Language economy and the pro-drop theory from a minimalist perspective

Erika Corbara

1 Introduction

The general literature concerning the pro-drop phenomenon tries to explain the internal mechanisms and properties of the concerned languages focusing mainly on morphological features, which could allow or not such languages to drop their subject pronouns. Actually, under such an assumption, the methodic analysis of morphological data leads to the systematic formulation of language specific parameterizations in order to confirm already established and currently accepted internal structures: Proposals like the existence of the lexical unspecified and phonological null element pro among the pronominal array of pro-drop languages, whose originally unvalued $\phi$-features are assigned by agreement (Agr) in the course of the derivation, are indeed widespread (Rizzi 1982; Chomsky 1981), but none of them is constantly capable to explain – and be validated by – all the syntactic constructions of the different languages of the world.

Target of this research is thus the determination of cross-linguistic factors, which could be inferred responsible for the appearance of pro-drop phenomena worldwide, giving thereby to the pro-drop theory a scope of application valid for all the cases in which subject pronouns are dropped. In order to reach such an extended explanation and avoid interferences with already existing theories the whole matter should be reconsidered from its initial data, taking into account only the very core properties which are shared by all the languages of the world: the human faculty of language (FoL) and the connected children's innate predisposition for language acquisition. 1 A parallelism between language faculty and biology is thereby undeniable and sets the premises for the proceeding development of this investigation. On the basis of such and further evidences, supporting a direct relationship between language and biology it will be then attempted to remark and consider the main principles regulating physical systems in order to apply them to language and draw the due logical conclusions.

2 Language as a biological system

Starting with the importance of the human genetic endowment for processes like language acquisition and production, there are many other pieces of evidence connecting the language faculty to the biology of the human being. 2 The following array is just a list of the main relationships between the two disciplines, but several recent studies validate the same assumption a well:

- FOXP2 (the language gene)
- Language as organon (language grows)
- Broca's aphasia (or expressive aphasia)

It will be aimed, at first, to analyze and deepen the data concerning such issues in order to observe to which extent language and biology are connected to each other and, as second step, to consider if the subsequent application of specific physical principles on language can be retained justified.

2.1 FOXP2

Forkhead box P2 (also known as FOXP2) is a gene, whose mutation has been proven to cause an inherited impairment of language and speech (Lai 2001, p. 520). Affected people have deficits in production, articulation and comprehension in a variety of grammar domains, together with difficulties in producing some ofrotal movements (Christiansen 2009, p. 142).

It was first discovered in 2001, when a molecular investigation study determined a link between the modification of this gene and severe developmental disorders disrupting the speech and language skills of many members of the same family. 3 Such connection represents a case of clear inheritance for a language disorder, and only by studying the DNA of affected and unaffected individuals of

---

1 Also referred to as LAD (Language Acquisition Device). It was first proposed by Chomsky in the 1960s.

2 Particularly relevant to this topic is the work of Lenneberg Biologische Grundlagen der Sprache (1977).

3 In the medical literature it is commonly indicated as KE family, but the real identity of such family does not appear to be publicly known.
the examined family it was possible to track down the specific mutation, which finally turned out to be an alteration of sequence in the FOXP2 gene. This means, in other words, that linguistic impairments can be transmitted from one generation to the other through the inheritance of a mutated gene and represents, on the same time, an irrefutable proof of the biological foundations of language. Because of this, FOXP2 has been also dubbed the "language gene".

2.2 Language as organon

The developmental proceedings of the faculty of language (FoL) can be compared to the growth process of an organ in the biology: if certain skills are not stimulated within the right period, then it won’t be possible anymore to learn them in the proper way. Such time limit for the acquisition of specific physical abilities is also known as critical period (Lenneberg 1967) and it is closely linked to the synaptic plasticity of the brain i.e. organ during its growth: cells getting an external input are stimulated und thus supplied with blood and those remaining inactive step by step get no longer nourishment.

Parallel to the development of the FoL is the development of the visual system: in this field were carried out several studies in order to define the critical period of some animal species. Thus, it has been proven, for instance, that kittens are susceptible to visual stimulation only up to three months of age. After that, even if the visual input is restored, their sight does not function anymore in the normal way and even their directional sensitivity is permanent affected. About language, the critical period hypothesis proposed by Eric Lenneberg in 1967, according to which the language susceptibility phase ends with the puberty, finds its consolidation in the few recorded cases of infant language deprivation and feral children. The consideration of the FoL as a growing organ plays moreover an important role in many critics against the behaviourist model: children are actually particularly resistant against corrections and a repeated grammatical stimulus does not produce any response as long as the corresponding grammatical categories are not acquired. It is thus not possible to regulate the process of children language acquisition before certain grammatical structures get identified and recognized, i.e. before the FoL has grown (Grewendorf – Hamm – Sternewald 1987, p. 22).

