

Kimberly Geeslin
Stanford, USA

A Typological Investigation of the Pro-drop Parameter

Introduction

Recently, linguists have placed a great deal of interest in linguistic universals and the principles and parameters theory. Because the theory of principles and parameters is based on linguistic universals the two theories together have the potential to be very powerful when typological evidence permits. Two areas where these hypotheses have great theoretical potential are historical reconstruction and acquisition research. In light of the benefits of these theories, the purpose of this paper is to use a typological approach to evaluate one particular parameter called the *pro*-drop parameter. After a general discussion of universals and this parameter in particular, language particular data collected for this study will be investigated. The data will be analyzed in terms of how it patterns with the established parameter. In addition, possible explanations for patterns and inconsistencies found will be sought in terms of other typological theories such as those by Greenberg (1963), Dryer (1991), Steele (1978), Jelinek (1984) and Nichols (1993) in contrast with syntactic theories like morphological uniformity (Jaeggli and Safir, 1989).

Universals, Principles and Parameters

An absolute universal is a property that all languages have in common, or a constraint that no language violates. Examples of this type of constraint are that all languages have vowels and all languages allow CV syllable structure. These absolute universals are scarce and the greater the field of research, the more likely it is that the universal will be proven not to be without exception. There are also frequency based universals such as those found in Greenberg (1963). These universals state general tendencies for languages of a particular type. One example of this type of universal is

his universal number four which states that "with overwhelmingly greater than chance frequency, languages with normal SOV order are post-positional" (Greenberg, p. 62, 1963). Since these universals are based on frequency, the way in which the sample of languages to be included is chosen and the methodology for categorizing each feature is very important. Due to these variants, these universals provide much weaker information about languages than do the absolute universals. Nevertheless, careful investigations can still provide researchers with a very good conception of how languages tend to behave.

Principles and Parameters theory makes use of absolute and probabilistic universals in a particular way. These tendencies were used to solve what was previously a great problem for theoretical linguists. When Chomsky (1965) defined the goal of explanatory adequacy in relation to linguistic theory, he discussed the problem of explaining acquisition. The basic difficulty has been described as the poverty of stimulus which notes that the evidence provided to learners is not sufficient to explain the infinite number of sentences they are able to create. The reasons for this are that learners often hear ungrammatical sentences, that experience is finite even though potential output is infinite, and that learners acquire rules even in the absence of positive evidence. (Lightfoot, 1991, p. 3) Thus, what is known as the triggering experience, or a reaction to input causing learners to formulate a rule, is insufficient explanation on its own. Since the poverty of stimulus could not be explained, explanatory adequacy was not achieved.

As a result, it was proposed that learners possess an innate grammar which is the same for all humans and thus not language specific. This grammar consists of universal rules such as the embedding principle, which states that a grammatical sentence can become subordinate in a complex sentence, and these principles are true of all natural languages (Haegeman, 1991). These principles consist of the same type

of information that linguistic universals do. They describe absolutes for all languages and thus provide a framework for historical and acquisition research alike.

In addition to these universal principles, the theory about grammars also consists of a set of parameters, or options which contain binary choices that describe not only language particular rules but also the variation between languages. With the concept of these parameters, it is possible to explain that a triggering experience affects one particular parameter in order to create the core grammar of a language by setting that parameter in one way or another. Rather than learning many individual rules, a learner merely chooses between two possible rule types and all of the other characteristics associated with that parameter follow. This means that not every cross-linguistic difference is evidence of a different parameter since many grammatical characteristics tend to group together into one parameter.

This framework not only solves the problem of explanatory adequacy, it serves as a useful tool for language historians who attempt to reconstruct languages for which a complete grammar is not available. The linguistic universals, or principles, state which properties must always be true and which constraints must never be violated. Even more importantly, the parameters relate features of the grammar which were previously seen as independent. This allows grammarians to reconstruct larger portions of the grammar with each discovery. It also enables them to evaluate the probability that their findings are accurate since a description of a grammar which violated many parameters should be less likely to be accurate than one which does not.

The *Pro*-drop Parameter

The *pro*-drop parameter describes languages that do not require an overt subject. According to the parameter, there are four characteristics of a language that should pattern together. Although data from many other languages will be included later in the discussion, only Spanish and English will be used here to illustrate each of these characteristics. This is because Spanish is a [+*pro*-drop] language according to

all four criteria and English is a [-*pro* -drop] language according to all four criteria. It will be seen later that this absolute obedience is not universal.

