The early processing of number agreement in the DP: evidence from the acquisition of Brazilian Portuguese*

Letícia Sicuro Corrêa, Marina Augusto & José Ferrari-Neto
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio/LAPAL*)

1. Introduction

The present study investigates the processing of number agreement in the acquisition of Brazilian Portuguese (BP). BP is a particularly interesting language for the verification of the extent to which young children are able to rely on the processing of syntactic agreement as a means of identifying the grammatical properties of the language under acquisition and the specificity of its morphology. BP exhibits two co-existing number systems: in the standard dialect (SD), plurality is morphologically marked by the inflection -s in the noun and in all elements that agree with it in the DP: determiners, possessives, and adjectives. The verb is also inflected for plural, expressing subject-verb agreement and subject-predicates agree with the subject as well. In non-standard dialects (NSDs), plurality is morphologically expressed in the determiner only (or the most leftward element in the DP) and sometimes in the verb. Children are exposed to both systems. It means that they have to be able to extract information concerning grammatical number exclusively on the basis of the morphological form of the determiner – marked or unmarked for number, when presented to the NSD. It is argued here that in order for children to do so, it is necessary that the DP is parsed as a phrase and that agreement between D and its complement NP is computed, on the basis of a model of agreement processing which incorporates a grammar characterized along the lines of the Minimalist Program (MP) (Chomsky, 1995; 1999).

In this paper, an experiment is reported which aimed at verifying the extent to which children would identify the plural morpheme in the determiner as information concerning the number of the DP and the extent to which they would be sensitive to the specificity of the language with regard to number marking. The working hypothesis underlying this study is that the computational system of the language operates as a language acquisition device, once its functioning is bootstrapped by the identification of the phonological and morphophonological distributional patterns (Corrêa, 2001; in press). Hence, if an inflection is identified in an element which is supposed to maintain an agreement relation with a noun due to their structural relation, the parser assumes that the value of the number feature of the noun is the same as the one signaled by the determiner’s morphology. Previous results concerning intrinsic gender in BP support this hypothesis (Name & Corrêa, 2002; Corrêa & Name, 2003; Augusto & Corrêa, 2005).

This paper comprises two parts. In Section 2, an exposition of the basic assumptions concerning the cognitive system of the language and the way syntactic computation can be incorporated into a parser is presented, followed by a discussion of the linguistic status of the number category and the characterization of the number systems of BP. In Section 3, a picture-identification experiment conducted with two year olds is described. Its rationale is presented in relation to a model of the processing of DPs as referential expressions which makes explicit how the processing of an

---

* This paper was presented as a poster at the 30th Boston University Conference on Language Development, 4-6 November 2005. The work reported here has been supported by the Brazilian National Research Council (CNPq) in the form of a grant to the first author and a graduate student scholarship to the third one, and by the Rio de Janeiro State Research Foundation (FAPERJ) in the form of post-doc scholarship to the second author. We would like to thank the children, their parents and the staff of the nursery schools (Creche M. Teresa Kling do Vale – Director: Alessandra D. do V. Pereira da Silva, and Creche Vila São José – Director: Kátia T.B. Beirão), where some of the tests were conducted as well as the comments obtained during the presentation of the poster.

* LAPAL: Laboratório de Psicolinguística e Aquisição da Linguagem (www.letras.puc-rio.br/lapal).
agreement relation within this unit would enable children to cope with the task. The results are discussed in relation to the working hypothesis. A brief conclusion is presented in the final section.

2. A minimalist approach to the acquisition of grammatical number from a psycholinguistic point of view

2.1. Grammar and parser in language acquisition

From the early studies on the relationship between the grammar and the parser, the way grammatical information can be incorporated in a processing model has been difficult to characterize (Fodor, Bever & Garrett, 1974). The conception of grammatical derivation in the MP facilitates, however, the establishment of this relationship is so far as the actions of the parser and of the syntactic formulator in processing models can be conceived of as the computation of syntactic relations upon the formal features of pre-selected lexical items, as in a minimalist derivation (Corrêa, 2002; 2005; Corrêa & Augusto, 2005; Fong, 2004; Phillips, 2004).

