ÍNDICE

1. METHODOLOGICAL UNDERPINNINGS

Language is difficult to define, but we may understand it as an intellectual activity that makes use of mental tools to produce speech throughout general structures. Ideally, Language needs to provide answers to: a) particular phenomena; b) learnability; and c) universality.

Linguistics is the science of Language where linguists attempt to provide theories of Language:

> it must be fully explicit and to be explicit, it must be formal - i.e. make use only of theoretical constructs which have definable formal properties. The use of formal apparatus (involving a certain amount of technical terminology) may seem confusing at first to the beginner, but as in any other serious field of enquiry (e.g. molecular biology) no real progress can be made unless we try to construct formal models of the phenomena we are studying. It would clearly be irrational to accept the use of formalism in one field of enquiry (e.g. molecular biology) while rejecting it in another (e.g. linguistics)». (Andrew Radford 1997:.99).

Noam Chomsky's put forward a particular theory of internal language, as the **initial state** of human language. The hypothesis of the initial state or **Universal Grammar** (UG) implies that the study of any particular external language (e.g. Chinese, Italian or English) is an instance of such a grammar.

For example, when studying the English language, the generative linguist is more interested in giving an account of general principles of UG apparent in that particular language. For instance when considering data as in (1), the general principle of UG that *anaphors corefer* with its *antecedent* is confirmed:

(1)   a.   John washes himself

     b.   Mary washes herself

Sentences in (1) are **grammatical** English sentences and are accepted by any native speaker of English. Sentences in (2), however, are **ungrammatical** or not accepted by native speakers. For convention a star (*) in front of the sentence indicates its ungrammaticality.

(2)   a.   *Herself washes Mary

       b.   *Himself washes John

When considering the ungrammatical sentences in (2), the linguist corroborates the general claim that any sentence has a complex hierarchical categorical constituent structure. The ungrammaticality of the sentences in (2) shows that *subjects* need to hierarchical dominate *object anaphors*. This can be seen as a **syntactic constraint** straightforwardly explained in structural terms.

In addition the linguist, when theorizing, cannot gather the infinite list of all possible sentences in a language. On the other hand, some **linguistic evidence** may be useful to analyse a particular phenomenon. The linguist needs to provide the data with explanatory adequacy, i.e. to explain why the native speaker accepts some sentences and rejects some others (e.g. 1 versus 2). To further provide support to the hypothesis, independent evidence should be found. The syntactic constraint mentioned before, for example, may imply that anaphors need to be coreferential with its **closest** nominal antecedent and so predicts the grammaticality contrast illustrated by the **minimal pair** in (3):

(3)   a.   John wishes that Mary looks at herself in the picture

       b.   *Mary wishes that John looks at herself in the picture

The ungrammaticality of the sentence (3b) indicates that the anaphor «herself» cannot corefer with their farthest nominal phrase «Mary», independently from the fact that it shares the same morphological features of gender (feminine) and number (+sing) with it. The sentence (3a) acts as the grammatical counterpart with respect to the sentence (3b). The minimal pair in 3 also serves to establish a difference between anaphors, on the one hand, and pronouns, on the other hand, given the fact that pronouns, in contrast with anaphors can corefer with their farthest antecedent, and in no way can they corefer with their closest antecedent, as illustrated by the minimal pair in (4). Recall that the star in front of a sentence means that the sentence is ungrammatical or not accepted by native speakers.

(4)   a.   *John wishes that Mary looks at her in the picture
             (her=Mary)

       b.   Mary wishes that John looks at her in the picture
             (her=Mary)

In effect, speakers do not accept that the pronoun «her» is coreferential with the nominal phrase «Mary» in (4a), while they accept that the pronoun «her» is coreferential with the farthest nominal phrase «Mary» in (4b), hence its grammaticality.

The amount of principles of UG considered with respect to particular linguistic evidence of a language constitute the grammar of that language. Notice, however, that not all principles of UG have to surface in all grammars. In the following section, we study how UG accounts for **language diversity** through the notion of **parameter setting**.

