

The syntax of a language without grammar*

"The language faculty appears to be a species property, common to the species and unique to it in its essentials, capable of producing a rich, highly articulated and complex language on the basis of quite rudimentary data."

Noam Chomsky (1988)

0 Introduction

In this essay, we will apply the theory of Government and Binding (Chomsky (1981)) to a rather peculiar field of study. For we will try to stretch the empirical domain of this theory so as to include a language which is normally considered to be non-natural and therefore not subject to the specific conditions for the human language faculty that GB researchers are after.

This essay consists of basically two parts: in the first we will examine the content of the notion *natural language* as it is in use in modern linguistics.

The second part of this essay, that consists of the chapters 2, 3 and 4, will be more linguistically technical in its nature. In it, we will explore some of the features of the language under examination here as being consequences of the human language faculty.

For this purpose, we chose three subjects of current interest within the linguistic sciences, viz. the Null Subject Parameter, functional prepositions and the binding of anaphors. We will try to make clear how current insights can be applied to our data.

* This essay is my Master's thesis. Thanks are due to Andrée Tingloo for improving my English and to Sylvia Pilger for her native judgements on the Esperanto facts. Henk van Riemsdijk, who acted as my supervisor, has given so many comments on earlier versions, that most of the good idea's in this essay are really his. All errors, on the other hand, are mine.

The syntax of a language without grammar

1 On the notion 'natural language'

1.0 Introduction

In most introductory textbooks on generative grammar, the object of study of linguistic theory is claimed to be the grammar of '*natural human language*.' This claim is usually followed by a definition of the latter notion. We cite a very instructive example from a Dutch introduction to general linguistics that is in use at several universities in the Netherlands (Dik and Kooij (1984), translated by us. The original text can be found in fn. 1):

The language that has been used in a human community from early days and that a child who grows up in such a community learns as its native tongue, is sometimes also called a 'natural language'. This draws the distinction with artificial and scientific languages. We can never say about a natural language that it is invented or designed by somebody or by some group of people at a specific moment of time. All natural languages are the products of a long-standing tradition and, however far we go back into history, there is no place to find any indication on the way they have originated.

Artificial languages like e.g. Esperanto on the other hand, are designed at a certain moment in order to serve as a replacement for natural languages. [...]"¹

In this short passage, there is notice of at least three different distinctions between natural and artificial languages. These are, we think, the three most important distinctions found in the scarce literature on the subject. We shall label them here in order to ease further discussion:

*A natural language has been used (in a human community) 'from early days' and one is uncertain concerning its origin, while an artificial language is invented at a specific time by a specific person or group of persons.

*A natural language is learned by children as their native tongue while an artificial language isn't.

*The function of an artificial language is to replace natural languages.

We will call the first distinction the *diachronic distinction here*; the second we will call the *synchronic distinction*; and the third we will call the *functional distinction*.²

It should be noticed that these three distinctions are not consistent with each other. Latin is diachronically and (probably) functionally a natural language, but synchronically artificial. There are no native speakers of Latin.

The syntax of a language without grammar

Modern Italian, on the other hand, is more or less *invented* by Dante in order to replace Latin (which was still living and therefore natural in every respect at that time) and some Italian dialects as a written language. This language could be argued to be diachronically (and functionally) artificial but synchronically it is, as a matter of course, natural.

Unfortunately, the question of what a natural human language really is, is not very extensively discussed in the linguistic literature except on the first pages of introductory text books like Dik and Kooij (1984). It is perhaps for this reason that artificial languages have not been the subject of any serious study within for example the GB framework³ until this day.

There is an inclination to think that a language like Esperanto is not a natural language, because it is made up by a Polish amateur linguist in the 19th century (i.e. on the basis of what we have just called the diachronic distinction).

We will try to show in this essay that this point of view is mistaken, or at least that one misses some interesting observations by just assuming that Esperanto is not a natural language. We will show that Esperanto, a language which was intended by its author to be as *logical* as possible, has some very illogical features which can only be explained if we assume that Esperanto, like every other natural language, is subject to the specific conditions of the human language faculty.

In this chapter, the following subjects are down for consideration. In section 1.1 we will give a critical discussion of the first two chapters of Chomsky (1986b) in order to make clear what exactly is the object of study for generative grammarians.

In section 1.2 we then examine the distinction between artificial and natural languages in this light. We shall see that this amounts to adopting a version of the synchronic distinction. Especially we shall argue that there is no valid reason not to implicate Esperanto into generative grammar.

Because of the rather exceptional sociolinguistic status of Esperanto (there are no monolingual speakers of the language), we have to consider the status of the judgments of bilingual speakers and second language learners with respect to the theory of UG. For this reason, we review some principles-and-parameters approaches to these phenomena in section 1.3. The last section of this chapter is devoted to a conclusion.

1.1 What is a natural language?

In this section we will try to set down what the relevant characteristics are for a language to be called a *natural* language. Yet we will not try to formulate a closely-reasoned definition of ourselves for the concept here. We just assume that '*natural language is what generative grammar studies*' and examine where this assumption brings us.

The syntax of a language without grammar

1.1.1 E-language and I-language

Perhaps the only serious recent study that has appeared within the GB framework, or even more broadly, within generative grammar, and that tries to answer the question what a natural language really is, is Chomsky (1986b).

In this work, a distinction is drawn between the notions 'E-language' and 'I-language' and between these two *scientific* notions and the *commonsense idea* about language.

The commonsense idea of language, Chomsky (1986b) argues, is chiefly a political-sociological one, independent of the individual mind and brain. We call Chinese a language and Portuguese and Italian two other languages, although some 'dialects' of Chinese are as different from each other as Portuguese and Italian are. On the other hand, Dutch and German are considered to be separate languages, even though people on both sides of the Dutch-German frontier understand each other better than e.g. speakers of the German dialects of Rhineland and Bavaria do. In scientific work that makes use of some notion of language, we have to abstract away from this aspect of language.

Another aspect of the commonsense concept that, according to Chomsky, is not taken into consideration in the linguistic sciences, is normative-teleological. In the view of Chomsky (1986b), we have no way of referring directly to what the person who is still learning English as a first or second language knows. We can not say that the person has a perfect knowledge of some language similar to English but still different from it, but we do say that the child or foreigner has '*a partial knowledge of English*'.

We find it very difficult to see what the difference between these two points of Chomsky's is. To us, they both amount to saying that notions like 'English', 'Chinese', 'Dutch' and 'German' can not really be the object of any scientific study. More importantly, we do not see in what direction the abstraction goes for the last point; is the English of a 22-year old Dutch student considered as a proper instance of a language or not? We would say that it is, although with some restrictions to which we will return in the next section. But we are not very sure that Chomsky would conclude in the same way. (We will return to this issue for a short while in section 1.3)

However this may be, both aspects are disregarded or abstracted away from in the Chomskyan view of language science in which only the language in an *idealized* language community is relevant⁴.

Subsequently, Chomsky makes a further abstraction in a very interesting passage, which we will quote here in full:

The syntax of a language without grammar

'We should also make note of a more subtle theory-internal assumption: The language of the hypothesized speech community, apart from being uniform, is taken to be a "pure" instance of U[niversal] G[rammar] in a sense that must be made precise, and to which we will return. We exclude, for example, a speech community of uniform speakers, each of whom speaks a mixture of Russian and French (say, an idealized version of the nineteenth century Russian aristocracy). The language of such a speech community would not be "pure" in the relevant sense, because it would not represent a single set of choices among the options permitted by UG but rather would include "contradictory" choices for certain of these options.'

(Chomsky (1986b:17))

We will return to this passage later, after we have seen what Chomsky's precise definition of UG is. Notice for now that it is rather difficult to conceive how an idealized version of the nineteenth century Russian aristocracy (*idealized* in the sense that its members are supposed to have had a uniformly working language faculty) could ever speak a language in which contradictory choices for certain options were included. This would in our opinion mean that everybody would choose exactly in the same way in the same situation between these options, and everybody would choose otherwise (but again all alike) in a different situation. Obviously this must be managed without using a version of UG. We do not see how this could be done.⁵

But we have now first to expose the distinction that Chomsky (op. cit.) makes between the two possible objects of scientific study, *E-language* and *I-language*.

E-language, 'externalized language', is a collection of actions or behaviours in connection with linguistic utterances. The *E-language* is independent of the human mind and brain. It is the object of study of the majority of structural and descriptive linguistics. The status of the notions *grammar* and especially *Universal Grammar* is disputable for a student of *E-language*. On the other hand, *E-language* doesn't seem to be an object of study for the plurality of generative grammarians. Hence, we will drop it here.

The notion *I-language*, 'internalized language', is of more interest for our objective. Chomsky (op. cit.) defines it in terms of the *notion of structure which is definite enough to guide [the mind of the speaker/language user] in framing sentences of his own [in particular] free expressions*, which Otto Jespersen (1924, cited in Chomsky (1986b:23) already found the most prominent notion of linguistics.

I-language thus is the psychological reality of language. It is that part of the human *mind* (which in itself is a theoretical construct in Chomsky's view) that makes up the knowledge of language. Seen this way, *grammar* is the theory about some person's *I-language*, and UG

The syntax of a language without grammar

"now is construed as the theory of human I-languages, a system of conditions deriving from the human biological endowment that identifies the I-languages that are humanly accessible under normal conditions. These are the I-languages L such that $R(H, L)$ may be true (for normal H, under normal conditions."
(Chomsky, 1986b:23)⁶

It is assumed that UG describes a certain initial state of the language faculty S_0 , which is common to all human beings. S_0 itself consists of a set of constants (called *principles*) and a set of variables over a limited range (called *parameters*). A child that is learning a language sets the values of the various parameters in his head in such a way that they are in line with the sentences he hears from his parents. After a certain period of time, the language faculty will reach a relatively stable steady state S_s , which incorporates a full-fledged I-language. A human (or natural) language, then, is an I-language which can be attained by setting the parameters in S_0 in harmony with some experience or another.

We want to remark here that it is assumed by most generative grammarians (cf. Chomsky (1986b:147) that normally a human language not only consists of the components listed above, which are usually called the *core language*, but also of a so-called *periphery*. In the periphery we can probably locate things like irregular verbs, idiomatic constructions, etc. These are also the issues that most attention is normally devoted to in grammars intended for second language learners, and that have to be 'learnt by heart'. Nevertheless we don't have any compelling reason to regard having a periphery as a necessary condition for a language to be considered as *human*. On the contrary, the part of the language faculty that we call the periphery shows an unusually large spectrum of crosslinguistic variation, probably also in size. For the importance of the core-periphery dichotomy for the theory of (second) language acquisition, we refer to section 1.3.

1.1.2 An artificial I-language?

We now ask ourselves whether, within the study of I-language, any reasonable diachronic distinction can be made between natural and artificial languages.

One can imagine a (for humanistic reasons impracticable) experiment in which a human language L is constructed by some linguist. The grammar of L violates all the principles of UG -e.g. it allows for 6-subjacency violations, it requires that names be bound, it has no vowels, etc. Every feature that is known to be universal across languages, is not included in L.

We start to teach this language to a group of adult human beings and, when these persons are fluent in L, we bring some 2-weeks old babies into the community. We let the adults only communicate in L to the children for at least eighteen years. No other people are allowed to have any form of contact with the group.

The syntax of a language without grammar

We predict that the outrageous grammar of L is unlearnable for these children.⁷ Now, there are two possibilities: either the children will not learn a language at all, or (more probably) the children will transform L into a different language L+. L+ will then be their native tongue; if our theories of UG are correct, L+, unlike L, will obey subadjacency and the Binding theory. Put otherwise: L+ will be an I-language like any other. At the same time, L+ will, in all likelihood, be cognate to L in its lexical material, periphery and -where possible- even in the settings of individual parameters.

Henk van Riemsdijk (p.c.) points out that L can be regarded as a language with an empty core grammar and an outrageously overloaded periphery.

As far as we are concerned, this view puts the question of differences in learnability between the core and the periphery, in an interesting light. We suppose we cannot assume that core and periphery are equally easy to learn. There would be no point in studying UG if a language obeying the principles of UG in all relevant ways was as easy to learn for a child as a language disobeying all these principles.

This last argument presupposes that there is no inherent internal structure within the periphery of natural language. If we found out some day that there was evidence for such a structure, we would have to take care in the construction of L that even this kind of secondary structuring is not available. The predicted outcome of our experiment would then be the same.

Now, following the diachronic distinction we would have to say that L+ is an artificial language, even after a period of many centuries (when the native speakers of our experiment have transmitted L+ to their children, who have themselves transmitted it, possibly somewhat deformed, to their own offspring, and so forth) because there would from now on always be a tolerably definite point in time when L+ had started and a fairly definite group of people (viz. the first native speakers) who 'invented' or initiated the language.

Yet for a student of UG in the sense of Chomsky (1986b), L+ could be of as much interest as any other I-language. It would even have some extra importance in comparison with L, because one could learn from the collation exactly what are the features of a learnable human language and what are not. At the same time, we don't want to take L itself into consideration, because that would make it impossible to say anything sensible about UG at all.

As is said before, we can't carry out the experiment which we have set out here, except perhaps in a totalitarian state with a linguistically trained brutal dictator, but these are not easily found nowadays. Nevertheless, there are a few remotely resembling situations in the real world.

First, we have the so called *creole languages*. A creole is the language of the children whose parents speak a pidgin, that is a *contact language* of people with different native tongues. These pidgins are, of course, not specifically designed to be as queer as possible and to violate all principles of UG. On the contrary, we can suppose that

The syntax of a language without grammar

the parents in a pidgin community will try to make their common language as realistic and as simple as possible, for instance by setting the parameters in the same way as they stand in their native tongue. At least, we don't expect serious violations of the principles underlying the language faculty such as $\bar{6}$ -subjacency violations to appear in a pidgin.

The same applies, we suppose, for the English of the Dutch student mentioned in the preceding section. True enough, he may make many errors in the periphery and in the parameter settings; but we don't really expect him to err consciously against the principles of UG.

It is now time that we shortly return to Chomsky's (1986b) passage about the homogenous community speaking a mixture of Russian and French (mentioned above on page 4). Chomsky (op. cit.) claimed that the language of this community doesn't represent a 'pure' instance of UG, in the sense we have now made explicit.

It still is not very clear to us, however, what Chomsky means with those words. It seems to us that a mixture M of French and Russian is either a pidgin or a creole.

If M is a pidgin, we presumably can hardly speak about a 'homogenous language community.' Besides, we observe that the term *contradictory choices* is ambiguous. Either it means that for two speakers A and B of M and one or more parameters P of UG holds that A thinks that P and B that not P. This would make the whole passage about homogeneity nonsensical. On the other hand, if the variant of M of a person A really has contradictory choices for various parameters of UG in it, we can not really expect that the variant of person B has exactly the same contradictory choices.

Yet if M is a creole we don't expect to find contradictory choices for parameters of UG, because there is no way in which such contradictions could arise, exactly for the same reason as we suppose ideally that two speakers of Russian don't make any contradictory choices in the setting of their parameters, namely that it is learned as a native language by children about who we assume that they will make one sole choice.⁸

A second group of languages approximating the experimental situation sketched at the beginning of this section consists of languages like Nynorsk and Modern Hebrew.

These languages are created (or *planned*) by a person or a committee on the basis of existing material of a group of dialects or of a dead language, respectively. Both of the exemplary languages now serve as an all-purpose language for a considerable group of native speakers. Also in these cases, we can assume that the grammar was not designed to be any different from other human languages. We think that there are very few linguists who would regard these languages as unsuitable for research.

A very interesting third example of diachronically artificial languages which probably have developed into synchronically natural languages, are sign languages. (For a survey of the discussion about the 'naturalness' of sign languages within psycho-, socio- and theoretical linguistics see Tervoort (1988)) Although there are, as a matter of course, quite a few differences between 'normal' oral languages and sign languages, many researchers claim that sign languages

The syntax of a language without grammar

are most fruitfully studied as a product of the same mental device as e.g. English or Chinese are.

An interesting point-of-view in this debate is offered by Fischer (1978), who argues that American Sign Language (ASL) can indeed be considered a creole language. Her best argument for this point of view is that most deaf children have to learn it from hearing parents with an other language than ASL as their native tongue. It can be argued, then, that these people use some form of pidgin.

However this may be, Sign language seems to be a very interesting phenomenon for a UG-based linguistic theory. Yet this is not what we would expect under the diachronic distinction, because ASL has come into existence, when two gentlemen (T.H. Gallaudet and L.Clerc) brought French Sign Language (FSL) to the United States (cf. James Woodward (1978)). FSL has been *invented* at the end of the eighteenth century by Abbé C.M. de l'Epée. This would mean that both FSL and ASL (which are not mutually understandable) are artificial languages, even if they both have several thousands of native speakers.

Coming to a conclusion, we can see that the diachronic distinction between artificial and natural languages doesn't make much sense in the generative programme. What we have called the *synchronic distinction* comes closer to the distinction I-language versus non-I-language, (as we can learn from the definitions of the two notions already) which is more relevant to generative grammarians. This means that it makes only sense to study a language which has one or more native speakers.

There is still another objection against any diachronic treatment of the question discussed here.

We can nowadays safely suppose that English is a natural language, because it is spoken *since early days* and nothing is known of an inventor. But let us imagine for a moment that a group of archaeologists finds some ancient documentary, consisting of a few specimens of grammar and a tentative English-Norwegian dictionary, which proves that English in fact is invented by the poet of Beowulf. From that moment on, we would have to assume that English is an artificial language and can be no longer the subject of linguistic inquiry. We presume that this result is rather undesirable for any linguist.

1.1.3 A brief overview of Esperanto history

Despite the preliminary arguments Esperanto certainly takes a unique position in the language spectrum. The language can in some way be called a pidgin/creole language but it has been historically submitted to far more planning than the languages of the preceding paragraph, because it has been designed with the specific purpose to be the simplest language in the world.

Yet this simplicity is difficult to describe in terms of modern linguistics. If anyhow, it can only be easily expressed in terms of nineteenth-century descriptive grammar. Or rather, in terms of *E-language*. The objective of the language designer was not to create something as *natural* as possible, but to make it as logical as possible. Whatever he meant with this notion 'logical', we cannot

The syntax of a language without grammar

assume that he could have foreseen the relatively deep underlying logic in the language system that researchers have found during the past thirty years.

The language was designed in the late 1880s by the young Polish-Jewish oculist L. Zamenhof (1887), who intended it to become an auxiliary language for international communication. The purpose was that all people of the world would learn Esperanto as their second 'neutral' language so as to use it in international communication. Despite the fact that there have been many competitors (of which Volapük, Interlingua and Ido are certainly the best known) at the time and later, Esperanto spread rather quickly through Europe.

The most important Esperanto association, the *Universala Esperanto Asocio* was founded in 1908. It now has approximately 50,000 members, mostly Eastern European. Each month a few books and several magazine issues appear in the language. There are broadcastings by among others Vatican, Polish and Chinese radio world services.

Estimations about the number of speakers of Esperanto vary between 50,000 and 10,000,000. More importantly, there are some native speakers, mostly children of enthusiastic binational couples of Esperantists. This group consists in all probability of less than 1000 persons around the world. As far as we know, all of these people are bilingual, which is a reason to treat the grammaticality judgements of these people with some care, but not to exclude them totally from linguistic research.

As far as we know, Esperanto is the only completely planned language which has attained a community of native speakers, however small and weird this community may be. Hence, we suppose it is an interesting case of study for generative grammar.

1.2 The principles of Esperanto syntax

Brandt Corstius (1988) has remarked that 'Esperanto has no syntax' because syntax was not really an object of very deep study at the time the scheme for the language was published in the 1880s. With this remark, Brandt Corstius rallied round to Zamenhof, who himself has once claimed that 'Esperanto has no grammar'. The 'fundamental grammar', which appeared in Zamenhof (1962), consisted of only 16 rules, most of which were about orthography, phonology, lexicon and semantics; only two of them can be considered syntactic with a little good will.'

It is, at least for a theoretician working within the GB framework, very difficult to imagine a human language with *no grammar* and with *extremely free word order*, which are two other traits that Esperanto has been claimed to possess. Especially we don't expect that the Esperanto of the native speakers has no syntax or an absolutely free word order with no compensation in other modules in the grammar (for instance, in the form of a complex morphological case system).

Both claims accordingly can not be maintained in reality; in the second part of this essay we will find that Esperanto has some complex syntactic features. Besides, the word order of Esperanto is certainly not even half as free as that of Warlpiri or other non-

The syntax of a language without grammar

Indo-European languages with a really free word order. In other words: Esperanto is a configurational language. Hale (1981) suggested a clustering of properties for non-configurational languages:

- (1) a. Free word order
- b. The use of split or discontinuous constituents
- c. Free or frequent pro drop
- d. The lack of NP-movement
- e. Lack of expletive elements (it, there)
- f. Use of a rich case-system
- g. Complex verb words
- h. The lack of VP-rules (preposing, deletion)

Later research has proved, however, that so-called non-configurational languages like Hungarian and Japanese displayed only a subset of these characteristics. (cf. Marácz (1989)); the same was true for configurational languages like Italian and Dutch.

This is not important, however; if Esperanto really was not subject to syntactic rules, we would predict that it would meet the requirements (1)a, (1)b, (1)c, (1)g, and (perhaps) (1)e, and that it would not meet the requirements (1)d, (1)h and (perhaps) (1)f. Most of these predictions are not borne out, as far as we can see.

For example, a sentence like (2)b (with a discontinuous form of the NP *tiun kangaruon*) cannot be found in Esperanto, while its counterpart is found in Warlpiri (Hale (1983)):¹⁰

- (2) a. (Warlpiri)
 Wawirrikapi-rna panti-rni yalumpu.
 kangaroo Aux spear nonpast that
 'I will spear that kangaroo.'
- b. (Esperanto)
 *Kangaruon mi batos tiun.
 kangaroo I beat-FUT that

Furthermore, Esperanto has only a limited form of pro-drop, no strong Aux node and no VP deletion, although it lacks overt expletives and a rich case system.

We can explain the fact that people like Zamenhof thought that Esperanto had no syntax in at least three different ways.