In the MP, language includes a universal computational system (C_{HL}) and a lexicon upon which this system operates. The lexicon is constituted of open class and closed class elements, that is, matrices of phonological, semantic and formal features (originally semantic or word-order properties that have been taken to be grammatically relevant in the language). The latter are the only lexical features that are accessible to the computational system, which is endowed with the following operations: Merge, which combines two syntactic objects recursively; Agree, which matches semantically interpretable and non-interpretable formal features in a c-command configuration1 resulting from Merge, and values the non-interpretable feature of the pair according to the value of the interpretable one, with possible consequences in the inflectional morphology; and Move or Copy, which accounts for displacement of constituents and linearization. The derivation of a linguistic expression is started once an array of lexical elements is displayed and two of them are selected to be merged on the basis of the argument structure of predicates and the properties of functional heads. The result of this syntactic computation is spelled out into two distinct but corresponding forms that are taken to be interface levels between the language system and the cognitive systems that are involved in language processing – a phonetic form (PF), containing information derived from the phonological features of the lexical items that will be legible by perceptual and articulatory systems, and a logical form (LF), containing a semantic representation that is legible by conceptual and intentional systems.

The syntactic computation can be assumed to take place during the syntactic formulation of an utterance and during its parsing. It means that from the point of view of the parser, once a DP is analyzed as such, agreement between the elements that share a formal feature of the same dimension in this configuration is automatically computed. For instance, given two elements sharing a formal feature with number dimension, their matching through the operation Agree results in the valuation of the element with the non-interpretable feature on the basis of the value of the element with the interpretable one (plural, for instance).

As for language acquisition, it can be assumed, according to the minimalist framework, that the computational system is available at the initial state of the acquisition of the language (either ready for use or depending on maturation). The task of the child is to acquire the lexicon of the language. In particular, his/her task is to identify which are the formal features of the language and the specific values of the parameters pertaining to them.

Language acquisition takes place on the basis of children’s processing the linguistic input. Assuming that processing units can be delimited on the basis of prosodic and distributional

1 A c-commands B iff: A ≠ B and the first branching node that dominates A also dominates B and neither A nor B dominates each other (Reinhart, 1976).
information (Morgan & Demuth, 1996; Christophe, 2002), that children are sensitive to functional elements (Shady, 1996; Hohle & Weissenborn, 2000), and that the computational system is bootstrapped by the identification of these elements, it is argued that children identify information concerning the properties of the formal features of the language by taking into account morphophonological variation in the elements within closed classes, whenever applicable, by automatically computing agreement once a hierarchical structure is parsed (Corrêa, 2001).

Number is a semantic feature that is taken as a formal feature in most known languages (though there are exceptions, e.g. Pirahã – a language of the Amazon region that seems to lack grammatical number (Everett, 1986; Corbett, 2000)). This formal feature may assume different values in a single language (e.g. singular and plural), and these values can vary across languages (singular, dual, trial, quadril, paucal and plural (cf. Corbett, 2000)). The expression of number agreement also differs considerably across languages. There are languages that exhibit an independent number morpheme (e.g. Yapese (Dryer, 1989)), there are languages that exhibit different number morphemes in the noun only (English), in the determiner only (French, BP in the non-standard dialect), in both determiner and noun as well as in other elements that maintain a c-command relation with the controller of agreement (European Portuguese, BP in the standard dialect, Spanish, Italian). Moreover, number morphemes may include suffixes (English, Portuguese, Spanish, Italian), prefixes (Swahili, Isangu) or even infixes (Sudanese) (see Comrie, 1999; Corbett, 2000; Robins, 1970). Thus, children acquiring languages in which number is a formal feature must recognize its morphological expression and identify the values that this feature assumes in the language on the basis of interface information. Morphological variation within closed class elements in the target language can be particularly informative in this regard, since it would result from the valuation of features by means of the operation Agree in the syntactic computation of the DP underlying the formulation of the linguistic stimuli children have at their disposal.

There is evidence that children are sensitive to determiners and that they parse DPs by the age of 12-14 month (Hohle & Weissenborn, 2000; Höhle, 2005; Name & Corrêa, 2003a, b; Shi, 2005). There is also evidence that children are sensitive to the morphological marking of gender in the determiner and that they ascribe the gender identified in the determiner to the noun in which gender is an intrinsic feature (Name & Corrêa, 2002; Corrêa & Name, 2003; Jonhson & Diks, 2005). It is possible then that the presumption of agreement between determiner and noun once a DP is parsed contributes to the setting of parameters pertaining to the number system of the language. Unlike intrinsic gender, however, grammatical number corresponds to conceptual number. Consequently, children have not only to establish an agreement relation within the DP on the basis of morphophonological information but also assume that the number feature is semantically interpretable. It means that children must acquire semantic information concerning number by processing the DP as a referential expression.

