Agreement and markedness in the ascription of gender to novel animate nouns by children acquiring Portuguese

LETÍCIA M. SICURO CORRÊA
MARINA R. A. AUGUSTO
ANA CASTRO

Abstract

This paper focuses on the role of agreement in the ascription of gender to animate nouns by children acquiring Portuguese. An elicited production task was used in which children were requested to refer to novel objects named by pseudo masculine/feminine nouns. It aimed at verifying the extent to which an agreeing element (the determiner), the noun-ending or a correlation between the gender of the determiner and the noun-ending would guide the ascription of a pseudo-noun to a masculine/feminine gender class. This study extends an earlier one, in which 2-4 year olds acquiring Brazilian Portuguese were shown to rely more on agreement than on correlational patterns, when ascribing gender to pseudo inanimate nouns (Corrêa & Name, 2003). 80 2-4 year olds acquiring Brazilian and European Portuguese were tested. The results suggest that reliance on agreement prevails, though children are sensitive to correlational patterns and the production of feminine DPs (determiner phrases) is particularly hard. It is argued that children rely on an algorithmic procedure for gender identification and that gender markedness in nouns with an optional gender feature increases the demands of DP production.

1. Introduction

This study focuses on the acquisition of gender from the point of view of a theory of language acquisition which explores the use of syntactic computation in the identification of the properties of formal features of the language by children at an early stage (Corrêa, 2009). The acquisition of gender has usually been taken as a peripheral aspect of grammar and, as such, regarded as a late developmental process heavily dependent on strategic cue-

-based learning (Karmiloff-Smith, 1979; Pérez-Pereira, 1991; Clahsen & Almazan, 1998; Rodina & Westergaard, 2008). This sort of theoretical account would not, however, explain the early production of gender agreement by children (Levy, 1983, 1988; Pinker, 1995; Name, 2002; van der Velde, 2004). The aim of this paper is to provide an algorithmic procedure based on the processing of agreement within the DP (Determiner Phrase), in order to explain both children's early identification of gender and long lasting effects that suggest their relying on correlational patterns in the ascription of gender to novel words.

The acquisition of grammatical gender involves the identification of the gender classes in the language, the ascription of nouns to those classes, and the identification of the gender agreement relations that are reflected in the morphology. Languages vary in the number and in the kind of gender classes they present, as they vary with regard to the sort of agreement relations that are morphologically expressed (Corbett, 1991; van Berkum, 1997; Comrie, 1999). Adults learning a second language have great difficulty in the assignment of nouns to gender classes, given that class membership is usually perceived as highly arbitrary (Harley, 1979; Scherag, Demuch, Roesler, Neville & Roeder, 2004; Unsworth, 2008; Arnon & Ramscar, 2009). Children, on the contrary, do not have difficulty in acquiring gender as attested by a number of studies on spontaneous language production across different languages, and possible semantic associations between conceptual and gender classes do not seem to guide the acquisition process (Levy, 1983, 1988; Pinker, 1995). By the age of three, gender classes have been distinguished, regardless of the complexity of the gender system of the language, as suggested by the acquisition of a number of languages (Levy, 1988) and, in particular, by Bantu languages in which there are more than three gender classes and gender agreement relations are expressed both in adjectives and in verbs with distinct prefixes for singular and plural (Comrie, 1999). In Sesotho, for instance, the noun class prefixes beginning with an unmarked (coronal) consonant may be optionally realized as null when followed by some form of agreement. Despite the variable input children receive and the agreement relations involved, by 2-3 years of age, even null prefixes have been identified (Demuth & Ellis, 2008).

Bilingual children do not have particular difficulty in acquiring gender, either. Cross-language studies with French-German and with Dutch-English bilinguals have shown that each language gender system develops independently and that the process is similar to monolingual children development, with gender agreement being attested by the end of the second year of life (de Houwer, 1990; Müller, 2000; Unsworth, 2007).

The acquisition of gender, though giving rise to a number of descriptive developmental accounts, has not been taken as a major issue in linguistically oriented theories of language acquisition since it appears to be out of core grammar, due to its idiosyncratic character. Taken as a peripheral aspect of

grammar, the acquisition of gender has not been object of a principled account. The idiosyncratic membership of nouns into gender classes has, in fact, been explored in the investigation of cue-based language learning. A number of experimental studies seem to support the view that gender acquisition is a long lasting learning process and that children rely on either morphophonological cues or frequency in the identification of the patterns that would characterize the form of the nouns in each gender class (Karmiloff--Smith, 1979; Pérez-Pereira, 1991; Rodina & Westergaard, 2008).

In Romance languages, Karmilloff-Smith´s pioneering study presents a series of experiments with French-speaking 3 to 12 year-olds exploring their use of morphophonological correlation between the gender of a noun and its ending (Karmiloff-Smith, 1979). Imaginary animate/inanimate objects named by pseudo-nouns were presented to children, who had to refer back to them. Their responses suggested reliance on a morphophonological strategy up until around the age of nine and the mismatch of phonological/morphological, syntactic and natural class cues to gender increased the difficulty of the task. Pérez-Pereira (1991) used a similar procedure in an investigation with Spanish-speaking 4 to 11 year-old children and found similar results.

More recently, a study conducted in Russian has favored the view that cue-based strategies dependent on frequency would be necessary for the acquisition of gender, when gender/noun-ending correlations cannot be observed (Rodina & Westergaard, 2008). In an elicitation experiment, male kinship terms in –a (*papa* 'daddy'), double gender nouns (*plaksa* 'cry-baby'), hybrids (*vrač* 'doctor'), and female names in –ok/–ik (*Lenok*), showing gender-agreement mismatch in the DP, were presented to children from 2;6 to 4;0 and from 5;1 to 6;5. The results show that younger children are less accurate with hybrids and female names, but the situation improves with age. Assuming Westergaard's (2006, 2008) extended cue-based model, the discrepancy was attributed to factors such as input frequency and consistency/saliency of the input information where the cue is expressed.

A long lasting reliance on strategic cue-based procedures captured in experimental tasks would be compatible with the observation of over-generalizations, sporadic errors and self-repairs suggesting attention to possible correlation between gender and semantic/morphophonological patterns, which occur early in acquisition and may persist till mid and late childhood (Figueira, 2003, for Portuguese; van der Velde, 2004, for Dutch). There is not, however, productivity enough to suggest that this sort of episodic production reflects a cue-based learning procedure, i.e., that children rely on noun/gender correlation patterns in order to identify the gender of nouns, as a sort of grammatical generalization that would be postponed to mid-childhood (cf. Tomasello, 1992).

