

Null Subjects and and Structure of Parametric Theory

This is a research project funded by the AHRB (AR14458) over 5 years starting 1 October 2002. It is a joint project between the Universities of Cambridge and Durham, led by Anders Holmberg (School of Linguistics and Language, University of Durham), and Ian Roberts, and David Willis (Department of Linguistics, University of Cambridge).

Project description (short version)

The null-subject parameter will be investigated in order to establish whether the correlations that it has been claimed to account for can be explained in terms of parameter interaction. The aim is to develop a typology of linguistic correlations and a theory of parameter interaction.

Project description (long version)

Research Context.

If we follow Chomsky in postulating Universal Grammar (UG) as a solution to the logical problem of language acquisition, then we must provide an adequate account for the attested range of variation among grammatical systems. Over the past twenty years, this has been done by appealing to the notion of “parameters of UG”. Associated with some aspects of the invariant UG are parameters specifying a limited range of variation. For example, the hierarchical structure and categorical nature of lexical categories such as the Verb Phrase (VP) may be fixed by UG, while the ordering of the verb and its complements is a matter for parametric variation: this may account for the existence of OV and VO languages.

Research Questions.

While it has often been observed that principles and parameters (P&P) model of comparative syntax offers a way of understanding typological correlations, the theory of parameters has not been subjected to the systematic and in-depth investigation it deserves. As a result, certain basic questions concerning the nature of parametric variation have not been adequately addressed. An important question concerns the matter of parameter interaction.

It is standardly assumed that a given parameter setting may have effects in different parts of the grammar, and therefore be associated with a cluster of grammatical properties. Correspondingly, clustering of properties across languages is taken to be diagnostic of a parameter. However, to what extent the cluster of properties is actually realized in any given language depends on how other parameters are set. Ultimately the form of a given language is the result of the interaction of the entire set of parameters and principles of grammar. We envisage essentially two types of relations among parameters:

(1) A parameter P is dependent on another parameter P'. A special case of this relation (perhaps the only case) is when setting P' to value V completely de-activates P. For example,

in languages lacking infinitival embedded clauses, any parameter exclusively concerning properties of infinitives would be de-activated. Call this Linked Parameter interaction.

(2) The setting of a parameter P to value V implies the setting of parameter P' to V', not by virtue of any hierarchic relation between P and P', but because of the way they interact with the principles of grammar. Call this Independent Parameter interaction.

Linked Parameter interaction is forced by UG and therefore the empirical correlations that it gives rise to, henceforth called Type I correlations, are exceptionless. Independent Parameter interaction yields empirical correlations which are exceptionless in grammars of a certain type (that is, as long as other parameter values are set alike), but need not hold in typologically more distant grammars. Call them Type II correlations.

Type II correlations must be distinguished from correlations which are entirely fortuitous from the point of view of UG, resulting from (recent) language contact, henceforth Type III correlations. These have no universal generality, and are subject to fortuitous local exceptions.

Our initial hypothesis is that all three types of correlations are found, each having the empirical status mentioned.

Aims and objectives.

The aim of this project is to develop the second-order typology of cross-linguistic correlations just sketched. This entails systematic work in two areas: first, on developing a clear picture of what implicational relations among parameters and parameter settings are conceivable; second, investigating very systematically the correlations that have been observed. We will pursue these two areas of investigation by focussing on one of the most controversial parameters, the null-subject parameter. We expect that our investigations will allow us to see which of the properties associated with this parameter are deeply linked, which only in certain types of systems, and which – despite contingent local correlations – are illusory. If we achieve this goal, we will have substantially deepened P&P theory, and gone a small way towards exploiting its potential.

Research questions in more detail.

The empirical issues associated with the null-subject parameter are fairly well known, yet remain badly understood. The language coverage is also relatively large, yet few languages have been subject to detailed investigation. For this reason this domain represents an ideal testing ground for the questions concerning the nature of parametric variation outlined above.

Examples of the empirical correlations we are interested in are the following, all of which are connected to the null-subject parameter:

(i) The presence of referential null subjects correlates with “free inversion,” as originally proposed on the basis of Italian examples like (1) by Rizzi (1982):

Ha telefonato Gianni.
“Gianni has phoned.”

This correlation holds in Standard Italian, numerous Italian dialects, Spanish and Greek, but it does not clearly hold in Portuguese, and Gilligan (1987) – the only thorough cross-linguistic survey to date – shows that it has little true cross-linguistic standing. It is therefore highly unlikely that this is a correlation of Type I, although it may be either a Type II or Type III case; we intend to determine which of these is correct.

(ii) If a language has referential null subjects, it does not have overt referential expletives.

It appears to hold in many Germanic and Romance languages, but the status of (ii) in Finnish and in Welsh needs close investigation. Again it seems unlikely, although not entirely impossible, that this is a correlation of Type I, but it could be either of Type II or Type III.

(iii) If a language is verb-second, it does not freely allow referential null subjects.

The only clearly attested verb-second languages which have referential null subjects at all are Old French, various Medieval Northern Italian dialects and some varieties of Rhaeto-Romansch. In these languages, null subjects are only found in verb-second clauses. However, the status of Middle Welsh and Modern Breton in this connection needs investigating. This correlation may, on the face of it, be of any of the types listed above. It is clear that the eventual conclusion on this point will have implications for our general understanding of verb-second phenomena, as well as of null subjects.