The first characteristic of a [+*pro* -drop] language is that overt subjects are optional in many cases. Due to the marking on the verb which is related to the subject, the subject will be understood even if it is covert except in the case where proper nouns are being introduced to a conversation. This is seen in the contrast between the English sentence "I go to the store" and the Spanish sentence "(Yo) voy a la tienda" where the parentheses indicate that the subject pronoun is optional. Nevertheless, in the sentence "John goes to the store" or "Juan va a la tienda", the subject 'John' can not be inferred unless it has been previously introduced in some way.

In addition to distinguishing languages which do not require an overt subject from those that do, this parameter implies other characteristics. One example is subject-verb inversion which is allowed in *pro* -drop languages like Spanish where both "Juan viene a casa" (John comes to house) and "Viene Juan a casa" (Comes John to house) are acceptable sentences. Allowing subject-verb inversion should not however imply that this is free word order between the subject and the verb. The basic word order for Spanish is still SVO since this type of inversion is restricted for some native speakers depending on verb choice (see Liceras, 1988).

Another characteristic which is part of the *pro* -drop parameter is the existence of expletives. English, which is not a *pro* -drop language, requires the expletives 'it' and 'there' since the subject position in a sentence can not be left empty. This is seen in the sentences "It rains in March" and "There are three men in the room". In contrast, a *pro* -drop language does not allow expletives, and therefore, the verb 'llueve' (it rains) in "Llueve en marzo" (rains in March) includes the 'it' subject in the conjugated verb and the subject position is left empty. If one were to fill the subject position as in "*Lo llueve en marzo" (*It rains in March) the result is an ungrammatical sentence. The same follows for the Spanish counterpart of the sentence "There are three men in the room",

"Hay tres hombres en el cuarto", where no expletive is permissible and the verb 'hay' (there are) stands alone.

The final characteristic involved in the *pro*-drop parameter is called the 'that-trace effect'. In languages which do not allow *pro*-drop, the complementizer 'that' may not be left in the sentence after *wh*-movement has taken place. Because of this the sentence "*Which book did you say that is interesting?" is not grammatical whereas the identical sentence "Qué libro dijiste que es interesante?" is grammatical in Spanish. This characteristic is somewhat controversial since for some native speakers of English the sentence above is grammatical.

The study

Seventeen languages are included in this pilot study. The languages surveyed were chosen only on the basis of availability of informants. As a result, the sample does not account for genetic nor areal relationships between languages. In fact, it appears to be quite heavily weighted towards the western world and the Indo-European language group. As a result, scientific conclusions can not be based solely on this sample. There is however one strong advantage to this sample, the proficiency level of the informants. The vast majority of the informants who were not native speakers of the language they described have a near native proficiency in that language. This is extremely important since the results of studies that rely entirely on grammar descriptions by authors who do not speak the language are less reliable. Nevertheless, as a result of the limited sample size and lack of diversity in it, the results can be taken only as preliminaries for future large scale studies.

Each informant was asked to provide a description of how their language behaves in relation to the four characteristics of the parameter described above. In some cases more than one informant reported on the same language. After categorization of the results, follow up questions were asked of each informant when necessary. Once each characteristic of the parameter is evaluated individually, their

relationship to each other is investigated. Finally, these results are followed by a discussion of other features more closely linked to morphology which have been used to explain certain phenomena in previous research.

Results

The complete chart of results can be found in the appendix. Each language is categorized according to their behavior with all four characteristics of the parameter. Below, each feature of the parameter will be analyzed individually and then patterns within the parameter will be assessed. Also included in the appendix are the features of word order, the degree or complexity of inflection and, the uniformity of this inflection. The definition of each of the latter group of characteristics will be included in the discussion below and their relationship with the parameter will be investigated.

Of the seventeen languages included in the study, seven of them allowed covert subjects. In addition to those seven, ASL and Hebrew allow covert subjects in some contexts. In the case of Hebrew, for example, the null subject is allowed in the past and future tenses in the first and second person. The covert pronoun is not allowed in combination with any other verb forms. None of the informants had trouble analyzing their language in terms of this feature. It is also the only feature of the parameter which did not produce an unsure response, marked by a question mark in the table in the appendix.