There are, therefore, conflicting data stemming from early spontaneous language production and experimental results obtained with children older than three. Corrêa & Name (2003), however, explored, on an experimental

basis, the presumption of agreement within the DP by children younger and older than three (mean ages: 2;7 and 4;6 years of age) acquiring Brazilian Portuguese, as the clue to an early ability of ascribing novel words to gender classes. In this language, as in other Romance languages, there are two gender classes (masculine/feminine) and, for historical reasons, there is some correlation between gender and noun-endings: *-o*, masculine/*-a*, feminine (a detailed description of the Portuguese gender system is provided in the next section). In the Corrêa & Name study, this correlation was manipulated in DPs with inanimate pseudo-nouns whereby invented objects were presented to children. The following conditions were contrasted: positively correlated (the vowel-ending of the noun being the same as the one of the determiner: *-o* for masculine and *–a* for feminine); negatively correlated (the vowel-ending of the noun being opposite to the one of the determiner: *-a* for masculine and *–o* for feminine) and neutral (nouns with the vowel-ending *-e*, which cannot be correlated with gender). An elicited production task was used, following the traditional paradigm.

A main effect of gender/noun-ending correlation was obtained with a greater number of errors in the negatively correlated condition. This effect was, nevertheless, due to the performance of the oldest group, since no significant effect was obtained when the data of the youngest children were analyzed separately. It was argued that as children's vocabulary increases, they become more sensitive to gender/noun-ending correlations in the language. Gender/noun-ending correlations cannot, however, explain the process whereby gender is acquired, since the youngest children appeared to rely in the gender information provided by the determiner in the DPs whereby the invented objects were presented to them. In sum, in spite of the effect of gender/noun-ending correlation, the majority of children's responses stemmed from their reliance on the gender of the determiner in the ascription of the value of the gender to a novel noun. That is, children apparently departed from the presumption of agreement in the DP and ascribed to the novel noun the gender value identified in the determiner, a closed class element in which morphological variations can be particular visible in the process of grammar identification.

The present study extends the experimental assessment of gender acquisition by children acquiring Portuguese by using animate nouns and two varieties of this language – Brazilian Portuguese (BP) and European Portuguese (EP). Animate nouns seem to impose different demands to both gender acquisition and the production of referring expressions. The gender of animate nouns can be either intrinsic, as in inanimate nouns, or optional, when there is a single nominal base for masculine and feminine pairs, being the latter morphologically marked (*gato*/*gat-a* [male cat/female cat]). In this case, grammatical gender expresses the conceptual distinction between male and female. Consequently, given an animate noun with *–a* ending, this ending can be perceived either as a thematic vowel, as in inanimate nouns, or as a gender

morpheme to be semantically interpreted. The production of a feminine DP as a referring expression may, therefore, involve a more complex conceptual planning (distinguishing the subclass of female individuals), a particular syntactic structuring of the DP (as it is argued in section 3) and more demanding morphophonological encoding than the unmarked masculine form. As for the use of the European variety of Portuguese, it may enable the conclusions so far obtained with BP to be generalized across varieties, since the differences between them (final vowel reduction in EP, for instance), which might affect a cue-based learning process, would not be relevant to a process of gender acquisition that is based on agreement.

This paper is organized as follows: in the next section, the Portuguese gender system is characterized. In section 3, the differential representation of intrinsic and optional gender in the lexicon is related to a formal distinction between gender as an intrinsic feature of the noun and gender as a feature which heads an independent GenP functional projection. In section 4, the elicited production experiment is reported and the results are discussed in relation to the formal distinction between intrinsic and optional gender, and markedness. In section 5, a procedural characterization of gender acquisition is proposed, which can account for the early acquisition of gender and for children's sensitivity to correlational patterns, which can give rise to occasional production errors. The last section brings the conclusion.

2. The Portuguese gender system

Gender can be characterized as a system of classification in which nouns are assigned to classes and this assignment is reflected in the forms taken by elements syntactically related with them (Comrie, 1999). In more formal terms, in the light of a minimalist conception of language (Chomsky, 1995, 2007), the lexicon is composed of matrices of features, which can be semantic, phonological and formal. The latter, though possibly having a conceptual basis, are distinguished for their role in syntactic computation. A universal computational system is assumed to operate on formal features. Lexical elements are concatenated and those sharing formal features of the same type are paired during the linguistic derivation, resulting in a linguistic expression. This is a pair of interface representations – a phonetic and a semantic one, which relate the language system with the performance systems, thereby enabling linguistic expressions to be produced and comprehended in actual language use.

In this context, gender is a formal feature of the lexicon, which can assume different values, corresponding to gender classes. In Portuguese, masculine and feminine are the values that the gender feature can assume. Portuguese has, therefore, a bipartite gender system. The pairing of features in a linguistic derivation may reflect in the morphology, giving rise to inflectional morphemes that express agreement. Language may differ with regard to the

extent to which the pairing of gender features is reflected in the morphology. In Portuguese, the noun is usually viewed as the controller of agreement (the element that presents the feature which defines the value to be assumed by the gender agreeing elements) and there is morphological expression of gender agreement between nouns, determiners, adjectives and participial phrases.

According to Chomsky (1995), formal features can be intrinsic or optional. The former can be considered to be arbitrary whereas the latter encodes a conceptual distinction, such as the gender distinction between male and female. In Portuguese, inanimate nouns (1), as well as some invariable animate nouns (2), have intrinsic gender. Among the latter, the intrinsic gender may coincide with the natural gender of the class of individuals named by the noun (2a); the gender invariant noun may apply to individuals of both sexes (2b) or it can be taken as a generic form, requiring the specification male or female, when necessary (2c).

(1)   ***o** livr**o**/ **a** mes**a**/ **o** comet**a**/ **a** trib**o**/ **o** pent**e**/ **a** ponte*
       the-M book/ the-F table/ the-M comet/ the-F tribe/ the-M comb/ the-
       -F bridge

(2)   a. ***o** homem/ **a** mulher; **o** boi/ a vaca; **o** carneiro/ **a** ovelha*
           the-M man/ the-F woman; the-M ox/ the-F cow; the-M sheep
           (male) / the-F sheep (female)
       b. ***a** vítim**a**/ **a** crian**ça***
           the-F victim/ the-F child/
       c. ***o** jacaré / **a** tartarug**a***
           the-M alligator (male or female) / the-F turtle (male or female)

Animate nouns with optional gender may be gender inflected (3). In this case, the feminine gender defines a subset of the class denoted by the masculine noun.