The first is that at the end of the nineteenth century the concepts of classical grammar, which were based upon the data of Latin and Greek only, were still thought to have a universal value across languages¹¹. The second explanatory fact is that (linguistically untrained) people in general seem naively to think that their own language has no grammar, or else that it has the simplest grammar possible (cf. the 'White' point of view in matters of second language acquisition, as discussed in section 1.3).

A third possible explanation comes from Marácz (1989), who argues against what he calls the *Hungarian-as-a-different-language-doctrine*, i.e. the view that Hungarian is a language with rather specific properties which do not turn up in other languages. (Marácz (1989:403)). With reference to Staal (1986), Marácz contends that the

The syntax of a language without grammar

western tradition with respect to the study of language has been *word oriented*:

In Hungary, a country belonging to the western cultural sphere, this view [that language is a collection of words] has been popular as well. This may be observed from the fact that grammar books on Hungarian mainly contain long lists of morphological paradigms. It is often claimed that this covers the whole language structure.

Marác (1989:403)

If we replace 'long lists of morphological paradigms' by 'extremely short lists of morphological paradigms', the same argument would apply for the *Esperanto-as-a-different-language-doctrine* as well.

1.2.1 Esperanto as a Slavic language

The attentive reader of the following chapters in this essay will find that Esperanto bears on an abstract level some striking similarity with the Slavic languages (in particular Russian) in the features discussed there.¹²

This should not be a reason for much astonishment. Zamenhof, the creator of Esperanto, has used Russian as his home language during most of his life (Boulton (1960)) and the first brochure on the language was in Russian (Zamenhof (1887)). Up until now, a large majority of the Esperanto speakers lives in Russia and Poland (Piron (1989)) and as we have outlined above, a naive language user may very well assume that the grammar of his language is at most minimal; he doesn't really perceive the underlying complexities. He will therefore expect that a language without a grammar behaves in many ways in the same way as his own language does.

On the other hand, certainly not all of the parameter settings are influenced by the Russian manner of speech - Esperanto is not just a variant of Russian under the disguise of a pile of French words ending in -o, -a and -e.

We will give two examples, one from morphology and one from syntax.

First, Esperanto morphology seems to obey the Right Hand Head Rule (Williams (1981))¹³ quite strictly, while Russian doesn't. Schubert (1989c) points out that this is so¹⁴, even while Zamenhof himself sometimes sinned against this principle: where Zamenhof created *mil-jaro* (thousand-year, millenium) to the analogy of Russian *tysjačelietie*, the common word now is *jar-milo*.

Also the Head Final Filter (Williams (1982)) doesn't seem to hold in Russian, while it does in Esperanto, i.e a sentence like (3)a is ungrammatical in Esperanto (as it is in English), while something like it can easily be said in Russian.

The syntax of a language without grammar

- (Esperanto)
- (3) a. *Ĉi estas agrabla por legi libro
It is nice for read book
b. Ĉi estas agrabla libro por legi
It is nice book to read
'It is a nice book to read'

One could inquire whether these differences between the two languages could be traced to some deeper underlying distinction, but we will not go into that topic here. Henk van Riemsdijk (p.c.) notes that also in English some restricted form of apparent HFF violation is allowed. A sentence like (4) seems to be rather good in that language.

- (English)
- (4) (He is) a nice to talk to colleague.

Because of the *preposition stranding*-like construction, we might conclude here, however, that there is connection with WH movement constructions. In that case, according to Van Riemsdijk, some sort of Reanalysis is possible, thus avoiding a violation of the HFF.

If this is correct, there might be some parametric variation in languages with respect to the cases where Reanalysis is possible and where it isn't. This issue deserves more discussion than we can give it here.

1.2.2 The simplicity of a language

A natural question that now arises, is: if Esperanto is a natural language, guided by the principles of UG as any other language is, can it be really as simple as its advocates claim it to be? The answer, of course, depends on the way in which one defines the notion *simplicity*.

Chomsky (1975, 1988, e.a.) has pointed out more than once that human language is not as simple in a computational sense of the word as it could have been. An example he often quotes to illustrate this point is question formation. In Spanish (and in English) we find declarative sentences like in (5)a and (5)b and more complex sentences as in (6).

- (Spanish)
- (5) a. El hombre está en la casa.
The man is in the house
'The man is at home.'
b. El hombre está contento.
The man is happy.
- (6) El hombre, que está contento, está en la casa.
The man who is happy is in the house
'The man, who is happy, is at home.'

Let us now consider for a moment the way in which the forms in (5) are made interrogative (again both in Spanish and in English):

The syntax of a language without grammar

- (Spanish)
- (7) a. Está el hombre en la casa?
Is the man in the house?
'Is the man at home?'
- b. Está el hombre contento?
Is the man happy?

Clearly the most simple rule to explain the difference between (5) and (7) is 'find the first occurrence of *está* (or other verbs like it) and move it to the front of the sentence.'

Subsequently, comparing (6) to (8)a and (8)b, we perceive that this rule can't work and that a different, more 'complex' rule is needed - probably one that refers to the structure of the sentence.

- (Spanish)
- (8) a. *Está el hombre, que contento, está en la casa?
Is the man , who happy, is in the house
- b. Está el hombre, que está contento, en la casa?
Is the man who is happy in the house
'Is the man, who is happy, at home?'

Esperanto, now, is by no means *simpler* in this respect than any known natural language; we could translate the inversion problem into this language without any significant problem¹⁵. This is also what we anticipated. For, given what we now know about the language faculty, we don't expect to be able to construct a language in which the interrogative form of a sentence like (6) will have a structure like (8) and which at the same time will be experienced by a (human) language learner as easy to learn.

If we assume that Esperanto is a natural language, and we have seen that we have little reason not to, we can on first sight not express in any way the fact that the core language would be much simpler. For we cannot really express that one set of parameter settings is *a priori* simpler than another. Yet this seems to us to be a problem (although certainly not a death-blow) for the parametric approach.

We take an example from phonology to illustrate this point. Since Hayes (1981), it is generally assumed in phonological work that the variation in the metric systems of the languages in the world can be expressed by a limited set of parameters. Now let us consider two of those systems:

- (French)
- (9) *Stress the last syllable of the word.

The syntax of a language without grammar

- (10) (Palestine Arabic, cf. Hayes (1987))
*If the penultimate syllable is heavy, stress it.
*If the word consists of exactly four light syllables, stress the first syllable.
*In all other cases, stress the antepenultimate syllable.

Under a parametric approach to grammar, we would now suppose that the French and the Arabic system could be accounted for in a system of, say, 16 parameters with different settings. Yet someone with a native language totally unrelated to either French or Arabic, say a Chinese, would probably find the French system much easier to acquire. It remains an open question why this is so.

The solution to this problem can in our view, if anywhere, only be found within markedness theory. Markedness has been introduced as a linguistic notion in the 1930s by the phonologist Trubetzkoy. During the last fifteen years, the concept has also received some attention within the UG-oriented framework, especially since works like Van Riemsdijk (1978) and Chomsky (1981) and Belletti, Brandi and Rizzi (1981).

There are at least three types of markedness within this framework, namely markedness within the core, the core-periphery dichotomy, and markedness within the periphery¹⁶. Furthermore, it is often assumed that a 'marked' construction is more difficult to learn than an unmarked one, at least in first language acquisition.

The same would apply however, for second language learning, if this would imply at least partly the same processes. As we shall see in the next section (1.3) there are at least two different schools of L2 researchers: those who believe that the acquisition of a second language is more or less influenced by the parameter settings of the first language (e.g. White (1985)), and those who believe that the process of learning a second language is an exact parallel of that of learning a first language (e.g. Mazurkewich (1985), (1988)). The last class of researchers believes that a second language learner at first assumes the unmarked settings for the new language.

This problem is perhaps also somewhat related to the *universalist versus substratist debate* in creolistics (cf. Muysken and Smith(1986)).

According to the universalist point of view in this debate, the grammars of creole languages take the unmarked settings for the parameters of UG. The substratist point of view, on the other hand, has it that every creole language inherits all of its features from the donor languages. The parameter settings are no exception to this.

Like for any creole language¹⁷, one could suppose that Esperanto takes the unmarked options in the parameters of UG. This would then explain the relative simplicity of the language, if the unmarked options of the language faculty are easier to learn. Unfortunately, it is very difficult to test this claim, because there is no real agreement among linguists about what is the unmarked option of what parameter. This lack of agreement has already led to some undesirable consequences; as Gair (1988) reports:

The syntax of a language without grammar

For example, White (1985) and Phinney (1987), in their studies of the 'Pro-drop parameter' in L2, come to different, in fact diametrical conclusions based on the same fundamental assumptions and non-conflicting data, as a result of assuming opposite markedness values for that parameter. [...] White claims that marked L1 settings persist in L2, on the assumption that the non-pro drop setting [...] is the unmarked one. Phinney, however, [...] claims that features reflecting unmarked settings are easier to learn on the assumption [...] that the pro-drop setting [...] is the unmarked one.

Hopefully it is clear that it isn't easy to sort out whether the parameter settings in Esperanto are the unmarked ones, given this controversy.

Yet, it would be rather surprising to us, if they would. As we have argued in the preceding section and as we shall see in the next chapters, Esperanto is very similar to the Indo-European (Slavic) languages in its syntactic features.

If we assume that L2 learning is influenced by the parameter settings in the first language, this would mean that Esperanto can only be simpler to native speakers of other Indo-European languages. If we take the other position and assume that a L2 learner is guided by the unmarked options of UG, Esperanto would only be simpler if its parameter settings were unmarked. But in that case, also Russian would be an extremely simple language for any L2 learner to acquire. We have found no evidence supporting this remarkable prediction.

But there is still another aspect in which Esperanto perhaps could be called simpler than other natural languages: the periphery of the language could be minimal. Remember that the *periphery* of a language consists mainly of the language's irregularities and that it is this periphery that most of the attention is devoted to in text books for second language learners. It also is not entirely precluded that the periphery can be left out in any language without any danger for the language to become unlearnable.¹⁸

We think that this lack of periphery is what makes the language unique and different even from other creole languages, which have come into existence almost by accident and inherit some of the irregularities of their donor languages.

On the other hand, we find it significant that even in Esperanto after 100 years of existence some sort of periphery is tentatively coming into being (see e.g. Piron (1986)). We suppose that the study of Esperanto (and that of creole languages) could shed some new light on the rather obscure question what the function of the periphery is in the language faculty. Unfortunately we have no space here to explore this issue.

1.3 The status of judgements of bilinguals and L2 speakers

The language of the 'pure' native speaker is the main concern of most GB theoreticians. It is often assumed that UG has some sort of 'critical period': a child can learn a language until a certain age,

The syntax of a language without grammar

because during those first years of life he has access to his internal principles and parameters. Adult second language learning in this view relies on other, more general, learning strategies. The supposition that these (unknown) general devices are less suited for the learning of a language, then explains the fact that most adults seem to be unable ever to speak a foreign language as fluently as their mother tongue.¹⁹

We will examine shortly examine this assumption here, because it has some relevance for the linguistic research of Esperanto. For, because, this language has no monolingual speakers, there would be no real point in this sort of research.

We will assume in this essay that there is only a gradual difference between 'native' bilingual speakers and second language learners. In more general terms, this assumption may not be totally justified; but we think that in this case it is, because even the real 'native speakers' of Esperanto can use the language only in a very restricted domain, namely in their family life. The language will not be spoken at school or by other children in the street. For an overview of bilingualism research in relation to linguistic theory and to L2 research see Hyltenstam and Obler (1989b) and Meisel (1989).

1.3.1 A 'poverty of the stimulus' problem

One of the strongest arguments in favour of a linguistic theory of UG is the so-called *poverty of the stimulus* argument: the fact that a child knows things about his language that have never explicitly been told to him, can only be explained by reference to some inheritable linguistic device. (We have chosen one of Chomsky's most recent formulations of this argument as a motto for this essay.)

White (1985), among others, has pointed out, however, that there is no fundamental difference concerning the logical problem of language acquisition in L2 or L1 learning. Also adults learning a second language, attain some form of grammatical knowledge that is underdetermined by the available evidence.

One very good example of this fact is shown by Felix (1988), who tested a group of native speakers of German that had learned English for some years for their ability to give correct English grammaticality judgements. The test sentences reflected linguistic phenomena that are, as a rule, not explicitly taught in the classroom and that have no direct counterpart in German, like parasitic gaps, superiority effects, control vs. ECM verbs, etc. We have given examples of two sets of such test sentences in (11) and (12).

- (11) (English -Case filter)
- a. John seems to love Mary
 - b. *Mary seems John to love
 - c. John seems to like Bavaria
 - d. *Bavaria seems John to like

The syntax of a language without grammar

- (English -SSC)
- (12) a. who did the man see pictures of
b. *who did the man see John's pictures of
c. who did he hear stories about
d. *who did he hear John's stories about

Now, whereas all the counterparts to the sentences in (11) are grammatical and all the counterparts to the sentences in (12) ungrammatical in German, and whereas the Case Filter and the Specified Subject Condition are subjects that most second language teachers are unaware about, the judgements of the German natives were not very dissimilar from those of an English speaking control group. Furthermore, whether or not an individual had lived in an English-speaking community for some time appeared to have very little effect on his judgements.

All of this would be a great mystery, if we couldn't resort to an UG based approach to L2 acquisition. On the other hand, if we do adopt a principles-and-parameters view in these matters, these facts can easily be explained, because the Case Filter and the SSC are (or reduce to) principles of the UG. We will for this reason adopt this point-of-view here.

It should be noted, however, that the issue is not without controversy. One of the main opponents of the UG based approach to L2 learning is Clahsen (Clahsen (1988), Clahsen and Muysken (1986), see also Obler (1988)), who argues that given the clusterings of acquisition in childhood that can be explained by a given parameter, components of which occur before or after each other in different (adult) L2 acquirers, we should conclude that L2 acquisition follows a different path.

Also the fact that many adult language learners never reach the level of proficiency of a native speaker, should warn us, that although L2 acquisition may be subject to the same principles and parameters as L1 learning, there might also be other cognitive processes involved. Still we think that there is reason to assume that the difference between a L1 learner and an adult L2 acquirer is only gradual, not fundamental.

1.3.2 The acquisition of a second language

There have traditionally been two ways of regarding second language acquisition: Contrastive Analysis (CA) and Creative Construction (CC). Although most authors in e.g. Flynn and O'Neil (1988) have assumptions that are somewhere halfway between CA and CC, the division is still visible.

Within Contrastive Analysis (Fries (1945), Lado (1957)) it is assumed that L2 acquisition is totally different from L1 learning. It is supposed that that the learner of a second language attempts to transfer the linguistic habits from his first language to the L2. Where the L1 and the L2 match *positive transfer* takes place: where they do not match there is a *negative transfer* of habits. At points of interference, the learner must acquire the new L2 habits through *deletion* or *addition*. As can be deduced from the terminology, this theory has its roots within the behaviorist framework.

The syntax of a language without grammar

Creative Construction (Dulay and Burt (1974)), on the other hand, supposes that the L1 and L2 acquisition processes follow the same set of innate principles and a prior language experience does not determine subsequent L2 acquisition. For a thorough comparison between CA and CC and a discussion of the drawbacks of both theories see Flynn (1987).

As we have seen above, both points of view are represented within the community of UG oriented L2 researchers. For example, Haegeman (1988) argues for a CA based approach on the basis of the use of modals in the English of Dutch speakers from Belgium and French speakers from Switzerland. Haegeman argues that her facts can only be explained, if it is assumed that both groups of speakers suppose English to have the same parameter settings as their own language.

Mazurkewich (1988) on the other hand, argues that L2 learners have a preference for unmarked constructions, even if they are uncommon both to their own and their target language. According to her, Inuit learners of English as a second language, acquired the use of the less marked infinitive before that of the more marked gerundic construction in sentences like in (13).

- (English)
- (13) a. Philip likes to buy Inuit prints.
b. Philip likes buying Inuit prints.
c. Nick offered to help Nora.
d. *Nick offered helping Nora.
e. *Robert finished to write an article.
f. Robert finished writing an article.

As we have, seen above, it is very difficult to decide in this question, given the lack of agreement in questions of markedness theory.²⁰ For example, the three phenomena of Esperanto syntax discussed here, can be explained in both ways: the pro drop parameter setting is similar to that of Russian, but it might be the 'unmarked' setting as well. The same applies for the Esperanto reflexivization process and, *mutatis mutandis*, for the Esperanto directional prepositions.

1.4 Conclusion

Concludingly, we can say that the distinction between artificial and natural languages within the framework of GB theory only makes sense in a synchronic way. The origin of a language can in no way be regarded of much interest to a generative grammarian. What matters, is whether a language has native speakers or not. Since Esperanto has a few native speakers and since second language acquisition involves at least partly some UG parameter setting, we should certainly regard it as a natural language.

This same hypothesis will be the starting point for some actual linguistic research in the next three chapters. We have more or less at random chosen three topical subjects of linguistic inquiry. We will show that Esperanto behaves like modern theories (under the assumptions made here) predict.

The syntax of a language without grammar

2 The Binding nature of *si(a)*

2.0 Introduction

In this chapter, we will explore the Binding theory features of the Esperanto reflexive pronoun *si* and its possessive variant *sia*²¹. We will find that these categories show a very strange behaviour, which can, at least partly, be explained under the modern Binding theory.

This chapter is divided as follows. In section 2.1 we will give an overview of the relevant facts. In section 2.2 we analyze these facts. Section 2.3 is devoted to a conclusion.

2.1 The facts

The basic rule about the Esperanto reflexive *si(a)* is informally stated in Esperanto text books and descriptive grammars as

- (14) *si(a)* can only have the (3rd person) subject of its clause as its antecedent.

Some of the consequences of this rule are illustrated in (15)^{22,23}. In these examples, *li* is the third person singular masculine pronoun and *ŝi* (pronounced as *she*) the third person singular feminine pronoun.

- (15) (Esperanto)
- a. Paŭlo_i ludas kun Petro_j kaj li_i vundas sin_{i/*j}-
/lin_{*i/j}
Paul plays with Peter and he injures
himself/him
- b. Anna_i parolas kun Maria_j en sia_{i/*j}/ŝia_{*i/j} ĉambro
Anna talks with Maria in her (own)/her
room
'Anna talks to Maria in her room.'
- c. Paŭlo_i rifuzis al Petro_j PRO_i fari sian_i/lian_j
devon.
Paul refused to Peter to do his
duty
- d. Paŭlo_i demandis al Petro_j PRO_j fari sian_j/lian_i
devon.
Paul asked to Peter to do his
duty

As can be noticed from (15)c-d the phrase *the subject* in (14) is not to be read as *the subject of the main clause* but as *the nearest subject*, as is indicated in the examples by using the PRO of standard Binding Theory. This nearest subject can even be the implicit or explicit subject in a noun or adjective:

The syntax of a language without grammar

- (Esperanto)
- (16) a. Paŭlo_i amis Martan_j, [PRO_i konsciante pri sia_i
amo al ŝi_j/si_i].
Paul loved Marta being-conscious-about his
love to her/himself
- b. Paŭlo_i amis Martan_j, [PRO_i konsciante pri ŝia_j
amo al si_j/li_i].
Paul loved Marta being-conscious-about her
love to herself/him
- c. Jam blindiĝanta, li volis viziti [la muzeon
[PRO_i faman pro siaj/liaj pentraĵoj]]
Already blind-getting, he wanted to-visit the
museum famous for its/his paintings.

Kalocsay and Waringhien (1935/85) note, however, that there is one clear exception to this rule²⁴, namely *sia* can relate to words like *ĉiu* 'everybody' and *ĉio* 'everything' even if they are not in subject position. According to Kalocsay and Waringhien (1935/85), the word has a different shade of meaning then, namely *sia propra*, 'his own'.

- (Esperanto)
- (17) a. Li metas ĉion en sia loko.
He places everything in its own place
- b. Por ĉiu afero estas sia tempo.
For every affair is its own time
'There is a good time for everything.'

Pretheoretically, the sentences in (17) mean a complication for the simple rule in (14). Again this complexity is inexplicable, if we cannot refer to any quality of the language faculty itself.

As far as we can see, there is no very clear equivalent to these facts in any (Slavic) language. Although subject binding is the default case also in e.g. Russian, we have found no evidence that QPs form a distinctive class of exceptions to this rule. But deeper research might be needed here, because the Russian reflexive Binding facts are themselves rather complex (see Timberlake (1980) for a discussion of those facts within the framework of Relational Grammar).

2.2 Analysis

The fact that the Esperanto reflexives have to be bound within a certain domain, is not very interesting from a theoretical point of view. It follows as a matter of course from the fact that they, as anaphors, are subject to Condition A of the Binding Theory:

- (18) An anaphor must be bound (at LF) within its governing category.

It is interesting that in Esperanto at first sight more projections of lexical categories seem to be able to perform the function of governing category than we are used to in for example English and Dutch. Put otherwise: it seems that in Esperanto the SPEC position of

The syntax of a language without grammar

the nominal and adjectival projection can also be filled with PRO much more freely than in those languages.

Chomsky (1986b) has argued, however, that also in English, we could sometimes assume an instance of PRO in an NP, in order to explain the various facts in (19).²⁵

- (English)
- (19) a. They_i told [stories about each other_i].
b. *They_i told [my stories about each other_i].
c. *They_i told [stories about them_i].
d. They_i told [my stories about them_i].
- e. They_i heard [stories about each other_i]
f. *They_i heard [my stories about each other_i]
g. They_i heard [stories about them_i]
h. They_i heard [my stories about them_i]

In (19)b and (19)f, the subject *my* blocks the binding of *each other* to *they* in the main clause, while it permits the binding of *them* in (19)d and (19)h. In (19)c and (19)g an instance of PRO is assumed. Control Theory is now responsible for the fact that PRO in (19)c must be coreferent with *they* in the main clause, thus forbidding the binding of *them* indirectly. In (19)g, on the contrary, Control Theory lets PRO obligatory be disjoint from the subject of the main clause.