2.3 Broca’s aphasia

Broca’s aphasia, also known as expressive aphasia, is characterized by the loss of the ability to produce language (spoken or written). People suffering from this form of aphasia exhibit the common problem of agrammatism: for them it is extremely difficult to start sentences and their speech is non-fluent and labored. Even writing results complicated, as well. Their language ability is thus reduced to disjointed words, and the sentence construction is poor cause of the omission of functional words and inflections (bound morphemes).

The grade of expressive aphasia varies from patient to patient but it is in each case characterized by a lesion of a specific region of the brain, the so-called Broca’s area. The fact that a physical damage in a very specific part of the brain leads to language impairments underlines the narrow nexus between FoL and biology: language functions and develops on biological bases.

2.4 General remarks

On the basis of the presented data it appears thus justified to regard the FoL as a biological system of the human being. In the same way of a physical organ, also the FoL is namely genetically inherited, grows and its functions can be attributed to the activity of a specific region of the human body. This leads to the subsequent logical assumption that linguistic processes originate in the human brain following the very same principles regulating physical systems as well: economy and optimality. Of course such considerations have to be applied only to the triggering mechanisms of language production and not to the usage which could be done of the FoL. In this view, it will be hence tried to analyze and explain the pro-drop languages on the basis of economy and optimality criteria.

4 Further information to this topic can be found in the following article by N. W. Daw and H. J. Wyatt (1976): Kittens reared in a unidirectional environment: evidence for a critical period.

5 The most famous cases of feral children who could not acquire language normally are Victor von Aveyron (1797) and Genie (1970). However, it is not excluded that these children were retarded from infancy and abandoned because of this, or that their inability to develop language was a consequence of the treatments they suffered. Isabelle (1938), found at the age of 6 – i.e. still within the critical period – was anyway able to reach at the age of 9 the normal development of a child of her age.

6 Language purposes are known to be extremely varied: human passions, poetry and social engagements do find through language their highest expressions and linguistic forms are therefore often pleonastic and redundant, i.e. not economic at all. This article is just intended to remark how syntactic operations take place in the human brain and which principles regulate such proceedings at their very initial stage.
3 Development of linguistic theories

The development of the generative linguistics was essentially triggered by a need of explanation and principally aims to answering the following two questions (Grewendorf 2002, p. 98–99):

- How is it possible that children perfectly acquire the grammar of their own language in such a short time, independently and often under bad conditions? (Plato’s problem)
- How could the variation of natural languages be explained?

On this purpose, in the modern linguistic theory there are two main models being mostly accepted and examined: the Government and Binding Theory (GB) and the Minimalist Program (MP). Both of them are based on the Chomskyan universal grammar (UG), which through its nativist assumptions provides an answer to Plato’s problem and at the same time it is flexible enough to allow the language variation as a natural parametric effect, but they do differ substantially in their modus operandi.

3.1 From GB to Minimalism

The GB-theory, based on UG, assumes that a large portion of the grammar of any particular language is common to all languages. In such concept can be distinguished four different levels of representation, as shown below in Fig. 1: D-structure, S-structure, PF and LF. From the lexical array words are combined together at an underlying stage called D-structure (deep structure); here, according to specific rules of each language (Move-α), they are moved and bound to each other to form syntactic sequences, which reflect the resulting surface order (S-structure) of each specific language.

```
PF ← S-structure → LF
           ↑ Move-α
D-structure

Lexicon
```

Fig. 1: The T-model of the GB-Theory (Grewendorf 2002, p. 107)

The S-structure itself cannot be interpreted, but it is factored into PF (phonological form), the interface with phonological features, and LF (logical form), the interface with semantic ones.

In contrast to the GB-theory the Minimalist Program (Chomsky: 1995), which is nothing but the logical development of the former GB-theory on the basis of economy principles, tries to eliminate the existing redundancy within the principle-and-parameter approach of GB and to reduce thus the operations to the conceptual necessary. In the minimalist architecture of language design (Fig. 2) the lexicon already contains all the features to be interpreted at the interfaces:

```
LF ← PF   (Spell-out)
         ↑ Derivation
Lexicon
```

Fig. 2: The minimalist model (Grewendorf 2002, p. 109)

With other words, every linguistic expression consists now of just a pair of representations $(\pi, \lambda)$, one for each representational level: $\pi$ for the PF and $\lambda$ for the LF, which are directly read at the interfaces. As a result, constraints can be motivated only by feature requirements directly at the interpretational levels PF and LF. The MP aspires thus to a maximal economy concerning the number and complexity of operations and rules: uneconomical or superfluous proceedings of the GB-theory, such as the intermediate stages D- and S-structure, are left off. Syntactic (i.e. cognitive) processes are so realized with the minimal effort.