(3)   ***o** coelh**o**/**a** coelh**a**; **o** gat**o**/ **a** gat**a**; **o** menin**o**/ **a** menin**a***
       the-M rabbit (male) /the-F rabbit (female)– the-M cat (male) /the-F
       cat (female) – the-M boy (male)/the-F girl (female)

Actually, this is the only group of nouns in which gender might be assumed to be an inflectional process. This view is not widely accepted, though. A number of proposals have argued for a derivational account of the feminine gender, either in general or in particular cases, as in the conversion from adjectives (as in nouns like *advogado*/*advogada* [lawyer masculine/feminine]) (Basílio, 1992; Vale, 2004). Masculine/feminine animate pairs such as *gato*/*gata* have also been viewed as independent forms in the lexicon (Villalva, 2008). In both cases –*a* would characterize morphological noun classes[1].

---

[1] Attributing to a lexical process the relation between animate nouns with optional gender implies that two distinct lexical entries are assumed, that is, the gender feature would be taken as intrinsic in these nouns, as it is commonly the case in

It should be noticed that in the corresponding feminine form of nouns having the last syllable with coda in /r/ or /s/, the feminine morpheme *–a* is added to the base (4). These nouns constitute a clear-cut evidence of feminine gender as the gender marked form in Portuguese.

(4)   a.   *o pastor/ a pastor-a; o professor/ a professor-a*
          the-M shepherd (male)/ the-F shepherd (female); the-F teacher (male); the-M teacher (female)
      b.   *o freguês/ a fregues-a*
          the-M client (male)/ the-F client (female)

For the present purposes, the nature of feminine optional gender as the result of a derivational or inflectional process is immaterial. The relevant point, for the present investigation, is that the feminine noun is morphologically marked, either by a derivational suffix or by a gender inflection morpheme. In the former case, the morphological process would take place pre-syntactically in the lexicon (see discussion in Alexiadou & Muller, 2008). In the latter, it would be the result of a syntactic computation (see next section). In both cases, feminine nouns with optional gender would be expected to be more demanding to processing and acquisition.

There are also a small number of animate nouns in which the value of the optional gender is only expressed in D and in other agreeing elements (5):

(5) ***o dentista/a dentista; o modelo/ a modelo; o comediante/ a comediante***
          the-M dentist (male)/ the-F dentist (female); the-M model (male)/ the-F model (female); the-M comic actor (male)/ the-F comic actress (female)

Given the properties of the gender system in Portuguese, it can be noticed that the noun-ending *–a* in animate nouns is particularly ambiguous for the language learner: it can be a thematic vowel, i.e., a vowel that ends a stem and precedes an inflection; or the feminine gender morpheme (an inflectional or

---

homem (man)/*mulher* (woman). However, there are empirical results suggesting that independent entries can be supposed when the feminine noun acquires more specialized meaning or becomes particularly frequent in the language (as *empregada* (employee) or *viúva* (widower)) (Corrêa, Almeida & Porto, 2004). Additionally, at least in BP, there are differences in grammaticality/acceptability judgments of the sentences such as (i) and (ii) (Zocca, 2003):

(i) ?Eu já peguei o meu filho na escola e você deve pegar a sua também. (filho/filha)
     I have picked up my son-MS at school and you should pick up yours-FS too. (son/daughter)
(ii) *Eu já visitei o barão de Cocais e você deve visitar a de Tefé também. (barão/baronesa)
     I have visited the barão-MS de Cocais and you should visit the one-FS from Tefé too. (baron/baroness).

derivational affix). Hence, the vowel –*a*, in *mesa* (fem) (table*), cometa* (masc) (comet) or *criança* (fem) (child), is a thematic vowel. In *menin-a* (girl) or *macac*-a (female monkey), it is an inflection morpheme, the gender marked form.

For historical reasons, most nouns having the thematic vowel –*o* are masculine, whereas most nouns presenting the thematic vowel –*a* are feminine. Being so, an associative pattern can be easily established between gender and the morphophonological ending of the noun, regardless of its actual morphological status. This correlation is not, however, reliable to the language learner, since there are feminine nouns ending in –*o,* masculine nouns ending in –*a*, and a large number of nouns ending in -*e*, or showing no thematic vowel. Table I presents examples of nouns distributed in formal classes.

Table I: Formal classes of nouns and gender (based on: Alcântara, 2003)

	MASCULINE	FEMININE	MASC./FEM.
-o	livr-o	trib-o	model-**o**
-a	map-a	**bol-a**	coleg-**a**
-e	pent-e	**pont-e**	client-**e**
athematic	armazém	nuvem	refém

Thus, in order to identify the gender of a noun in Portuguese, the most reliable information comes from the elements agreeing with it.

3. Intrinsic and optional gender in the lexicon: implications for language acquisition and production

The differences between intrinsic and optional gender attested above must be represented in the lexicon.[2] There is a trend in generative linguistics which favors the projection of formal features whenever they have semantic content – a tendency that stems from the pre-minimalist research. In this context, some proposals focusing on independent projections for number – NumP – and gender – GenP – have been put forward (Picallo, 1991; Ritter, 1991). Following minimalist requirements, however, the number of functional projections should be minimal. In line with this view, Chomsky (1999)

---

[2] It should be noticed that a strict correspondence cannot be established between the lexicon as formally characterized in a grammar and the mental lexical, accessed during language processing. The formal treatment of the lexicon may, in any case, provide some insights to be taken into account in a theory of language processing and acquisition.

assumes that N bears the so called phi-features (gender, number and person). The relevance of maintaining independent projections within the DP in a minimalist framework would concern the DP internal agreement phenomenon and word-order issues (Cinque, 1994; Magalhães, 2004; Augusto, Ferrari--Neto & Corrêa, 2006).

In Portuguese, as in other Romance languages, intrinsic gender and noun classes are confounded. Noun classes, for not having impact on syntax, have been considered to require a pre-syntactic morphological layer, where material that will be used in the syntactic derivation would be assembled. That is, a noun would enter derivation with its noun class (and intrinsic gender) previously defined (Alexiadou & Muller, 2008). Intrinsic gender is totally arbitrary whereas optional gender is necessarily related with conceptual gender classes. Consequently, it can be assumed that intrinsic and optional gender must be distinguished. Intrinsic gender may be taken as a feature of N whereas optional gender, for defining semantic classes, may head its own projection, as shown in Figure 1.
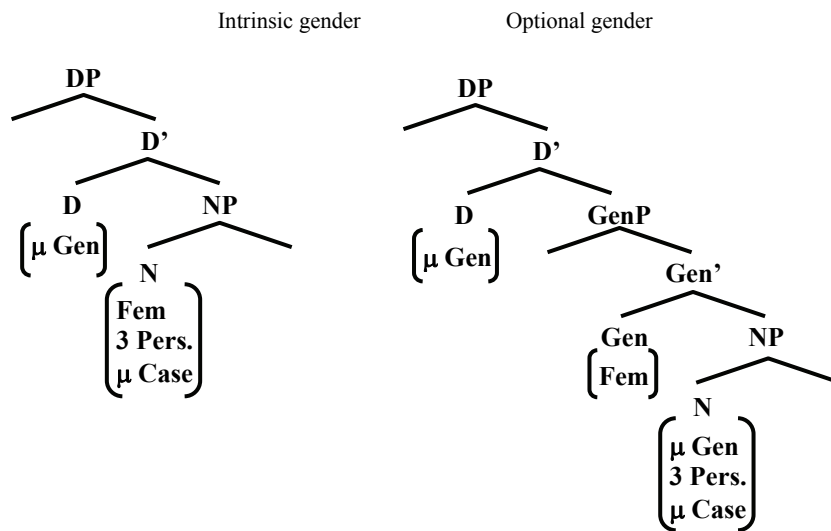