In contrast with the certainty involved with the use of null subjects, informants were less sure about whether their language allowed subject-verb inversion or not. Hebrew was described with an uncertain yes, both German and Korean received conflicting information from different informants, and Swedish and French were described as allowing this type of inversion but only in a specific context. Thus, of the seventeen languages included in the sample only five of them (Russian, Czech, Arabic Spanish and Portuguese) allowed definite free subject-verb inversion. This is interesting since both Russian and Czech and Spanish and Portuguese are closely

related and since Arabic and Spanish have been in contact for thousands of years. This means that even though five languages represent over one fourth of the sample, they do not represent a diverse group of languages.

None of the seven languages which allow covert subjects in all positions allow expletives. This means that for [+*pro* -drop] languages, the parameter patterns as predicted without exception. In addition, four languages, Esperanto, Hindi, Marathi, and Assamese, all of which require overt subjects, do not have expletives either. While ASL only allows only some covert subjects it does not have expletives, whereas Hebrew does. In fact, expletives were found to exist in only five of the languages in the sample. It is interesting that while all languages which allow covert subjects do not permit expletives, the inverse is not true.

The final characteristic of the parameter, the that-trace effect, produced the most questionable results. Of the seven languages that allow null subjects, four of them allow an overt complementizer after *wh*-movement. These four were Spanish and Portuguese, Czech and Arabic. Russian was not included in this group despite the fact that it patterned with these languages in terms of subject-verb inversion. Of the seventeen languages in the sample, only three informants were certain that their language did show that-trace effects. These languages were English, French, and Russian. The rest of the informants were either unsure, or their languages did not have overt *wh*-movement and the feature did not apply.

Discussion

At this point in the analysis, the study has confirmed some of the problems with this parameter. The most obvious problem is that only the lack of expletives patterns without exception with complete allowance of null subjects. Thus all seven of the [+*pro*] languages lack expletives. On the other hand, of the 10 languages which do not always allow null subjects, only five of them have expletives. This shows that even this one feature can not be viewed as a complete correlate with the use of null subjects. One of

the languages which violates this, Hebrew, allows some null subjects but also has expletives. Since Hebrew is split between allowing null subjects and not allowing them, it necessarily violates the parameter. Four other exceptions are Esperanto, Hindi, Marathi and Assamese which do not allow null subjects but also lack expletives. A better way to describe these last four languages is that they only allow null subjects for impersonal, non-theta marked subjects.

The problems become more serious when investigating the other two features associated with the parameter. The fact that all languages which allow null subjects and lack expletives do not necessarily follow the other two characteristics associated with the parameter, raises an important question. That is, does the parameter actually describe all languages, or does it just describe a tendency within Indo-European languages? This question will be discussed in light of two particular groups of languages included in the sample.

The first group of languages which do not pattern with the parameter are Japanese and Korean. These languages do not allow free subject-verb inversion nor do they show that-trace effects. This is because these languages have SOV word order and they do not allow overt wh-movement making the that-trace feature inapplicable. The second group of languages, all spoken in India, are Hindi, Marathi and Assamese. Since these languages also have an SOV word order, subject-verb inversion is not at all possible. If this parameter is universal, than no language which has an SOV word order should allow null subjects. This seems to be too broad of a claim. Both of these language groups point to a failure to include non-Indo-European languages in the parameter. This means that the description of the parameter is either too 'euro-centric' to include other *pro*-drop languages, or the *pro*-drop feature found in Asian languages is somehow a different type of process.

A final piece of evidence that weakens the strength of these two characteristics as part of the parameter is the confusion that was raised by asking informants about

them. While free subject -verb inversion is the only type of inversion to be associated with the use of null subject, many informants referred to various types of obligatory inversion and the responses had to be re-interpreted. In addition, the responses to the question of the that-trace effect produced eleven answers that were left just as question marks or were answered and then followed by question marks. It is highly possible that these features are related to other processes or that they are only relevant for a small group of languages.

Since the purpose of the parameter is to describe a situation where characteristics of a language are harmonious (Croft, 1990), this is actually the best test of the parameter. A language which is [+*pro* -drop] harmonious will allow null subject, have free subject -verb inversion, lack expletives and will allow an overt complementizer. In contrast, a language which is [-*pro* -drop] harmonious will have no null subjects, no free inversion, will require expletives and will show that-trace effects. Of the seventeen languages included in the sample, only ten are harmonious. These languages are English, Spanish, French, German, Czech, Arabic, Portuguese, Russian, Swedish, and Esperanto. Since this is just over half the sample, the parameter is not yet adequate for universal description.