2.2. The number system in Brazilian Portuguese and the functional category Num

In BP, number is predominantly an optional formal feature, displaying two values: singular (unmarked) and plural (marked by the affix –s). The morphological expression of number agreement varies as a function of social and regional dialects as well as a function of the degree of formality in the speech of a single speaker. It is, therefore, liable to say that children are exposed to two number systems. Taking into account that there are singular nouns in the language that end in /s/, children are crucially dependent on the number inflection in the determiner, which results from agreement, in order to semantically interpret the number information relative to the DP.

Table 1 summarizes the distinctions between the expression of number agreement in the DP of both dialects of BP.
Table 1: Expression of number agreement in the DP in two variants of BP

<table>
<thead>
<tr>
<th>Standard</th>
<th>Os</th>
<th>meu</th>
<th>carros</th>
<th>novos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-standard</td>
<td>Os</td>
<td>meu</td>
<td>carro</td>
<td>novo</td>
</tr>
<tr>
<td>The (pl)</td>
<td>my(pl)</td>
<td>car(pl)</td>
<td>novo(pl)</td>
<td></td>
</tr>
</tbody>
</table>

Although number agreement in the DP is expressed by means of a plural affix attached to the
determiner and other possible elements c-commanded by D, linguistic analyses of the structural
relationships pertaining to grammatical number have led to the proposal of an independent
functional category NumP, which is located between NP and DP (Picallo, 1991; Bernstein, 1991).
The presence of such functional category receives support from morphological, syntactic, semantic
and interface considerations. As already mentioned, the presence of an independent number
morpheme in some languages (e.g.: Yapese and Tagalog) is taken as evidence for the independent
number projection (Dryer, 1989). In syntactic terms, different word order patterns between
adjective and noun in the DP has given rise to the idea that an intermediate projection is to be
assumed between D and N, and that this projection may possibly be NumP (Cinque, 1994). In
semantic terms, the presence of a NumP projection has been related to the distinction between
massive and countable nouns (Vangsnes, 2001; Borer, 2005). From a formal point of view, it has
been argued that functional projections must be semantically motivated. Number, on a par to Tense,
fulfills such requirement. Thus, in more general terms, it could be assumed that formal features
which bear some semantic import related to reference project independent functional categories. It
would allow for a distinction to be established between intrinsic and optional features during the
processing of agreement in so far as the latter would project independent functional categories.
As far as acquisition matters are concerned, it can be assumed that children are predisposed to
represent as a functional element heading its own projection those closed class items to which some
semantic interpretation can be ascribed. It follows that in the processing of the linguistic stimuli
children’s parser would create a functional projection as soon as they relate closed class elements
detected in PF with semantic information pertaining to a formal feature.

3. Solving reference on the basis of number agreement: parsing the DP

3.1. The rationale of the experiment

From a psycholinguistic point of view, children acquiring BP must segment elements from the
lexicon in the input phonetic string. They must parse a DP within a processing window by
identifying a functional element belonging to a D class and a lexical element belonging to an N
class, on the basis of phonetic and distributional regularities. It is assumed here that once the
computational system has been bootstrapped, children are capable of keeping track of the elements
from the input by assuming that their linear order reflects a hierarchical structure resulting from
Merge (according to the Linear Correspondence Axiom, Kayne, 1994). Thus, given a sequence D-
inflection N, a branching root node is generated giving rise to a sisterhood relationship to be
established. The recognition of systematic variation within the closed class affixes in D enables the
identification of the number morpheme which is taken as an independent functional feature of the
lexicon (Num). The c-command relationship established between D, Num, and N guarantees that an
agreement relationship is computed among the elements of the DP.
The identification of elements in the open class N is highly dependent upon the delimitation of a DP on the basis of the position of closed class elements (Höhle, Weissenborn, Kiefer, Schulz & Schmitz, 2002). Therefore, the acquisition of novel nouns may depend on morphophonological information concerning number and gender provided by closed class elements. It is on the basis of this possibility that the rationale of the experiment to be described here was conceived of. Given a novel noun and tokens of three invented object types that could be named by it – one single token of two object types and a multiple token display of another object type --, the task of the child is to identify the referent of a plural DP on the basis of the only relevant semantic information available, namely, the information provided by the number morpheme. In BP, this information can be either in N and in D, or only in D.

It is considered that once the number information that is identified in the number morpheme is ascribed to the DP as a whole, agreement has been computed within the DP. Consequently, it can be argued that the processing of agreement relations within the DP not only enables children to set parameters pertaining to properties of the number functional category in the language but also facilitates the acquisition of novel nouns.