Taking the MP as point of reference, it is thus possible to state as follows: both physical and grammatical procedures operate in the same way and are regulated by the same economy principles.
3.2 Ungrammaticality as interface effect

As the input arriving at the interfaces is not checked anymore through intermediate stages (D- and S-structure are as mentioned eliminated), the information has to be read directly at the interfaces; either are their conditions satisfied and the formulated expression is interpretable or not. Everything is thus shifted to the two interpretational levels PF and LF and it can be stated as follows:

7 This representation was originally created by Prof. Dr. Peter Kosta, to whom I owe my deepest and sincerest gratitude for all his explanations and support in writing this article. For a complete understanding of the graphic it has to be taken into consideration as follows: S-M: senso-motoric system (interpreting π, the phonological component of the spell-out) C-I: conceptual-intentional system (interpreting λ, the semantic component of the spell-out)

Erika Corbara

PERCEIVED UNGRAMMATICALITY → INTERFACE EFFECT

Let's now analyze the following sentences as examples of lacking interpretation, i.e. perceived ungrammaticality:

(1) * Kolja dal mne.
   Kolja_{nom} gave_{pref} to-me_{dat}
   'Kolja gave me.'

(2) * Ging sofort zur Schule
   went_{3sp, past} immediately to school
   'went immediately to school.'

Both in (1) and in (2) is missing a verbal complement, which is necessary for the valence theory. In the first case (1) it is the direct object of a trivalent verb (to give) and in the second case (2) it is the subject of the sentence, a verbal complement as well: both expressions, because of the lack of a syntactic constituent, are perceived as ungrammatical.

Let's now look at the following examples of grammaticality:

(3) Šel hned domů.
    went_{3sp, m} immediately home_{acc}
    'he went immediately home.'

(4) Andó subito a casa.
    went_{1sp} immediately to home
    'he/she went immediately home.'

Here some complements which are necessary for the valence theory are missing as well, but nevertheless both sentences are interpreted as absolute grammatical and even usual. In other words, in (3) and (4) it is possible to drop the subject, i.e. a verbal complement, without causing any ungrammaticality.

Moreover, comparing the examples in (2), (3) and (4) are the translation of the same sentence with the very same syntactic structure in three different languages: German (2), Czech (3) and Italian (4). What in German results ungrammatical because of the lack of the subject, for Czech and Italian people is completely normal and correct. From this evidence it is possible to draw the following remarks:
4 Pro-drop theory

The peculiarity of the sentences in (3) and (4) is typical of the pro-drop languages (from pronoun-dropping), whose subject personal pronouns, if pragmatically inferable, are generally not expressed.

The precise conditions for the omission vary from language to language, but it is possible to generalize saying that most of the Romance\(^8\) and Slavic languages are pro-drop. Here are some further examples:

(5) Lo conosco. (IT) 'I know him.'
(6) Tak mne skazali. (RU) 'I was told that.'
(7) Gdzie mogę wymieniać pieniędzy? (PL) 'Where can I change money?'
(8) Jmenuji se Erika. (CZ) 'My name is Erika.'

\(^8\) With the exception of French

4.1 The pro-drop parameter

In order to explain the pro-drop phenomenon and the consequent violation of the valence theory, it has been introduced the pro-drop parameter (Rizzi 1986) that has been introduced, which assumes that the apparently empty argument position is actually covered by a lexical unspecified element without phonological realization: pro.

(9) Pro-drop parameter:
   a. pro is governed by X\(^5\) (Rizzi 1986, p. 524)
   b. X\(^5\) is the head, licensing occurrence of pro: pro has the same
      grammatical properties of the features of X\(^5\) (Rizzi 1986, p. 520)
The idea is that in languages with sufficiently rich verbal \(g\)-feature (person, number, gender) agreement morphology, pronominal arguments can remain without phonological realization: who or what is \textit{pro} can be deduced from the verbal inflection. In this way it is found an explanation for the missing realization of a verbal argument (subject), whose presence is actually fundamental for the valence theory: the subject is present, but phonologically empty and receives its necessary information through feature movements from other syntactic categories.