Before discussing other possible correlates with the use of null subjects, it is important to point out some reasons why this parameter has not been successful. The most serious weakness is that the only languages which are harmonious are those which are spoken in Europe. Although Arabic is not actually spoken in Europe, it is geographically proximate and has a great deal of contact with Spanish. This points to a failure to describe languages that do or do not allow null subjects from other geographical areas. Since this parameter is intended to describe universal behavior of languages, it must be modified to include others.

This type of modification can be accomplished in one of two ways. First, it can be said that the tendency of the parameter is not very strong, and therefore languages

do not always pattern completely. This would imply that the geographical effect was coincidence. Since this is unlikely, a better solution must be sought. A second possibility is that the parameter has been mis-defined and that some other characteristic is actually a better predictor of the use of null subjects. The rest of the discussion will assess the relationship of the use of null subjects with other characteristics which have been important predictors of other types of relationships.

Word Order

Many of the Greenbergian (1963) frequency based universals correlate with order of certain constituents in a sentence. Thus word order served as a statistical predictor of other features in the grammar. Steele (1978) used similar methods to examine the possible variation of word order and how languages with free, rigid and mixed word order might be classified. Through her studies, she further developed the typology of word order correlates. The way in which word order was used as a predictor was expanded upon by Dryer (1991) where correlates were studied in relation to the order of the verb and the object of the verb. Dryer uses this division to improve upon some of the previous predictions and to disprove other correlates that were found previously.

All of the research mentioned above points to a possible solution for the *pro*-drop parameter. If word order can predict the use of null subjects, as shown by a patterning in the data collected for this study, than it may prove to be a more relevant characteristic of the parameter than subject-verb inversion or that-trace effects. Nevertheless, it is clear from the discussion above that basic word order points to a flaw in the parameter rather than improving it. This is because within the sample there are many SVO languages which are [+*pro*-drop] and more which are not. In addition there are five SOV languages, three of which are [-*pro*-drop] and two of which are not. The sample has only one VSO language. These results show that no evidence has been

found that basic word order contributes to the description of the null subject parameter in isolation.

If Steele's (1978) distinctions between rigid and flexible word order are helpful, then this characteristic may be able to distinguish between languages that obey the parameter and those that do not. This however is difficult to support. Arabic has flexible word order such that the basic word order VSO is controversial and it is harmonious with the *pro*-drop parameter. ASL is in the process of change from SOV to SVO (Jacobson, 1995, personal communication), and it thus flexible, and it only allows *pro* in certain instances. In regards to the Asian languages, both Korean and Japanese allow *pro*, but the former is rigid verb-final and the latter also allows OSV. It appears that word order, even qualified by the characteristic rigid or flexible is still not sufficient.

A final possibility is that word order can be used to distinguish between types of languages in order to create Greenbergian type universals. This does not seem all that implausible. An example of this type of universal would read: A language which is SVO and allows free subject-verb inversion, lacks expletives and shows that-trace effects, will also allow null subjects. Although this universal is correct, it has provided nothing more than a technical means for including all of the Indo-European languages in the sample and excluding the others. A larger sample would probably refute this universal. Even if further research supports this universal it does not serve the same function as the previous parameter since it cannot include all languages and describe them in terms of a binary choice. Consequently, it is prudent to explore other possible correlates in hopes of finding more conclusive results.

Relationship to Morphology

Current syntactic theory has placed a great deal of emphasis on the morphological features of languages which allow null subjects in an effort to distinguish them from those that do not. This type of research has been incorporated into

typological investigations in similar ways. Following a brief description of the progress in syntactic theory these particular typological approaches will be explored in terms of their relevance to the *pro*-drop parameter.

The initial constraint on *pro*-drop languages in Government and Binding Theory was that all instances *pro*, or the null subject, must be licensed by a rich INFL node. This constraint adequately distinguishes between languages like Spanish which are highly inflected from languages like English which lack such inflection in most verb forms. Therefore, in *pro*-drop languages the licensing head (the inflection node) shares the grammatical features with *pro* and thus serves to supply its content (Haegeman, 1991, p.457). One complication with this type of explanation, is that it does not explain why languages such as German, which are highly inflected and thus capable of supplying the content of *pro*, do not allow null subjects. As a result, the licensing constraint was broken into two separate parts. First, the null subject must be licensed in a particular language, and second rich inflection must be available to recover the content of that null subject. A rich inflection system is one that provides information for person and number on the verb form. German is not able to license *pro* and since the first requirement is not met, its ability to recover the content never comes into play.

