as a condition on the subject position (i.e. an EPP-requirement) and must vary cross-linguistically. This conclusion is not unwelcome, since the EPP is known to be satisfied in widely differing ways in different languages, either as a PF-condition (simply requiring the subject position to be overtly filled) or as an LF-condition (requiring subjects to have specific properties such as definiteness or referentiality), or a combination of the two (requiring, for example, referential subjects always to be overt). If there could be shown to be a correlation between the EPP-conditions of a language type and its behaviour with respect to *pro*-drop, this could provide the basis of a solution to the apparent anomaly of discourse *pro*-drop.

Tomioka's Discourse *Pro*-drop Generalisation, which states that "[a]ll languages which allow discourse *pro*-drop allow (robust) NP arguments" Tomioka (2003:336) constitutes such a correlation. Expressed in the terms of the discussion above, the EPP of a language such as Japanese requires only an NP rather than the full DP required in agreement *pro*-drop and non-*pro*-drop languages in the subject position. If, as is commonly assumed, the D-head is the position where ϕ -features, and hence pronouns both overt and null are hosted, then the absence of this head relaxes the condition on null subjects that they be licensed by rich agreement, since, if there is no *pro* in the subject position, then there are no ϕ -features requiring values. This is not an option in Scandinavian, if subjects are always DPs, since the ϕ -features of the D-head will remain unvalued unless an overt pronoun is merged. Neither is this possible in rich agreement languages, since there would be no head to which the verbal ϕ -features could move to become interpretable.

Consequences and Conclusions

The most significant consequence of adopting this model is that *pro*-drop ceases to be a parameter, at least in the traditional sense of being a property that the child selects from a range of options on the basis of limited primary linguistic data. Instead, the possibility of agreement *pro*-drop follows from independently motivated properties of feature interpretability, and the (presumably universal) availability of a fully underspecified D-pronoun *pro*. This admits the possibility of a mixed paradigm with *pro*-drop allowed only in those cases where the agreement can identify the subject, something difficult to account for if a language has a single setting for each parameter. The possibility of languages with obligatory overt subjects despite rich agreement and discourse *pro*-drop languages is admitted by reference to the EPP-condition, which, being a property of a syntactic position, may not apply selectively according to the features of individual lexical items. (That the EPP may be satisfied differently in different languages is hardly controversial, indeed so great is the cross linguistic variation that it is tempting to speculate that more than one parameter may be involved.) This precludes the possibility of a language with degenerate agreement allowing null subjects throughout the paradigm, a prediction that seems to be borne out empirically. It furthermore predicts the correlation between discourse *pro*-drop and bare NP arguments identified by Tomioka.

The different forms the EPP may take have not been discussed in detail here, but it is clear that there are a number of unresolved issues, not least precisely what it means for a language to project a DP/NP-subject, since languages with determiners may allow bare NP arguments under certain semantic conditions (Tomioka identifies plurals and mass nouns in English), while discourse *pro*-drop languages allow subjects with determiners such as demonstratives. If this is a matter of parametric choice rigidly admitting only one setting, it may simply be that the relative

frequency of DP and NP subjects leads a child to analyse its language as having NP subjects with optional Ds or DP subjects with null determiners. Such assertions can, however, be no more than speculative at this stage.

References

- Alexiadou, Artemis, and Anagnostopoulou, Elena. 1998. Parametrizing AGR: Word order, V-movement, and EPP-checking. *Natural Language and Linguistic Theory* 16:491-540.
- Benedicto, Elena. 2004. Non-traditional uninterpretable D features in the verbal functional domain: Parasitic D[class]. Paper presented at *GLOW*, Thessaloniki.
- Brown, Robert. 1981. Semantic aspects of some Waris predications. In *Syntax and Semantics in Papua New Guinea Languages*, ed. K.J. Franklin, 93-123. Ukarumpa: S.I.L.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, eds. R. Martin, D. Michaels and J. Uriagereka, 89-156. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Gilligan, Gary. 1987. A Cross-Linguistic Approach to the Pro-drop Parameter, University of Southern California.
- Holmberg, Anders. 2003. Null subjects and uninterpretable features: Evidence from Finnish. *Durham Working Papers in Linguistics* 9.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17:501-557.
- Roberts, Ian. 2003. Some Consequences of a Deletion Analysis of Null Subjects. Paper presented at *Meeting of the North-East Syntax Symposium*, York.
- Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In *The interfaces. Deriving and interpreting omitted structures*, eds. K. Schwabe and S. Winkler. Amsterdam: John Benjamins.