

Early acquisition of gender agreement in the Spanish noun phrase: starting small*

SONIA MARISCAL

Universidad Nacional de Educación a Distancia, Madrid

(Received 3 December 2005. Revised 2 November 2006. First published online 2 September 2008)

ABSTRACT

Nativist and constructivist accounts differ in their characterization of children's knowledge of grammatical categories. In this paper we present research on the process of acquisition of a particular grammatical system, gender agreement in the Spanish noun phrase, in children under three years of age. The design of the longitudinal study employed presents some variations in relation to classical studies. The aim was to obtain a large corpus of NP data which would allow different types of analysis of the children's productions to be carried out. Intra-individual variability in early NP types was analyzed and measured, and an elicitation task for adjectives was used. Results show that the acquisition of NP and gender agreement is a complex process which advances as the children gradually integrate different pieces of evidence: phonological, distributional and functional. The reduction of variability as the grammatical process advances is a key feature for its explanation.

INTRODUCTION

In the Romance languages – Spanish, French, Portuguese, Italian etc. – all words referring to beings, objects and events are of masculine or feminine gender. Speakers of such languages learn to use different masculine and feminine forms for a range of associated words every time they want to say something about a particular entity. For nouns referring to inanimate objects, meaning usually provides no clues as to the grammatical gender of the word. Gender agreement is mainly an arbitrary linguistic process.

[*] This research was funded by the Ministerio de Educación y Ciencia (DGICYT). Project PB91-0352. Thanks to the following people for their comments to previous drafts on this paper: Susana López-Ornat, Pilar Gallo and Alexandra Karousou. Special thanks to the two anonymous reviewers of this paper, as well as to the action editor. Address for correspondence: Sonia Mariscal, Departamento de Psicología Evolutiva y de la Educación, Universidad Nacional de Educación a Distancia (UNED), C/ Juan del Rosal, 10, 28040 Madrid, Spain. e-mail: smariscal@psi.uned.es

For languages with gender agreement, speakers acquire the relevant information that enables them to assign gender mainly through two sources. The primary source of information is lexical and distributional – it comes from the other words with which the noun occurs systematically and which mark gender unambiguously. The most informative of these items is the definite article (in Spanish, *el/los* ‘the (masc., sing./pl.)’ and *la/las* ‘the (fem., sing./pl.)’), which occurs more frequently than the indefinite one (*un/unos* ‘a (masc., sing./pl.)’ and *una/unas* ‘a (fem., sing./pl.)’). Other less frequently occurring form classes are adjectives (for example, *pequeño/a* ‘small (masc./fem.)’), personal pronouns *él/ellos* ‘he/they (masc.)’, *ella/ellas* ‘she/they (fem.)’ and other kinds of determiners and pronouns (for example, demonstratives *este/esta* ‘this (masc./fem.)’, *otro/a* ‘another (masc./fem.)’). A second source of influence is sublexical (phonetic), and it derives from the fact that the endings of many nouns in Spanish are more often associated with one gender than with the other: an ending in *-a* is usually associated with feminine gender and an ending in *-o* usually indicates masculine gender. However, this is not a one-to-one correspondence: there is a significant number of masculine nouns ending in *-a*, and a very small number of feminine nouns ending in *-o*. There are also nouns with no overt gender marking.

Similarly, many adjectives show gender agreement by means of the same endings found in nouns (*-o* and *-a*), but there is also an important group of adjectives lacking these two endings; for example, *grande* ‘big’ is invariant.

According to some linguistic proposals (see Harris, 1991), *-o* and *-a* are word markers, rather than gender markers, because they are not confined to lexical items that have gender, being found also in adverbs. However, looking at the process of language acquisition, these sublexical cues seem to be treated as gender markers for both L1 and L2 learners.

Acquisition of gender agreement by L1 and L2 learners

In both French and Spanish, there is evidence that when children aged between 4;0 and 11;0 are presented with competing cues (semantic, morphophonological and syntactic) in order to decipher the gender of nouns in noun phrases, they rely primarily on morphophonological and syntactic information (Karmiloff-Smith, 1979; Pérez Pereira, 1991). There is also evidence (Hernández Pina, 1984; Mariscal, 1996, 2001; López-Ornat, 1997; Lleó, 1997) that article–noun agreement in Spanish appears to be acquired by the age of three or even earlier. Similarly, the process of article acquisition by Italian- and French-speaking children is complete by the age of three, and their acquisition appears to be relatively problem-free (Pizzuto & Caselli, 1992).