Evidence from languages like Japanese, Korean and Chinese created the need for another modification in the above restriction. The reason for this is that these languages do not have a rich inflectional system and it was not clear how they could license a null subject and successfully recover its content. This led to another modification which relied on morphological uniformity (Jaeggli and Safir, 1989). This means that any language which is consistent, either by including overt person and number markers on each form of a verb or by never including them, could license *pro*. In order to explain how the content of *pro* could be recovered, Jaeggli and Safir (1989), describe a process of identification by agreement (p.32). This allows other verbs or

nouns to provide the content for *pro* in languages which lack overt morphology. This allowed theorists to include Asian languages and Indo-European languages under the same classification.

It is clear that the last constraint which relies on morphological uniformity accomplishes the most adequate description of the data, and it is the most widely accepted within the Government and Binding framework. Nevertheless, it can not explain why German and ASL which are uniformly inflected do not always allow free subjects. It also does not explain why Hebrew which is not uniform is able to license *pro* in certain contexts. Finally, the way in which Asian languages recover content remains unclear. In light of these weaknesses, there is good motivation to investigate typological approaches to the same problem.

One attempt to incorporate *pro*-drop languages into a typological theory was made in Jelinek (1984). This paper was written to explain why non-configurational languages like Warlpiri allow covert arguments in addition to other properties. Jelinek proposes that AUX (or INFL) be interpreted as the subject of the verb since it always carries the person and number information. She claims that since only the third person singular is marked by the absence of phonological material that this absence was equivalent to using some other morpheme since there was no ambiguity in interpretation. Her theory is supported by the fact that while the *pro*-drop parameter cannot account for Warlpiri, her analysis can account for both.

Jelinek (1984) provides interesting data from Spanish to support her proposal that nominal subjects are never arguments of the verb. Instead, they serve the purpose of contrasting or emphasizing the relationship marked by the verbal suffix which provides the information for the subject. These data, such as the sentence "las mujeres tenemos esperanza" (women, we have hope) (p.48), show that the noun phrase that would normally be interpreted as the subject does not agree with the verbal inflection markers. The ability to explain these examples is one strength of the theory.

Her analysis leads to a description of W-type non-configurational languages which she defines as those languages where AUX clitic pronouns always mark all verbal arguments and co-occur with optional nominals. She arrives at a parameter for configurationality which states that if a language is configurational then the object nominals are properly governed by the verb and if the language is a w-type non-configurational language, nominals are not verbal arguments, but are optional adjuncts to the clitic pronouns that serve as verbal arguments.

This analysis uses the inflection on the verb (or AUX) to license *pro*. It can accurately describe the difference between Spanish and English since the latter lacks sufficient inflection in AUX. Another advantage is the description of 'las mujeres' (women) in the example in Spanish "las mujeres, tenemos esperanza" (women, we have hope), as an adjunction. This seems like the only logical possibility since it does not agree with the verbal suffixes and is still grammatical. A final strength of this analysis is its description of the function of pronouns. It seems quite likely that clitic pronouns in combination with overt nominals serve exactly the contrastive or emphatic purpose Jelinek describes.

Nevertheless, this approach does not account for all of the languages in this study. The first problem is the incorporation of the Asian languages. In Jelinek (1984) it is claimed that Japanese may not be non-configurational which is controversial. In addition, she claims that *pro*-drop should be recoverable by syntactic properties yet in Japanese this is not always the case. Since Japanese can be ambiguous when omitting noun phrases, she claims that this is discourse related and thus it is not relevant to the discussion at hand. This is supported by the fact that Chinese is configurational and it behaves much like Japanese.

Although it is possible that the null subject in Asian languages is a different phenomena from null subjects in other languages, this approach does no better to include all languages and arrive at a universal than do previous ones. Furthermore, it is

quite possible in Spanish for a sentence with a null subject to be ambiguous and require qualification. In fact, it is unlikely that any language has fully free null subject use even though languages which also mark gender may be close. If the goal of this study is to find a syntactic or grammatical element which determines the use of the null subject than this goal has not been met with this analysis. It may, however, add to the evidence that such a syntactic factor may not exist and perhaps the context itself plays a role in all languages.