(19)a and (19)e under this theory prove that PRO is optional in NPs.²⁶

Another argument in favour of the hypothesis that NPs can have PRO in their subject position is the contrast in (20) (already discussed by Ross (1967))

- (English)
- (20) a. *[the knowledge that John_i might fail] bothered him_i
b. [the possibility that John_i might fail] bothered him_i.

It should be observed that there is a reading in which (20)a is grammatical, namely the reading where there is some sort of *general knowledge* that John might fail; the sentence is not possible, however, under the reading that John knows that John might fail, for the same reason as why (21) is ungrammatical:

- (English)
- (21) *[PRO_i knowing that John_i might fail] bothered him_i.

Both in (21) and in (20)a we assume a PRO in the subject position. Both of these sentences fall then out immediately, because they are violations of Principle C of the Binding Theory. In (20)b, on the other hand, we have no reason to assume the existence of a PRO within the NP; in fact there even seems to be no room for any subject at all:

The syntax of a language without grammar

- (22) (English)
*Our possibility that John will come, bothers Jane.

Thus, we do not seem to encounter grave difficulties if we assume that the Esperanto NP has a subject position which can sometimes be filled up with PRO. We can now turn to the subtleties and the details of the Esperanto reflexive system.

First of all, notice that there is something very *illogical* in the formulation of (14). Since *si(a)* must always reflect the subject of the clause, we expect that it cannot be part of the subject itself. This expectation seems to be affirmed by the examples in (23); *si* is never found in the nominative case.

- (23) (Esperanto)
a. **Si* batas Johanon
 himself beats John
b. **Sia* patrino amas Johanon

At the same time, from the examples in (16) we learn that the subject position of an NP is also relevant for the binding of reflexives inside the NP. Combining these two observations, we would expect that *sia* were impossible in the subject position of NP, and consequently, that *sia* did not exist at all.

We could try to relate to this point the observation that English doesn't have a possessive reflexive. This latter observation is open to question, however, because, as is noted by Higginbotham (cited in Chomsky (1986b)), the English phrase *his own* has the same anaphoric qualities as a possessive reflexive. This is illustrated in (24)

- (24) (Esperanto)
a. [La patrino de Maria_j]_i ŝatas siajn_{i/*j} amikojn pleje.
(English)
b. [Mary_j's mother]_i likes her own_{i/*j} friends best.

Both *her own* in (24)b and *sia* in (24)a can only refer to the whole NP [*Mary's mother*], not just to Mary, because of the c-command condition which is included in the definition of governing category, so in English we have at least the same problem.

We think that the same kind of argument can be set up for the literal translation into Dutch of 'his own', *zijn eigen*. This phrase can in some Dutch dialects even be used as a reflexive in other positions than that of [SPEC, NP] as we can see in

- (25) (Dutch)
Hij houdt van zijn eigen.
He loves his own.
'He loves himself.'

The syntax of a language without grammar

And even if we don't accept the English *her own* as a good counterpart to the Esperanto *sia*, there still is the reciprocal *each other* to consider, as in (26).

- (26) (English)
These boys like each other's friends.

We have a real problem here, which we can divide into two subproblems. Firstly, why is the Esperanto reflexive system strictly subject-oriented in its binding features? Secondly, why is the subject of an NP different from the normal subject?

The fact that Esperanto reflexives can normally only be bound by the subject, not by the object, is in itself crosslinguistically not a reason for amazement, because the language shares this feature with many languages of the world. In fact, it is supposed by some linguists that languages like English in which a reflexive can also be bound by (in-)direct objects is exceptional on this point (but see also Szabolsci (1986) who contends that '*this crosslinguistic argument is not watertight*').

Anaphoric binding sometimes is strictly subject-oriented even in English, namely in the case of long-distance binding, as is illustrated in (27), which is taken from Chomsky (1986b).

- (27) (English)
a. They told us that [[pictures of each other] would be on sale]
b. They told us about each other.

In the normal case in (27)b, *each other* can be ferent both with *us* and with *they*, but in the long-distance case in (27)a only the subject of the main clause *they* is available as an antecedent for the reciprocal. This restriction is explained by Chomsky (1986b) by supposing that anaphors undergo LF-movement to the INFL-position, leaving a trace. The LF representation of (27)b then looks more or less like (28)

- (28) (English -LF)
they each other_i-INFL [_{VP} tell us about e_i]

This movement is supposedly the same as the one that moves the reflexive clitics to their preverbal position at S-structure in Romance languages like Spanish or Slavic languages like Russian:

The syntax of a language without grammar

- (29) (Spanish)
Juan se afeitada.
John self-shaves.
'John shaves himself.'
- (30) (Russian)
Deti mojucs'a teper'.
children wash-self now
'The children are washing themselves now.'

In fact, we suppose that the same rule can apply in Esperanto optionally (or even preferredly) at S-structure with *sin* although it cannot with emphatic *sin mem*. For more discussion of the treatment of Esperanto pronominals and reflexives as clitics, we refer to chapter 4 (page 64).²⁷

- (31) (Esperanto)
a. Johano sin 'a as.
John self-shaves.
b. ??Johano sin mem 'a.as.
John himself shaves.
'John shaves himself.'

Reflexive movement would be obligatory at S-structure in the Romance languages, obligatory at LF in English and optionally at S-structure or LF in Esperanto (and probably Russian), like wh-movement is obligatory at S-structure in English, obligatory at LF in Chinese and optionally at S-structure or LF in French.

It now follows that *each other* can only be bound by the subject in both (27)a and (27)b. Object binding can in some cases be accomplished by adjunction to VP and a '*slight revision of the notion c-command that may well be required for other cases not discussed here.*' (Chomsky (1986:fn. 43))

Thus we see that this theory, although it is devised explicitly only to account for the English facts, inherently predicts the relative crosslinguistic markedness of object binding. Chomsky did not elaborate on the way in which object binding would work in English, but one could imagine some sort of parametric variation there.

Subsequently, we have to explain why the subject position of an NP, unlike that of an S, can be filled with a reflexive. We can also here try to let Chomsky do the hard work for us. Chomsky (1986b) observes that, in English, both (32)a and (32)b are grammatical.

- (32) (English)
a. The children like [each other's friends]
b. The children like [their friends]

Chomsky (1986b) explains this fact as a consequence of the difference in the binding nature of anaphors which *must* be bound and pronominals which *may* be free. Hence it follows that the relevant local domain is somewhat different for anaphors and pronominals. He now reformulates the Binding Theory in such a way that for each anaphoric or

The syntax of a language without grammar

pronominal expression E the relevant governing category is the least domain in which E could satisfy the binding theory with some indexing. This indexing may be different from the actual indexing in the expression under investigation. (We will not go into the details of this theory.)

In Esperanto, however, only the counterpart to (32)a is grammatical.

- (33) (Esperanto)
- a. La infanoj_i ŝatas [siajn_i amikojn]
The children like their (own) friends.
- b. *La infanoj_i ŝatas [iliajn_i amikojn]
The children like their friends.

It is obvious that a Binding Theory along the lines of Chomsky (1986b) can't easily explain this contrast. We therefore reject this theory and return to an older formulation of the Binding Theory, which makes use of the notion Accessible SUBJECT.²⁸

- (34) *SUBJECT* (formulation of Van Riemsdijk and Williams (1986:275))
The SUBJECT of a clause is [AGR_i, S] if there is one. otherwise [NP_i, S] or [NP_i, NP].

- (35) *Accessibility* (formulation of Van Riemsdijk and Williams (1986:276))
A is accessible to B iff A c-commands B and the assignment of the index of A to B doesn't lead to a violation of the *i*-within-*i* Condition.

Under this formulation, the Esperanto facts fall out immediately. The NP in principle doesn't count as a governing category for its own subject position (because this would lead to a violation of the *i*-within-*i* Condition). Both in (33)a and (33)b, it is the S that counts as the governing category and in which the anaphor can whereas the pronominal can not be bound.

But now we would expect the same contrast in the English sentences in (32). We tentatively suggest here, that perhaps there is such a contrast, but this is made invisible that the English possessives are ambiguous between a pronominal and an anaphoric reading, just as the Esperanto first and second person pronouns are (cf. (36)) (where *vin* is the accusative form of *vi*).

- (36) (Esperanto)
- a. Vi amas vin.
You love you
'You love yourself.'
- b. Vi diras ke vi amas min.
You say that you love me.

The English possessives can be thought to be ambiguous in the same way as the Esperanto pronouns.

The syntax of a language without grammar

In Dutch we have two forms for the third person singular masculine possessive, viz. *z'n* and *zijn*. When we assume that *z'n* is preferably read as a reflexive and *zijn* as a pronominal, we see that Dutch shows the same behaviour as Esperanto:

- (Dutch)
- (37) a. Hij_i leest zijn_{??i/j} boeken.
 He reads his books
- b. Hij_i leest z'n_{i/j} boeken.
 He reads his own books

We now turn our attention to the very remarkable examples of (17). We propose to explain them with reference to the theory of *Generalized Binding*.

Aoun (1986) proposed to extend the theory of Binding so that the ECP could partly or perhaps even totally be reduced to it.²⁹ One of the crucial ideas in his work was that there are two kinds of anaphora, viz. *A anaphora* and *A bar-anaphora*. An *A anaphor* is bound by an antecedent in an Argument (subject, object,...) position, an *A bar anaphor* is bound by an antecedent in a non-argument position. Both *A* and *A-bar* anaphora can be found in an *A* or an *A-bar* position, so that we have the following possibilities (which we copy from Aoun (1986))

<i>Antecedent</i>	<i>Anaphor</i>	
a. A-position	A-position:	r e f l e x i v e s , reciprocals, NP- traces
b. A-position	A-bar-position:	f l o a t i n g quantifiers, <i>l'uno</i> in reciprocal con- structions in Italian
c. A-bar-position		A-position:wh-trace
d. A-bar-position		A - b a r - position:traces in COMP

In this scheme, a. and b. are the *A* anaphors and c. and d. are the *A-bar* anaphors. It is noteworthy that while *A* anaphors can apparently be either visible or null, the only *A-bar* anaphors in this scheme are empty categories. This raises the question whether this is a mere coincidence or there is any principled way in which overt *A-bar* anaphors can be excluded.

The latter seems not to be the case: Aoun mentions in passing one example of a lexical *A-bar* anaphor himself, namely *l'altro* in Italian reciprocal constructions. Reciprocals can in Italian be expressed by the discontinuous expression *l'uno ... l'altro*. The members of this expression seem to be in a Binding relation to each other, as is illustrated in (38).

The syntax of a language without grammar

- (Italian)
- (38) a. Quei reporters ammiravano l'uno [_{NP} le foto dell'altro]
 those reporters admired the one the pictures of the other
 'Those reporters admired each other's pictures.'
- b. *Quei reporters ammiravano l'uno [_{NP} le tue foto dell'altro]
 those reporters admired the one your pictures of the other

The contrast between (38)a and (38)b can be explained by the SSC; the association between *l'uno* and *l'altro* is blocked in (38)b by the subject *tue* of the embedded NP. This means that *l'uno* stands in an anaphoric relation to *l'altro*. Furthermore, as Aoun argues, *l'uno* is in an A-bar position; for, if *l'uno* is in an A-position, as in (39), the association between *l'uno* and *l'altro* is no longer constrained by the binding theory.³⁰

- (Italian)
- (39) L'uno ammira le tue foto dell'altro.
 the one admires your pictures of the other
 'They admire your pictures of each other.'

Yet another example of overt A-bar anaphors we find in Aoun (1985) where it is claimed that *ne* in *ne...personne* constructions in one French dialect is in a A-bar position, while *personne* is in an anaphoric relation to it. This means that in this particular dialect both (40)a and (40)b are ungrammatical.³¹ (In standard French, only (40)b is ungrammatical because the negative polarity item here has the same distribution as a variable.)

- (French)
- (40) a. *Je ne veux que tu parles à personne.
 I NEG want that you speak to nobody
 'I don't want that you speak to anybody.'
- b. *Je ne veux que personne vienne.
 I NEG want that nobody comes
 'I don't want that anybody comes.'

In this dialect *personne* apparently has to be bound by *ne* within the subordinate clause; the latter is of course in an A-bar position, so that nothing prevents us from presuming *personne* is an A-bar anaphor.

Now we have established that overt A-bar anaphors are indeed possible (at least at A-positions), we propose to analyse the Esperanto reflexives *si* and *sia* as ambiguous between an A-bar anaphor and an A anaphor. As we mentioned before, the reflexives can in the normal cases be analyzed as a relatively unmarked kind of anaphor.

The fact that the reflexive has a slightly different meaning if it is associated with *ĉio* and *ĉiu* could indicate already that we have to do here with an ambiguous word. Furthermore, also the other forms of

The syntax of a language without grammar

the reflexive are ambiguous. Although we can't treat the first and second person reflexives as A-bar anaphors, they are phonologically identical to the first and second person pronominals, as has been illustrated in (36).

So we have no cogent argument not to consider *sia* as an ambiguous anaphor. Because the (Generalized) Binding Theory applies at LF, the anaphor could be bound by the universal quantifier in COMP³². For an example, we give the Logical Form of the relevant part of sentence (17)_a in (41) (where the quantifier is represented as moved to the SPEC of CP although it is sometimes argued that the landing site is the SPEC of IP. This question is totally irrelevant to our present discussion.)

(Esperanto -LF)
(41) [CP $\hat{c}ion_i$ [CP [li metas \underline{e}_i en *sia*_i loko]]]
 everything he puts in its place

Note that in fact we predict the sentence to be ambiguous: in one reading *sia* is an A-bar anaphor (this is the reading in (41)), bound by the universal quantifier, and in the other *sia* is an A anaphor bound by the subject. This last reading, in which somebody puts everything on *his* own place is indeed also available.

We could of course now wonder why (41) doesn't show a weak crossover effect, because in this sentence two elements are bound at the same time by the universal quantifier in its A-bar position. We assume here that for some reason anaphors are never subject to WCO, just like PRO isn't (see section 2.2.1 on page 49); we could perhaps generalize the observation to all [+anaphoric elements]. Riny Huybregts has recently suggested in a class lecture on the basis of Arabic data in unpublished work that we can reduce WCO to the requirement that an antecedent can only bind two elements that are similar in the relevant features. Perhaps one of these relevant feature here is [+/-anaphoric]; hence an (overt) pronominal and a variable cannot be bound the same way, but an overt A-bar anaphor and a variable can.

2.3 Conclusion

In this chapter we have seen that the Esperanto reflexives *si* and *sia* are ambiguous in a very interesting way. As in any other natural language, this behaviour is not very 'logical' in a superficial sense of that word, but at the same time it patterns with other natural languages in a way that makes us suspect that principles of UG are at work.

Although we don't have one consistent theory which can explain all the facts (and this is also the case with English), at least we have seen that the facts are not different from those in other languages if they are considered at a level of sufficient abstraction. On the other hand, the exceptional status of the word we call *quantifiers* is unexpected in a theory that doesn't treat Esperanto as a natural language.

The syntax of a language without grammar

3 An all-purpose accusative?

3.0 Introduction

In this chapter, we will examine the behaviour of Esperanto and some other languages with respect to the case theory. We will see that at first sight the (morphologically not very rich) case system of Esperanto has a lot of work to do within Esperanto syntax. In particular, we will examine the interaction between accusative case marking and prepositional phrases. A theory will be developed for the Esperanto facts and similar constructions in German, Hungarian, English, Russian and Dutch.

This chapter is structured as follows. In section 3.1 we will discuss the relevant facts in Esperanto, German and Hungarian and we will look what standard Case theory has to say about them. In section 3.2 we will try to analyze the facts that are left unexplained by standard theory. The last section is, like in the other chapters, devoted to a conclusion.

3.1 The case system of Esperanto

Given the relatively free word order in Esperanto, we would perhaps expect the language to have a rich declinational paradigm. This happens not to be the case. There are no more than two distinct forms for the majority of nominal, adjectival and pronominal elements.³³ These are traditionally called a *nominative* and an *accusative* case marker respectively.

The accusative is marked by the ending -n, whereas a nominative is marked by a zero ending. The (definite) article doesn't show any case marking at all.

		(Esperanto)		
			NOM	ACC
(42)	a.	<i>pronoun:</i>		mi-min'I/me'
	b.	<i>noun</i>		domo-domon 'a house'
	c.	<i>adjective:</i>		b e l a - belan'beautiful'
	d.	<i>NP:</i>		la bela domo - la belan domon 'the beautiful house'

When other thematic relationships occur in a sentence, they are normally represented by a preposition such as *al* ('to'), *de* ('from'), or *je* (which seems to have only a case assigning quality but no specific meaning). Since Esperanto is an SVO language, or even a head-initial language, we assume case assignment is to the right, although the constituents can move around rather freely.

The syntax of a language without grammar

- (43) (Esperanto)
- a. La knabo donas kison al la knabino.
the boy gives a kiss to the girl
 - b. ?La knabo donas al la knabino kison.
the boy gives to the girl a kiss
 - c. ??Kison la knabo donas al la knabino.
a kiss the boy gives to the girl
'The boy kisses the girl.'

Interestingly, we find accusative NPs also as the complements of nouns (and adjectives), although these seem rather marked constructions. Nominatives are clearly not allowed at all.

- (44) (Esperanto)
- a. ?Via amo Johanon estas netolerebla.
Your love John-ACC is unbearable
 - b. *Petro amo Johanon estas malpermesata.
Peter love John is prohibited

There seems to be no independent principle of UG prohibiting this form of accusative case assignment, especially because in Esperanto, nouns and verbs are very closely related morphologically.

Furthermore, if we assume that nominatives can only be assigned by AGR, the subject-object dichotomy in (44) can also be explained. What we have to explain now is why there is no similar accusative case assignment by nouns and adjectives in English -and the same applies for Dutch, German, French, etc. as well.

Chomsky (cited in Lasnik and Uriagereka (1988)) has argued that in English nominal elements can indeed assign case to their complements as well. It only is not an accusative, but some sort of genitive, realized with the semantically empty *of*. Under this assumption, English shows the same paradigm as Esperanto.

- (45) (English)
- a. Your love of John is unbearable.
 - b. *Peter love of John is prohibited.

Apart from indicating the complement of a head, the accusative form seems to be able to act also as some sort of default case, for adjunct NPs.

- (46) (Esperanto)
- a. Mi laboris la tutan tagon.
I worked the whole day.
'I worked during the whole day.'
 - b. Li kuris la duonan vojon
He ran the half way
'He ran for half of the way.'

Esperanto is very similar to German (and Dutch) in this respect. (46) can be translated literally into German.

The syntax of a language without grammar

- (47) (German)
a. Ich habe den ganzen Tag gearbeitet.
b. Er ist den halben Weg gerannt.

We will not discuss this form of case assignment here, but immediately turn to a more interesting parallel between these two languages.

Oblique case in Esperanto is, at least phonologically, equal to nominative case. All prepositions assign this case uniformly:

- (48) (Esperanto)
Li parolas [pri la afero] [en la domo] [kun sia amiko] [dum la morgaùo],
He speaks about the question in the house with his friend during the morning.

Some locational prepositions, however, can also assign an accusative case to their complement. If they do this, they get a directional meaning. Prepositions with only a directional meaning and no locational one, most of the time do not assign this form of case:

- (49) (Esperanto)
a. Li promenas en la strato.
'He walks in the street.'
b. Li promenas en la straton.
'He walks into the street.'
c. Li promenas al la strato.
He walks to the street.
d. *Li promenas al la straton.

Again, we find more or less the same type of construction in German; only the 'default' case for the German locational preposition is the dative, not the nominative:

- (50) (German)
a. Er geht in der Strasse.
'He walks in the street.'
b. Er geht in die Strasse.
'He walks into the street.'

As far as we know, it has always been assumed in the literature that there are two prepositions in German: one with a directional meaning, assigning an accusative case and one with a locational meaning, assigning a dative case.

The same would have to apply for other German prepositions, such as *auf*, *hinter*, *neben*, *über*, etc. This seems in itself a rather undesirable situation, because now we miss the generalisation that dative = locational and accusative = directional.

Besides this, we also miss some important cross-linguistic generalisations, because German and Esperanto are not the only languages in which a directionally used locational preposition shows deviant behaviour:

The syntax of a language without grammar

-in English the suffix/preposition *-to* is attached to the preposition in many cases: *into*, *onto*.

-in Dutch, the preposition (sometimes) suddenly becomes a postposition: *in het huis* = in the house, *het huis in* = 'into the house'

-in Hungarian, the postposition gets some morphological form, analogous to that of a locative or lative case on nouns, respectively. (cf. Marác (1989); we will discuss these facts in more detail later on).

-in Slavic languages like Russian (cf. Neidle (1988)), Polish (cf. Brooks (1975)) and the variety of Kashubian that is spoken by an allochthone Polish community in the US (Perkowski (1969)), the preposition assigns the accusative meaning if it has a directional meaning and the instrumental case if it has a locational meaning. Sometimes, the preposition can also get a different morphological form, depending on its meaning (for instance, Russian *po* , 'to under' vs. *podóbio*, 'under').

Both in Hungarian and in the Slavic languages³⁴, we find even a third variant besides the locational and the directional forms, namely an 'ablative'. In Hungarian, this kind of preposition gets a distinctive morphological form. In Russian and Polish, it 'governs' the genitive case, although (as we see in (52)a) we also find some slight morphological distinction for some Polish prepositions.

- (51) a. (Russian)
 (Ona vošla) v komnatu
 (she entered) in room-ACC
 'She entered into the room.'
 b. (Ona vyšla) iz komnaty
 (she went) out room-GEN
 'She went out of the room.'
- (52) a. (Polish)
 (Nie widać go było) zza drzew
 (He wasn't seen) behind trees-GEN
 'He wasn't seen from behind the trees'
 b. (schować się) za dom
 (to hide oneself) behind house-ACC
 'to hide oneself behind the house'
 c. za górą (leży miasto)
 behind mountain (there is a city)
 'There is a city behind the house.'
- (53) a. (American Kashubian)
 Piotr postawił stół między okno a drzwi
 Peter placed table between window-ACC and door-ACC
 'Peter placed the table between the window and the door'

The syntax of a language without grammar

case mentioned above: apparently case can be realised as some sort of preposition-like element.