For children acquiring Spanish as their first language, gender agreement errors are few when compared to other kinds of morphological errors, such as tense and person agreement in the verbal system.ⁱ Data obtained from the Spanish corpus 'María' (López-Ornat, Fernández, Gallo & Mariscal, 1994), coded in CHAT and incorporated to the CHILDES database (MacWhinney, 1991), indicate that there is only a percentage of 8.8% gender errors, including late errors which affect clitics (Mariscal, 1997). Some of these errors are over-regularizations such as **la fantasma* 'the (*fem.) ghost (masc.)' (2;04).

In contrast to these data, learners of Spanish as a second language (L2) produce a high percentage of gender agreement errors (Fernández-García, 1999). It seems clear that the reasons for these differences lie in the learning process followed by L1 and L2 learners.

Gender is, typically, explicitly taught in Spanish second-language courses because of its relative transparency, given the above-mentioned sublexical cues. But this explicit information seems to be insufficient and children learning Spanish as L1 are much more successful than L2 students. As we have seen, gender agreement acquisition happens very early and few mistakes are produced. Therefore, looking at the process of L1 acquisition in detail appears to be a good strategy in order to account for the differences between L1 and L2 learners.

The main purpose of this paper is to analyze this particular process of grammatical acquisition, which seems to be so relatively 'easy' for children: the acquisition of gender agreement. We will first review studies on noun phrase acquisition in Romance languages, including not only the determiner+noun structure, but also the qualifying adjectives that should agree in gender with the aforesaid nouns. Then we will present an empirical study designed to obtain early and critical data on this process, given that previous data with Spanish children are either not abundant or have been obtained from children over three years of age (Pérez Pereira, 1991).

Studies on the acquisition of the determiner category in Romance languages

Several studies on Romance languages carried out mainly in the nineties (French: Heinen & Kadow, 1990; Italian: Pizzuto & Caselli, 1992; Bottari, Cipriani & Chilosi, 1993; Catalan: Capdevila & Serra, 1996; Spanish: Hernández Pina, 1984; Aguirre, 1995; Mariscal, 1996; López-Ornat, 1997; Lleó, 1997) found a high frequency of determiner omissions in obligatory contexts during the initial phases of the process of NP acquisition. Children gradually 'fill in' the positions before nouns with phonological forms more

[i] Verb morphology has been investigated to a much greater extent than has the system of gender agreement in Spanish, as it is reflected in publications on these topics.

and more approximate to adult determiner forms. However, from very early on, even from the so-called 'one-word stage' (Lleó, 1997), children already produce some syllables or vowel-like elements before nouns. An example is *ð pé*, an utterance produced by a child, under two years old, pointing to a fish (in Spanish, *el/un pez* 'the/a fish (masc.)'). This kind of evidence is parallel to that found first in English (Bloom, 1970; Dore, Franklin, Miller & Ramer, 1976; Peters, 1977) and then in other languages: the so-called 'fillers'. For NP structures, fillers found before nouns occupy what would be a determiner slot. Descriptively, in Romance languages these linguistic units are under-specified phonologically and appear randomly in combinations with both masculine and feminine nouns. Gradually, their phonological form and their distribution converge with the adult-like model. They are, therefore and without a doubt, equivalent to articles and other determiners.

From a theoretical point of view there are different proposals on the kind of knowledge underlying children's productions of filler + noun structures. Taking a generativist position, Lleó (1997, 2001) stated the availability of the functional category D from the beginning of the acquisition process, 'but it is undifferentiated because the phonological component is still immature' (1997: 255). Along the same lines, Aguirre (1995) explains the high percentage of determiner omissions solely through performance factors.

From a constructivist position, Pizzuto & Caselli (1992) found interesting imbalances in Italian between the uses of masculine and feminine articles: the feminine *la* reached Brown's acquisition criterion (i.e. correct use in at least 90% of the contexts in which the morpheme is clearly required; Brown, 1973) before the masculine *il*. Similar results were found for Spanish in the single-case longitudinal study analyzed by Mariscal (1996; 1997) and López-Ornat (1997). It was found that feminine nouns tended to be combined more frequently with the article *la* or the vowel-form *a*, whereas masculine nouns were preceded (less frequently) by variable vocal forms and reached Brown's criterion later than feminine ones. In both studies the authors attributed this imbalance to the frequency, phonological simplicity and clarity of the feminine form of the article in both Italian and Spanish. Therefore they accounted for the process without relying on innate knowledge of grammar or preformed linguistic categorial schemes.

The studies in Spanish mentioned above highlighted another empirical phenomenon which characterizes the early phases of language acquisition: the co-occurrence of non-grammatical determiner omissions and the production of forms with different proximity to the NP structure (filler + N included). For example, María, the child in these studies, produced the same noun, *pies* 'feet (masc.)', in three different forms (*apes*, *pes*, *epes*), in the same session, at 1;07. In Mariscal (1997), quantitative analysis showed that these variable productions constituted the most frequent nominal type

during the early phases. This variability, even though found in only one subject, could not be accounted for by rules at any linguistic level. Variability is, of course, a phenomenon found and accounted for by different researchers, such as Peters & Menn (1993) and López-Ornat (2003), but in our opinion it has not been sufficiently operationalized and quantified.