Figure 1: Representation of intrinsic and optional gender

Assuming that language production involves syntactic computation, it could be argued that computational complexity can be measured as a function of the number of functional projections it requires (Jakubowicz, 2002, 2003). It can be expected then that the production of optional gender nouns is computationally more demanding than the production of intrinsic gender nouns. Some results in a comprehension task point to this direction (Augusto & Corrêa, 2005).

As for the acquisition of the gender of novel animate nouns, it should be noticed that it is not immediately clear whether the gender is optional or intrinsic. The optimal learning procedure would be to assume that gender is intrinsic by default, and to represent it as optional, provided positive evidence (a single root admitting two genders) is available. Animacy may, nevertheless, induce the association of grammatical gender with a conceptual class. In this case, the gender of the noun would be underspecified in the lexicon and additional cost would be expected for the production of feminine nouns, in so far as feminine is the gender marked form (Halle & Marantz, 1993; Noyer, 1997).

4. An elicited production task

The present experiment aimed at verifying how children acquiring Portuguese ascribe gender to novel animate nouns. Two different approaches to this process are contrasted, namely, the cue-based strategic one, which have been widely assumed in the literature, and an algorithmic procedure dependent on the processing of agreement in the DP, as suggested in Corrêa & Name (2003). According to a cue-based view of gender acquisition, children would heavily rely on gender/noun-ending correlations. According to a syntactic view, although correlational patterns may affect children's performance, reliance on an agreeing element (D in particular) as the source of information concerning gender would prevail in the ascription of novel nouns to a gender class. In this experiment, animate nouns were used, which might be taken as having optional gender. Unlike inanimate nouns, for which gender is invariably an intrinsic feature, the gender of animate nouns may be optional, that is, the value of the gender feature of the noun remains underspecified in the lexicon. Its value can only be ascribed in the production of a DP, when grammatical gender encodes the sex of the referent. Given that gender did not affect the ascription of inanimate nouns to a gender class in Corrêa & Name (2003), this study aimed at verifying the extent to which gender would affect children's performance in the presence of animate novel nouns.

An elicited production task was used. Children were presented a short story with two characters, one of them being an invented figure named by either masculine or feminine pseudo-nouns in a sort of comic strip. Their task was to refer back to the newly presented invented character. The pseudo-nouns had three different endings (-*o*, -*a* and -*e*). and the only reliable information concerning their gender was provided by the determiner in the DPs used to present (indefinite articles *um* (masculine)/ *uma* (feminine)) and to refer back to the invented character (definite articles *o* (masculine)/ *a* (feminine)). The correlation between the gender of the determiner and pseudo-noun endings was manipulated.

Design:

The independent variables were: 1) *gender/noun-ending correlation* (correlated, non-correlated, and neutral, that is, there is no possible morphophonological correlation between the gender expressed in the determiner and the noun ending); 2) *grammatical gender* (masculine and feminine); 3) *age* (2 and 4 year olds); and 4) *language variety* (Brazilian and European Portuguese). The design was, therefore, a multi-factorial one: 3 (gender/noun-ending correlation) X 2 (grammatical gender) X 2 (age) X 2 (language), where the first factor is a repeated measure. The dependent variable was the number of responses maintaining the gender expressed in the determiner. These responses could include a full DP, a deictic pronoun or a DP with noun ellipsis, since in all of these forms, the gender ascribed to the noun would necessarily be expressed. Table II displays the types of stimuli presented to children in each condition.

Table II: Examples of the linguistic stimuli in each condition

	Correlated	**Non-correlated**	**Neutral**
Masc.	**o** dab**o**	**o** dab**a**	**o** dab**e**
Fem.	**a** bid**a**	**a** bid**o**	**a** bid**e**

The working hypotheses were: (i) children rely on an optimal procedure for the identification of the gender of novel nouns, even though gender noun-ending correlation is a factor that may affect this process; and (ii) animate nouns favor gender to be taken as optional and feminine animate DPs to be morphologically marked. Being so, the hypotheses tested were: (a) the gender class identified in the determiner (an agreeing element) that introduces novel nouns defines the value of the gender of the newly acquired noun; (b) gender/noun-ending correlation is not a major strategy; (c) gender marking imposes additional demands to the production of DPs.

Given (a), it would be expected that the majority of children's responses, when referring to the newly introduced character, maintain the gender of the determiner that introduced the novel noun. Given (b), more target responses are expected in the gender-noun ending correlated condition. Given (a) and (b), more alterations in the noun-endings than in the determiner are expected, in order for the correlation between the gender in the determiner and the noun-ending to be maintained. As for (c), feminine DPs are expected to be harder to produce.

Method:

Participants: A total of 80 monolingual children, being 40 middle-class Brazilian children, living in Great Rio de Janeiro, and 40 middle-class Portuguese children, living in Great Lisbon, were tested. All children were attending a nursery school. Two age groups were created: 2 year olds and 4 year olds. The younger group was formed by 20 Brazilian children (12 girls) (M_{age} = 2;6, age range: 2;0-2;11 years) and by 20 Portuguese children (12 girls) (M_{age} = 2;6, age range: 2;2- 2;11years). The older one by 20 Brazilian children (10 girls) (M_{age}= 4;2, age range 3;6-4;11 years), and by 20 Portuguese children (13 girls), (M_{age}= 4;0, age range: 3;6-4;5 years). Half of the children in each language and age group were assigned to a *grammatical gender* condition. That is, 10 children of each language/age group were presented to masculine stimuli and the other 10 to feminine stimuli, in order to avoid too many stimuli per child. None of the children was reported to have family history of language deficits or be affected by factors that might impair language development. Children who did not engage in a verbal interaction with the experimenter prior to the test were not included in the group of participants.