The linguistic research initiated by Noam Chomsky goes parallel with other **cognitive sciences**. He claims in favour of UG which produces a unique **computational system** of the human mind, generating a number of strings of words derived from certain morphological and syntactic structures. This account may also explain **language variation**. Languages do not differ with respect to the computational system, but with respect to specific **morphological properties** included in the **lexicon** of each particular language. The generative grammar abandons the traditional idea that language variation responds to different grammatical systems as the set of rules specifying each particular language. According to this alternative approach, we cannot speak of different systems of rules to explain apparent contradictory data found in the variation of languages. Rather, a set of specific **linguistic parameters** allowed by UG characterizes the particular grammar of a language.

What traditionally has been known as grammatical constructions turns to be taxonomic epiphenomena that respond to particular structures containing properties that respond to the interaction of **principles and parameters**. In this sense, we may speak of a **derivational component** equal for all languages in which a language L sets a system of grammar allowed by UG parameterised for that particular language. In this respect, Noam Chomsky points out:

> I will assume the familiar Extended Standard Theory (EST) framework, understood in the sense of the Principles and Parameters approach. We distinguish the lexicon from the computational system of the language, the syntax in a broad sense (including phonology). Assume that the syntax provides three fundamental levels of representation, each constituting an «interface» of the grammatical system with some other systems of the mind/brain: Deep structure, Phonetic Form (PF), and Logical Form (LF). (Chomsky: 1996:130)

Each lexical unit in the lexicon contains a consistent system of **features** according to some **phonetic and syntactic properties** that determine its

sound, meaning and syntactic roles through more general principles parameterised for each language. For example, the English verb «hit» is specified for its vowel quality, for the properties proper of an action verb, as well as for the requirement of taking an object since it is a transitive verb and accordingly forms part of verbal phrase (VP).

In this theory, we may only speak of a single internal language and languages differ from each other with respect to features parameterised in the lexicon. The task of **language acquisition** is reduced to acquisition of those particular features.

SUMMARY

At the end of this section, you should have a clear idea of the distinction between Universal Grammar and the grammar of a particular language. You should also have understood the concept of «grammaticality» and how to analyse linguistic phenomena through minimal pairs. Language diversity should also be understood in terms of parameters and universal principles.

NOW YOU ARE READY TO DO EXERCISES 1, 2, 3, 4 AND 5

2. COMPETENCE AND LANGUAGE ACQUISITION

According to generative grammar, **linguistic competence** is the set of rules that conform the speaker's knowledge of language, thanks to which it is possible to learn and produce language. The English speaker's competence explains the ability to render **grammaticality judgements** with respect to an unlimited number of English sentences. It also allows the generation of new strings of language according to English grammar.

> *La competencia (competence) se opone a la actuación (performance), definida ésta por el conjunto de las restricciones que se ejercen sobre la competencia para limitar su uso: la actuación da cuenta de las diversas utilizaciones de la lengua en los actos de habla. Se distingue una competencia universal, constituida por reglas innatas que sustentan las gramáticas de todas las lenguas, y una competencia particular, constituida por las reglas específicas de una lengua, aprendidas merced al entorno lingüístico.* (Dubois y otros:1994,competencia, p 119)

> Our ability to speak a language is based partly on the innate principles and parameters available in UG, partly on the triggering experience of exposure to a specific language. On the basis of these components we develop a grammar of one (or more) specific languages: the core grammar of such a language. Schematically we can represent the generative view of language acquisition as follows

> Triggering experience Language X $\longrightarrow$ UG (with parameters) $\longrightarrow$ Core grammar Language X «
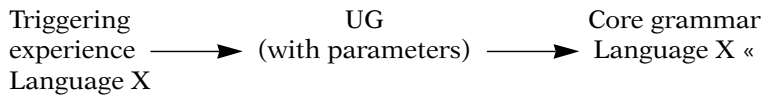
> (Haegeman 1994:16)

Turning back to the example «**John washes himself**», we showed that the anaphor «himself» necessarily refers to the nominal phrase «John». Now we argue that this is due to the fact that there is an internal rule of **binding** that applies to anaphors. The English child will not have to learn such a rule because it is part of her linguistic competence: what she only needs to learn is that «himself» belongs to the category of anaphors in her language.