This whole idea might remind us of an old analysis of Fillmore(196-8). Van Riemsdijk (1978) argued against this proposal with the following two syntactic arguments (we omit details which we feel are irrelevant for the current discussion):

First, there is no cumulative case-marking found in any language of the world. That means that we never find an NP which has both a dative and a genitive case marking. We can't see this in Esperanto of course, because there are only two different markings in this language, one of which is a zero mark (and it isn't very easy to see the difference between e.g. *la domo-0-n* and *la domo-n*). But since we find no cumulative case marking in any other languages (although we do find a complex form of case marking in some of these as we shall see later on), we have to suppose there is no such thing in Esperanto either.

Secondly, there is no equivalent to case agreement or case attraction with prepositions. Prepositions can never be distributed inside the NP onto the determiner, the adjective, the noun, etc.³⁵

The second objection wouldn't have to be a problem for us, if we could state, that ACC is simply an empty preposition, assigning accusative case. But then a lot of new problems arise. First of all, Van Riemsdijk's (1978) first objection becomes more urgent, because now two prepositions would assign case to the same category. If we could manage to gather evidence in some way or another that *en* in (56) doesn't assign case, a second problem arises, because, in other languages it is, in the case of a P with a [_{PP} P NP] complement, always the innermost P that case-governs the NP.

For instance in the German example in (57) (cf. Van Riemsdijk(19-90)), *bis* is a preposition that governs the accusative exclusively, whereas *unter* can govern either the dative or the accusative. For this reason, the NP can only have received its case from *unter*.

- (57) (German)
[*bis*_P [*unter*_P *dem Tisch*]_{PP}]_{PP}
until under the table (dative)

Thirdly, it would be very strange if there were only one preposition governing the accusative case; if at the same time this preposition could violate the normal minimality constraint for prepositional case governing; and if this exceptional preposition would at the same time also be phonologically empty. That seems to much exceptional behaviour for just one category.

Fourthly, we would still have no explanation for the related English, Hungarian and Dutch facts. For instance in English, we would want to assume an underlying structure like (58), for the same semantic reasons that have lead us to a structure like (55) in Esperanto, and possibly German.

- (58) (English)
[*to*_P [*in*_P *the house*]]

The syntax of a language without grammar

- b. Stoĵ staĵ miĉdy oknem a drzwiami
 table stands between window-INSTR and door-
 INSTR
 'The table stands between the window and the
 door.'

A third and last objection against the standard analysis, arises within the study of Esperanto itself. In this language, the -n ending can in itself function as a directional marker, not only on adjunct NPs, but even on categories with the adverbial ending -e:

- (54) a. (Esperanto)
 Ni iru Berlinon!
 We go Berlin-to
 'Let's go to Berlin!'
 b. Ni iru Berlinen!
 'Let's go to Berlin!'
 c. Li falis antaŭen.
 He fell forward.

For this reason, we would like to be able to say that the accusative assigns the meaning 'direction' almost independently from the preposition. This would mean, that we assigned the structure in (55) to the PP in (49)b. More or less the same structure would be given to the English counterpart, but with the preposition *to* in stead of the accusative marker. This could be explained by reference to the lack of clear morphological case marking in English.

- (55) [_{PP} enP [_{PP?} ACC [_{NP} la strato]]]

Unfortunately, this structure poses quite some problems. First of all, there seems to be something wrong semantically, because *en la straton* is not equal to [*en [al la strato]*], but rather to [*al [en la strato]*]. The same applies for English as well: *into the street* is not the same as [*in [to the street]*], but rather to [*to [in the street]*].

So semantically we seem better off with a structure like (56)

- (56) [_{PP?} ACC [_{PP} enP [_{NP} la strato]]]

Also this structure is not without problems, however, as we will see in the next section, where we will try to solve this syntactic bracketing paradox.

3.2 Analysis

First of all, we have to ask ourselves what is the categorial status of the element we have called ACC in (56). Clearly, the structure as a whole behaves exactly like a PP; there seem no differences in the syntactic distributions of locational and directional PPs (cf. also (49)a and (49)b). This would mean that ACC -the case assigner- is a preposition. This doesn't seem a very unreasonable assumption, given Chomsky's proposal about *of*-insertion as a way of assigning

The syntax of a language without grammar

But we have no way to invert the prepositions in order to acquire the surface word order.

A final argument against this hypothesis is, that the structure of the adverbial constituents in (54) (except for (54)a) would also still be incomprehensible. For what element would receive the case in something like (59)?

(Esperanto)
 (59) [ACC_P [kie]_{PP}]

Thus, we have to find a different structure for (49)b.

In order to accomplish this, we propose to adopt a (somewhat revised) theory of Van Riemsdijk (1990). Van Riemsdijk (1990) proposed an extension of the theory of functional heads in order to account for some phenomena concerning circumpositions, mainly in German. If we assume that there is a 'verbal' functional head (namely I) and that there is a nominal functional head (namely D), we have little reason to assume that a prepositional functional head would be impossible in principle. Every p would have a PP as its complement³⁶, just like an I has a VP as its complement and a D has an NP as its complement. This means that the typical prepositional phrase would look like (60), at least in a head-initial language.

(60) [_{PP} p [_{PP} P NP]]

Given this analysis, Van Riemsdijk proposed to assign a structure like (61) for German circumpositional phrases.

(German)
 (61) [[auf_P den Berg]_{PP} hinauf_P]_{PP}
 on the mountain onto
 'onto the mountain'

For a more thorough argumentation in favour of this analysis, we refer to Van Riemsdijk (1990). Here, we will simply assume that this approach is (basically) correct. Notice that it is plausibly assumed that the functional preposition is on the right of the PP in German, since this language is (for the larger part) a head-final language.

Another interesting aspect of the pP in (56), from the case assignment phenomena illustrated in it. Although Van Riemsdijk (1990) argues extensively that it is the innermost preposition that assigns the case (as we have already seen in (57)), at the same time there has to be also some influence from the prepositional element in the p position. A construction like (62)a is ungrammatical in German, although the proposition *auf* could in principle assign the dative case, as has been demonstrated in (62)b.

(German)
 (62) a. * [[auf_P dem Berg]_{PP} hinauf_P]_{PP}
 on the mountain-DAT onto
 b. [[auf dem Berg]_e]
 on the mountain

The syntax of a language without grammar

Van Riemsdijk (1990) explains this by stating that the relation between P and p does perhaps involve the same sort of agreement as the relation between D and N (where in many languages we find agreement in number, gender and case). We should notice, however, that this analysis is not entirely without problems, since the case assigning property seems to be of a different nature than the case receiving property.

Another question is, why *hinauf* in (61) is in the p position. There are two logically possible answers: it could be base-generated there, or otherwise, it could be moved to that position. We assume, extending the theory of Van Riemsdijk (1990), that this is a case of movement, not unlike V-to-I or N-to-D movement. Yet it is not very clear why only *hinauf* would be subject of this type of lexical-to-functional head movement, i.e. why we don't find a structure like (63) in German?

(German)
 (63) * [[_{t_i} den Berg]_{pp} [auf hinauf]_{pt}]_{pP}
 the mountain on onto

This phenomenon can be explained if we assume that only substitution, but no adjunction, is possible for head-to-head movement, as has been argued by Van Riemsdijk (1988). Consequently only one position would be available in p position.

Given this extension, we will, again, assume that Van Riemsdijk's (1990) approach is basically correct. Furthermore, Van Riemsdijk suggests that we could perhaps treat the Dutch postpositions in the same way. Thus, in Dutch we have structures like (64). (Dutch, like German, is a (perhaps non-uniformly) head-final language.)

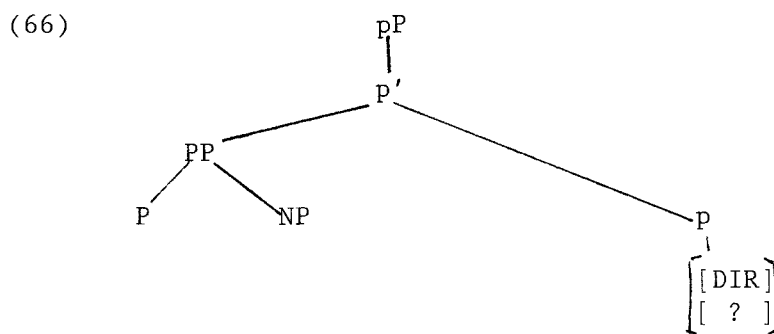
(Dutch)
 (64) [[[_{t_i} de berg]_{pp} op_t]_{pP}
 the mountain on
 'onto the mountain'

As we have seen before, this movement intuitively seems to be triggered by the directional meaning of the pP. The same applies for German; in structures like (63) we find very often some element with the elements *hin-* or *her-*, implying directionality. Yet in German, we also find structures like (65) (Van Riemsdijk (1990)):

(German)
 (65) [[auf dem Berg]_{pp} oben]_{pP}
 on the mountain up
 'up on the mountain'

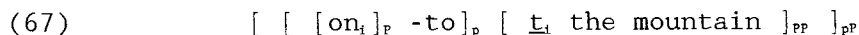
But in these cases, we would like to suggest that *oben*, *unten*, etc. have an explicit negative setting for the feature [+DIR]. (Prepositions like *auf*, *in*, *neben*, etc. would have no explicit marking for this feature.) The feature would reside within the functional prepositional element. We would thus get a structure like the one in (71) for Dutch and German:

The syntax of a language without grammar



The ?-mark in(66) indicates that there might be more features in p, like we also have both AGR and INFL in I. We will see later on, that perhaps in Hungarian AGR is also (sometimes) in p. English would also have the structure in (66), although the order of the p' and the PP would be inverted since English is a uniform head-initial language.

We have seen that P moves to p in Dutch (and German) iff the DIR feature has a positive value (related to the fact that V2 (or V-to-I) only applies in finite clauses, i.e. if the feature [\pm finite] has a positive value). The same might apply to English of course. But now, we can assume that the [\pm DIR] feature is realized in this language as an 'inflectional' -to on the preposition, i.e. the English directional pP has the following structure:



Let us now consider Hungarian. According to Marác (1989), there are two different types of Ps in this language; what he calls *naked* and *dressed* Ps. *Dressed* Ps may be inflected for AGR (they allow for a form of pro-drop, when they are) and assign structural case to their NP complement; that is the nominative case in Hungarian.

Naked Ps, on the other hand, are more like the prepositions that we know in German, Dutch, English and other (Indo-European?) languages. They show no AGR and they assign thematic case to their complements.

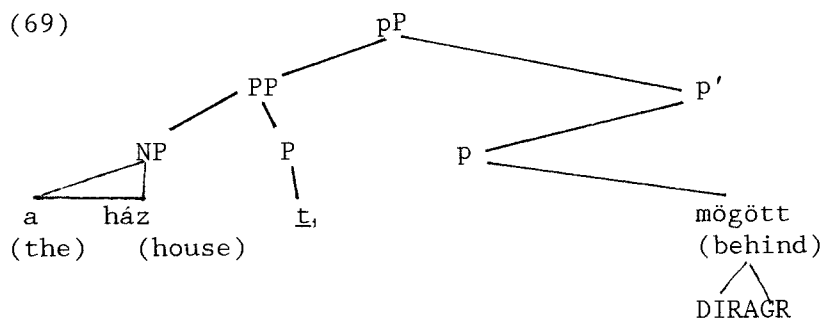
As we have seen in the preceding section, there is another interesting characteristic of Hungarian postpositions (Hungarian is a head-final language), namely a large subset of them can be divided into three parts, each part showing its own typical morphology: -á/-é has a directional (Marác (1989) uses the term *latíve*) meaning, -tt has a locational meaning and -l has an ablative meaning. For instance we find the following six prepositions:

(68)

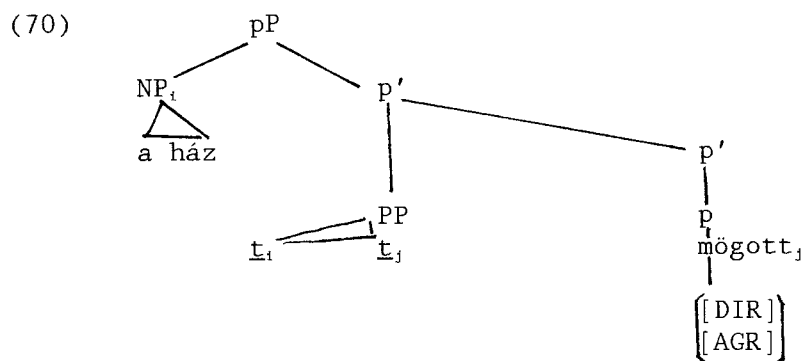
	(Hungarian)	
a.	alá	aláttalól
	(to) under under	(from) under
b.	elé	előttelől
	(to) before before	(from) before

Although, Marác (1989) doesn't take explicit note of this, it is quite remarkable to see that we find only such a tripartition within the class of *dressed* Ps.³⁷ Now, if we assume that Hungarian has three different possible values for the feature DIR, we can propose the following structure for the Hungarian *dressed* pP:

The syntax of a language without grammar



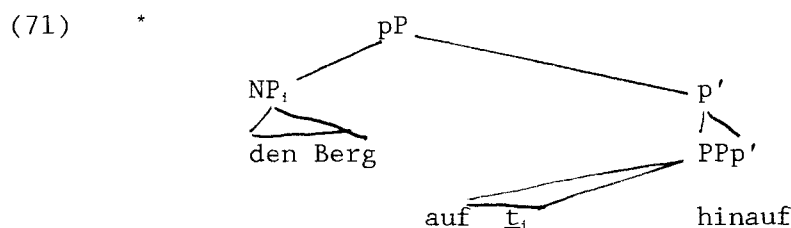
We have assumed here that AGR is also in p in Hungarian, and that the lexical dressed preposition has always to move to the functional p position. Perhaps this is so, because it has to move to AGR. But now we have to explain the fact that *dressed* Ps like mögött only can assign structural case and why it allows for pro-drop. Let us assume that dressed Ps can't assign case for themselves (this could also be related to the fact that they have to move); the case assigner is the functional preposition. This functional preposition assigns nominative case, like also I assigns nominative case. This case is normally assigned to the SPEC position³⁸, so the NP has to move there. In that case we have a perfect explanation for both the structural case and the pro-drop, because there is no structural difference between the relation p-NP in (70), and the relation I-[NP, IP] in an IP.



On the other hand, Hungarian *naked* Ps would probably remain in situ within the PP; or at least they would be able to assign case for themselves.

If we now turn our attention back to German, we find that also in this language structural (accusative) case is assigned with directional prepositions, but lexical (dative) case is assigned with locational prepositions. In this case we cannot assume however, that the NP has to be moved, because we would end up with the wrong structure (cf.(71)) and because, although the accusative is the *default case* in German it cannot be assigned to a SPEC position.

The syntax of a language without grammar



The term *default case* can be the source of some misunderstanding here. We could also say that the dative case is the default within the pP in a certain way; if nothing peculiar happens, the lexical preposition assigns this case to its complement.

What we mean here, is that the accusative case apparently can be assigned both in German and in Esperanto to an NP by a case assigner. This remains true even if this case assigner doesn't assign a theta role to that NP (cf. (47)). This applies both for adjunct NPs and for ECM constructions. Accusative non-thematic NPs can also occur within NPs (which by *default* assign genitive case to argument NPs) and APs. The following examples are taken from Haider (1985):

- (German)
- (72) a. die Ereignisse letzten Sommer
the events last summer-acc
'the events during the last summer'
- b. der Flirt vorigen Dienstag
the flirtation last tuesday
- c. die diesen Sommer sehr günstige Witterung
the this summer very favourable weather
'the weather that is very favourable this year'
- d. die diesen Sommer sehr teuren Urlaubsreisen
the this summer very expensive holiday trip
'the holiday trips that are very expensive this year'

Haider (1985) contends that, for some reason or other, the NP in a directional PP is in an A-bar position and that there is some independent device distributing the accusative case to NPs in A-bar positions. We will not adopt this proposal here, firstly because we find it not very clear to see why an NP in a directional PP would be in an A-bar position whereas an NP in a locational PP wouldn't; and secondly because it is not easy to explain the facts in English and Hungarian along these lines.

We will assume that, for some reason or another, although in German like in Hungarian, the *dressed* P cannot assign a case of its own, the former language applies a different strategy. In this language, there would be some sort of Exceptional Case Marking by the functional head -perhaps the NP would have to move to some position [NP, PP] (which would have to be at the left of the P'), perhaps it won't, because what we have called a PP, is called P' in Van Riemsdijk's (1990) original analysis, so there is no barrier to (case) government in this structure anyway, although it would have to be prohibited in Hungarian of course.

The syntax of a language without grammar

Because of the lack of phonological case marking in Dutch and English, we cannot easily see which of the strategies is applied in these languages.

Yet we find that Esperanto is exactly like German, except for the headedness parameter. Our final analysis for the pP in (49)b is given in (73)

(73) [[en_i]_p [t_i la straton]_{PP}]_{PP}

Like in German, the accusative is the default case in Esperanto. Besides that, the zero ending is shared by the nominative case and the oblique case in this language.

Although we will not give an explicit analysis of the Slavic facts here, we assume that in essence the same strategy is used in these languages as in German and Esperanto. The only difference is that in these languages, like in Hungarian, there is a tripartition concerning their [DIR]-feature instead of a dipartition.

The last phenomenon we now have to explain, is the occurrence of the -n marking on adverbial elements in (54) above. First of all, we have to make clear that we can analyze these elements as PPs (or pPs). Dasgupta (1989) has proposed that the Esperanto adverbial ending -e acts as some sort of functional preposition. He proposes a structure like the one in (74) for the adverb *varme*, 'warmly'.

(74) (Esperanto)
 [PP [A varm-] -e]
 warm ly

This may seem very odd to people who are not acquainted with Esperanto morphology. In Esperanto one can change the grammatical function of a word by changing its ending. For instance: *parol-i* means to speak, *parol-o* means a speech, *parol-a* oral, and *parol-e* orally. One is, in principle, totally free in this changing of functions: from the word *student-o* one could derive *studenta*, *studente* and *studenti*, even if the last word wouldn't have a very clear meaning.

Yet, at the same time, it is assumed since the work of De Saussure (1910, 1914) that each stem also has its own inherent functional quality. An example that is often quoted as an illustration for this statement is the pair *martel-* (hammer) / *seg-* (saw). The first stem is inherently a noun and the second a verb. Therefore *martel-o* means: hammer, but *seg-o* doesn't mean a saw, but sawing. The word for hammering is *martelado*, and the word for a saw is *segilo*, where *-il-* is a affix that marks an instrument, and *-ad-* a affix that marks an action.

Now Dasgupta (1989) uses this morphological theory in order to resolve the problem of how adverbs could fit into the tetramorous categorial system (N, V, A, P) that is generally assumed in linguistics. While there are stems which belong inherently to any of these four categories, we do not expect to find inherently adverbial stems³⁹. For this reason, Dasgupta (1989) draws a distinction between the categorial status of a word and the functional status of it -

The syntax of a language without grammar

inspired by the N vs. D and V vs. I dichotomy. According to him, an adverbial adjunct has the categorial status of an AP (that is: its stem -or the head of its lexical maximal projection- has inherently an adjectival category) while it has the functional status of a PP, i.e. it can occur in exactly the same positions as a PP. This means that the adverbial ending *-e* has to be analysed as a functional prepositional element.

If we try to translate this analysis to our approach of functional prepositions, we may assume that a functional *p*, apart from taking a PP as its complement can also take an AP as its complement. The head of the AP would move to the *p* position, where the functional element would realise itself morphologically as the ending *-e*. The same would apply, *mutatis mutandis*, for English. We give an example structure for both languages in (75)

- (75) a. (Esperanto)
 [[[*varm(a)*_i]_A -e]_p [*t*_i]_{AP}]_{PP}
 (English)
 b. [[[*comfortable*_i]_A -ly]_p [*t*_i]_{AP}]_{PP}

In Esperanto, the positive value for the DIR feature could the in turn realise itself as the ending *-n*, not unlike the realisation of directionality features on Hungarian 'adverbs' (cf. Marácz (1989))⁴⁰

- (76) (Hungarian)
 bentre bentről
 to inside from inside

In Esperanto, we find the endings *-e* and *-n* also if the moved element is an intransitive lexical preposition, e.g. *antaŭ* in (77) (remark that *-ward* in the English gloss is possibly a (relic of a) comparable element, given words like *homeward*, *landward*, *inward*, etc.).

- (77) a. (Esperanto)
 [[[[*antaŭ*_i]_P -e] -n]_p [*t*_i]_{PP}]_{PP}
 before -ADV -DIR
 'forward'
 (English)
 b. [[[*for*_i]_P -ward]_p [*t*_i]_{PP}]_{PP}

We find some evidence for this analysis also in Russian where we find inflectional forms for adverbs like *dóma* 'at home', *domój*, 'homeward'. (cf. Stankiewicz (1986)).

3.3 Conclusion

In this chapter, we have studied the behaviour of Esperanto with respect to case theory. In particular, we have tried to give an account of the 'accusative' case with directional prepositions, by extending a recent proposal concerning functional prepositions by Van Riemsdijk (1990). We have seen that there is reason to assume that a locational P can move to a functional position in order to get a directional

The syntax of a language without grammar

meaning. After it has moved to this position, it can only assign structural case.

We have also seen that our extended theory can explain some related phenomena in English, Dutch, German and Hungarian as well. One could ask what the relevance is of this chapter within this essay as a whole. Unlike the preceding chapter, this chapter has been devoted for the larger part to the development of a new theory. Given that this theory is developed partly on the basis of the Esperanto facts it accounts for, we could ask ourselves whether this theory proves that Esperanto is a natural language not unlike any other. It is very important to notice, however, that we didn't really introduce any new formal devices in this chapter; in particular no concepts were introduced that weren't necessary independent of Esperanto, on the basis of English, Dutch, German and Hungarian facts alone.

It should be noticed that the theory as we have presented it here is not yet complete. Although it has been explained why sometimes directional prepositional elements move, it isn't very clear why in other cases prepositions do not move. For example, why do we not find sentences like (78)b in Dutch?