In the following chart, the steps in the parsing of a DP in the accomplishment of the experimental task are considered.

```
PF: [^[1^u^dabu(∫)]] = Os dabo(s) (The(pl) novel noun(pl))

Speech segmentation
in prosodic units

Lexemes recognition

/PF[\oslash]\oslash\dabo(s)/

Parsing 1: morphological/
word order analysis and
access to lemmas

D Num N
  Mascal Pl ?
```
If children are able to identify the display of multiple tokens of a given type of unknown object on the basis of the Plural DP in both the SD and NSD variants of BP, they demonstrate that number has been recognized as a formal feature, the number morpheme has been semantically interpreted and number agreement in the DP has played a role on the acquisition of novel nouns.

3.2. The processing of number agreement: an experiment with Brazilian 2 year-olds

The present experiment was intended to verify the extent to which the information provided by the number morpheme –s is taken into account by young children in identifying the referent of a DP, containing a pseudo-noun, among invented objects in a picture-identification task. Children were requested to show a picture to a puppet called Dedê ([de´dε]). The dependent variable was then the number of responses to the multiple token picture.

The independent variable initially considered was The morphological expression of number in the DP: grammatical and ungrammatical.

In the grammatical condition, plural DPs in both the SD and NSD variants of BP were presented. Hence, the –s plural morpheme appears as a suffix of both D and N (SD), and as a suffix of D only (NSD). In the ungrammatical condition, possible morphological expressions of number in human languages were explored, namely, plural marking exclusively in a noun suffix, as in English, and plural marking as a noun infix, as in Sudanese.

In Portuguese, the phoneme /s/ is a possible syllable coda. It means that there are singular nouns with an /s/ ending, as lápis (pencil) and pires (saucer) and singular nouns with an /s/ at the end of a non-final syllable, as in words such as mosca (fly) and casca (peel). An /s/ ending in the noun can be mistaken as a plural morpheme in the language if the number information in D is not taken into account. By the same token, an /s/ ending in a word initial syllable can also be mistaken as a plural infix — there is a single case in Portuguese (the plural form of the indefinite pronoun qualquer (whatever) as quaisquer) --, or as the plural suffix of a monosyllabic noun, if the information in D is not taken into account. Being so, if children distinguish grammatical from ungrammatical expression of number in the language under acquisition, it is expected that they will rely more on the morphophonological information in D than in the noun. Hence, the condition grammatical plural is expected to present a higher number of multiple token responses.

The type of grammatical plural DPs was taken as an independent variable in order to verify whether children prefer one or another expression of plurality in the language. If children prefer redundant information, more multiple token responses would be expected in the standard condition. If, on the other hand, children crucially rely on the information provided by D, no difference would be expected between the two variants they are exposed to. Ungrammatical plural DPs, which are possible singular DPs in BP, were also contrasted with singular DPs without an /s/ at a syllable ending in order to verify the extent to which the presence of /s/ in critical positions would affect children’s performance. If children rely exclusively on the information provided by D in order to distinguish possible plural and singular nouns, a similar number of multiple token responses would be expected in these two conditions.

Table 2 displays the types of stimuli presented to children in each condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Grammatical plural</th>
<th>Ungrammatical plural</th>
<th>Singular control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>Mostra os dabos pro Dedé.</td>
<td></td>
<td>Mostra o dabo pro Dedé</td>
</tr>
<tr>
<td></td>
<td>Show the dabos (pl) to Dedé</td>
<td></td>
<td>Show the dabo (sg) to Dedé</td>
</tr>
<tr>
<td>Non-standard</td>
<td>Mostra os dabo pro Dedé.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Show the dabo (pl) to Dedé</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo plural</td>
<td>Mostra o dabos pro Dedé.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-s suffix</td>
<td>Show the dabos (sg) to Dedé</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo plural</td>
<td>Mostra o dasko pro Dedé.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-s infix</td>
<td>Show the dasko (sg) to Dedé</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Method:**

Participants: A total of 18 middle-class children from 18 to 30 months (mean age – 25 months), living in Petrópolis, a mountain town nearby Rio de Janeiro, took part in the experiment. All children were acquiring BP as their first language and had no systematic contact with a foreign language. None of them was reported to have family history of language deficits nor was affected by factors that might impair language development.
Material:
The material consisted of four lists of sentences recorded in a CD, a picture-book and a puppet. The lists of sentences contained 12 test stimuli, three for each experimental condition; 6 control sentences and 6 fillers with nouns from the vocabulary of the children. Gender and animacy were balanced. The order of presentation of the trials in each list was randomized though care was taken to prevent consecutive test conditions from occurring.