Material:

10 lists of short stories (5 masculine and 5 feminine lists to the two groups, defined by *grammatical gender*, respectively), corresponding to sequences of slides. Each list contained 3 pre-tests, 9 test stories (3 trials for each experimental condition) and 9 fillers, totalizing 21 stories. In each of the lists, the order of presentation of the test stimuli was randomized with the provision that fillers were maintained in the same position across lists. The corresponding visual material consisted then of 21 power-point slides. Each slide/story consisted of 4 pictures displaying an invented and a real animate figure as the characters of 4 different scenes, creating a sort of comic strip, as shown in Figure 2. The real character was always an animal, easily recognizable by 2 year olds, and the invented character was a sort of fancy animal. The visual material was the same for the groups defined by *grammatical gender*. The lists included the same pseudo-nouns to refer to the invented characters, but their gender was defined by the determiner according to the list: in the masculine lists, masculine determiners were used in the DPs referring to the invented character, and in the feminine lists, feminine determiners were used. The linguistic material was the same across language groups. The only difference was in the use of gerundive/infinitival forms of the verb for describing the event, in order to meet language variety specificities (see example in Figure 2). The apparatus used in the BP group was a Toshiba lap-top and a Sony tape-recorder. In the EP group, a Macbook (lap-top) and an Ipod nano with a recording device were used. There was also a response sheet per child in order to guarantee that possible non--audible responses in tape would not be lost. An example of a test story is provided below:
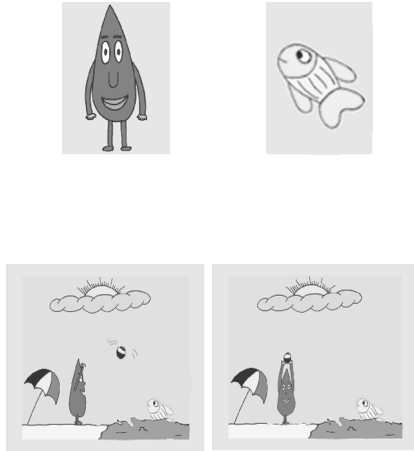
	Aqui tem / está um dabo e aqui tem / está um peixe. O dabo e o peixe estão jogando / a jogar bola na praia. Olha o peixe! Olha o dabo! Ih, o que aconteceu? Quem pegou a bola? Here is a dabo and here is a fish. Look! The dabo and the fish are playing with a ball. Look at the fish! Look at the dabo! Hey, what happened now? Who has the ball?

Figure 2: An example of a test story (characters and story pictures,
and a description of what was told to the children)

Procedure:

Children were invited to play a computer game in which they would be presented to stories and asked a question at the end of them. Half of the 2 year olds and half of the 4 year olds were presented to masculine lists; the other halves, to feminine lists. The stories were presented orally by the experimenter (a native BP speaker and a native EP speaker, for the BP group and the EP group respectively), with the support of the slides displayed in the lap-top screen. The 4 pictures of each slide were gradually displayed as the story was told. After the fourth picture, a WH question was presented and the child had to refer back to the invented character in order to answer it by saying it aloud. Pointing was discouraged. If it occurred, children were told that they should tell the experimenter the answer aloud. If the child insisted in pointing, the response was considered to be nil. Children were tested individually in a quiet room at their nursery school. The warming up pre-test consisted of three stories involving known animals and an invented character. In this pre-test, children were only requested to refer back to the known animal characters. After the pre-test, children were invited to continue playing

the game and were told that from that point onwards more different characters, which might be unknown to them, would be presented. Children were encouraged to continue at the end of each trial. The whole session took 8-10 minutes.

Results:

The scores were submitted to a 2 (*language variety*) X 2 (*age*) X 2 (*grammatical gender*) X 3 (*gender/noun-ending correlation*) ANOVA. As expected, there was no significant main effect of *language variety* (p =. 88). The other factors gave rise to significant main effects: *age* ($F(1,72) = 21.3$ p<. 00001), with more target responses in the oldest age group ($M = 1.84$, SD = 1,08 vs. M= 2.62, SD = 0.64, respectively for 2- and 4-year olds); *grammatical gender* ($F(1,72) = 5.91$ p <.02), with more target responses in the masculine condition ($M = 2.43$, SD = 0.79 vs. M = 2.02, SD = 1.11, respectively for masculine and feminine); *gender/noun-ending correlation* ($F(2,144) = 3.95$ p =.02), with more target responses to correlated DPs ($M = 2.39$, SD = 0.9 (correlated) vs. M = 2.11, SD = 0.86 (non-correlated) vs. M = 2.19, SD = 0.82 (neutral)). A four-way significant interaction was also obtained ($F(2,144) = 4.08$ p =.02). Table III presents the means and the SDs.

Table III: Means (standard deviations) as a function of gender/noun-ending correlation, grammatical gender, age and language (max score =3)

Age	Gender	Correlated		Non-correlated		Neutral	
		BP	EP	BP	EP	BP	EP
2 years	Masculine	2.5 (1.08)	2.3 (1.05)	2.2 (0.63)	1.9 (1.28)	1.8 (1.14)	2.2 (1.03)
	Feminine	1.4 (1.17)	2.0 (1.41)	1.2 (0.92)	1.4 (1.26)	1.8 (1.03)	1.4 (0.97)
4 years	Masculine	2.7 (0.67)	2.9 (0.32)	2.5 (0.71)	2.9 (0.32)	2.6 (0.52)	2.7 (0.67)
	Feminine	2.8 (0.42)	2.5 (1.08)	2.6 (0.70)	2.2 (1.03)	2.5 (0.71)	2.5 (0.53)
Total mean		2.35 (0.84)	2.45 (0.97)	2.13 (0.74)	2.1 (0.97)	2.18 (0.85)	2.2 (0.8)

The main effects were in the predicted direction. Performance improved with age, the production of feminine DPs was more demanding and children were sensitive to gender/noun-ending correlation. The latter effect does not indicate, however, that children rely on it in the identification of the gender of novel words. Target responses (that is, responses in which the gender of the determiner is maintained) predominated in both age and language groups,

except for the feminine non-correlated condition in the Brazilian 2 year old group, which may account for the four-way interaction obtained.

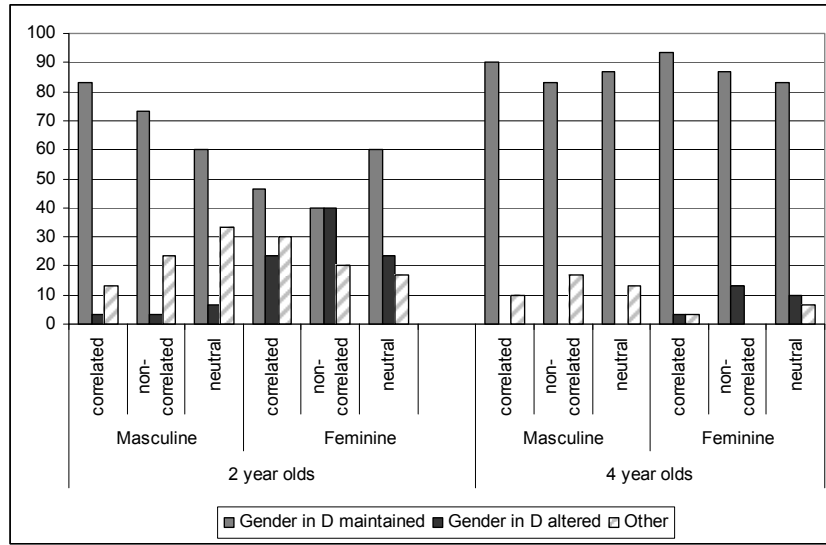Figures 3 and 4 present the breakdown of children's responses in each language group.