### 4.2 Morphological uniformity

The relation between a rich verbal agreement morphology and the occurrence of \textit{pro} is not always confirmed empirically. The idea that just languages with an abundant inflectional system are able to license empty subjects was already relativized by Jaeggli and Safrí (1989, p. 29–44): they pointed out that it is not the richness of the verbal morphology to license \textit{pro}, but the morphological uniformity of the verbal paradigms. This points out that just languages as Spanish, Italian or Czech, whose verbal agreement morphology shows a one-to-one relationship between verbal agreement and person, can allow \textit{pro}.

In this way, the morphological uniformity clarifies the concept of the pro-drop parameter, at least for the European linguistic area. Such generalization, as it will be shown in the next paragraphs, is anyway not confirmed by the syntactic structure of various Asiatic languages.

### 4.3 Chinese

The pro-drop theory reveals actually to be quite arbitrary if we compare European languages with Asiatic ones. Chinese, for instance, is a strong isolating language i.e. words in most cases consist of just a single morpheme and are not changed by inflection.

The verbal conjugation is therefore absent: verbs are always in their infinitive form and grammatical features such as number, aspect or tense are expressed through separate particles or adverbs. The most striking fact about Chinese is that subject personal pronouns can be dropped as well, even though it is not possible to deduce them from the verbal morphology. Let’s now consider the following sentences:

(12) Xia yǔ le.
fall\(_{\text{INF}}\) rain le.

"It rains"

As in all isolating languages, in Chinese there is a Fixed syntactic position, which reflects the SVO word order: the subject is always at the beginning of the sentence, followed by verb and object. After verbs that don’t require any subject (12) or that just drop it (13 and 14), only the object position is covered and the first position remains thus empty. Should therefore Chinese be considered as a pro-drop language? According to the pro-drop definition in section 4 it is, but according to the pro-drop parameter in section 4.1 it could not be in any way, as it does not have any verbal morphological feature to license and identify empty subjects. Nevertheless Chinese does show the tendency to omit its subject pronouns.

### 4.4 Pro-drop worldwide

According to a worldwide overview from WALS\(^9\) it results that many other languages having neither subject nor object verbal congruence do regularly allow pro-drop patterns: from the 248 examined languages, 26 of them showed the tendency to omit their subject pronouns without exhibiting any verbal agreement.\(^{10}\)

Such languages are listed below:

<table>
<thead>
<tr>
<th>Language</th>
<th>Language Family</th>
<th>Country</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mongolian, Halti</td>
<td>Altiic</td>
<td>Mongolia</td>
<td>2,337,995</td>
</tr>
<tr>
<td>Khu</td>
<td>Austro-Asiatic</td>
<td>Laos</td>
<td>479,739</td>
</tr>
<tr>
<td>Malari</td>
<td>Austro-Asiatic</td>
<td>Thailand</td>
<td>324</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>Austro-Asiatic</td>
<td>Vietnam</td>
<td>674,399,139</td>
</tr>
<tr>
<td>Maori</td>
<td>Austronesian</td>
<td>New Zealand</td>
<td>60,000</td>
</tr>
<tr>
<td>Rapanui</td>
<td>Austronesian</td>
<td>Chile</td>
<td>3,392</td>
</tr>
<tr>
<td>Ipuna</td>
<td>Choco</td>
<td>Columbia</td>
<td>8,050</td>
</tr>
<tr>
<td>Yoruba</td>
<td>Niger-Congo</td>
<td>Nigeria</td>
<td>19,327,000</td>
</tr>
<tr>
<td>Lengzi</td>
<td>NortheastCaucasian</td>
<td>Russia</td>
<td>451,112</td>
</tr>
</tbody>
</table>

\(^9\) Even expletives (filler words) are absent in Chinese.

\(^{10}\) World Atlas of Language Structures.

\(^{11}\) The data refer to 2007. For further information http://n-true.livejournal.com/478571.html>.
Language economy and the pro-drop theory from a minimalist perspective