In addition to excluding discourse related languages from the analysis, the parameter described in Jelinek (1984) also claims that all configurational languages will have an object nominal which is properly governed by a verb. Despite the many possible interpretations of this, it also appears inadequate. If the interpretation of the parameter is taken to state that all configurational languages have *pro* whereas all non-configurational languages have no such element, this contradicts the use of Spanish to support the argument. If on the other hand, the parameter is taken to describe only objects in configurational languages, it would be necessary to explain why subjects in Italian are carried on AUX but null objects are represented by *pro*. No matter what the interpretation of this parameter is, it does not seem to account for the Indo-European languages with any more simplicity than do the other accounts. In fact, the only advantage in describing overt subjects as optional in some languages is to explain the sentences where the subject does not agree with the verb. In this case, there is no evidence to show that this overt NP does not actually precede the null element *pro*. As a result, the use of adjunction to describe this process is quite helpful, but it is not necessary to propose that Spanish has optional NPs instead of *pro* in order to do so.

A final analysis which is worth mentioning here, due to its relation to the morphological elements in a language, is the work in Nichols (1992). Nichols uses Type Features to refer to four particular salient morphosyntactic features (p.46) that she employs to search for correlations between languages of particular lineage or location.

One of these features is called morphological complexity. This is calculated by adding the number of head marks, dependent marks and free markers. These morphological markers are defined as "any form of inflection, affixation, cliticization, or other overt morphological variation that signals some relevant relation, function, or meaning."

(p.48) The maximum number for complexity for one pair such as a VP and a noun is 3. This means that individual morphemes which mark for example, gender and number, on a dependent head are counted as one and not as two. In addition, PPs were excluded since languages which lacked this information in their descriptions would be mis-represented. Although Nichols acknowledges that this count overlooks other important factors, she states "It is my impression that complexity as measured here, partial and artificial though it is, correlates straightforwardly with overall morphological complexity and can hence be used as an index of something real," (p.64).

In keeping with the idea expressed by researchers previously discussed, it might be hypothesized that Nichol's morphological complexity is related to whether or not a language can license *pro*. One interesting fact is that the Ancient Near East, Northern Eurasia, Europe and the Ancient Near East are the four areas she claims have more highly complex languages than other types. Interestingly, this more or less includes all of the *pro*-drop languages in the sample. Another important piece of information is that there is no correlation between complexity and word order. This is important since the above discussion was able to rule out word order as a predictor of the use of the null subject.

Unfortunately, it is not possible to replicate Nichol's evaluation of the complexity of the languages in this sample since many of the languages in each do not overlap and this sample collected insufficient information for such a count. Of the five languages that do overlap between this study and hers, four are *pro*-drop and one is not. The languages which allow null subjects are Colloquial French, Japanese, Korean, and Russian and their complexity values are 11, 9, 8, and 11 respectively. The one

language which does not allow *pro* is English and it has a complexity level of 4. Notice that since the French in her sample is colloquial it is considered verb initial and therefore *pro*-drop and thus it does not agree with the information in this study. So far as is possible, a comparison of the *pro*-drop feature with the levels of complexity assigned to each language by Nichols produces very positive results.

Conclusion

The purpose of this study was to conduct a survey of languages in regards to their *pro*-drop features in order to evaluate the accuracy of the *pro*-drop parameter. It was shown that only the use of expletives patterns with languages that allow null subjects, while subject-verb inversion and that-trace effects produce mixed results. This lead to the suggestion that the current description of the parameter might not be as accurate as previously hoped since in only accounts for languages with a particular SOV word order which are spoken in one geographical area. In order to better describe what properties of language correlate with null subjects, generative syntactic theory and typological theory was examined in order to find more appropriate correlates. The word order studies were inconclusive in this respect. Generative theory will its principle of morphological uniformity also falls short of complete description. The work by Jelinek(1984) which claims that AUX is actually the subject of the sentence produced the same sort of problems as the syntactic theory. Only the work by Nichols which evaluates morphological complexity was able to explain the data, but information was only available for five languages in this sample.