In accordance with Pine & Lieven (1997), we consider the necessity of designing new measures, apart from analysis of error and non-error patterns, which would allow us to obtain critical evidence for the evaluation of alternative models for the acquisition of grammatical categories. Accordingly, in this study we analyze intra-individual variability in children's utterances and propose a way of measuring this developmental phenomenon.

Gender agreement between determiner, noun and adjective

As we mentioned earlier, adjectives in Spanish are marked for gender or show gender agreement with the noun they determine/modify. Following Harris (1991), only adjectives belonging to the prototypical group ('inner core of prototypes', in Harris' terminology) have the ending *-o* for masculine and the ending *-a* for feminine.

The majority of adjectives used by children belong to the prototypical group, for example *tonto/-a* 'silly (masc./fem.)', *rojo/-a* 'red (masc./fem.)', *guapo/-a* 'pretty (masc./fem.)', but we also find exceptions, which are invariants, for example, *grande* 'big' and *triste* 'sad'. In contrast to determiners, the position of adjectives is generally postnominal.

As it has already been discussed, data on the acquisition of gender morphology in adjectives is scarce in Spanish. The experimental study by Pérez Pereira (1991), based on Karmiloff-Smith (1979), produced interesting and valid results, but the subjects in his sample were older than four years of age. He found a strong tendency towards the use of masculine adjectives with unknown nouns, and interpreted this pattern of errors in agreement with Greenberg's theory of marked-unmarked grammatical features. Hernández Pina (1984) found overgeneralization errors such as **mota rota* instead of *moto rota* 'motorcycle (irregular fem.) broken (fem.)' and *tierra *azula* instead of *tierra azul* 'soil (fem.) blue (invariant for both genders)' between 1;09 and 2;01 for the only subject of her study. She stated that errors were found until 2;08, but no quantitative data were offered, nor information about the error types provided.

Mariscal (1996) performed qualitative analysis of errors in adjectives for María. The pattern found was an initial production of non-analyzed units followed by minimal productivity and errors of commission.ⁱⁱ From 2;02

[ii] López-Ornat (1994), in her model of acquisition of (bound) morphology, differentiated 'omission errors', 'commission errors' and 'overgeneralization errors'.

TABLE 1. *Formal characteristics of definite articles which favour (+) and hinder (–) acquisition*

(+)	(–)
High frequency in the input.	Definite articles are non-stressed units.
Fixed position (ART+N).	Definite articles are monosyllable.
ART+N form an intonational unit.	Subtle phonetic differences between masculine and feminine forms.

onwards, errors disappeared, and later only a few overgeneralizations were found. From a quantitative point of view, errors were scarce, partly because the spontaneous production of adjectives was infrequent.

From a theoretical position which conceives of language acquisition as a complex construction process (i.e. constructivist position) (Lieven, Behrens, Spears & Tomasello, 2003; Pine & Martindale, 1996; López-Ornat, 1994; Peters & Menn, 1993) the distributional and phonological characteristics of the linguistic items or units to be learned need to be considered. As a result, the frequency, phonological saliency and distributional properties of the linguistic units to be acquired are relevant, even though in different degrees as the development proceeds.

In Spanish, determiners (mainly definite articles) are amongst the most frequent linguistic units in the input.ⁱⁱⁱ Articles are almost always compulsory in an NP, but from a phonological point of view they have low saliency: they are monosyllables and non-stressed elements. In contrast to determiners, adjectives have relatively low frequency. These linguistic units have morphological gender markers, the vowels *-o* and *-a*, which are suffixed to the root as in *tonto* ‘silly (masc.)’ and *tonta* ‘silly (fem.)’. From a distributional point of view, adjectives can appear in structures such as [DET+N+ADJ] and also in DP structures consisting of a determiner and an adjective but no overt noun [DET+ADJ], such as *la roja* ‘the (fem.) red (fem.)’, *el tonto* ‘the (masc.) silly (masc.)’. In these last cases, appropriate interpretation of the so-called null nominals depends on the gender and number features present in adjectives and determiners; such ellipsis is a feature of Spanish.

Tables 1 and 2 summarize the above-mentioned formal properties which could favour (+) or hinder (–) the acquisition of gender in articles and

[iii] Definite articles, mainly singular forms, are the second (*la*) and the fifth (*el*) most frequent words in Spanish: data taken from *Diccionario de Frecuencias* (Alameda & Cuetos, 1995). Data obtained from the longitudinal corpus ‘María’, in CHILDES (López-Ornat *et al.*, 1994), situate these articles between the ten most frequent words in adult input directed to that child.