The picture book had 26 pages, two of them presenting two very common pictures for children to identify in the pre-test phase and the remaining 24 presenting four pictures each – the target-picture, in which multiple tokens of a novel object/creature were displayed, and 3 distracters, containing two singletons of novel objects/creatures and one familiar object. An example of a picture-book page is provided below:

![](image)

The apparatus consisted of a Sony Compact Disc Player, for the auditory presentation of the linguistic stimuli, and a JVC Digital Video Camera for the recording of the session.

Procedure:
The experiment was conducted in a quite room of a nursery school or in the child’s home environment. If required, a person familiar to the child remained in the environment without interfering in the task. The whole session takes about 10 minutes and is video recorded. The steps of the procedure are described below:

Warming-up: the experimenter interacts with the child and introduces the puppet “that can speak” (Dedé). Some recorded greetings are presented (Oi, tudo bem? – Hi, is everything all right? – Que lugar legal! – What a nice place! – Vamos brincar? – Let’s play a game? Do you want to play a game?). If the child shows interest, the book is presented and the pre-test session begins.

Pre-Test: The two initial pages of the picture-book constitute the pre-test sentences. If children respond accordingly and keep interested in the task, the test phase is initialized.

Test-Phase: The game goes on with Dedé asking children to point to some picture in the picture-book. Children’s responses are positively commented upon, regardless of their correctness.

Results and discussion:
The multiple-token picture responses obtained in the grammatical and in the ungrammatical conditions were submitted to a t-student test. A significant effect was obtained, with more multiple-token responses to grammatical plural DPs (t(17) = 5.65 p <.0001). It suggests that children take into account the information provided by D as far as number inflection is concerned. It implies that children are already tuned in to the grammatical system of the language (BP). Graph 1 presents the distribution of these responses.
In a second analysis, the Type of DP in the grammatical condition was considered. There was no significant difference between standard and non-standard plural DPs ($t(df_{17}) = 1.16 \ p=0.26$). It means that children deal with the two co-existing systems in a similar fashion, making it clearer that the number information provided by D is relied upon when it is the only means of enabling them to identify the referent in the task. Graph 2 presents the distribution of these responses.

No significant difference was obtained between the suffix and infix types of DP in the ungrammatical condition ($t(df_{17}) = 0.97 \ p=0.34$). Nevertheless, the frequency of the multiple-token responses to the singular condition was lower than to the ungrammatical suffix and infix conditions ($t(df_{17}) = 2.97 \ p <0.01$), suggesting that there was some interference of the presence of the phoneme /s/ in the pseudo-noun. Graph 3 presents the distribution of multiple-token responses in the two types of ungrammatical condition and in the singular control condition.
In sum, the effects obtained suggest that 2 year-olds are sensitive to the morphological expression of number in BP, that the crucial information concerning number is extracted from D since they do not distinguish the two co-existing systems as far as the extraction of number information is concerned. The results also show that 2 year olds are still sensitive to the presence of a phoneme /s/ in contexts in which it may function as a number morpheme. It appears, therefore, that children analyze the DP as a unit, establish agreement relations between its constituents and assume that a formal feature can be morphologically expressed in whatever element of an agreement relation. The results are, thus, compatible with a model of language acquisition in which the operations of a language independent computational system function as a language acquisition device as soon as this system is bootstrapped by information in the phonetic interface.

4 Conclusion

This paper aimed at providing experimental evidence in support to a theory of language acquisition in which the operation of agreement plays a role in the identification of language specific formal features and morphological patterns. The computation of agreement in early language acquisition is, however, a matter of controversy. Research on spontaneous production data causes into question the availability of functional categories or suggests the underspecification of uninterpretable features (Hyams, 1996; Müller, 1994; Radford, 1990). The present results add evidence in favor of the view that children are able to rely on grammatical information provided by elements of a functional category, as data concerning the early parsing of a DP and the processing of gender agreement in the DP suggest (Hohle & Weissenborn, 2000; Name & Correa, 2003). The presumption that a universal computational system operates as an instrument of the acquisition of a particular language is compatible with the view that this system is ready to be implemented once it is bootstrapped by information in the phonetic interface. The fundamental role of agreement in the process of language acquisition is also compatible with a minimalist perspective in so far as it is not necessary to ascribe to the initial state of language acquisition more than a universal computational system and a disposition for the construction of a lexicon with formal features defined in terms of parameters to be fixed with interface information, provided that the principle underlying the operation of language is one that guarantees visibility of semantic/syntactic relations at the interfaces.
References