- (Dutch)
- (78) a. De jongen zit naast het huis.
the boy sits next the house
'The boy sits next to the house.'
- b. *De jongen loopt het huis naast.
the boy walks the house next

The same question arises for almost all of the other languages which we studied in this chapter.

A related question that is not answered in the current essay, is why we have no evidence that a preposition with an inherent directional meaning (like *to* in English) moves to the functional position. *nach* in German assigns the dative case, *naar* in Dutch stays in front of the NP,

**toto* is a non-word in English, *al* cannot assign accusative in Esperanto, etc. This is certainly a remarkable fact, but we do not yet have an explanation for it.

Finally, our theory also doesn't explain why there is a class of prepositions that can only assign an accusative case in German. Many of these prepositions (for instance, *für* 'for', *ohne* 'without') have no directional meaning at all. The same question applies for the Slavic languages and for Hungarian, because not all dressed Ps have a directional meaning in the latter language. We will leave these questions open for future research.

The syntax of a language without grammar

4 Expletive Pro-drop

4.0 Introduction

In this chapter we will discuss the behaviour of Esperanto with respect to the so-called Null Subject Parameter. In order to attain this objective, we will first try to set forth a theory that has a large enough empirical coverage to explain the facts in languages as diverse as English, Italian, Russian, Dutch, German and Chinese. We will find that, although Esperanto doesn't pattern exactly with any of those natural languages on the surface, it very probably shows just another possible setting for the parameters at stake.

Because a relatively high number of languages is taken into consideration here, examples are quoted from authors with very differing opinions and theories concerning the pro-drop parameter. To discuss all those proposals to a worthy extent, in itself an interesting enterprise, would have meant a considerable lengthening of this essay. For this reason, we have chosen to restrict ourselves and to review only those approaches which we found we could use best in our own exposition. Alternative theories the interested reader can find of course in the references to this chapter.

This chapter is structured as follows. In the first section we will describe the implicit rule of Esperanto with respect to the Null Subject Parameter. In section 4.2 we then develop a theory of pro-drop which is strong enough to cover the languages mentioned above. In section 4.3 we will also consider an alternative analysis for the facts presented here while studying the seemingly odd behaviour of German and Dutch. Section 4.4 is entirely devoted to the analysis of the Esperanto facts, and we end this chapter with a conclusion in section 4.5.

4.1 An odd case of pro drop

It is a well-known and widely studied fact since Perlmutter (1971) that there is a contrast between languages (like Italian) that allow for phonetically null subjects in tensed clauses, and languages (like French) that don't. We give the relevant paradigm for French and Italian in (79) and (80) respectively.

- (79) (French)
- a. Mon père fume.
My father smokes.
 - b. Il fume.
He smokes.
 - c. *Fume.
*Smokes.

The syntax of a language without grammar

- (80) a. (Italian)
 Il mio padre fuma.
 My father smokes.
 b. Egli fuma.
 He smokes.
 c. Fuma.
 Smokes
 'He/she Smokes'

Because of the Projection Principle, we have to assume that in (80)c there is an invisible constituent which fills up the thematic position of the subject; we will call this constituent *pro* or *the null subject* here, in correspondence with tradition. Languages like Italian, which allow for null subjects, we will call *pro-drop*⁴¹ *languages* or *Null Subject Languages* (NSLs).

These differences between Italian and French are generally considered as stated in a parameter of UG, the Null Subject Parameter (Rizzi (1982)). Because a lot of *pro-drop* languages (besides Italian, we can mention Spanish, Portuguese, Polish and Serbo-Croatian here) show a very 'rich' verbal morphology -compare for example the Spanish paradigm in (81) with its English counterpart in (82) - this property of a so-called rich Agreement or rich Inflection has until very recently been seen as an essential licenser for the setting of the NSP.

- | | | |
|------|--|-------------------------------------|
| (81) | (Spanish)
habl-o
habl-as
habl-á
<i>speak</i> | habl-amos
habl-áis
habl-an |
| (82) | (English)
I speak
you speak
he speaks | we speak
you speak
they speak |

Intuitively, the Spanish declination system of the verb *hablar* seems to be sufficient in itself to identify the relevant subject, because all the forms of the paradigm are different. Other well-known *pro-drop* languages like Italian, Portuguese, Polish and Serbo-Croatian have an equally 'rich' paradigm. On the other hand, the English pendant of *hablar*, *speak*, has one phonological form that is on its own ambiguous between (at least) five different readings.

If we now consider the Esperanto verbal system, the first thing we observe is that it is extremely 'poor'; it doesn't show any markings of number or person agreement⁴²:

- | | | |
|------|---|---|
| (83) | (Esperanto)
mi parolas
vi parolas
li parolas
<i>speak</i> | ni parolas
vi parolas
ili parolas |
|------|---|---|

The syntax of a language without grammar

Given our first intuitive idea of the NSP, we would now predict that the setting of Esperanto for the Null Subject Parameter has to be negative. On first sight, this prediction seems to be borne out. Esperanto behaves exactly like French with respect to the given paradigm:

- (84) a. (Esperanto)
 La patro fumas.
 The father smokes
 'My father smokes.'
 b. Li fumas.
 'He smokes.'
 c. *Fumas.
 smokes

There are, however, a few cases to be found where a subject might be missing in a tensed clause. We give some of the relevant examples in (85) - (88). It should be noticed that the Italian translations of these sentences are like the Esperanto originals, whereas French patterns with the English glosses.⁴³ I.e.: in French there is absolutely no possibility whatsoever of dropping the subject pronoun; whereas in Italian all subjects can be freely dropped.

- (85) a. (Esperanto)
 Pluvas
 rains
 'it rains.'
 b. Estas varme en la parko.
 is warm in the park
 'It is warm in the park.'
 c. Mal-lum -iĝas.
 un -light-becomes
 'It is getting dark.'
- (86) a. Ne indas eĉ paroli pri tio.
 not worth-is even speak about that
 'It isn't worth even to speak about that.'
 b. Estas bone ke Johano iris hejmen
 Is good that John went home-to
 'It is good that John went home.'
 c. Ŝajnas nur al vi ke ni perdis la vojon.
 seems only to you that we lost the way.
 'It seems only to you that we lost our way.'
- (87) a. Temas pri via honoro.
 Theme-is about your honour
 'Your honour is at stake.'
 b. Dum plaĉas al vi, mi atendas
 While pleases to you I wait-will
 'I will wait your pleasure.'
- (88) a. ('Vi) iru hejmen!
 >('You) go home!'

The syntax of a language without grammar

- b. ('Vi) ne ŝiru la rozojn!
you not tear the roses
'Don't (you) tear up the roses!'

Leaving aside the cases in (88) for a moment, which are exceptions (to our knowledge hardly noticed and never studied in the literature before) even in *hard core* non-NSLs like English and French, we must ask ourselves what the relevant difference between (84) and (85) - (87) is.

We can immediately see that there clearly is a distinction concerning Theta Theory. In (7), some specimens of what is traditionally called a weather verb, are used. We assume here that *pluvas* (like the English *rain* (cf. Chomsky (1981, 1982))) has an empty theta grid. This assumption, and the fact that also the counterparts to (86) and (87) are ungrammatical in English, has led Chomsky (1982) to the formulation of his Extended Projection Principle, paraphrased by us in a somewhat simplified way in (89)

- (89) *Extended Projection Principle*
The Projection Principle must hold and every S must have a NP in its structural subject-position.

The EPP can also be called into use in order to explain the obligatory *it* in the English counterparts to (86) and (87), where the real subject seems to be the extraposed tensed or untensed clause. Furthermore, Esperanto also shows some of the features that are standardly related to the NSP. For instance, on the surface it seems to violate the *that-trace filter*.

- (Esperanto)
(90) Kiu_i vi diris ke e_t venas?
Who you said that comes?
'Who did you say comes?'

We have now in principle two approaches to our disposal -apart from putting the Esperanto paradigm aside as a mere caprice of its author- to account for the contrast between English and Esperanto: we can reject the EPP and try to account for the English overt expletives in another principled (and parametrized) way, or we can accept the EPP and then try to formulate a satisfying theory of pro-drop which can explain why only expletives can be null in Esperanto. We shall basically adopt the second line of research here. An approach along the former lines will be briefly discussed in section 4.5.

Before we dig ourselves into purely theoretical considerations, however, we first have to make four more observations which shall have to be accounted for by any (ideal) theory of Esperanto syntax.

First, there seems to exist no expletive pronoun at all in Esperanto. Whereas the real *pro* example in (80)c has an overt counterpart in (80)b, this is not the case with e.g. sentence (85)a:⁴⁴

The syntax of a language without grammar

- (91) (Esperanto)
 *Ĝi pluvas.
 it rains

On the other hand, under specific circumstances sometimes a fully specified NP can act as the overt subject of clauses like those in (85) and (87):

- (92) (Esperanto)
 a. ?Benoj pluvas sur ilin.
 blessings rain on them-to
 'It rains blessings onto them.'
 b. ?Kugloj hajlis sur la malamikon.
 Bullets hailed on the enemy-to
 'Showers of bullets came down on the enemy.'
- (93) La libro temas pri politiko.
 'The book is about politics.'

The third observation we have to make is that there is a distinction between the form of the predicative adjective in sentences with and sentences without an overt subject:

- (94) (Esperanto)
 a. Estas varme en la parko (= (85)b)
 'It is warm in the park.'
 b. *Estas varma en la parko
 c. La suno estas varma.
 The sun is warm.
 d. *La suno estas varme
 e. La plaĝoj estas varmaj.
 The beaches are warm-plural

In (94)c *varm-* has taken the adjective ending *-a*; moreover it shows number agreement with the subject, as is illustrated in (94)e. In (94)a however, *varm-* has taken the ending *-e* which is normally reserved for adverbs.

Our last observation here is that, whatever the status of the empty category in (85)-(87) may be, it cannot be PRO because PRO in Esperanto (like in any other known language in the world, see Jaeggli and Safir (1989b)) is not allowed to be expletive. Thus, we find no empty expletive in untensed clauses.

- (95) (Esperanto)
 a. *Estas neklare kiam PRO pluvi.
 is unclear when to-rain
 b. *Li trovas PRO ne indi eĉ paroli pri
 tio.
 he finds not worth-be even speak about
 that
 c. *Mario ne scias kial PRO verŝajni ke Johano
 ne helpos nin.
 Mario not knows why probable-be that John
 not will-help us.

The syntax of a language without grammar

Concludingly, we can formulate our observations concerning the occurrence of the subject in Esperanto as follows.

- (96) Every sentence contains an overt subject in subject position, with the exception of sentences in which the main verb has no thematic role to distribute, or in which the role of the subject is performed by an extraposed clause. In clauses without an overt subject the predicate has the adverbial form.

It is clear that this rule is not very obvious or very simple in a mathematical sense of those words. Nevertheless, it is a rule which seems to be accepted without a grumble by Esperanto speakers both with non-NSLs and NSLs as their mother tongue; moreover it is a rule which happens to be acquired without any apparent problem by children who have learned Esperanto as one of their native languages. This makes us suspect that the irregularity might have to be connected to underlying principles of UG.

4.2 A parametric theory of null subjects

Perhaps because it is the parameter which is supported by the largest amount of crosslinguistic evidence, the Null Subject Parameter is one of the most widely studied phenomena in current principles-and-parameters-approaches to grammar. An interesting collection of articles (Jaeggli and Safir (1989a)) has recently given an extra impulse to this kind of research; most of the work presented in this and the following section 4.3) can be seen as an extension upon work of some of the authors represented in that volume.⁴⁵

The following subjects will be discussed in this section. We will first give some attention to the question whether there is any difference at all between *pro* and other empty categories, especially PRO. We then will treat the question what role AGR plays in the licensing of *pro* and finally we shall treat the theory of identification and so-called expletive *pro*-drop.

4.2.1 *pro* as an empty category

In the literature we regularly find the following scheme of empty categories:

(97)

	pronominal	anaphoric
NP-trace	-	+
<i>wh</i> -trace	-	-
PRO	+	+
<i>pro</i>	+	-

We will assume that this scheme is correct here. Especially, we will adopt the view that *pro* and PRO are distinct entities.

The syntax of a language without grammar

Although it has been argued by quite a number of linguists that there are no real differences between *pro* and PRO, Jaeggli and Safir (1989b) list a number of features that the two elements don't seem to share. Among the most important of them is of course the fact that PRO can't be expletive in any language, which we showed to be a quality also of Esperanto above. Yet there are still other reasons to distinguish the two categories:

**pro* (like an overt pronoun) yields a *weak crossover effect*, while PRO doesn't. Compare for example the sentences in (98)a, with an overt pronoun, (98)b, with *pro* and (98)c, with PRO.

- (98) a. (Spanish)
 ?*A quién_i acusó [la mujer_j que_j [e]_j bailó con él_i] [e]_i
 whom accused the woman that danced with him
 'Who did the woman who danced with him accuse?'
- b. (Spanish)
 ?*A quién_i acusó [la mujer_j con quien_j *pro*_i bailó [e]_j] [e]_i
 whom accused the woman with whom danced
 'Who did the woman with whom he danced accuse?'
- c. (English)
 Who_i did [_s [_{NP} PRO_i washing his_i car] upset e_i]?

*PRO can't act as a resumptive pronoun, whereas *pro* can:

- (99) a. (Spanish)
 *Ese es el tipo que no sabíamos si sería posible PRO nadar.
 that is the guy who not we-know whether it-was possible to-swim
- b. (Spanish)
 Ese es el tipo que María conoce a la mujer con quien *pro* se casó.
 that is the guy who Mary knows the woman with whom *pro* married.

We can summarize these findings by saying that *pro* can occur in the same positions as an overt pronominal, whereas PRO is in complementary distribution with overt pronominals. The logical consequence of this, is that PRO and *pro* are in complementary distribution.

4.2.2 AGR as a licenser of *pro*

It has been a standard assumption within probably every theory of grammar that the inflectional morphology of a language has to be rich enough in order to allow for null (thematic) subjects. Jaeggli and Safir (1989b) point out, however, that this notion *richness* is far from precise.

The systems of Spanish (shown in (81) above) and Italian as compared to that of French or English (given in (82)) suggest that

The syntax of a language without grammar

richness has something to do with the question whether a language has a different ending for every person/number combination or not.

Consider now the Irish paradigm which is given in (100). Irish has so-called *analytic* and *synthetic forms* (according to McCloskey and Hale (1984)). Although a typical paradigm of Irish seems to be rather poor, because most of the forms are analytic, this language is at least partly a pro-drop language; where the verb is synthetic (in this case the first person singular) a null subject is allowed.

(100)	(Irish)	
	cuir-im	cuir-eann
	cuir-eann	cuir-eann
	cuir-eann	cuir-eann
	put (simple present)	

But if one distinct form is enough to license pro-drop for that one form, we would expect that English could be a pro-drop language for its 3d person singular, too.

Henk van Riemsdijk (p.c.) notes that we find a vaguely similar phenomenon in Züritütsch and Rhaeto-Romance. Züritütsch has the following paradigm:

(101)	(Züritütsch)	
	ich rede	mir reded
	du redsch	ir reded
	er red(e)t	si reded
	<i>speak</i>	

This paradigm is certainly not richer than that of Standard German:

(102)	(Standard German)	
	ich rede	wir reden
	du redest	ihr redet
	er redet	sie reden
	<i>speak</i>	

At the same time, Züritütsch shows some restricted form of pro-drop, while Standard German doesn't: in the former language the 2nd person singular can be dropped in subject-inversion structures (e.g. questions):

(103)	a.	(Züritütsch)	
		redsch (du) ?	
		(Standard German)	
	b.	redest *(du) ?	
		Do you speak?	

Matters grow even more complex, if we also take restricted pro-drop languages like Chinese (Huang (1989)) or Japanese (Hasegawa (1985)) into consideration; these languages show no person or number agreement whatsoever, although Japanese (but not Chinese) has an inflectional system to express tense, mood and aspect.⁴⁶

The syntax of a language without grammar

- (104) (Japanese)
- | | |
|------------------------------|--------|
| yom-ru | yom-ru |
| yom-ru | yom-ru |
| yom-ru | yom-ru |
| <i>read</i> (simple present) | |

Although these sentences do not allow for the subject to be null under all circumstances (roughly said, they only allow for the subjects of embedded clauses to be null), it has been argued quite convincingly by Huang (1989) that at least the Chinese null element is pro.

On the other hand, Icelandic, which has an rather complex, *rich* paradigm is only a limited pro-drop language. The same applies for Russian.

- (105) (Icelandic)
- | | |
|-----------------------------|--------|
| segi | segjum |
| segir | segid |
| segir | segja |
| <i>say</i> (simple present) | |

It is clear that it would be very difficult to maintain a crystal clear notion of *richness* for AGR which could also explain the contrast between e.g Japanese and English, or Russian and Italian.

For this reason, Jaeggli and Safir (1989b) propose a different analysis of pro-drop, repeated here in (106)-(120)d. (We will postpone our exposition of their theory of identification to the next section.)

- (106) *The null subject parameter* (Safir and Jaeggli(1989b))
Null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms.
- (107) *Morphological uniformity* (Safir and Jaeggli (1989b))
An inflectional paradigm P in a language L is morphologically uniform iff P has either only underived inflectional forms or only derived inflectional forms (where an inflectional form is *underived* if it is equal to the stem and *derived* otherwise).
- (108) *Identification requirement*
A thematic empty pronoun must be identified.

The reader can check for himself that the paradigms of Japanese, Irish, German and Spanish are all morphologically uniform: they either consist of only the stem for every form, or otherwise of always the stem plus any affix. The paradigms of both French and English on the other hand show a mixture of forms consisting of a mere stem and forms consisting of a stem plus an affix. Moreover, it is clear that AGR could still play part of its role under the

The syntax of a language without grammar

disguise of the Identification requirement, although identification in Chinese and Japanese still cannot be explained by this principle.

4.2.3 Expletive pro drop

It was noted already by Rizzi (1982) that in NSLs like Italian, null subjects could occur also in infinitival and gerundic clauses; the only restriction seemed to be that such subjects had to be non-theta bearing or 'dummy' subjects:

- (109) a. (Italian)
Essendo piovuto per tutto il pomeriggio, non siamo usciti.
Having rained for the whole afternoon, we didn't go out.
'Because it had rained for the whole afternoon, we didn't go out.'
- b. Ritengo [esser nevicato anche sotto i mille metri]
I-believe to-have snowed even below 1000 metres
I believe that it has snowed even below 1000 metres.'
- (110) a. Essendo molto improbabile che Mario ci aiuti, dovremo cavarcela da soli.
Being very unlikely that Mario helps us, we must get-off by ourselves.
'While it is very unlikely, that Mario will help us, we must get off by ourselves.'
- b. Suppongo [esser molto improbabile che Mario ci aiuti]
I-consider to-be very unlikely that Mario helps us.
'I consider it to be very unlikely that Mario helps us.'
- (111) a. Avendo telefonato tuo fratello, io sono rimasto a casa.
Having telephoned your brother, I stayed at home
'After your brother telephoned, I stayed at home.'
- b. Ritengo [esser successo qualcosa di molto spiacevole]
I believe to have happened something very unpleasant.
'I believe something very unpleasant has happened.'

That the equivalents to (109)-(111) are not grammatical, led Rizzi (1982) to the conclusion that these facts are somehow related to the Null Subject Parameter.⁴⁷

Notice that, under the given theory of pro-drop, the strange thing about the Italian paradigm is not so much that the sentences in for

The syntax of a language without grammar

example (109)-(111) are grammatical, but that a sentence like (112) (where there is an overt thematic pronoun missing) is ungrammatical.

- (112) (Italian)
*Avendo telefonato, io sono rimasto a casa
'After he telephoned I stayed at home.'

Another observation Rizzi (1982) made, is that -as we have seen earlier- PRO can never be a null expletive, as can be seen in (113), so that we have to assume the empty category in the sentences (109)-(111) is *pro*.

- (113) (Italian)
a. *Gianni vorrebbe aver telefonato tuo fratello
Gianni would like to have telephoned your brother
(=that your brother telephoned)
b. *Non è chiaro [quando piovere]
It is unclear when to rain
c. *Mario non sa [perché esser probabile che Gianni non ci aiuti]
Mario doesn't know why to be likely that G. not help us
d. *Sembra [esser successe cose terribili]
Seems to have happened terrible things

We will now explain this feature of PRO with the following (rather *ad hoc*) condition of Safir (1985) (where the acronym *emex* stands for empty expletive.)

- (114) *The Emex Condition*
An expletive empty category must be governed.

Safir (1985) now claimed that German tensed clauses behaved -with some provisos- like Italian gerunds and tensed infinitives, in that it allowed for expletive, but not thematic, pronouns to be null, on the basis of i.a. the contrast in (115)

- (115) (German)
a. *Er sagte, dass e den Hund getötet hat.
He says that the dog killed has
b. Er sagte, dass e ihm scheint dass Hans den Hund getötet hat.
He says that him seems that Hans the dog killed has
'He says that it seems to him that Hans killed the dog.'

German like Italian (and English) doesn't allow for PRO to be expletive:

The syntax of a language without grammar

- (116) a. (German)
 Er_i sagte, PRO_i den Hund töten zu wollen
 He said the dog kill to want
 'He said he wanted to kill the dog'
- b. *Er sagte, PRO_{arb} ihm zu scheinen dass Hans den
 Hund getötet hat.
 He said him to seem that Hans the
 dog killed has.

We will discuss German (and the related Dutch facts) in more detail later on in this chapter. For now, it should be noted that German is not really the expletive pro-drop language without further preface; expletives with weather verbs or extraposed sentences, or expletives in main clauses can in general not be null:

- (117) a. (German)
 *Er sagte, dass e regnet.
 He said that rains
- b. *Er sagte, dass e ist schade dass du so spät
 kommst.
 He said that is pity that you so late
 come
- c. *(Es) scheint ihm, dass Hans den Hund getötet
 hat.
 It seems to-him that Hans the dog killed
 has

That German doesn't really pattern immediately in the right fashion doesn't need to have a devastating effect upon the theory, however, because there is at least one natural language which does. This language is Russian as has been noted by Franks (1990)⁴⁸.