TABLE 2. *Formal characteristics of gender marking in adjectives which favour (+) and hinder (–) acquisition*

(+)	(–)
Ending position: Located at the end of the stem.	Relatively low frequency. Non-syllabic elements (assimilated to the root).
Non-stressed, but inserted in a trochaic pattern.	Subtle phonetic differences between masculine and feminine prototypical morphemes (-o/-a).
Repetition of the vocalic element in sentences as: <i>la casa roja, el niño tonto, otra casa, otro niño</i> .	There are also non-prototypical, but frequent cases (e.g. adjective <i>grande</i> 'big'), and exceptions (<i>mano pequeña</i> 'little (fem.) hand'; <i>otra mano</i> 'another (fem.) hand').

TABLE 3. *List of the most frequent Spanish determiners^a*

Subclass	Masculine gender		Feminine gender	
	singular	plural	singular	plural
Definite articles	<i>el</i>	<i>los</i>	<i>la</i>	<i>las</i>
Indefinite articles	<i>un</i>	<i>unos</i>	<i>una</i>	<i>unas</i>
Demonstratives	<i>este</i>	<i>estos</i>	<i>esta</i>	<i>estas</i>
	<i>ese</i>	<i>esos</i>	<i>esa</i>	<i>esas</i>
	<i>aquel</i>	<i>aquellos</i>	<i>aquella</i>	<i>aquellas</i>
Quantifier 'other/another'	<i>otro</i>	<i>otros</i>	<i>otra</i>	<i>otras</i>
Possessives (1 st , 2 nd , 3 rd person singular)	<i>mi, tu, su</i>	<i>mis, tus, sus</i>	<i>mi, tu, su,</i>	<i>mis, tus, sus</i>

^a From Alameda & Cuetos, 1995.

adjectives. Table 3 presents a list of the most frequent determiners in Spanish.

AIMS AND PREDICTIONS OF THE PRESENT STUDY

The main aim of the present study was to assess the adequacy of a constructivist approach to (L1) grammar acquisition that would account for the learning process of gender agreement in Spanish. In order to reach this general aim, more specific objectives were formulated.

The first objective was to replicate initial results obtained from the single-case longitudinal study (Mariscal, 1996, 1997) regarding the pattern of acquisition of NP and gender agreement affecting determiners, nouns and qualifying adjectives. The second objective was to develop a non-classical longitudinal design incorporating the specific aim of obtaining different productions of the same nominal items during every session (see Method below). This design would make it possible to develop (new) variability

measurements, whose relevance for the explanation of the process will be discussed. The final objective was to enrich spontaneous data with elicited data for adjectives, whose frequency in early child language is very low.

Taking into account these objectives, and having considered the main aim of this study, our general assumptions are as follows:

The acquisition process of NP and gender agreement will be a gradual and progressive construction, yielded by the interaction between formal properties of the input, processing abilities (i.e. extraction of regularities) of the learner and the particular knowledge state of the system.

The early abilities of formal analysis and extraction of regularities, on the part of the children, will allow them to build their first (pre)grammatical constructions. On these first partial representations, and based on a self-organizing activity of the system and experience with the input, new and more complex representations will be developed (Veneziano & Sinclair, 2000; Peters, 2001). As is characteristic of self-organizing systems, linguistic representations at a particular instant (t_1) will influence new advances (in t_2 , t_3 , ...) and learning-system changes. A detailed analysis (micro-analysis) of spontaneous and elicited productions will enable us to distinguish between different degrees or levels of grammatical knowledge. The acquisition process, therefore, will not be of a discontinuous none-all type, but, rather, it will give rise to partial or incomplete representations, without the high degree of abstraction which characterizes adult linguistic categories. These will be the results of the process, and not the starting point.

This view is in contrast with the generativist approach to language acquisition. Within this model, the acquisition of gender agreement requires a computational system that operates upon the formal features of the functional category D from an early age (see, for example, Sicuro Corrêa & Name, 2003). The availability of that functional category is assumed from the beginning of the acquisition process (Lleó, 1997, 2001).