> Furthermore, even if a homogeneous speech community existed, we would not expect its linguistic system to be a «pure case». Rather, all sorts of accidents of history would have contaminated the system, as in the properties of (roughly) Romance versus Germanic origin in the lexicon of English. The proper topic of inquiry, then, should be a theory of the initial state that abstracts from such accidents, no trivial matter. For working purposes (and nothing more than that), we may make a rough and tentative distinction between the core of a language and its periphery, where the core consists of what we tentatively assume to be pure instantiations of UG and the periphery consists of marked exceptions (irregular verbs, etc). Note that the periphery will also exhibit properties of UG, although less transparently. A reasonable approach would be to focus attention on the core system, putting aside phenomena that result from historical accident, dialect mixture, personal idiosyncrasies, and the like. As in any empirical inquiry, theory-internal considerations enter into the effort to pursue this course, and we expect further distinctions to be necessary. (Chomsky, 1996: 20)

2.1. Insufficient external stimuli

As we saw, knowledge of language may be seen as an internal ability the speaker has, although he is unaware of it. It is formally represented by UG with parameters for each specific language. The grammar of a language contains all the possibilities to generate unlimited possible sentences providing each derivation with their syntax, semantics and phonetics. The linguist's task is reduced to formalize the **theory** of that language.

The linguist can take specific external data to develop such a theory. However this might be insufficient. To offer a feasible characterization of any production of language, the linguist should also study the speaker's competence, by analysing the way the speaker himself assesses his grammar. For example, English speakers know unconsciously that the sentence in (5) is grammatical because they accept it, as opposed to the ungrammatical sentence in (6):

(5)   Pall looks at her because he loves her

(6)   *Pall looks at her because loves her

It is only the task of the linguist to explain the speaker's grammaticality judgements with respect to the sentences (5) and (6). At first sight, the linguist may postulate that the sentence (6) is ungrammatical because the embedded clause lacks a subject. However this is not a universal rule, because other languages (Spanish) allow *embedded sentences* without *overt subjects*, cf. (6) versus (7):

(7)   Pall la mira porque la quiere

The Spanish sentence in (7) is indeed grammatical and therefore we find the first difference between a language like English and a language like Spanish. The question is then whether the linguist needs to exclusively look at one particular language to provide a consistent theory of that language. From the generative perspective, the linguist needs to look at both UG and the particular grammar of a language. By doing so, the linguist shows how the principles of UG are fixed for each particular parameter. In the generative work, **English** has been characterised as part of the **obligatory subject parameter**, whereas **Spanish** belongs together with other languages (Italian, Portuguese, and many more) to the **null subject parameter**.

> Work in generative linguistics is therefore by definition comparative. Generative linguists often do not focus on individual languages at all: they will use any human language to determine the general principles of UG and the choices it allows. Data from a dialect spoken by only a couple of hundred people are just as important as data from a language spoken by millions of people. Both languages are human languages and are learnt in the same way. (Haegeman, 1992, p.18)

2.2. UG and parameters

In addition, the linguist needs to explain how each language is acquired. The null hypothesis is that the child living in an English community will

be exposed to sufficient data to fix her own obligatory subject parameter. In this way, grammar should offer **explanatory adequacy** for particular linguistic phenomena rather than being merely **descriptive**.

The problem of language acquisition cannot be solved by assuming pure imitation of stimuli, since the child will never be exposed to all possible sentences of her language. On the other hand, she needs to be exposed to those crucial examples, which will lead her to fix her own language parameter. The linguist's task is to identify those crucial examples, which actually **trigger** language acquisition. Let us suppose that the child hears the following sentences:

(8) a. I like the boy who I see in that picture

   b. I like the boy that I see in that picture

   c. I like the boy I see in that picture

According to the sentences in (8), the child may think that the *relative pronoun* «who» may freely alternate with «that» and may also be omitted in *relative clauses* with a human antecedent. However, this hypothesis would be incorrect since she will never hear sentences like the one in (9), because it is ungrammatical and adults never produce ungrammatical sentences:

(9) I like the boy speaks Japanese

The sentence in (9) is ungrammatical because in English the relative pronoun cannot be omitted with *subject relative clauses* regardless the human antecedent. It is difficult to see how the child acquires this knowledge without assuming that she does not start from scratch but from UG which allows her to make this difference: object relative clauses allow the three options that/who/- whereas subject relative clauses do not. Obviously, the adult does not have to explain this since he may not be aware of it either.