In light of the discoveries above, future research may move in two directions in order to better describe the *pro*-drop parameter. First, since the calculations of Nichols appear to agree with the data presented here, one might expand the calculations of morphological complexity to include the languages presented here. In this way, it might be found that while the intuition that verbal morphology is related to the use of *pro*, it is necessary to evaluate this through new methods. It is possible however, that no matter

how complexity is calculated, languages like German may cause a problem. This leaves a second possible direction. That is, perhaps some other correlate exists which is not related to morphology at all. In this way, the 'discourse factors' that recover the content may be discovered to explain all *pro*-drop languages. Regardless of the direction taken in future studies, it is clear from this data that previous accounts have not achieved the goal of universal description of the *pro*-drop parameter.

Appendix A - Full data report

Key:

pro: Does the language allow null subjects? (yes or no)

inversion: Does the language allow free subject-verb inversion? (yes or no) (free in parenthesis indicates that some other type of subject-verb inversion is allowed)

Expletives: Does this language have expletives? (yes or no)

that -trace: Does this language allow overt complementizers? (yes or no)

W.O.: What is the word order of this language? (? = in disagreement about basic form)

INFL: Does this language have a rich inflection system? (R(ich) or W(eak))

Uniform: Is the inflection morphologically uniform? (yes or no)

Harmony: Is this language harmonious for the first 4 features in the table? (yes or no)

Language	<i>pro</i>	inversion	expletives	<u>that</u> -trace	W.O.	INFL	uniform	Harmony
English	no	no	yes	no (?)	SVO	W	no	yes
Spanish	yes	yes	no	yes	SVO	R	yes	yes
Swedish	no	no (free)	yes	resumpt.	SVO	W	no	yes/no
French	no	no (free)	yes	no	SVO	R/W	no(?)	yes/no
German	no	no (free)	yes	y-2/n-1		R	yes	yes
Japanese	yes	no	no	no	SOV	W	yes	no
Czech	yes	yes	no	yes	SVO	R	yes	yes
Korean	yes	n/y	no	no	SOV	W	yes	no
ASL	some	no	no	?	sov/svo	R	yes	no
Portuguese	yes	yes	no	yes	SVO	R	yes	yes
Russian	yes	yes	no	no		R	yes	no
Hebrew	some	yes ?	yes	yes ?		R/W	no	no
Arabic	yes	yes	yes	yes ?	VSO ?			yes
Esperanto	no	yes	no	?	SVO	W	yes	yes
Hindi	no	no	no	?	SOV	R	no	no
Marathi	no	no	no	?	SOV	R	no	no
Assamese	no	no	no	?	SOV	R	no	no

Appendix B - Following its use of *pro*, does the language obey the predictions of the parameter?

Language	<i>pro</i>	sub-verb	expl.	<u>that</u>
English	(N)	yes	yes	yes
Spanish	(Y)	yes	yes	yes

Swedish	(N)	yes	yes	?
French	(N)	yes	yes	yes
German	(N)	yes	yes (?)	yes(2/1)
Japanese	(Y)	no	yes	no
Czech	(Y)	yes	yes	yes
Korean	(Y)	no	yes	no
ASL	(Y/N)	x	x	x
Portuguese	(Y)	yes	yes	yes
Russian	(Y)	yes	yes	no
Hebrew	(Y/N)	x	x	x
Arabic	(Y)	yes	no	yes(?)
Esperanto	(N)	no	no	?
Hindi	(N)	yes	no	?
Marathi	(N)	yes	no	?
Assamese	(N)	yes	no	?

Bibliography

Croft, W. (1990). *Typology and Universals* .

Dryer, M. (1991). The Greenbergian word order correlations. *Language*, 68, 1, 81-138.

Greenberg, J. (1963). Some universals of grammar with particular reference to the

- order of meaningful elements. In *Universals of Language*, MIT Press.
- Haegeman, L. (1991). *Introduction to Government and Binding Theory*. Second Edition. Oxford, UK: Basil Blackwell.
- Jacobson, H. (1995). Information on ASL, personal communication.
- Jaeggli, O. & Safir, K. (1989). *The Null Subject Parameter*. Dordrecht: Kluwer.
- Jelinek, E. (1984). Empty categories, case and configurationality. *Natural Language and Linguistics Theory*, 2, 39-76.
- Liceras, J. (1988). Syntax and stylistics: More on the pro-drop parameter. In Pankhurst, J. & Sharwood-Smith, M. (Eds.) *Learnability and second languages*. (pp. 71-93). Dordrecht: Foris.
- Lightfoot, David. (1991). *How to set parameters: Arguments from language change*. Cambridge, MA.: The MIT Press.
- Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Steele, S. (1978). Word order variation: a typological study. In Greenberg, et al.