For the present study, and assuming a constructivist approach, our specific predictions are:

- (1) The general pattern of acquisition of NP will be consistent with the one revealed in Mariscal (1996). Initial phases will be characterized by high percentages of determiner omissions, fuzzy and variable forms of pre-noun vocalic elements and differences between particular nominal types as to their distribution with fillers and non-analyzed forms of adjectives.
- (2) The expected development from initial phases will be a decrease in variability between nominal types, together with a gradual production of different subtypes of determiners and a decrease of (gender)

TABLE 4. *Subjects: Ages and Mean Length of Utterance (MLU) in the first and last session of each cycle*

Subject	Cycle	Age	MLU	No of recordings (duration in minutes)
1 CARLOS	1	(1;10·15-1;11·05)	1·17-1·45	8 (240)
	2	(2;02·05-2;02·27)	2·20-2·31	8 (240)
	3	(2;05·12-2;06·08)	2·47-2·58	8 (240)
2 CLARA	1	(2;00·09-2;01·23)	1·23-1·48	9 (250)
	2	(2;05·24-2;07·07)	1·99-2·08	7 (240)
3 ANDREA	1	(2;01·20-2;02·29)	1·45-1·86	7 (240)
	2	(2;05·26-2;06·25)	2·46-2·24	9 (250)
4 ARTURO	1	(2;01·23-2;03·10)	1·49-1·63	8 (240)
	2	(2;05·14-2;06·14)	2·35-2·34	7 (240)

morphological errors in adjectives. Variability is expected because if children were operating with an abstract (categorical) representation, they would productively combine the determiners they produce with all of the nouns in their vocabularies. By contrast, if the children's use of determiners was limited, a high variability between the different noun types will be observed; that is, some of them will appear combined with determiners (or precursors) more frequently than others, with no grammatical reason for it. The latter case will constitute evidence in favour of lexical specificity or exemplar dependency that is not compatible with a categorical definition of noun targets as N, and of determiners (or precursor forms) as D.

- (3) Performance factors, as processing limitations in production derived from the length of the sentences, will not be the only explanation for the omissions of determiners in NP structures.

METHOD

Participants

The subjects were four middle-class Spanish children: two boys (Carlos and Arturo) and two girls (Clara and Andrea), with ages starting between 1;10 and 2;01 (see Table 4). Their families were contacted through a kindergarten in the north of Madrid, and they allowed the audio-recording of their children's language at home. All the children showed normal development and their only mother tongue was Spanish.

Procedure

The children were visited at their homes twice a week for one month by the main researcher, who had become well acquainted with each child through

TABLE 5. *Adjective elicitation task: Items and structure*

Masculine items	Feminine items
<i>oso</i> 'bear', <i>plato</i> 'dish', <i>peine</i> 'comb', <i>coche</i> 'car', <i>pez</i> 'fish', <i>sol</i> 'sun'	<i>silla</i> 'chair', <i>taza</i> 'cup', <i>pelota</i> 'ball', <i>cuchara</i> 'spoon', <i>flor</i> 'flower', <i>llave</i> 'key', <i>mano</i> 'hand'
Structure of the task	
1. Child: (ask for an item in the shop; for example <i>un oso</i> 'a (masc.) bear').	2. Adult: <i>De qué color es el oso?</i> 'What colour is the (masc.) bear?'
3. Child: (response) (<i>rojo</i> 'red (masc.)').	4. Adult: (gives the item to the child) <i>Toma</i> 'take it'.

previous visits to their school. Seven to nine recordings were made during this one-month period, defined as a 'cycle'. The first recording consisted of spontaneous language produced during child-mother interactions in free-play situations. From the second to the last session the main researcher interacted with each child, using the same set of toys. Some toys were duplicated and had different colours/sizes in order to elicit the production of adjectives and the quantifier *otro/a* 'another (masc./fem.)', which is one of the earliest non-article determiners produced by Spanish children. Spontaneous and elicited language (see below) was recorded.

Once the first cycle was completed, a second cycle was initiated after ten to twelve weeks (less than three months). If the analysis of the data showed that children did not accomplish acquisition criteria, then a third cycle was carried out, with the same scheme, after another ten- to twelve-week interval. Table 4 shows the ages, Mean Length of Utterances (MLU) in words and number of recordings for each child.

As we mentioned in the Introduction, due to the low frequency of adjectives in children's spontaneous language, 'the shop', an ADJECTIVE ELICITATION TASK was developed. The items selected for this task were nouns of masculine and feminine gender which represent prototypical and non-prototypical Spanish gender endings. Table 5 shows these items and the structure of the task inserted in a play situation in which the child went shopping for the objects matching the selected items.

The study is longitudinal because data were obtained in at least two developmental periods (two or three cycles). But it is not a classical longitudinal study because recording sessions were concentrated in order to obtain a dense corpus of nominal productions in a single period.