As is shown in (118)a, Russian is, different from its cognates Polish and Serbo-Croatian, not a NSL.

- (118) a. (Russian)
 ja student/ ustal
 I student tired
 'I am a student/ tired'
- b. (Serbo-Croatian)
e student/ umoran **sam**
 student tired be(1SG)
- c. (Polish)
e **jestem** studentem/zmęczony
 be(1SG) student tired

This is unexpected under the uniformity hypothesis, since the Russian verbal paradigm is morphologically uniform; it even has a 'rich' paradigm for normal verbs, although there are, as can be seen in (118), no copula.

It is, on the other hand, clear that the contrast between (118)a and (118)b-(118)c can be explained under the old assumptions along

The syntax of a language without grammar

the same lines as the contrast between English and Italian. Where Polish and Serbo-Croatian have some device other than the pronominal system to express number and person agreement at their disposal, such a system is lacking altogether in Russian, at least in phrases with a copula.

But now we find that Russian is not a rigid non-null subject language. In some cases the subject pronoun is omitted; these cases seem to be more or less parallel with the Italian gerundic facts. It is, again, in the case of weather verbs and extraposed subjects (or subject sentences) that the pro-drop behaviour becomes apparent.

- (119) (Russian)
- a. Ivan/ on/ *e kupil gazetu
John/ he bought a newspaper
 - b. e temneet
'it is getting dark'
 - c. e ne stoit daže i govorit' ob ètam
'it is not worth it even to speak about that.'

We thus find that the Russian paradigm strongly resembles the Esperanto facts which are under study in this essay. Another example (from (Rizzi (1982))) is the Italian dialect of Padua, with the following paradigm.

- (120) (Italian -dialect of Padua)
- a. Piove
rains
'It rains.'
 - b. Vien Giorgio
comes Giorgio
'Giorgio comes.'
 - c. *Vien
comes
 - d. El vien
He comes.

What becomes clear from the Russian and Paduan (and possibly Dutch / German) paradigms and the Italian gerundic and tensed-infinitival facts, is that NSLs do not always allow for thematic subjects to be invisible, although they seem to allow for null expletive subjects unconditionally. We stated the condition on null thematics informally in (108) (cf. Jaeggli and Safir (1989b)).

- (108) A thematic null pronoun must be identified.

It should be noticed here that the Esperanto paradigm seems to be devised precisely to fit into this theory for now; it certainly has a *morphologically uniform* paradigm, all of its forms being derived in exactly the same way; thus it should be a NSL according to (106). But because a thematic null subject can not be identified in any way in an Esperanto sentence, the only *pro* we would expect to occur is the expletive one. This is exactly what happens.

The syntax of a language without grammar

Unfortunately, there are not very many natural languages which fit into the theory as neatly as Esperanto does, as we shall try to show now.

But first, we have to develop a theory of identification. An interesting variant of such a theory has been proposed by Jaeggli and Safir (1989b). Identification under this theory can be accomplished in two ways by different languages: by a rich agreement system on the verb (which is the strategy of Italian, Spanish, Portuguese, etc), or non-locally (which is the strategy of Chinese and Japanese). Identification by Agreement is defined as in (121); implicit in this definition is that AGR has to be 'able' to identify the subject, i.e. it has to be *rich* enough.

- (121) *Identification by Agreement*
AGR can identify an empty category as thematic *pro* iff the category containing AGR Case-governs the empty category.
(Jaeggli and Safir (1989b), p. 35)

Jaeggli and Safir (1989b) argue that the notion *case-govern* is needed here, because it is the only concept that can explain why German and Icelandic, which have a relatively rich verbal inflection, cannot identify thematic null subjects. Jaeggli and Safir argue that in these languages Tense is located in the head of CP, whereas AGR is located in the head of IP; this has to do with their being Verb Second languages. In Italian AGR and Tense really are in the same node (probably INFL). Moreover, Jaeggli and Safir (1989b) assume that one of the functions of Tense is to assign Case. Whence the formulation of (121).

Jaeggli and Safir (1989b) also explain the fact that in West Flemish thematic *pro* can only show up if COMP is cliticized and the fact that in French null objects can arise if a clitic is attached to the verb with the definition in (121).

- (122) a. (West-Flemish)
... dase *pro* komt.
that-3-F-SG comes
'that she comes.'
b. *... da *pro* komt.
that comes
- (123) (French)
Je le crois [pro ridicule]
I him-believe ridiculous
'I believe him to be ridiculous.'

In Chinese and Japanese, there is no overt AGR at all, so that thematic null subjects have to be identified in a different way. Jaeggli and Safir (1989b) and Borer (1989) assume that in these languages the AGR nodes are morphologically empty. Therefore, *pro* can not be identified locally by AGR; but AGR can inherit features from a

The syntax of a language without grammar

c-commanding NP, usually a subject, and then these features can identify *pro*.

Perhaps an example will clarify this. In Chinese we find sentences like (124)

- (124) (Chinese)
Zhangsan shuo [e lai le]
Zhangsan say come ASP
'Zhangsan_i said that he_i came.'

It has been argued (by Huang (1989)) that the empty category in (124) is governed and therefore cannot be PRO. We assume, then, that it has to be *pro*.⁴⁹ Moreover, it is an instance of thematic *pro*, because *lai* usually has an agent role to assign. Identification by AGR now cannot work here directly, because, as is argued earlier. But if we now suppose, with Borer (1989), that AGR is anaphoric, and furthermore that it is coindexed with *Zhangsan* in (124) we could identify the *pro* in the subordinate clause with this indexed clause. *pro* has then become coindexed with *Zhangsan* in an indirect way

The same approach can probably be extended to Japanese, although this is not a totally uncontroversial issue.

It can now be seen, however, that Russian and Esperanto facts are a problem under this approach.

In particular, why does Esperanto, which has the same kind of verbal paradigm as Japanese (tense inflection but no agreement inflection) not identify thematic null subjects in the same way? I.e. why are sentences like (124) not allowed in this language?

- (125) (Esperanto)
*Johano diris ke *pro* venas
John said that comes

We shall try to formulate an answer to this question at the end of the next section, after we have tried to solve the problems of this analysis with German and Dutch.

4.3 German and Dutch expletive *pro*-drop

Quite a different approach to null expletive phenomena, especially in Dutch, than the approach presented here, is demonstrated in Bennis (1983, 1986).

According to Bennis, the Extended Projection Principle has to be rejected, so that subjectless clauses containing a main verb without a theta grid can be analyzed without any empty category. It is clear that the Russian, Italian gerundic and Esperanto facts now fall out immediately; what remains to be explained, are for example the obligatory *there* and *it* in English and *er* and *het* much in evidence in Dutch. We will briefly outline the theory of Bennis here.

German and Dutch have for a long time posed some problems for the theory of e.g. the *that-trace filter*: although these languages clearly are not NSLs as Italian is, they still seem to allow for violations of this filter (and hence for the ECP, since the *that-*

The syntax of a language without grammar

trace filter is derived from that principle): a sentence like (127) is fairly acceptable for many native speakers of Dutch, although nobody will hold that (126)c is grammatical in that language.

- (Dutch)
- (126) a. Mijn vader rookt.
My father smokes
b. Hij rookt.
He smokes
c. *e rookt.
smokes
- (127) Wie denk je dat e morgen op bezoek komt?
Who think you that tomorrow on visit comes
'Who do you think will come and see us tomorrow?'

We have seen before, however, that Safir (1985) tried to explain the fact that German allows under certain circumstances for expletive null subjects (see the examples in (115)). Exactly the same argument goes for Dutch:

- (Dutch)
- (128) a. (*)Ik denk dat e gedanst gaat worden.⁵⁰
I think that danced goes be
'I think one is going to dance'
b. Gebleken is dat dit meisje geen kans onbenut
laat.
Appeared is that this girl no chance
unutilized lets.
'It has appeared that this girl makes the most
of all of her chances.'

As we have learned from (i.a.) Bennis (1983) there is an essential distinction between (128)a and (128)b, because the former sentence is considered to be ungrammatical by a considerable group of native speakers of Dutch, whereas the latter is considered to be grammatical in all dialects.

We differentiate here, following Bennis (1983), between Dutch A (where (128)a is grammatical) and Dutch B (where (128)a is ungrammatical) and assume that Dutch A and B are completely similar in all other respects.⁵¹

Bennis (1983) contends that also other non-NSLs like Icelandic or Old-English show a lack of that-trace effects. These languages probably also have the same form of limited pro drop.

- (Icelandic)
- (129) a. Hver sagdi þú, ad t vaeri kominn til
Reykjavíkur
Who say you that was come to
Reykjavík
'Who did you say has come to Reykjavík?'

The syntax of a language without grammar

(Old-English)

- b. Ac hwaet saegst du donne daet sie forcodrie
 donne si ungesceadwisnes
 But what say you then that is worse
 than is stupidity
 'But what then do you say is worse than
 stupidity?'

Bennis (1983) then proposes that there are in fact two parameters at work: besides the NSP (which he calls the PRO-drop parameter) he proposes a that-trace parameter and he gives the following scheme for it:

(130)	Null Subject that-trace	+ -	- +	- -
		Italian Spanish	Dutch A Icelandic	English French Dutch B

In order to explain the asymmetry in (130), Bennis (1983) proposes that the that-trace parameter be made dependent upon a negative setting of the NSP: a positive setting would somehow imply a negative setting for *that-trace*. We will not go into these matters any further, although we do not think Bennis (1983) gives very convincing linguistic arguments for this position.

It is clear that we still don't have an explanation for the subjectless clauses in Dutch. In later work Bennis (1986) has argued that the Dutch 'dummy subjects' *het* and *er* are no subjects at all. In order to maintain this position he had to abandon the EPP.

The problematic fact, up until now not mentioned in this paper, that Dutch weather verbs always seem to require for a subject, even in contexts comparable to (128)b, Bennis (1986) solves by stating that the subjects of such verbs are indeed thematic. Independent evidence for this can be found among other things in the fact that in Dutch sometimes fully specified NPs can act as the subject of a weather verb (cf. (92))

(Dutch)

- (131) In het park regent *(het).
 In the park rains it
 'It rains in the park.'

- (132) a. Het druppelt in het park.
 it drops in the park
 'It is spitting in the park.'
 b. De kraan druppelt.
 The tap drops.

It is now clear, however, that this approach, however elegantly it may do its work for the Dutch facts, is not very easy to be used crosslinguistically.

The syntax of a language without grammar

Especially, it is not easy to explain the Russian facts within Bennis' (1986) account. On first sight, these facts seem to plead rather in favour of abandoning the EPP, because sentences without a subject are allowed for in this language as long as the Projection Principle is obeyed. We could assume then that Russian is negative for the NSP and positive or negative (according to the facts) for the that-trace parameter.

But if we reject the EPP and at the same time suppose that the subject of weather verbs gets a thematic role, something very mysterious is going to happen: Russian and Paduan Italian will turn out to be very limited pro-drop languages, in which only *pro* in clauses with weather verbs can be null.

Besides, the facts of Standard Italian gerunds and tensed infinitives also get rather inexplicable, because also in this case we would have to assume that only the thematic subjects of *piovere* and the like are to be dropped in such constructions.

Another objection to an approach along the lines of Bennis (1983) is more theory-internal. For a parametric variation of that-trace effects is certainly unexpected, if one has reduced the that-trace filter to the ECP. As is known the ECP is supposed to be a principle of UG.

On the other hand, the Dutch and German facts mean a severe problem for any pro-drop analysis. Jaeggli and Safir (1989b) have it in a footnote, following Travis (1984)) that there might be a (crosslinguistic) implicational hierarchy as to which expletives can be null. They also suppose that the theory of identification does not play a role in determining this hierarchy. There seems to be little reason to expect that such a theory could serve for an explanation for an hierarchical theory about expletives indeed.

But now it is very difficult to see what accounts for the hierarchy, if identification doesn't. Expletives under the uniformity+identification hypothesis seem to be just 'the subjects that can be null because they do not have to be identified'. Because German and Russian are certainly and Dutch is probably morphologically uniform, all three of these languages are NSLs; because all these three languages also do not have sufficient means to identify their thematic subjects, they can only allow for expletive pronouns to be null. We don't have any reason to suspect, however, that German and Dutch would only allow for part of the expletive hierarchy (that is, for every expletive but the ones in construction with a weather verb) to be *pro*.

This is just one problem a pro-drop analysis has to face. Still another problem is that German and Dutch, as opposed to Russian, also seem to restrict the configurations in which an expletive may be null: the clause in which *pro* arises has to be (at least in Dutch A) subordinate, or else (both in Dutch A and Dutch B) the first position of the sentence (which is normally in unemphatic sentences distributed to the subject) has to be filled. We believe the German paradigm to be exactly parallel, although we are not sure whether one could not also make a difference between some German A and German B.

The syntax of a language without grammar

- (133) a. (Dutch A and B)
In de huiskamer werd gedanst.⁵²
In the livingroom was danced
'One danced in the living-room.'
- b. *Werd in de huiskamer gedanst.
was in the living-room danced
- c. (Dutch A)
Ik denk dat e/ er gedanst werd.
I think that there danced was
'I think one danced yesterday.'
- d. (Dutch B)
Ik denk dat *e/ er gedanst werd.

For this reason we will adopt a different approach here. Koster (1987) has argued that there is a no difference between Dutch A and Dutch B at all, or that at least this difference is negligible, because of a sentence like (134) which is acceptable for all speakers of both variants of Dutch.

- (134) (Dutch)
Wie denk je dat het gedaan heeft?
who think you that it done has
'Who do you think has done it?'

We will now assume that Koster is basically correct in stating that there is no real difference between Dutch A and Dutch B and that a sentence like is correct for all speakers of Dutch⁵³. This means that the *that-trace filter can be violated and that the (slight) contrast in (133)d - (133)c is (for some inexplicable reason) limited to intransitive verbs.*

Besides, we take here the standard position that in both Dutch and German main clauses, Verb Second first moves the inflected verb to the head of CP and then (obligatory) one of the remaining constituents to the SPEC-position of that projection.

Now we can generalize the observations in (133) to the statement in (135)

- (135) In Dutch and German, the [NP, S] position can be filled with a null expletive iff the head or the specifier of CP is filled.

At the same time we have to take into account that (135) does not apply to Russian, Esperanto or Italian, so that there seems to be a question of parametric variation here.

Notice also that (135) seems to have an intuitive connection with the observation we made concerning the non-possibility of PRO to be expletive. We have supposed before that this had to do with PRO being ungoverned and stated that a null expletive has to be governed. The parametric variation thus seems to have to do with the question as to

The syntax of a language without grammar

when an expletive is sufficiently governed to be null. We can now try to connect this to a better known parameter.

- (136) A null expletive has to be governed by an element in the CP/IP containing it.

Now, because in Verb Second languages like Dutch and German the presence of a CP seems to be always obligatory, at least in main clauses, it is not very strange to assume that a null subject has to be governed in main clauses by an element in the COMP-projection in those languages. On the other hand, it has been argued for Italian that its main clauses can be IP. In these languages government by the head of IP is sufficient.⁵⁴

In all these languages, expletive PRO also would have to be governed within the IP, which is known to be impossible.

A different approach can perhaps be found if we look to the theory of identification. Remember that Jaeggli and Safir argued that AGR could only identify a thematic *pro* iff the category containing AGR case-governed the empty category. We could extend this definition, so that in every NSL Tense had at least to govern *pro*. In Verb Second languages, where Tense is in COMP, this can only be accomplished by moving something into the projection of COMP. In languages like Russian and Esperanto, however, Tense as well as AGR is in INFL. The distinction between Russian/Esperanto and German/Dutch thus will fall out in an independent way. Notice that also under this approach the distinction amounts to a parametric distinction between CP and IP.

We can now, under any of the two approaches outlined here, ask ourselves why a sentence like (137) is excluded in Dutch under the given analysis, i.e. with *pro* moved to the SPEC of CP and governed there by the verb in the head, exactly as *pro* is governed in Italian by the verb in the head of its projection:

- (Dutch)
(137) *_{[CP pro_i wordt _{[IP e_i ... _[VP gedanst]]]]}}

This problem could be solved by stipulating that the relevant notion of *government* for the definition in (136), is *government under strict c-command in Dutch*. It is not clear why this stipulation should be made in Dutch, but not in Italian.

Fortunately, there is another explanation. We observe that in Dutch affirmative sentences, the [SPEC, CP] position has to be filled by some lexical element. In the default case, this is the subject NP, but it can in principle be the (in)direct object or an adjunct as well. On the other hand, in a non-WH question, the first position is normally left empty.⁵⁵

Let us now suppose that every CP has a feature [+/- question]. The values of this feature have to be realized in one way or another. This realisation will probably be submitted to parametric variation across languages.

In Dutch main clauses, for example, [+ question] is realized by either a [+WH] word in the [SPEC, CP] or else by a phonologically

The syntax of a language without grammar

empty [SPEC, CP]. [- question], the unmarked option, is realized by a [-WH] element in the SPEC of CP. In subordinate clauses [+ question] can also be realized by the complementizer *of*.

In English, [+ question] is realized by either inversion or [[+WH], CP]; in this language, [- question] is the unmarked option with no positive form feature related to it. In subordinate clauses, we also have the [+ question] complementizer *whether*.

Finally, in Esperanto [+ question] is realized by either a [+ WH] word in [SPEC, CP] or else by the complementizer *ĉu* both in main and subordinate clauses. Like in the English CP, [- question] has no positive form feature in this language.

In many languages (English, Dutch, French, Indonesian) the [+ question] can also feature can also be expressed by a phonological device such as (rising) *intonation*.

At what level does all of this take place? On the one hand, from the formulation we would like to say that the realisation of [+/- question] is a PF Filter (e.g. because empty categories are invisible even when they get case). On the other hand the question feature certainly has a role in the interpretation of a sentence and consequently at LF. At an abstract level, we find the feature even at the other levels of syntax. It is a well-known fact that some verbs subcategorize for a [+ question] CPs, others for a [- question] CPs and yet another class for [+ or - question] CPs:

- (English)
- (138) a. He wonders whether we sleep.
b. *He wonders that we sleep.
c. He wonders who sleeps.
- (139) a. *He believes whether we sleep.
b. He believes that we sleep.
c. *He believes who sleeps.
- (140) a. He knows whether we sleep.
b. He knows that we sleep.
c. He knows who sleeps.

We propose the following analysis. Suppose that the [+/- question] feature on CPs is similar to the [+/- case] feature on NPs. Furthermore, let us adopt the theory of Baker (1988) concerning case.

Baker (1988) proposed to differentiate between Case and case. *Case* (with a capital letter) is an abstract feature, that can be assigned thematically at D-structure and structurally at S-structure. At S-structure, the Case filter applies in order to see to it that every lexical NP receives Case. Subsequently, at LF, Case is used to identify the relevant thematic relations and at PF, it is realized as case (with a lower case 'c'). Languages may differ with respect to the way in which the realization of Case takes place. Some languages show a very rich morphology on the nominal elements (e.g. Latin), others show a rich agreement on the verb, whereas in a third class of languages (e.g. Winnebago), Case is realized by requirements of ad-

The syntax of a language without grammar

jacency and direction (e.g. in English). Baker proposes an explicit filter which requires that every instance of Case is identifiable as case at PF.

If we suppose that [+/- question] on CPs is something like [+/- case] on NPs, we may assume that we also have an abstract feature Question and a concrete realisation of this feature, question. Question is assigned at D-structure. At S-structure (after WH-movement), a Question Filter applies, that filters away the incorrect sentences in (138)-(140), and at LF, Question is used for the interpretation. AT PF, *question* realisation can cross-linguistically make use of very differing (morpho-phonological) means.⁵⁶

The last question we have to answer concerning Dutch and German expletive *pro*, is why weather verbs do not allow for null subjects in these languages. We assume here, with Bennis (1986), that *het* in sentences with weather verbs in these languages have in fact to be regarded as a thematic subject, whereas they don't have to be seen like that in Esperanto and Russian. Consequently, there would be some low-level lexical parametric variation between the different languages. The low-levelness of this parameter is also illustrated by the fact that in Esperanto poetry sentences like (92) the verb can take a lexical subject.

Finally we want to note that, if we would adopt the strongest version of the theory of identification by Jaeggli and Safir (1989) outlined above, thematic null subjects also are in some cases at least marginally grammatical in Dutch, if they can be identified. This identification is only possible by a subject clitic. As Dutch is a Verb Second language this clitic has to appear on (the head of CP). The Dutch clitic paradigm is very poor, because it consists only of a 3d person singular form. Except for these considerations, the thematic paradigm is exactly the same as the expletive paradigm (cf. the Flemish paradigm in (122) above)⁵⁷:

- (141) (Dutch)
- a. Gisteren kwam-ie *pro* ook al te laat op school.
 Yesterday came-3MS also too late at school
 'Yesterday, he was also late for school.'
- b. ... dat-ie *pro* gisteren te laat op school kwam.
 that-3MS yesterday too late at school came
- c. 'Teruggekeerd is-ie *pro* naar zijn vaderland.
 Returned is-3MS to his fatherland.
 'He has returned to his fatherland.'
- d. **pro* kwam-ie gisteren ook al te laat op school.
 came-3MS yesterday also too late at school

Notice that also here *pro* seems to require a c-commanding governor in COMP, which lets us decide in favour of the second approach to the

The syntax of a language without grammar

Dutch and German facts discussed above. Notice also that this approach is inconsistent with the standard movement account of clitics. This raises many problems.

There are some indications that we could perhaps also treat the Esperanto nominative pronominals as clitics, although they would probably cliticize at the left side of the verb. It is noted by Waringhien and Kalocsay (1935/85) that nothing can come between a pronominal subject and the verb, although the two constituents are separable when the subject consists of a full NP (cf. (142)) and the pronominal subject cannot be as freely extraposed as the full subject (cf. (143)). This would also explain the *difference* noted in the preceding section between Esperanto on the one hand and Japanese on the other in the way the languages acquire identification: the Esperanto agreement slot would not be empty but it would (have to) be filled with a clitic at the relevant level.