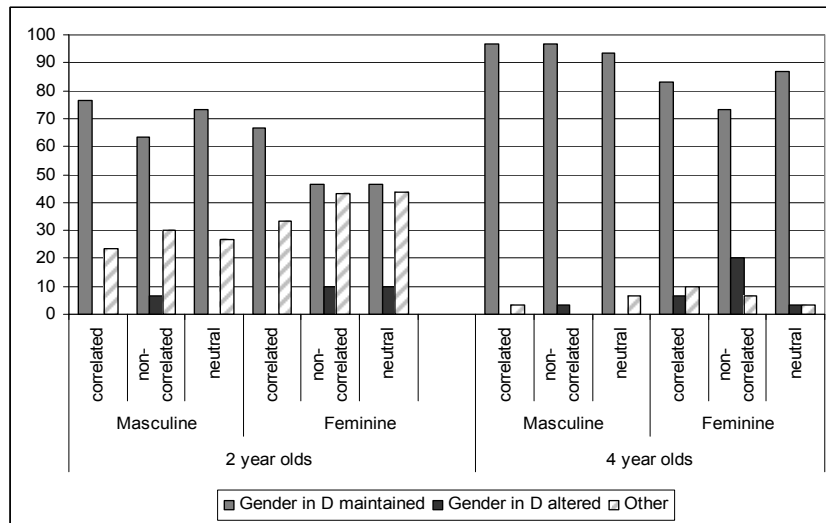Figure 3: Distribution (%) of responses per condition in the BP group (n=30)

Figure 4. Distribution (%) of responses per condition in the EP group (n=30)

The subset of target responses corresponding to a full DP (a DP presenting determiner and noun) in each language group was further analyzed. It was intended to verify the extent to which the noun-ending would be altered to match the gender of the determiner (See Table IV for BP and Table V for EP). In these tables, percentages are provided (for each condition in each age group), considering the total number of alterations in the age group, even though this number is small, particularly in the BP 2 year old group. In spite of that, the pattern of alterations is informative. The same pattern of responses can be observed in the two language groups. Alterations in the noun-endings predominated in the non-correlated gender/noun-ending condition, suggesting that it is the noun-ending that is changed in order to match the gender of the noun, ascribed on the basis of the information in D. Some alterations can also be observed in the correlated and neutral condition, possibly due to difficulties in the recall of the novel noun. This pattern is clearer in the EP group in which a greater number of full DP responses were obtained.

Table IV: Percentage target responses with noun-ending alterations in BP

Age	Gender	Correlated	Non--correlated	Neutral
			Gender/noun-ending correlation	
2 years (Total = 7)*	Masculine	14.3(7)	57.1(7)	0
	Feminine	0	28.6(7)	0
4 years (Total = 20) *	Masculine	10(20)	15(20)	5(20)
	Feminine	5(20)	50(20)	15(20)
Total mean		7.4	37.7	5

* Total number of responses with noun-ending alterations per age group

Table V: Percentage target responses with noun-ending alterations in EP

Age	Gender	Correlated	Non--correlated	Neutral
			Gender/noun-ending correlation	
2 years (Total = 22) *	Masculine	4.5(22)*	22.7(22)	22.7(22)
	Feminine	0	45.5(22)	4.5(22)
4 years (Total = 20) *	Masculine	5(20)	20(20)	10(20)
	Feminine	10(20)	40(20)	0(20)
Total mean		4.9	32.1	9.3

* Total number of responses with noun-ending alterations per age group

Discussion:

The results as a whole are compatible with the view that children rely on the gender information in the determiner when referring to the newly introduced character. Children appear to make use of an algorithmic procedure that applies to both inanimate nouns, as suggested in Corrêa & Name (2003), and to animate nouns, as the present results indicate. As in the previous study with inanimate nouns, children are sensitive to a gender/noun--ending correlation, though animate nouns make the feminine condition more vulnerable for an effect of correlational patterns, since the age of 2. It is possibly due to the representation of gender as an optional feature, which would raise the expectation for a gender marked noun. The pattern of children's responses suggests that feminine animate DPs impose additional demands for production possibly due to morphological marking. Moreover, the occurrence of some noun-ending alterations in masculine and neutral targets support the view that animate DPs favor the representation of optional gender.

The pattern of the results obtained in Portuguese does not conflict with the one obtained by Pérez-Pereira (1991) in Spanish. In both cases, there was an effect of correlational patterns but the majority of the responses were based on the determiner and the feminine condition imposed greater difficulties to children. A different explanation is nevertheless provided. While the effect of gender-noun ending suggested cue-based learning in that study, in the present one, it was taken to reflect some awareness of the gender/noun-ending correlation in the language, after the gender identification process has been successfully established on a syntactic basis. As for the French data (Karmiloff-Smith, 1979), it may take longer for children to succeed in a task that requires them to rely exclusively on the information in the determiner. In this language, gender is morphologically expressed only in singular determiners and these morphological distinctions may be neutralized by a subsequent noun or adjective beginning with a vowel. In any case, the effect of correlational patterns would not be indicative of a learning process. No matter how strong these correlations can be, the fact that children produce gender agreement since an early age is compatible with their reliance on the information provided by agreeing elements, D in particular. It should be emphasized that children must give priority to gender information in the determiner when there is inconsistence in the gender noun-ending pattern and this preference was attested in the present results.

5. A procedure for gender acquisition in Portuguese

In this section, a procedure for gender acquisition heavily dependent on syntax is proposed, which may account for the identification of the gender of both inanimate and animate nouns. It reconciles the phonological bootstrapping hypothesis (Gerken & McIntosh, 1993; Morgan & Demuth,

1996; Christophe, 2002) with a minimalist conception of language (Chomsky, 1995, 2007) by assuming that infants process information available at the phonetic interface, thereby identifying prosodic and distributional cues which enable them to delimit closed and open class elements (Corrêa, 2009). This initial categorization would constitute the basic differentiation between functional and lexical classes in the lexicon under acquisition. There is evidence that children, by the age of 10-12 months, distinguish these classes perceptually and that by the beginning of the second year of life, they distinguish the phonological and distributional properties of determiners and start parsing DPs, by relying on determiners as closed class elements (Höhle & Weissenborn, 2000; Name, 2002; Johnson & Diks, 2005; Corrêa, Bagetti & Name, 2008). There is also evidence that children by the end of their first year of life are sensitive to the morphophonological form of functional elements (Corrêa, Bagetti & Name, 2008; Bagetti, 2009) and that by the age of two, gender mismatch between determiner and nouns can be perceived (Name & Corrêa, 2001; Corrêa & Name, 2003).