It is also tenable to assume that the child will only hear grammatical sentences and therefore will always be exposed to **positive evidence**. This goes against the idea that children learn from negative feedback or uncontradictory data. Rather, the child will always be exposed to grammatical sentences and from here she will fix her own parameter.

> We assume that the system described by UG is a real component of the mind/brain, put to use in the complex circumstances of ordinary life. The validity of this assumption is hardly in question. To reject it would be to assume either (1) that non-homogeneous (conflicting) data are required for language acquisition, or (2) that the mind/brain does indeed have the system described by UG, but it is not used in language

acquisition. Neither assumption is remotely plausible. Rejecting them, we accept the approach just outlined as a reasonable approach to the truth about humans, and a likely prerequisite to any serious inquiry into the complex and chaotic phenomenal world. (Chomsky, 1996:19)

Since the child does not learn from conflicting data, the problem of language acquisition is reduced to fixing parameters allowed by UG. Recall Haegeman's schema above, where UG acts as a **filter** between language experience and the actual grammar of a particular language. In fact UG was proposed by Chomsky to explain how children acquire language, which is a complex task, at such an early age and with such speed and efficiency. The term «universal» suggests that this theory is used to explain language acquisition in general (i.e., it's not language specific). It would take a lifetime to learn all the rules and different possible ways words can be used, so making language an inefficient way of communicating, therefore evolutionary would not have stood the test of time.

However, language is perhaps one of the things that make us human, and is indeed a universal phenomenon amongst humans, in its many forms. So, how does a child, by the age of four, have grammatical rules in place, which it could not have been possible to be learnt through Skinnerian type conditioning, in that time? This is where UG comes in; it is proposed that UG is an innate, unconscious ability present at birth, knowledge of grammar. This is not suggesting that a child does not make grammatical errors, as we all know, children do, but it seems that they only make irregular type errors, such as «he holded» instead of «he held», so somehow they have the ability to accept these rules and apply them. It can be argued, by people such as Skinner, that these rules are simply learnt through conditioning, but how can this be so? As the child never hears anyone make these mistakes, so does not learn them that way, this is the poverty of the stimulus theory i.e. there is not enough, or indeed any, of this kind of stimulus to learn from. With UG set in place at birth, the child is able to take on whichever language it is exposed to, as all languages have common elements and are inter- translatable.

SUMMARY

If we want to provide our theory of language with explanatory adequacy we need to assume UG. By adopting linguistic competence as the core grammar, language acquisition is the task to learn those marked options that characterise a particular parameter.

1. We cannot count with all possible linguistic data of a particular language to formalize a grammar. In addition, if we only pay attention to linguistic performance, we will sometimes obtain contradictory data caused from other factors such as slips of the tongue, insecurity, nervousness, etc.
2. Linguistic competence is unlimited and in principle we may learn any particular language. Third, the speaker may produce new sentences, never heard before, and be sure that such sentences are grammatical in the grammar of his language.

NOW YOU ARE READY TO DO EXERCISES 6, 7 and 8

3. LEXICAL AND FUNCTIONAL CATEGORIES, SUBCATEGORIZATION AND CONSTITUENT STRUCTURE

The **lexicon** constitutes the basis for all grammatical relations. In fact, this grammar level needs to satisfy «external» constraints of the **interface** relation with other levels of grammatical representation. The **constituent structure** and the so-called **subcategorization** give account of the organization of **syntactic categories** reflected in the structure of the semantic universe. In this chapter, we look at the main categories found in the English lexicon and describe them considering their structure and how they relate to each other.

The lexicon of English as of any language is composed of words and lexical units that define it in a particular way. Each word can be classified into categories such as *nouns*, *verbs*, *prepositions*, *adjectives*, and *adverbs*. Whereas we may find a limit set of categories, new words are freely incorporated as long as they are created. In order to classify each lexical unit we may take advantage of **features**. For example a word like «pen» is a lexical unit with the features [+N] and [+sing]. A word like «ugly» is another lexical unit with the exclusive feature [+Adj] and no number feature in this language.

Likewise, there are other elements, which although cannot be created as freely as we saw before, are part of constituent structure and are subject to subcategorization requirements. They are devoid of lexical content and since they exclusively play a grammatical role, they are interpreted as **functional categories**. For example, the conjunction «which» cannot