- (142) a. (Esperanto)
La viro antaŭ la nokta kunveno la
nigran hundon batis.
the man before the nocturnal conference the
black dog beat.
'The man beat the black dog before the
nocturnal conference.'
b. *Li antaŭ la nokta kunveno la nigran
hundon batis.
he before the nocturnal conference the black
dog beat.
- (143) a. Batas la hundon la dika viro.
beats the dog the fat man
'The fat man beats the dog.'
b. *Batas la hundon li.
beats the dog he

If this is true, we can suppose that Esperanto even allows for thematic *pro-drop*:

- (144) (Esperanto)
pro li-fumas
he smokes

In this case, we would have to suppose of course that thematic *pro-drop* is obligatory in Esperanto for there would be no pronominal subject available and a sentence like (145) is ungrammatical.

- (145) (Esperanto)
*La viro li-fumas
The man he-smokes

But we had already to assume that expletive *pro-drop* is obligatory in Esperanto, and there is evidence also from other languages that there is some further parametric distinction to be drawn here. In Russian, for example, expletive *pro-drop* is optional⁵⁸, while it is obligatory

The syntax of a language without grammar

in Italian, Polish and Serbo-Croatian. In Galician Portuguese pro-drop is absolutely optional for all relevant cases:

- (146) (Galician Portuguese)
- a. nós/ĕ comimos o caldo.
we ate the soup
 - b. el/ĕ chovia.
it rained
 - c. el/ĕ parecía que o patrón andaba canso.
it seemed that the boss went around tired.

All of this makes us suppose that there might be some more subtle parametric differences at stake than hitherto assumed in the literature (and here). It remains unclear to us what this difference could be.

Concludingly, we can say that Dutch and German are Null Subject Languages. Because AGR is not available at the right (Case-marking) node to identify thematic null subjects, normally only expletive pronouns are allowed. Dutch and German therefore show the Russian setting of the NSP. In these languages, however, *pro* has to be governed by an overt element in COMP; that requirement doesn't hold in Russian and Esperanto; in both these languages we also don't have to suppose the main clause to be a CP, since they are no Verb Second Languages.

Furthermore, we have seen that it is possible that Dutch and Esperanto have the possibility of allowing thematic *pro* if certain conditions are met.

4.4 Analysis of the Esperanto facts

Now that we have a crosslinguistically sufficiently adequate theory of pro-drop, we can try to explain the peculiarities of Esperanto syntax mentioned in section 4.1 in terms of the new theory. As a matter of course, we will skip the subjects that we have already discussed in the preceding sections here.

4.4.1 The form of the adjective

The adjectival predicate of a clause with expletive *pro* takes the ending -e, which is normally reserved for adverbs, as is illustrated in the examples in (94), repeated here as (147).

- (147) (Esperanto)
- a. Estas varme en la parko.
'It is warm in the park.'
 - b. La suno estas varma.
The sun is warm.
 - c. La plaĝoj estas varmaj.
The beaches are warm-PL

This certainly is a strange fact which at first sight we do not find in any other (Null Subject) language.

The syntax of a language without grammar

How can we explain this phenomenon? As we have seen in the preceding chapter, Dasgupta (1989) has argued on independent grounds that the Esperanto adverb has in fact a prepositional structure.

If we assume this theory, our problem reduces to the question: why is it that an expletive *pro* cannot be in the SPEC of an AP in Esperanto, while it can be in the SPEC of PP? We propose to link the explanation to the fact that expletive empty pronoun cannot in an IP with a Tensed I (where it would be PRO), while it can't be in an IP with an untensed I (where it would be *pro*). We have seen above, that there is much evidence that an empty pronoun has to be in a head-SPEC relation to AGR.

Now we have also seen in chapter 3 that a functional preposition can have an AGR-like element in it, although it need not be realized; this means there could also be a *pro* in principle. We have also seen that in Esperanto, the adverb probably moves to the head of the pP, and thus to the AGR-like element. We have no such evidence for the AP, however. On the contrary, we have seen in chapter 2, that Esperanto allows for PRO in [SPEC, AP] quite freely, and we have seen in chapter 3 that the latter position is not case-governed by its head.

Thus, we propose, that (148)a is parallel to (148)c, and (148)b to (148)d. (We abstract from the exact structural position of the extraposed sentence *ke li venas* here.)

- (Esperanto)
- (148) a. *Estas [_{AP} PRO verŝajna] (ke li venas)
 is probable that he comes
- b. Estas [_{PP} pro [_P AGR [verŝajne_i]] [_{PP} t_i]] (ke
 li venas.)
 is probably that he comes
 'It is probable that he comes.'
- c. *Marta supozas [_{TP} PRO [_{VP} verŝajni]] (ke li
 venas)
 Marta supposes probable-be that he
 comes
- d. Marta supozas [_{CP} ke [_{TP} PRO [_I AGR [verŝajnas_i]]
 [_{VP} t_i]]] (ke li venas.)
 Marta supposes that probable-
 is
 that he comes.
 'Marta supposes that it is probable that he
 comes.'

We could now extend this analysis to other sentences with an expletive, thus obtaining the desired result.

4.4.2 Pro-drop in imperatives?

As we have already briefly remarked in section 4.1, the imperative seems to be a mood in which a null subject is even allowed in non-NSLs like English or French:

The syntax of a language without grammar

- (149) a. (English)
 e go away!
 b. Don't (you/ e) leave me!
- (150) a. (French)
 Veuillez en prendre!
 Want from-it take
 'Please take some of these.'
 b. Va en pèlerinage!
 Go on pilgrimage

Given the (Extended) Projection Principle, we must assume that these sentences contain an empty category in the subject position. Since *movement* is clearly out of the question here, we must take it that the relevant category is *PRO* or *pro*.

Now we do not on first sight have to take this as a case of pro-drop. We could assume that the empty category in sentences like (149) is *PRO* and that the imperative mood -being somehow non-finite- is like the infinitive in that it doesn't Case-mark the subject. This analysis is rather problematic, however, because the empty category always has a very specific meaning (namely *you* for most of the time) and never an arbitrary one, as is demonstrated in (151)

- (151) a. (English)
 e wash yourself/yourselves!
 b. *e wash oneself!

Also the fact that the empty category can be replaced by an overt pronoun seems to be a fairly serious problem to any *PRO* analysis.

On the other hand, an analysis with *pro* is not without problems either, as the empty category can never be an expletive in imperatives even in languages which otherwise allow for expletives to be null. (Although the Dutch facts in (152) we might be able to explain independently along the lines of section 4.4)

- (152) a. (Dutch)
 *is duidelijk dat Jan ondeskundig is!
 is clear that John inexperienced is
 b. *wordt gevoetbald!
 is soccer-played
- (153) (Russian)
 *Temneet!
 is getting dark

We can however perhaps explain these facts pragmatically: because verbs with expletive subjects do not assign any theta role at all to their subjects, they in particular do not assign an agent role. And there seems to be little use in being imperative if there is nobody who can fulfill the orders.

The syntax of a language without grammar

Seen this way, the ungrammaticality of (152) and (153) can be explained in the same way as the relative ungrammaticality of the English unaccusative and passive sentences in (154) is.⁵⁹

- (English)
(154) a. *?Own a house!
b. *?Be slain in a battle!

There are some potential counterexamples to this pragmatic rule, namely sentences like (155)a and its Dutch translation in (155)b. We are not aware, however, of examples like these without the biblical/magical shade of meaning.⁶⁰

- (English)
(155) a. Let there be light!
b. Er zij licht!
There be light

That the imperative is also otherwise an exceptional mood is noted in Jaeggli and Safir (1989b, fn. 17) where it is observed that the imperative mood has to be excluded when considering whether a language is morphological uniform or not, because in languages like Danish, German and Turkish (and we can certainly add Dutch to this series) some (second person singular) imperative forms are the only underived forms without this having any effect on the NSP setting for these languages.⁶¹

We propose here to modify the requirement of Morphological Uniformity (which we gave in (107)) to explain both these points. What we propose is that uniformity is not considered in a language for the whole of its verbal paradigm, but separately for each mode.

- (156) *Morphological uniformity* (revision)
An inflectional paradigm *P* for a mode *M* in a language *L* is morphologically uniform iff *P* has either only underived inflectional forms or only derived inflectional forms (where an inflectional form is *underived* if it is equal to the stem and *derived* otherwise).

As can be seen, this is only a slight revision of the original formulation, and all of the explanations proposed earlier remain untouched by it. The Scandinavian examples we can now explain by taking recourse to the same escape hatch we already needed for the Dutch paradigm: at most one form can be equal to the stem if the rest of the paradigm is of a derived form.

Of course, we have to see to it that also *identification* can handle the imperative cases, but this seems not a very hard task to accomplish. The mood consists in many languages (English, German, Dutch, Esperanto) consists of only the second person singular and plural. In French it consists of three persons: second person singular and first and second person plural. Furthermore it is a peculiarity of the second person that it doesn't always have to make a distinction between singular and plural (cf. *you* in English) so that

The syntax of a language without grammar

agreement will be rich enough to identify a missing subject tolerably soon.

4.5 Conclusion

In this section we have seen that Esperanto shows a rule of pro-drop. Like in the preceding chapter, this rule on first sight is very intricate and very 'illogical'. At the same time, it is a rule which can be explained by reference to a parameter of UG.

It seems that this theory, most of which is developed only very recently on the basis of other human languages is needed to explain the Esperanto facts. Because it seems not very reasonable to think that an amateur linguist in the nineteenth-century could foresee the theoretical work of professional linguists in our time, or that it is simply a matter of accident that Esperanto behaves the way it does.

It also doesn't really make sense to suppose that Esperanto is just like Russian and that Zamenhof has simply imitated that language in this respect, because the verbal inflections of the two languages are too different from each other and verbal inflection, as we have seen, is a very important in matters of pro-drop. We have seen that Esperanto is more like Japanese in this respect -even if those two languages probably differ with respect to the pro drop parameter.

In the mean time, we have proposed a theory of pro-drop that is a partial extension/revision of the theory proposed by Jaeggli and Safir (1989) and that we can sum up in the following way. The pro-drop parameter doesn't have to be fixed for all paradigms in a language. It rather is mode-specific. At the same time there is a strong connection between the allowance of pro-drop in a certain mode in a certain language and the verbal paradigm of that mode in that language (definition (156)). In certain cases pro-drop is also possible within prepositional phrases; probably more or less the same requirement counts for those cases although we have done no research on this question.

There are at least two kinds of pro-drop languages. The verbal paradigm plays a rather crucial rule here also. Requirement (120)d is, we think, a principle of UG. Identification is most of the time done by the AGR system at the verbal (or prepositional) paradigm, but also other types of identification exist (for instance in Chinese). There is some evidence that we should differentiate pro-drop languages in still another way; in some languages pro-drop seems to be obligatory, whereas it is optional in others.

At least two other principles of UG interfere with these parameter settings.

Firstly, we have the *emex* condition (114), which we need independently for the explanation of the non-occurrence of expletive PRO and which can account for the behaviour of the Esperanto adjectival paradigm.

Secondly, we have a revised form of the PF Visibility Condition, which requires that the feature [+/- question] has to be realized at PF. This condition accounts for the fact that *pro* in Dutch and German can't be in sentence initial position.

The syntax of a language without grammar

The syntax of a language without grammar

5 Conclusions

In this essay after a short excursion to the philosophical study of the notion 'natural language', three topical subjects of generative linguistics, namely null subjects, functional prepositions and binding of anaphors, are treated, with special concern to the facts in Esperanto. We have the strong impression that this field of research is not yet dried up. On the contrary, many issues could have replaced the three considered here.

As mentioned in the first section, we could for example also have treated the *free word order* question. Although the word order is claimed to be completely free in this language⁶², there are many indications that the language has in the course of time grown into a quite regular Head-First language in which the word order is certainly not less restricted than in a relatively free language like Dutch.

Also the Esperanto participle system seems to be interesting to us in the light of the recent discussion on the status of participles (see among others Borer (1989), Milsark (1988) and references cited there). Not every Esperanto participle does, at least on first sight, in all respects *walk like an adjective* (cf. Borer (1989)). For instance, it cannot occur most of the time with the modifier *tre*, 'very', except when it is a participle of a psychological verb. Compare e.g. the sentences in (157) which have only a differing *shade* of meaning and the sentence in (158) with a psychological verb.

- (Esperanto)
- (157) a. *La tre dormanta knabino estas bela.
 *the very sleeping girl is beautiful.
 b. La tre dormema knabino estas bela.
 the very sleepy girl is beautiful.
- (158) La tre ravanta knabino estas bela.
 the very ravishing girl is beautiful.

It is interesting also that the Esperanto modifier *tre* shows the same sort of behaviour as the Hebrew *me'od* (discussed by Borer (1990)) and the Dutch *zeer* (discussed by Bennis and Wehrmann (1990)), not as the English *very* (*ibidem*):

- (Esperanto)
- (159) a. *Ŝi dormas tre.
 She sleeps very
 b. Ŝi ravas (min) tre.
 She charms me very
 c. *tre dormanta knabino
 very sleeping girl
 d. tre ravanta knabino
 very charming girl

The syntax of a language without grammar

We see that Esperanto shows the same behaviour here as the English glosses. Whatever the explanation for these remarkable data may be, they were certainly not explicitly known century ago.

We believe that all of the facts that are discussed in this essay are very difficult to explain without reference to the principles and parameters of UG.

Although it could be assumed that the setting for the pro-drop parameter in Esperanto is the most *logical* one also at a superficial level, it can't be easily explained why the behaviour of the adjectival/adverbial predicate in subjectless sentences is the way it is. The same applies for our analysis of directional/locational verbs. One could assume that this feature is *borrowed* from a language like German, but that doesn't explain the -n marking on adverbs in Esperanto. As for the Binding phenomena discussed here: it evidently is very strange that words like 'everything' and 'who' should be exceptional with respect to the Binding theory.

We don't know of course whether these issues will ever be picked up and elaborated, but we hope that we have at least shown that the apprehension of this (unquestionably sociologically peculiar) language within modern linguistics is unjustified.

Although facts in Esperanto can perhaps hardly be the only reason to totally change the scientific model of UG, it can at least bring up new external evidence for existing theories of Universal Grammar.

Marc van Oostendorp, August-November 1990.

The syntax of a language without grammar

Notes

1."De taal die in een menselijke gemeenschap van oudsher gebruikt wordt voor algemene communicatie en die een kind, opgroeiend in zo'n gemeenschap, als zijn moedertaal leert, wordt ook wel 'natuurlijke taal' genoemd. Dit ter onderscheiding van kunstmatige talen en wetenschappelijke talen. Van een natuurlijke taal kunnen we nooit zeggen dat hij op een bepaald moment door iemand of door een groep mensen is bedacht of ontworpen. Alle natuurlijke talen zijn produkten van een lange traditie en hoever we ook in de geschiedenis teruggaan, nergens treffen we gegevens aan die duidelijk zouden kunnen maken hoe ze ontstaan zijn.

Kunstmatige talen daarentegen, zoals bijvoorbeeld het Esperanto, zijn op een bepaald moment ontworpen om als vervanging van natuurlijke talen te dienen.[...]"
(Dik and Kooij (1984), p.12)

2.We will not go into this last distinction any further here, because we don't feel that the *function* of a language is in any way relevant for the formal branch of linguistics that we adopt.

In fact, we don't even see how one can determine what the function of a given language may be. We are not aware of any linguistic framework in which a similar determination can be made either.

3.With the exception of Dasgupta (1989), an article of only 14 pages of length. Besides, these pages are almost entirely devoted to the exposition of a theory of functional projections, but the author uses Esperanto as a language of illustration in them. We will discuss the main ideas in this article in chapter 3 and 4.

4.This is the Bloomfieldian (1928) concept of an 'idealized' homogenous linguistic community.

5.The same sort of questions are posed with reference to exactly the same passage by Liceras (1988). Yet I found this article after I had written this passage.

6.The H in this quotation stands for a human being, and the R for the '*knowing, having or whatever*' (Chomsky (1986b)) relation between a person and his (native) language.

7.At least, it can't be learned in the almost thoughtless way in which a child can learn Chinese or French. There is a possibility that it can be learned in the way a subject like mathematics is learned -that is also the way in which the adults of our experiment shall have to learn the language.

8.For a discussion of the relevance of creole languages for the theory of UG, we refer to the work of Pieter Muysken, e.g. the introduction to Muysken (1981).

We believe that much of Muysken's arguments in favour of the study of creole languages within generative linguistics could be also used to advocate the study of Esperanto.

On the other hand, Schubert (1989) contends that there are also differences between creoles and Esperanto, since native speakers have a minor influence in the latter language, but we believe that Schubert's interest is not as much in I-language as in E-language. He claims himself to be in the 'structuralist tradition.'

The syntax of a language without grammar

9.* rule 12

'If there is a negative in a clause, a second is not admissible.'

* rule 13

'In phrases answering the question 'where?' (meaning direction) the words take the termination of the objective case; e.g. *kie'n vi iras?*, 'where are you going?'; *dom'o'n*, 'home'; *London'o'n*, 'to London', etc.'

We will discuss the implications of rule 13 in chapter 3.

10. We have adopted the convention of giving a literal translation directly under sentences in languages other than English and subsequently, in quotes, a grammatical translation; unless the literal translation is by accident grammatical also in English.

11. For instance, Schubert (1989) points out that in the sixteen rule grammar it is stated that 'the past participle active ends in *-int*', without any explanation of what a past participle is or how it may serve in the language structure. These facts are apparently supposed to be known by the reader of the grammar, or as Schubert puts it, *in 1887, you had been told by your Latin teacher what a participle is.*

12. In fact we could say that Esperanto is in some way the Chomskyan (1986b) mixture of Russian and French because the lexical material is for the greater part derived from the Romance.

13. The RHHR even has a status as number 11 of the sixteen rules of grammar:
"Compound words are formed by simple junction of roots (the principal word standing last)."

14. Although he doesn't use the term Right Hand Head Rule.

15. Under the proviso that question formation itself cannot serve as a ready-made example in this language, because it isn't accomplished by any kind of inversion, but by placing a question marking particle, *ĉu*, at the beginning of the interrogative clause.

16. We will not discuss markedness within the periphery, because we find this notion very difficult to apply. There is also hardly any discussion of this concept in the literature.

17. Or even with more justification, since Esperanto was assumed to have no syntax at all.

18. Interestingly, most of the so called *naturalistic* competitors of Esperanto had a large periphery full of irregular verbs and intricate morphological rules applying to nouns, which made them *closer to real languages* according to their authors.

One of the most outstanding linguists of the twentieth century, Otto Jespersen, himself has devised once a language scheme -called *Novial*- which was far more naturalistic than Esperanto was. Jespersen extensively argued that the latter language could never be a success because it was so terribly *unnatural*. His own language has received very little support.

19. The phonetic component of the language faculty probably has to be treated as a special case here: the phonetic incompetence of a second language speaker can at least partly be explained with reference to physiological factors.

The syntax of a language without grammar

20. Notice that the experiment with the language L, described above, would also be a perfect test to decide in this controversy; first of all, we would predict that L would be very hard to learn for the adults. They would probably make many 'mistakes', resulting in a language L*. If L* would have all the parameters set in the unmarked way, the 'CC' approach would win.

21. We could make a distinction here, following Everaert(1986), between the *weak* reflexive *sin* and the *strong* reflexive *sin mem*, but we shall not do that here for matters of brevity.

22. It should be noted that in (14) *li*, being a pronoun, is free in its reference. This sentence therefore can also have a reading in which Peter (or anyone else but Paul) is the one who does the injuring. This is of no further importance to our discussion, however, for the other readings can be derived in the same way as these ones without significant complications.

23. All these examples are drawn from Waringhien and Kalocsay (1935/85), a valuable standard work in descriptive Esperanto syntax.

24. Apart from the fact that *si* in the formula's *en si* ('in itself') and *per si* ('by itself') and *sia* in *siatempe* ('at a convening time') have idiomatic readings.

25. Traditionally, there are two approaches within GB to the sentences in (19): we can treat them syntactically, as we have done here, or we can treat them pragmatically. The 'ungrammatical' sentences (19)b and (19)c seem to be pragmatically not very plausible. The same does not apply, however, to sentences like (19)f and (20)a, or to most of the Esperanto examples; that is the reason why we adopt Chomsky's syntactic approach.

26. This optionality is one of the features that has also grown in the course of time into the normal Esperanto usage, as we can learn from Waringhien and Kalocsay (1935/85), who note that, while (1)a, where an instance of PRO is included in the NP, was the standard in pioneering times, nowadays (1)b, without PRO, which was originally thought to be ungrammatical, is found in an equal quantity, or even preferred.

(Esperanto)

- (1) a. La dukino atendis la reporton de ŝia braceleto.
The duchess awaited the bringing-back of her bracelet.
a. La dukino atendis la reporton de sia braceleto.
The duchess awaited the bringing-back of her own bracelet.

27. It should be noticed, however, that in the analysis in that section it is assumed that clitic pronouns are base-generated at their S-structure position whereas they are moved to that position in the analysis of Chomsky (1986b) which is under discussion here.

28. Despite the justified theory-internal objections that Chomsky (1986b) makes against this approach.

Chomsky remarks himself repeatedly that the facts are more complex than the theory - that is: than any theory developed until now- can explain; thus it seems more or less justified in a modest work like this one to take the elements of the different versions of Binding Theory that suit the data best.

The assumption that an NP can have PRO in subject position, for that matter, doesn't seem to us to be inconsistent with an Accessible SUBJECT approach. Whether this theory

The syntax of a language without grammar

20. Notice that the experiment with the language L, described above, would also be a perfect test to decide in this controversy; first of all, we would predict that L would be very hard to learn for the adults. They would probably make many 'mistakes', resulting in a language L*. If L* would have all the parameters set in the unmarked way, the 'CC' approach would win.

21. We could make a distinction here, following Everaert(1986), between the *weak* reflexive *sin* and the *strong* reflexive *sin mem*, but we shall not do that here for matters of brevity.