Recording and transcribing

The sessions were recorded using a wireless microphone (Shure System, model T connected to a TP-80 Aiwa recorder) that ensured high-quality

TABLE 6. *Coding system of NP productions in the contexts of the obligatory use of the determiner*

*0N:	The utterance includes a non-grammatical omission of the determiner; for example, <i>pé</i> instead of <i>el/un pez</i> 'the/a (masc.) fish'.
0N:	The utterance includes a grammatical omission of the determiner; for example, <i>quiero agua</i> 'I need water', asking for water.
vN:	The utterance includes a prenoun vowel, which can take different forms; for example, <i>epé, opé, (a/e)pé</i> for <i>el pez</i> 'the (masc.) fish'.
ART+N:	The utterance includes the adult form of an article, this being definite or indefinite; for example, <i>el pé</i> 'the fish' or <i>un pé</i> 'a fish'.
ODET+N:	The utterance includes the adult form of any other non-article determiner; for example, <i>ete pé</i> 'this fish', <i>oto pé</i> 'another fish'.
?	The utterance is an amalgam (non-classifiable) or includes an ambiguous form; for example, <i>te...tete</i> for <i>el chupete</i> 'the (masc.) dummy' or <i>este chupete</i> 'this dummy'.

sound. Transcriptions of the recording sessions were made, for all children, in standard Spanish orthography, and include all the children's and adults' utterances, as well as information on the non-linguistic context of interaction. It must be noted that the Spanish spelling system bears an almost one-to-one correspondence between phonetic and graphemic units. However, in order to transcribe fuzzy segments, especially prenoun elements that were not very well defined phonetically, we separated these two categories: (1) 'precise vowel' when it was adjusted to any Spanish normative vowel; for example, *e pé* (instead of *el pez* 'the (masc.) fish (masc.)'); and (2) 'non-precise vowel' when it was fuzzy or not adjusted to the norms; these cases were transcribed in parentheses, where the two closest normative vowels were included (for example, *(a/e) pé*) (see López-Ornat, 1997).

Coding and analysis

After the transcriptions were finished, all spontaneous utterances which included nouns produced in contexts of the obligatory use of a determiner were extracted. Then a coding system following CHILDES rules (MacWhinney, 1991) was developed. A dependent tier (%cod) was included in the transcriptions after every (pre)NP production. The coding system is shown in Table 6.

It is important to highlight that the use of two codes for the adult-like NP productions (ART+N and ODET+N) does not have a theoretical nor a linguistic basis, but a methodological one. Given the differential frequency of both types of utterances in children's early language (see Results below) it seemed pertinent to separate articles from the other subtypes of determiners.

Using the CLAN program COMBO (MacWhinney, 1991), data were extracted and included in a database for each subject and cycle. Data were ordered so that the coded productions of the same noun were grouped together and imitations were excluded from analysis.

Adjectives produced in the context of the elicitation tasks were extracted and coded according to their grammatical correctness. Three types of productions were possible: correct adjective, error of commission and overgeneralization error. For example, for the noun *mano* 'hand (fem.)', *(r)oja* 'red (fem.)' is the correct form, but **(r)ojo* '*red (masc.)' would be an overgeneralization; for the noun *pelota* 'ball (fem.)', *rojo* 'red (masc.)' would be an error of commission. The same codes were applied for adjectives produced spontaneously.

Production measures and acquisition criteria

The form and evolution of (pre)NP children's productions were analyzed, taking into consideration only those utterances produced in the contexts of compulsory use of a determiner. The following quantitative measures were computed:

- (1) Percentage of determiner omissions (*0N) and alternative structures (non-omissions: vN, ART+N, ODET+N) for every subject and cycle.
- (2) Quantitative analyses of the form and distribution of pre-noun vowels (vN) for masculine and feminine noun types.
- (3) The number of different determiners used by every child in each cycle to establish if the syntactic position of D was occupied by the different subtypes (articles, possessives ...) which form part of this category.

The percentage of non-grammatical omissions (*0N) allowed for the application of Brown's criterion (Brown, 1973). Therefore, determiner category (and NP) was considered to be acquired if it was produced in at least 90% of the contexts of compulsory use.

Also, a new production measure was developed as an index of the generality (or of the opposite, an index of lexical specificity) in the use of determiners in children's utterances. This Measure of Variability (4) – applied exclusively to utterances where the use of determiners is compulsory – was calculated for nouns produced by every child with a frequency of five or more times in the same cycle. The proportions of non-omissions or alternative productions $[vN + (ART + N) + (ODET + N)]$ to *0N were computed for every noun type produced per child and cycle. Taking this set of proportions as a base, means and standard deviations were obtained, and variability was calculated obtaining the quotient between both measures (SD/mean). This quotient is the Coefficient of

TABLE 7. *Evolution of omissions (*0N) and of alternative structures (vN, ART+N, ODET+N)*