Given this evidence, the algorithmic procedure for gender identification would follow two steps: (i) identification of morphophonological variations pertaining to gender within the closed class of the determiner, which would enable children to identify possible gender classes or the values of a gender feature; (ii) parsing of the DP, under the assumption that agreement relations hold between syntactically related elements. Once a gender class is recognized in the determiner, the corresponding value of the gender feature is ascribed to the novel noun and represented in the lexicon as its intrinsic gender feature. As a result, the noun can control agreement in production, and a matching between the gender of the noun and its agreeing elements can be expected in comprehension (Corrêa, 2001; Name, 2002; Corrêa & Name, 2003). Reliance on agreement as an algorithmic learning procedure would account for gender acquisition across languages, regardless of the number of gender classes and the complexity of their morphology. Children would, in any case, become sensitive to frequency and correlational patterns, which would account for occasional errors and a reflexive behavior in relation to language (Figueira, 2003). This kind of association would occur independently of the language module involved in syntactic computation, giving rise to an interference effect rather than to a productive analytical procedure.

As for optional gender, two possibilities can be considered. Intrinsic gender is ascribed to the noun by default. Given positive evidence, that is, the occurrence of gender inflected nouns, which share a single nominal root, the gender value of the root becomes underspecified. Alternatively, animacy biases the representation of gender as an optional feature. The value of the gender feature informed by the determiner is taken to inform the gender of the DP, corresponding to the sex of the referent. The gender of the noun remains underspecified in the lexicon and the value to be assumed by it is defined at every single production of a DP, as a function of the sex of its referent. The

production of optional feminine gender requires access to the gender morpheme in the morphophonological encoding of the noun. This differentiation between the representation of intrinsic gender and optional gender nouns is compatible with the formal distinction between intrinsic gender feature and a gender functional projection characterized in section 3. Hence, from a conceptual, syntactic and morphological point of view, the production of feminine nouns would be harder. The present results are compatible with this second option.

This procedure would be expected to apply to all languages in which nouns are classified as a function of gender and there are agreement marks. Variations in the procedure would nevertheless be required in order for it to be adapted to the kind of gender agreement information the morphology of the language provides.

6. Concluding remarks

The aim of this paper was to provide an algorithmic procedure for the acquisition of gender in Portuguese that could account for the early identification of the gender of novel words and long lasting effects of correlational patterns. It was argued that even though such effects can be detected experimentally, by no means they justify a cue-based learning procedure. Children acquiring Portuguese, regardless of the variety of this language, take into account predominantly the gender information provided by the determiner when referring back to a newly introduced character presented by a novel noun. Mismatch between the gender in the determiner and the noun-ending is, nevertheless, more disturbing in feminine DPs. A similar gender effect was not obtained with inanimate novel nouns (Corrêa & Name, 2003). It was argued that animacy favors the representation of gender as an optional feature. Assuming that from a linguistic point of view, optional gender involves an additional functional projection in the DP and that feminine optional gender is morphologically marked in the noun, in the majority of the cases, feminine DPs were expected to be more demanding in a production task. Moreover, a feminine determiner creates the expectation for a corresponding marked morpheme, which would explain children's predominant altering the noun-endings to –*a*, in order to match the gender of the determiner in the non-correlated condition. A cue-based learning procedure would not be able to account for the unbalance between the production of masculine and feminine DPs.

In sum, by the age of 2, children acquiring Portuguese have identified the gender system though may not be able to cope easily with the demands of the production of newly acquired feminine DPs. This difficulty may still be noticed by the age of 4. An algorithm procedure based on agreement and the effect of markedness can better account for the acquisition of gender.

References

Alcântara, C. (2003) As classes formais do português e sua constituição: um estudo à luz da teoria da Morfologia Distribuída. Unpublished Doctoral Thesis, Pontifícia Universidade Católica do Rio Grande do Sul.

Alexiadou, A. & Müller, G. (2008) Class features as probes. In *Inflectional Identity* (A. Bachrach & A. Nevins, editors), pp. 101-155. Oxford: Oxford University Press.

Arnon, I. & Ramscar, M. (2009) Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. Talk presented at CogSci 2009.

Augusto, M.R.A. & Corrêa, L.M.S. (2005) Marcação de gênero, opcionalidade e genericidade: processamento de concordância de gênero no DP aos dois anos de idade, *Lingüística*, **1**, 207-234.

Augusto, M.R.A., Ferrari-Neto, J. & Corrêa, L.M.S. (2006) Explorando o DP: a presença da categoria NumP, *Revista de Estudos da Linguagem*, **14**, 245-276.

Bagetti, T. (2009) Um estudo experimental do processamento na interface fônica e da análise sintática inicial: o papel de elementos funcionais na aquisição da linguagem. Unpublished Doctoral Thesis, PUC-Rio.

Câmara Jr., J.M. (1976) *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão.

Cinque, G. (1994) On the evidence for partial N-movement in the Romance DP. In *Paths towards Universal Grammar* (G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini, editors), pp. 85-10. Washington DC: Georgetown Universisty Press.

Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, N. (1999) *Derivation by Phase*. MITWPL.

Chomsky, N. (2007) Of Minds and Language, *Biolinguistics*, **1**(1), pp.9-27.

Christophe, A. (2002) The role of phonological phrases in early language acquisition. In *Proceedings of the GALA'2001 Conference on Language Acquisition* (J. Costa & M. J. Freitas, editors), pp. 180-188. Cambridge: Cambridge Scholars Press.

Clahsen, H. & Almazan, M. (1998) Syntaxe and Morphology in William Syndrome, *Cognition*, **68**, 167-198.

Comrie, B. (1999) Grammatical gender systems: a linguist's assessment, *Journal of Psycholinguistic Research*, **28**(5), 457-466.

Corbett, G. (1991) *Gender*. Cambridge: Cambridge University Press.

Corrêa, L.M.S. (2001) Uma hipótese para a identificação do gênero gramatical com particular referência para o português, *Letras de Hoje*, **36**(3), 289-295.

Corrêa, L.M.S. (2009) Bootstrapping language acquisition from a minimalist standpoint: On the identification of phi-features in Brazilian Portuguese. In *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese* (A. Pires & J. Rothman, editors), pp. 35-62. Berlin: Mouton de Gruyter.