22. It should be noted that in (14) *li*, being a pronoun, is free in its reference. This sentence therefore can also have a reading in which Peter (or anyone else but Paul) is the one who does the injuring. This is of no further importance to our discussion, however, for the other readings can be derived in the same way as these ones without significant complications.

23. All these examples are drawn from Waringhien and Kalocsay (1935/85), a valuable standard work in descriptive Esperanto syntax.

24. Apart from the fact that *si* in the formula's *en si* ('in itself') and *per si* ('by itself') and *sia* in *siatempe* ('at a convening time') have idiomatic readings.

25. Traditionally, there are two approaches within GB to the sentences in (19): we can treat them syntactically, as we have done here, or we can treat them pragmatically. The 'ungrammatical' sentences (19)b and (19)c seem to be pragmatically not very plausible. The same does not apply, however, to sentences like (19)f and (20)a, or to most of the Esperanto examples; that is the reason why we adopt Chomsky's syntactic approach.

26. This optionality is one of the features that has also grown in the course of time into the normal Esperanto usage, as we can learn from Waringhien and Kalocsay (1935/85), who note that, while (1)a, where an instance of PRO is included in the NP, was the standard in pioneering times, nowadays (1)b, without PRO, which was originally thought to be ungrammatical, is found in an equal quantity, or even preferred.

(Esperanto)

- (1) a. La dukino atendis la reporton de ŝia braceleto.
The duchess awaited the bringing-back of her bracelet.
a. La dukino atendis la reporton de sia braceleto.
The duchess awaited the bringing-back of her own bracelet.

27. It should be noticed, however, that in the analysis in that section it is assumed that clitic pronouns are base-generated at their S-structure position whereas they are moved to that position in the analysis of Chomsky (1986b) which is under discussion here.

28. Despite the justified theory-internal objections that Chomsky (1986b) makes against this approach.

Chomsky remarks himself repeatedly that the facts are more complex than the theory - that is: than any theory developed until now- can explain; thus it seems more or less justified in a modest work like this one to take the elements of the different versions of Binding Theory that suit the data best.

The assumption that an NP can have PRO in subject position, for that matter, doesn't seem to us to be inconsistent with an Accessible SUBJECT approach. Whether this theory

The syntax of a language without grammar

could be made compatible with the movement analysis is unclear -but it is unclear anyway how Binding by a subject inside an NP works in an analysis of movement to INFL would work.

29. We will not deal with this reduction here. For a critical discussion see Lasnik and Uriagereka (1989). The question arises whether the Binding theories of Aoun (1986) and Chomsky (1986b) are compatible. We believe that they are for the parts that are relevant to this discussion, i.e. we believe that A-bar anaphora could be incorporated into Chomsky's framework.

30. This exposition of course raises a lot of new questions which, unfortunately, are not answered by Aoun (1986) and which, even more unfortunately, we don't have the space to answer here, because they are not really relevant to us.

31. We have made up these sentences by ourselves. Aoun (1986) gives only an abstract scheme. We have chosen here the lexical material of Standard French, because we don't know the particular dialect.

32. Of course, we now also would expect that *sia* could be bound by other constituents which are raised to the SPEC-position of CP at some syntactic level or another. Although Waringhien and Kalocsay make no explicit note of this, my informant has asserted to me that to her ear (1) sounds better than (2).

(Esperanto)

- (1) ??Kiu vi prezentas al si?
whom you introduce to himself
'Whom do you introduce to himself?'
- (2) *Vi prezentas Johanon al si.'
you introduce John to himself

33. Except that we find some relic (*sic*) genitival form in words like *kies* 'whose', *ĉies* 'everybody's' and *ies*, 'somebody's.'

34. At least this is true for Russian and Polish. We have found no evidence for this phenomenon in Perkowski's Kashbukian data.

35. This might be not quite true. In Hungarian, we find constructions like (1)a (cf. Marácz (1989)), where the demonstrative receives a P of its own. This form of 'agreement' is obligatory. See Van Riemsdijk (1990) for a discussion of this construction in a functional-head approach.

(Hungarian)

- (1) a. a mögött a ház mögött
that behind the house behind
'behind that house'
- b. *az a ház mögött
that the house behind

36. We will adopt the notation of Van Riemsdijk (1990) here, and present a functional preposition as 'p' (in lower case) and a lexical preposition as 'P' (in upper case) - although we are not very happy with this rather confusing form of notation.

The syntax of a language without grammar

37. The only exception is the naked P *szem-* 'opposite to', but this P shows a different morphology (Marácz (1989:364)):

- (Hungarian)
- (1) szembe szemben szemközt
 (to) opposite to opposite to (to) opposite to

38. We assume, with Van Riemsdijk (1990), that the SPEC position is the position immediately under the FP (where F = n, v, p, ...), adjacent to F'.

39. In fact, there are a few words which at first sight seem to be inherent adverbs, like *neniam*, 'never', and *tiom*, 'so much', but Dasgupta (1989) shows that also these words can be analysed as PPs.

Dasgupta (1989) assigns the structures in (1)a and (1)b, ? to the synonymous *tiel multe* and *tiom* respectively (translated to our approach of functional projections). In these structures Deg stands for the (functional) head Degree (the internal structure of *tiel mult* is not relevant to our discussion)

- (Esperanto)
- (1) a. [[[_{DEG}[tiel] mult_A]_i -e]_p [t_i]_{DEGP}]_{PP}
 b. [[[_{DEG}[ti-] om_A]_i 0]_p [t_i]_{DEGP}]_{PP}

For a justification of this analysis, see Dasgupta (1989).

40. Marácz argues that adverbs form a different category from postpositions, i.a. on the basis of these facts. Notice however that the two adverbs that he shows to illustrate this behaviour, namely *bentre/öl* '(to) inside' and *fentre/öl* '(to) above' seem to be related morphologically to the postpositions *belül* 'inside' and *felül* 'over', respectively.

41. The term *pro-drop* is only used for historical reasons here. No serious recent analyses, in which pronouns are deleted, are known to us.

42. The only purpose of Esperanto inflection seems to be the marking of mood and tense features: *parolos* = shall speak, *parolis* = spoke, *parolus* = would speak, etc. In this respect, it shall turn out to be like Japanese.

43. We could have included examples like in (1) here, following Rizzi (1982). The reason we didn't do this is that it is always affirmed in descriptive grammars that 'Esperanto word order is totally free', so that we can possibly analyze the sentences in (1) as base generated, without any null expletive or dummy pronoun at all.

- (Esperanto)
- (1) a. e venis sinjoro kun longa barbo
 came a gentleman with a long beard
 b. e okazis io tre malagraba
 happened something very unpleasant

44. The same observation applies for Italian and Spanish as well, but not for Galician Portuguese, as we shall see later on in this chapter.

45. Especially Jaeggli and Safir (1989b), Huang (1989) and Borer (1989).

The syntax of a language without grammar

46. Japanese is thus in its verbal morphology exactly like Esperanto as we have noted in an earlier footnote (fn. 42).

47. Again, Rizzi (1982) had a theory about these facts which we are not able to discuss here.

48. Franks (1990) offers a theory of pro-drop which is quite different from the theory developed here; we shall not discuss it, although it is quite an original approach, that might offer some answers to questions that we are not able to tackle.

49. This is not precisely the point of view of Huang (1989), who contends that there is no real distinction between PRO and *pro*.

50. In Dutch, like in German, it is possible to passivize intransitive verbs. In main clauses, expletive *er* has to appear in the first position of the sentence:

- (Dutch)
- (1) *(Er) wordt gedanst.
 there is danced
 'One dances.'

51. The terminology is in fact much older than Bennis (1983), but in the earlier work (Maling and Zaenen (1978) for example) a correlation between Dutch *that*-trace and sentences without a subject was more or less assumed. Pesetsky (1979) even claimed that Dutch B had the same setting for the NSP as English and French. It was Bennis who showed that such a correlation did not really exist.

Bennis (1983) also showed that the rule of *er* insertion of Maling and Zaenen (which would be optional in Dutch A and obligatory in Dutch B) could not be correct, i.a. because also a rule of *het* insertion would be needed, where it wasn't clear what the connection between the two forms of insertion would be, and more generally because insertion is in itself not a well-liked operation within the GB framework.

52. For some reason or other, sentences like this with a locative at the sentence-initial position, are better than sentences like the following, with some other adverbial (or prepositional) phrase.

- (1) *Gisteren werd gedanst.

The difference is only marginal to us, but there seem to be speakers who feel it rather sharply. We don't have an explanation for this.

53. On the other hand, both dialects of Dutch show a sharp distinction with so-called *wat voor* - constructions, as in (1)

- (Dutch)
- (1) a. *Wat zeg je dat e voor studenten klagen?
 What say you that for students complain
 'What kind of students complains, according to you?'
 b. Wat denk je dat die man e voor boeken leest?
 What think you that that man for books reads
 'What kind of books does that man read, according to you?'

This means a problem for the theory proposed here.

The syntax of a language without grammar

Jaeggli and Safir (1989b) tentatively propose that up to one stem identical form, excluding imperatives, is permitted. We adopt this position, with one modification to be discussed later.

62. We quote here a very instructive passage by Waringhien and Kalocsay (1985:362), which we have translated for ourselves:

"In Esperanto the word order is free, i.e. there don't exist special rules of word order in it, one should just strive for clarity and euphony. Yet, also in Esperanto some principles apply, which one should better not disobey, because they conform to the natural flow of thought [*sic*]. And there even exist some rare cases when one has to hold oneself to a strict word order."

The syntax of a language without grammar

References

- AOUN, J.
(1985) *A grammar of anaphora*. Cambridge (Mass.): The MIT Press
(1986) *Generalized Binding*. Dordrecht: Foris.
- BAKER, M.
(1988) *Incorporation; A theory of grammatical function changing*. Cambridge (Mass.): The MIT Press.
- BELLETTI, A.
(1988) 'The case of unaccusatives.' *Linguistic Inquiry* 19, 1-34
- BELLETTI, A., L. BRANDI and L. RIZZI (eds.)
(1981) *Theory of markedness in generative grammar*. Pisa. Proceedings of the GLOW conference.
- BENNIS, H.
(1983) 'De PRO-drop parameter en subjektloze zinnen in het Nederlands.' *Spektator* 12, 409-428.
(1986) *Gaps and dummies*. Doctoral dissertation, University of Tilburg.
- BENNIS, H. and P. WEHRMANN
(1990) 'Participles in Dutch'. Ms., University of Leiden.
- BLOOMFIELD, L.
(1928) 'A set of postulates for the science of language.' *Language* 2.
- BORER, H.
(1989) 'Anaphoric AGR.' In: Jaeggli and Safir (1989a).
(1990) 'V+ing: It walks like an adjective, it talks like an adjective.' *Linguistic inquiry* 19, 611-634.
- BOULTON, M.
(1980) *Zamenhof; Creator of Esperanto*. London: Routledge and Kegan Paul. 2nd edition.
- BRANDT CORSTIUS, H.
(1988) 'Malesperanto'. Elsevier.
- BROOKS, J.
(1975) *Polish reference grammar*. 's Gravenhage: Mouton.
- BURZIO, L.
(1986) *Italian syntax*. Dordrecht: Reidel.
- CHOMSKY, N.
(1975) *Reflections on language*. New York: Pantheon.
(1981) *Lectures on government and binding*. Dordrecht: Foris.
(1982) *Some concepts and consequences of the theory of government and binding*. Cambridge (Mass.): The MIT Press. Linguistic Inquiry Monograph 6.
(1986a) *Barriers*. Cambridge (Mass.): The MIT Press.
(1986b) *Knowledge of language; Its nature, origin and use*. New York (etc.): Praeger Publishers.
(1988) *Language and problems of knowledge; The Managua lectures*. Cambridge (Mass.)/ London: The MIT Press.
(1989) 'Some notes on the economy of derivation.' In: I. Laka & A. Mahajan (eds.): *Functional heads and clause structure*. Cambridge (Mass.): MIT. MIT Working papers in linguistics 10.

The syntax of a language without grammar

- CLAHSEN, H.
(1988) 'Parametrized grammatical theory and language acquisition.' In: Flynn and O'Neil (1988).
- CLAHSEN, H. and P. MUYSKEN
(1986) 'The availability of universal grammar to adult and child learners. A study of the acquisition of German word order.' *Second language research* 2, 93-119.
- DASGUPTA, P.
(1987) 'Towards a dialogue between the sociolinguistic sciences and Esperanto culture.' *Language Problems and Language Planning* 11. 305-334.
(1989) 'Degree words in Esperanto and categories in UG.' In: K.Schubert (ed.) (1989a)
- DIK, S.C. and J.G. KOOIJ
(1984) *Algemene taalwetenschap*. Utrecht/Antwerpen: Uitgeverij Het Spectrum.
- DULAY, H. and M. BURT
(1974) 'Natural sequences in child second language acquisition.' *Language learning* 24, 37-53.
- EVERAERT, M.
(1986) *The syntax of reflexivization*. Dordrecht: Foris.
- FELIX, S.
(1988) 'UG-generated knowledge in adult second language acquisition.' In: Flynn and O'Neil (1988).
- FILLMORE, C
(1968) 'The case for case'. In: E. Bach and R. Harms (eds.): *Universals in linguistic theory*. New York: Holt, Rinehart and Winston.
- FLYNN, S.
(1987) *A parameter-setting model of L2 acquisition; Experimental studies in anaphora*. Dordrecht: D. Reidel. *Studies in Theoretical Psycholinguistics*.
- FLYNN, S. and W. O'NEIL (eds.)
(1988) *Linguistic theory in second language acquisition*. Dordrecht: Kluwer academic publishers.
- FISCHER, S.
(1978) 'Sign languages and creoles.' In: Patricia Siple (ed.): *Understanding language through sign language research*. New York (etc.): Academic Press
- FRANKS, S.
(1990) 'On the status of null expletives.' *Lingua* 81: 1-24.
- FRIES, C.
(1945) *Teaching and learning English as a foreign language*. Ann Arbor (MI): University of Michigan Press.
- GAIR, J.W.
(1988) 'Kinds of markedness', In: Flynn and O'Neil (1988).
- HAEGEMAN, L.
(1988) 'The categorial status of modals and L2 acquisition.' In: Flynn and O'Neill (1988)

The syntax of a language without grammar

- HAIDER, H.
(1985) 'The case of German'. In: J. Toman (ed.): *Linguistic theory and the grammar of German*. Dordrecht: Foris. *Studies in Generative Grammar* 21.
- HALE, K.
(1981) *On the position of Warlpiri in a typology of the base*. Bloomington: Indiana University Linguistics club.
(1983) 'Warlpiri and the grammar of non-configurational languages.' *Natural language and Linguistic Theory* 1, 1-43.
- HASEGAWA, N.
(1985) 'On the so-called 'zero pronouns' in Japanese'. *The Linguistic Review* 4, 289-243.
- HAYES, B.
(1981) *A metrical theory of stress rules*. New York: Garland.
(1987) 'A revised parametric metrical theory.' *Proceedings of NELS* 17. Volume I. pp. 274-289.
- HYLTENSTAM, K. and L.K. OBLER
(1989a) *Bilingualism across the lifespan*. Cambridge (Britt.): Cambridge University press.
(1989b) 'Bilingualism across the lifespan: an introduction'. In: Hyltenstam and Obler (1989a).
- HUANG, C.T.J.
(1989) 'Pro-drop in Chinese: A generalized control theory.' In: Jaeggli and Safir (1989a). 185-214.
- JAEGGLI, O. and K. J. SAFIR (eds.)
(1989a) *The null subject parameter*. Dordrecht/Boston/London: Kluwer Academic Publishers.
(1989b) 'The null subject parameter and parametric theory.' Introduction to Jaeggli and Safir (1989a). 1-44.
- JESPERSEN, O.
(1924) *The philosophy of grammar*. London: Allan and Unwin.
- KALOCSAY, K. and G. WARINGHIEN
(1935/85) *Plena analiza gramatiko de Esperanto*. Rotterdam: UEA. 5th edition.
- KENSTOWICZ, M.
(1989) 'The null subject in modern Arabic dialects.' In: Jaeggli and Safir (1989a)
- KOSTER, J.
(1978) *Locality principles in syntax*. Dordrecht: Foris.
(1987) *Domains and dynasties*. Dordrecht: Foris.
- LADO, R.
(1957) *Linguistics across cultures*. Ann Arbor: University of Michigan Press.
- LASNIK, H. and J. URIAGEREKA
(1988) *A course in GB syntax; Lectures on Binding and Empty Categories*. Cambridge (Mass.): The MIT Press.
- LICERAS, J.
(1988) 'L2 learnability: Delimiting the domain of core grammar as distinct from the marked periphery.' In: Lynn and O'Neil (1988)

The syntax of a language without grammar

- MALING, J. and A. ZAENEN
(1978) 'The non-universality of a surface filter.' *Linguistic Inquiry* 9, 475-497.
- MARACZ, L.K.
(1990) *Asymmetries in Hungarian*. Doctoral dissertation, University of Groningen.
- MAZURKEWICH, I.
(1985) 'Syntactic markedness.' *Studies in second language acquisition* 7.
- MCCLOSKEY, J. and K. HALE
(1984) 'On the syntax of person-number inflection in modern Irish'. *Natural language and linguistic theory* 1, 442-487.
- MEISEL, J.M.
(1989) 'Early differentiation of languages in bilingual children.' In: Hyltenstam and Obler (1989a).
- MILLER, P.
(1990) 'A lexical analysis of French pronominal clitics'. Unpublished ms., Université libre de Bruxelles.
- MILSARK, G. L.
(1988) 'Singl-ing.' *Linguistic inquiry* 19, 169-180.
- MUYSKEN, P. (ed.)
(1981) *Generative studies on creole languages*. Dordrecht: Foris.
- MUYSKEN, P. and N. SMITH (eds.)
(1986) *Substrata vs. Universals in Creole Genesis*. Amsterdam: John Benjamins Publishing Company. Creole Language Library, Volume I.
- NEIDLE, C.
(1988) *The role of case in Russian syntax*. Dordrecht (etc.): Kluwer Academic Publishers. Studies in Natural Language and Linguistic Theory.
- OBLER, L.K.
(1988) 'Neurolinguistics and parameter settings.' In: Flynn and O'Neil (1988).
- PERKOWSKI, J.L.
(1969) *A Kashkubian idiolect in the United States*. 's Gravenhage: Mouton & Co.
- PERLMUTTER, D.
(1971) *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.
- PESETSKY, D.
(1979) 'Complementizer-trace phenomena and the Nominative Island Condition.' *The Linguistic Review* 1, 297-344.
- PIRON, C.
(1989a) 'Who are the speakers of Esperanto?'. In: K.Schubert (ed.) (1989a)
(1989b) 'A few notes on the evolution of Esperanto.' In: K. Schubert (ed.) (1989a).
- POLLOCK, J.-Y.
(1989) 'Verb movement, UG and the structure of IP.' *Linguistic Inquiry* 20, p. 350-394.

The syntax of a language without grammar

- RIEMSDIJK, H.C. van
(1978) *A case study in syntactic markedness*. Dordrecht: Reidel.
(1988) 'Against adjunction to the head.' McGill Working Papers in Linguistics 239-259.
(1990) 'Functional prepositions'. In: H. Pinkster and I. Genee (eds.): *Unity in diversity; Papers presented to Simon C. Dik on the occasion of his 50th birthday*. Dordrecht: Foris.
- RIEMSDIJK, H.C. van, and E. WILLIAMS
(1986) *Introduction to the theory of grammar*. Cambridge (Mass.): The MIT Press.
- RIZZI, L.
(1982) *Issues in Italian Syntax*. Dordrecht: Foris.
- ROSS, J.
(1967) 'Constraints on variables in syntax'. Doctoral dissertation MIT.
- SAFIR, K.
(1985) 'Missing subjects in German.' In: J. Toman (ed.): *Linguistic theory and the grammar of German*. Dordrecht: Foris. *Studies in Generative Grammar* 21.
- SAUSSURE, R. de
(1910) *La logika bazo de vortfarado en Esperanto*. Saarbrücken: Anton Iltis.
(1914) *Vortteorio de Esperanto*. Saarbrücken: Anton Iltis.
- SCHUBERT, K.
(1989a) (ed.) *Interlinguistics; Aspects of the science of planned languages*. New York: Mouton-De Gruyter. *Trends in linguistics studies and monographs* 42.
(1989b) Introduction to Schubert (ed.) (1989a)
(1989c) 'An unplanned development in a planned language.' In: Schubert (ed.) (1989a)
- STAAL, F.
(1986) *Over zin en onzin in filosofie, religie en wetenschap*. Amsterdam: Meulenhof.
- STANKIEWICZ, E.
(1986) *The Slavic languages; Unity in Diversity*. Berlin (etc.): Mouton De Gruyter.
- SZABOLSCI, A.
(1986) 'Bound variables in syntax (are there any?)'. In: Groenendijk and Stokhof (eds.): *Proceedings of the XVII Amsterdam conference on language and logic*.
- TERVOORT, B. T.
(1988) 'Gebarentaalonderzoek; Zijn gebarentalen natuurlijke talen?' *Nederlands Tijdschrift voor de psychologie* 43: 98-104.
- TIMBERLAKE, A.
(1980) 'Oblique control of Russian reflexivization.' In: C. Chvany and R. Brecht (eds.): *Morphosyntax in Slavic*. Chelsea (Michigan): Slavica Publishers.
- TRAVIS, L.
(1984) *Parameters and effects of word order variation*. Doctoral dissertation MIT.

The syntax of a language without grammar

- WHITE, L.
(1985) 'The pro-drop parameter in adult second-language acquisition.' *Language learning* 35: 47-62.
- WOODWARD, J.
(1978) 'Historical bases of American Sign Language.' In: P. Siple (ed.): *Understanding language through sign language research*. New York (etc.): Academic Press.
- ZAMENHOF, L.L.
(1887) [pseudonym: Esperanto]. *Meždunarodnoj jazyk*. Varšava: Kel'ter.
- (1963) *Fundamento de Esperanto*. Marmande: EFE. 11th edition.