Subject	Cycle	*0N (%)	vN (%)	ART+N (%)	ODET+N (%)	N total
1 Carlos	1	80.58	18.16	1.25	—	479
	2	35.25	62.27	2.47	—	607
	3	12.46	68.51	12.97	6.05	578
2 Clara	1	62.04	37.04	0.92	—	216
	2	9.25	47.98	33.53	9.24	173
3 Andrea	1	49.16	46.64	3.78	0.42	238
	2	8.65	35.58	52.40	3.36	208
4 Arturo	1	90.93	1.67	5.73	1.67	419
	2	51.38	4.97	36.74	6.91	362

Variation (CV), which allows for the comparison of variability in different samples (one per subject and cycle in this study). So, for example, the noun *pulpo* 'octopus (masc.)', produced by subject one 30 times in cycle one, had the form *0N 27 times, and the form vN 3 times; 3/27 (0.11) is the proportion of times this noun presents an alternative form to omissions. But for the same child, the N types *bruja* 'witch (fem.)' and *león* 'lion (masc.)' had the proportions 0/22 (0) and 22/46 (0.49), respectively. The comparison of these proportions (0.11, 0 and 0.49) can offer a 'flavour' of the variation that seems to characterize children's early production of NP. The CV obtained as the 'sum-up measure' of this variation considers all NP utterances per child and cycle and allows comparisons to be made between them. In contrast to the children's utterances, the adult-like productions of any of the target nouns would have always been a non-omission, so they would have led to a value of 1, and consequently variability, expressed in the CV, would have been 0.

RESULTS

(1) *NP structure and development: omissions of determiners (*0N) and alternative structures (vN, ART+N, ODET+N)*

Table 7 shows that the decrease of omission errors (*0N) was a common phenomenon for each of the four children. The percentage of omissions was higher during the first cycles, mainly for subjects one and four; at the end of the study (except for Arturo) Brown's criterion is reached.

For alternative structures to *0N, the following results were observed:

- (1) preNoun vowels (vN) were produced from early on, co-existing with omissions, although in variable proportions for each child.
- (2) Articles were the first kind of determiners used by the children (see ART+N column in Table 7). Differentiation between definite and

TABLE 8. *Relation between noun gender and type of NP production (% in parentheses)*

Subject	Cycle	*0N		vN + (ART + N) + (ODET + N)		CC (Coefficient of contingency)
		Nmasc.	Nfem.	Nmasc.	Nfem.	
1 Carlos	1	239 (76.36)	147 (88.55)	74 (23.64)	19 (11.44)	CC = 0.145 ($p < 0.01$)
	2	134 (29.52)	80 (52.29)	320 (70.48)	73 (47.71)	CC = 0.202 ($p < 0.01$)
	3	47 (13.95)	25 (10.37)	290 (86.05)	216 (89.63)	C = 0.054 n.s.
2 Clara	1	69 (61.06)	65 (63.11)	44 (38.94)	38 (36.89)	CC = 0.02 n.s.
	2	9 (10.46)	7 (8.04)	77 (89.53)	80 (91.95)	CC = 0.04 n.s.
3 Andrea	1	85 (61.15)	32 (32.32)	54 (38.84)	67 (67.68)	CC = 0.27 ($p < 0.01$)
	2	8 (8.16)	10 (9.09)	90 (91.84)	100 (90.91)	CC = 0.016 n.s.
4 Arturo	1	236 (95.16)	145 (84.79)	12 (4.84)	26 (15.20)	CC = 0.17 ($p < 0.01$)
	2	97 (48.74)	89 (54.60)	102 (51.25)	74 (45.40)	CC = 0.058 n.s.

indefinite articles was the first thing to be accomplished within the determiner's category. But, at first, articles appeared combined with particular nouns; that is, ART+N productions were linked to particular nouns produced by each child, and its use did not generalize immediately to other tokens. For example, for Clara *un zó* 'a sun (masc.)' and *un mono* 'a monkey (masc.)' were the only two productions with non-definite articles in cycle one; both were very frequent and early acquired nouns, learned and used in specific contexts.

- (3) Other kinds of determiners (ODET + N column) were produced much less frequently and later on.

In order to test if these alternative structures were produced productively with nouns of both grammatical genders, feminine and masculine nouns were analyzed separately. Table 8 presents data on this analysis where *0N productions are compared to utterances coded as vN, ART+N and ODET+N, grouped together. This data organization allows for comparisons to be made between non-grammatical utterances and productions considered (globally) more advanced.