Corrêa, L.M.S. & Name, M.C.L. (2003) The Processing of Determiner-Noun Agreement and the identification of the gender of Nouns in the early acquisition of Portuguese, *Journal of Portuguese Linguistics*, **2**(1), 19-43.

Corrêa, L.M.S., Bagetti, T. & Name, M.C.L. (2008) Distinção entre classes abertas e fechadas no processamento da fala ao fim do primeiro ano de vida. In *XV Congresso Internacional de la Associación de Linguistica e Filologia de América Latina (ALFAL)*, Montevidéu, CD, pp.70-85.

de Houwer, A. (1990) *The Acquisition of Two Languages from Birth: A Case Study*. Cambridge: Cambridge University Press.

Demuth, K. & Ellis, D. (2008) Revisiting the acquisition of sesotho noun class prefixes. In *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin* (J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura & S. Ozcaliskan, editors), pp. 131-148. London: Psychology Press.

Figueira, R.A. (2003) La Propriété Réflexive du Langage dans le Parler de l'Enfant. Quelques Manisfestations du Fait Autonymique dans l'Acquisition du Langage. In *Parler des Mots. Le Fait Autonymique en Discours* (J. Authier-Revuz, M. Doury & S. Reboul-Touré, editors), volume 1, pp. 193-204. Paris: Presses de la Sorbonne Nouvelle.

Gerken, L-A. & McIntosh, B.J. (1993) Interplay of function morphemes and prosody in early language, *Developmental Psychology*, **29**(3), 448-457.

Harley, B. (1979) French gender 'rules' in the speech of English-dominant, French--dominant and monolingual French-speaking children, *Working Papers in Bilingualism,* **19**, 129–156.

Höhle, B. & Weissenborn, J. (2000) The origins of syntactic knowledge: recognition of determiners in one-year-old German Children. In *Proceedings of the 24th Annual Boston Conference.* pp. 418-429. Sommerville, Mass: Cascadilla Press.

Jakubowicz, C. (2002) Functional Categories in (Ab)Normal Language Acquisition. In *The Process of Language Acquisition* (I. Lasser, editor), pp. 165-202. Berlin, Germany: Peter Lang Verlag.

Jakubowicz, C. (2003) Computational Complexity and the Acquisition of Functional Categories by French-speaking children with SLI, *Linguistics*, **41**(2), 175-211.

Johnson, E.K. & Diks, M. (2005) On-line processing of grammatical gender in Dutch--learning toddlers. Talk presented at the *Xth International Congress for the Study of Child Language*, Berlin, Germany.

Karmiloff-Smith, A. (1979) *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge Univ. Press.

Levy, Y. (1983) It's frogs all the way down, *Cognition*, **15**, 75-93.

Morgan, J. & Demuth, K. (eds.) (1996) *Signal to Syntax*: *Bootstrapping from Speech to Grammar in Early Acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Müller, N. (2000) Gender and number in acquisition. In *Gender in grammar and cognition* (B. Unterbeck, M. Rissanen, N. Nevalainen & M. Saari, editors), pp.351--399. Berlin: Mouton de Gruyter.

Name, M.C. (2002) *Habilidades perceptuais e lingüísticas no processo de aquisição do sistema de gênero no português.* Unpublished Doctoral Thesis, PUC-Rio.

Name, M.C.L. & Corrêa, L.M.S. (2003) Delimitação perceptual de uma classe correspondente à categoria Funcional D: Evidências da aquisição do português, *Fórum Lingüístico*, **3(**1), 55-88.

Pérez-Pereira, M. (1991) The acquisition of gender: what Spanish children tell us, *Journal of Child Language*, **18**, 571-590.

Picallo, C. (1991) Nominals and nominalization in Catalan, *Probus* **3**(3), 279-316.

Pinker, S. (1995) *The language instinct*. New York: Harper Perennial.

Ritter, E. (1991). Evidence for number as a nominal head, Talk presented at *GLOW-14*.

Rodina, Y. & Westergaard, M. (2008) A cue-based approach to the acquisition of grammatical gender in Russian. Talk presented at *Formal Descriptions of Slavic Languages 7.5*.

Scherag, A., Demuth, L., Roesler, F., Neville, H. J. & Roeder, B. (2004) The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects, *Cognition*, **93**, B97-B108.

Tomasello, M. (1992) *First verbs: a case study of early grammatical development*. Cambridge: Cambridge University Press.

Unsworth, S. (2008) Age and input in the acquisition of grammatical gender in Dutch**,** *Second Language Research*, **24**(3), 365-395.

Unsworth, S. (2007) Age and Input in Early Child Bilingualism: The Acquisition of Grammatical Gender in Dutch. In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (A. Belikova, L. Meroni & M. Umeda, editors), pp. 448-458. Somerville, MA: Cascadilla Proceedings.

van Berkum, J.J.A. (1997) Syntactic processes in speech production: The retrieval of grammatical gender, *Cognition*, **64**(2), 115-152.

van der Velde, M. (2004) L'acquisition des déterminants en L1: une étude comparative entre le français et le néerlandais, *Acquisition et Interaction en Langue Etrangère*, **21**, 9-46.

Westergaard, M.R. (2006) Triggering V2: The Amount of Input Needed for Parameter Setting in a Split-CP model of Clause Structure. In *Language Acquisition and Development: Proceedings of GALA 2005* (A. Belletti, E. Bennati, C. Chesi, E. DiDomenico & I. Ferrari, editors), pp. 564-577. Cambridge: Cambridge Scholars Press.

Westergaard, M.R. (2008) Acquisition and Change: On the Robustness of the Triggering Experience for Word Order Cues, *Lingua* **118**(12), 1841-1863.

Letícia M. Sicuro Corrêa
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Laboratório de Psicolingüística e Aquisição da Linguagem (LAPAL)
Departamento de Letras
Rua Marquês de S. Vicente, 225, sala K-121
Gávea – Rio de Janeiro
CEP 22453-900
Brazil
lscorrea@puc-rio.br

Marina R. A. Augusto
Universidade Estadual do Rio de Janeiro (UERJ)
Rua São Francisco Xavier, 524 – 11 andar
Rio de Janeiro – RJ
CEP 20550-900
Brazil
&
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Laboratório de Psicolingüística e Aquisição da Linguagem (LAPAL)
marinaaug@uerj.br

Ana Castro
Escola Superior de Saúde do Instituto Politécnico de Setúbal (ESS-IPS)
&
Centro de Linguística da Universidade Nova de Lisboa (CLUNL)
Avenida de Berna 26-C
1069-061 Lisboa
Portugal
acastro@fcsh.unl.pt