(iv) Null-subject languages have “rich” agreement.

This correlation appears to be spectacularly disconfirmed by the presence of null subjects (and objects) in many East Asian languages which are totally lacking in agreement inflections. However, it has been suggested (by Rizzi *inter alia*) that the East Asian facts represent a Type I correlation: The null subject parameter doesn't come into play in a language totally devoid of agreement morphology. There appear to be counterexamples to this correlation as well, though: Papiamentu has no agreement, yet does not allow null subjects (according to Gilligan). Furthermore, there are languages which have poor agreement, yet allow null subjects (Chamorro, Sao Tome Creole), and there are languages with apparently rich agreement which do not allow null subjects (Icelandic). But there is also striking support for the correlation, including languages which have agreement in parts of the verb paradigm, and allow null subjects just when the verb form agrees (Irish, Hebrew). A precondition for determining the type of the correlation is determining what constitutes “richness” of agreement.

In this connection we need to also consider the typology of null arguments. It has been claimed that agreement, rather than licensing a null pronoun (as in Rizzi's original theory), actually IS a pronominal argument. There are reasons to believe that both types of null argument constructions exist, and possibly other types. We need to investigate the possibility that correlation (iv), or any of the other correlations, holds universally, but only for a specific type of null subject construction.

(v) If a language has null subjects, it has infinitive movement. Kayne (1991) establishes the validity of this generalisation for Romance languages. Its status in other families, particularly Germanic, is not clear, however. In these languages, infinitive-movement interacts in complex ways with restructuring/clause-union phenomena (see (vi)). Again, the issue for the typology of correlations given above is whether this is a Type II or a Type III correlation.

(vi) If a language has restructuring/clause-union phenomena, it has null subjects. Again, this correlation holds fairly well across Romance, but appears to be entirely falsified in the West Germanic languages. Sophisticated analysis is required here, however, in order to determine the precise status of the cross-linguistic correlations, especially as (iii) may play a role in explaining why this correlation does not hold in German. English and North Germanic obey the correlation in being non-null-subject and lacking restructuring, while Celtic obeys the implication as stated, in that these languages are null-subject languages but lack restructuring.

Research methods

We propose to systematically analyse and investigate these correlations across mainly, but not exclusively, a range of European languages (Romance, Germanic, Celtic, Slavic and Finnic). This work will be divided among the members of the research team as follows: Holmberg will work primarily on the Finnic and Scandinavian languages, Roberts on language change among the Germanic and Romance languages, Willis on the Celtic languages, including earlier stages of these languages. The Postdoctoral Research Associate (PDRA) will have special responsibility for the Germanic languages, RS1 primarily for the Romance languages, while RS2 will focus on the role of agreement in relation to the null-subject property. The research will proceed in four stages.

Stage I (Oct 2002-Oct 2003): At this stage the research team will consist of the three senior researchers and the PDRA. The aim at this preparatory stage is

- (a) to investigate what the situation is globally as regards null-subjects. This will be done by the PDRA on the basis of Gilligan's (1988) survey, extended by a set of languages sampled so that all the major genetic and areal groupings in the world are represented. The aim is not primarily to put forth statistical generalizations, but to build a database from which we can draw languages against which hypotheses can be tested.
- (b) To begin to collect and systematize data from the European languages;
- (c) To clarify the theoretical issues involved.

Stage II (Oct 2003-Oct 2004): Collection and systematization of data. At this stage the research team includes the two research students (RS1, RS2); The status of all relevant languages in relation to the correlations listed above must be properly established. The work at this stage will be based on available descriptions of the languages in question, but also work with informants, in order to collect data from less well described dialects of the languages which are the primary focus of our research. This will include some travel, but we count on relying partly on contact by e-mail or indirectly, through the network of linguists in Europe and outside. The culmination of this stage will be a clear *prima facie* picture of the empirical status of the correlations across a range of languages and language types.

Stage III (Oct 2004-Oct 2005): preparing full-fledged theoretical analyses of the constructions relevant to the correlations. As adumbrated under (iv) and (vi) above, this may change the status of certain correlations. At this point we envisage the investigation focusing on a subset of the correlations listed above. A conference on null-subjects will be organized.

Stage IV (Oct 2005-Oct 2006): assignment of both correlations and non-correlations to the typology sketched above, with concomitant development of the theory of parametric variation. By the end of Stage IV, RS1 and RS2 will finish their dissertations.

Stage V (Oct 2006-Oct 2007). Preparation of articles and other publications.

The research team, otherwise divided between Durham and Cambridge, will meet three times a year throughout the period. The PDRA will be stationed in Cambridge, RS1 and RS2 in Durham.

Justification of staff.

The number of languages and dialects that need to be investigated and the depth of analysis required to attain the objectives motivates engaging three persons in addition to the three senior researchers: a PDRA (Theresa Biberauer) and two research students.

Theresa Biberauer is currently finishing her PhD at the University of Cambridge. She is a specialist on Germanic languages with an already extensive background of research in theoretical syntax focusing on Afrikaans. She is also a native speaker of German, and has wide general knowledge of other Germanic varieties. She is therefore ideal as a PDRA for this project.