<table>
<thead>
<tr>
<th>Language</th>
<th>Language Family</th>
<th>Country</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guguyimdjir</td>
<td>Pama-Nyungan</td>
<td>Australia</td>
<td>25</td>
</tr>
<tr>
<td>Kayardild</td>
<td>Pama-Nyungan</td>
<td>Australia</td>
<td>6</td>
</tr>
<tr>
<td>Martuyhunira</td>
<td>Pama-Nyungan</td>
<td>Australia</td>
<td>5</td>
</tr>
<tr>
<td>Yidiny</td>
<td>Pama-Nyungan</td>
<td>Australia</td>
<td>12</td>
</tr>
<tr>
<td>Awtuw</td>
<td>Sepik-Ramu</td>
<td>Papua New Guinea</td>
<td>506</td>
</tr>
<tr>
<td>Burmese</td>
<td>Sino-Tibetan</td>
<td>Myanmar</td>
<td>32,301,581</td>
</tr>
<tr>
<td>Chinese, Mandarin</td>
<td>Sino-Tibetan</td>
<td>China</td>
<td>873,014,298</td>
</tr>
<tr>
<td>Garo</td>
<td>Sino-Tibetan</td>
<td>India</td>
<td>650,000</td>
</tr>
<tr>
<td>Kayah, Eastern</td>
<td>Sino-Tibetan</td>
<td>Myanmar</td>
<td>360,220</td>
</tr>
<tr>
<td>Ladakhi</td>
<td>Sino-Tibetan</td>
<td>India</td>
<td>114,000</td>
</tr>
<tr>
<td>Lahu</td>
<td>Sino-Tibetan</td>
<td>China</td>
<td>577,178</td>
</tr>
<tr>
<td>Meitei</td>
<td>Sino-Tibetan</td>
<td>India</td>
<td>1,261,000</td>
</tr>
<tr>
<td>Thai</td>
<td>Tai-Kadai</td>
<td>Thailand</td>
<td>20,229,987</td>
</tr>
<tr>
<td>Imonda</td>
<td>Trans-New Guinea</td>
<td>Papua New Guinea</td>
<td>250</td>
</tr>
<tr>
<td>Gilyak/Nivkh</td>
<td>isolated</td>
<td>Russia</td>
<td>1,089</td>
</tr>
<tr>
<td>Japanese</td>
<td>isolated</td>
<td>Japan</td>
<td>122,433,899</td>
</tr>
<tr>
<td>Korean</td>
<td>isolated</td>
<td>Korea</td>
<td>67,019,690</td>
</tr>
</tbody>
</table>

**Fig. 4: Pro-drop languages without verbal agreement**

It follows, that pro-drop without verbal agreement is not an exception of Chinese, but a relatively widespread phenomenon, which cannot be explained just with the former theory of the pro-drop parameter. Strong agreement features are thus a tendency for the occurrence of pro, at least in the European languages, but not an absolutely necessary condition for that.

### 4.5 Internal contradiction

Besides that, the existence of pro constitutes a conceptual contradiction: the acquisition of elements which are superfluous for the communication (pro) and the correlated feature movements (g-features) from a syntactic category to the other are, from a minimalist point of view, absolutely uneconomic. Furthermore, according to the general remarks in section 2.4, it is expected that linguistic processes generate on the basis of the very same principles regulating physical systems as well, i.e. economy and optimality.

### 5 Summary and outlook

In order to economize the process neither superfluous elements nor empty arguments should be interpreted at the interfaces. Everything, which is pronounced at the spell-out, already has to contain all the necessary information for the interpretation: this implies a different and much deeper consideration of those properties which are intrinsic of the lexicon.

According to the minimalist assumption that during the language acquisition process words are learnt and stored with all their internal language-specific features and mental implications, it is to deduce that such input does shape and regulate the subsequent development of syntactic structures. Such considerations about the role of the lexicon and its language-specific features are the basis to explain the pro-drop theory from a minimalist background: the explanation of the pro-drop phenomenon is not to be searched in some universal constraint of syntactic or morphological structures, but in the lexicon of the specific language i.e. in the mental implications that certain words can evoke.

All the necessary features for the interpretation are therefore already contained in the lexicon and the whole process can be seen as interface effect: this provides a clearer and more economic explanation of pro-drop patterns, which leads to a shift from syntax to semantics.

---


13 Cf. Corbara (2012: 49-50)
The research concerning pro-drop should hence focus on the properties of the mental lexicon, i.e. the verb semantics: verbs could incorporate their subjects (pro-drop language) or not (non pro-drop language).

With other words, certain grammatical features could be carried out also by different word categories which in the original language would never be expected to. For this reason, it is just necessary to observe and analyze the core properties of the respective target language, without being influenced a priori by parameters of the classical grammar theory or already established rules.

References


Erika Corbara


STARKE, M. (2002). The day syntax ate morphology. Class taught at the EGG summer school, Novi Sad.


Online references


Language economy and the pro-drop theory from a minimalist perspective

This paper aims at a new reconsideration of the pro-drop theory, focusing on the properties and requirements of language economy principles, whose satisfaction constitutes the central point of the Minimalist program (MP). Starting and fundamental approach of the research is the consideration of the language faculty as a biological system of the human being: The pro-drop languages (PDL) and their syntactic and semantic features will be thus investigated from a strongly minimalist point of view in order to find a clearer and cross-linguistic explanation of the pro-drop phenomenon.