Examining the data in Table 8, a clear imbalance can be observed in the first cycles for the majority of subjects (Carlos 1, Carlos 2, Andrea 1 and Arturo 1), in relation to the distribution of masculine and feminine nouns

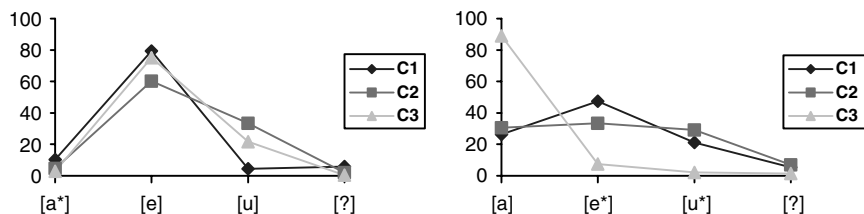


Fig. 1.1 & 1.2. Subject 1: Form and distribution of 'vN' for masculine (left) and feminine (right) nouns (C1=cycle 1; C2=2; C3=cycle 3).

with determiners or precursor forms. The coefficient of contingency (CC) was used as a measure for the relationship between the variables 'gender of nouns' and 'type of production' (*0N vs. vN/ART+N/ODET+N). The last column in Table 8 shows values and level of significance, computed for each child and cycle. The direction of this relation is not the same for the children. For subject one masculine nouns were distributed more frequently with the more advanced NP constructions, whereas for subjects three and four, the opposite relation was observed (i.e. noun phrases with feminine nouns were structurally more advanced than the ones with masculine nouns). The interpretation of these inter-subject differences will be treated in the Discussion of this paper.

Distribution and form of vN, ART+N and ODET+N

The difference between *0N productions and ART+N and ODET+N is clear; the former are non-grammatical, while the latter are correct, and when sufficiently generalized they indicate that the children's linguistic system does already have a morphosyntactic organization. However, vN utterances are more difficult to interpret. What is the representational status of these early forms? Why do they disappear later on? Is this due only to phonological reasons? In order to obtain an answer to these questions from our study, an analysis of their distribution and form was undertaken.

On the one hand, the position occupied by these forms is structurally correct. But on the other, the most frequent forms are vowels /a/, /e/, /u/, which could have been extracted from the adult forms *la* 'the (fem.)', *el* 'the (masc.)' and *un/una* 'a (masc./fem.)'; however, children also produce fuzzy forms (coded as '?'). Figures 1.1 to 3.2 show the distribution of these pre-noun vowels with masculine and feminine nouns. The symbol * (as in a*, e* and u*) marks an erroneous use of these vowels with masculine and feminine targets. So, for example, the distribution of vowel /a/, typical of feminine determiners, with masculine nouns is erroneous (see Figure 1.1.), whereas the vowel /e/ would be appropriate for these targets.

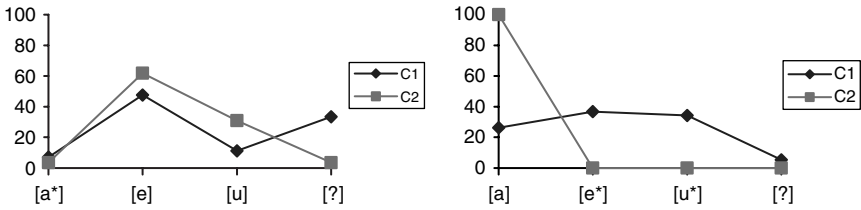


Fig. 2.1 & 2.2. Subject 2: Form and distribution of 'vN' for masculine (left) and feminine (right) nouns.

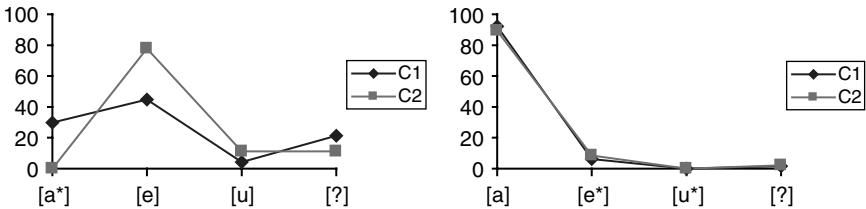


Fig. 3.1 & 3.2. Subject 3: Form and distribution of 'vN' for masculine (left) and feminine (right) nouns.

Inspection of the full set of figures shows that the distribution of these prenoun vowels with masculine and feminine nouns is not always adjusted to their grammatical gender. For subject one, it is only during the last cycle that prenoun vowels could be considered allophones of definite and indefinite articles. During cycles one and two this child used vowel /e/, with both masculine and feminine nouns (for example: *e ión* for *el león* 'the lion (masc.)', *e uma* for *la goma* 'the rubber (fem.)' (see Figures 1.1 and 1.2). Moreover, one single noun frequently appeared distributed with different vocalic forms (for *bruja* 'witch (fem.)' the utterances *e úja*, *a úja*, *? úja* were produced).

For subject two, vowel /e/ was also the most frequently uttered, both with masculine and feminine nouns. This child also used vowel /u/ in combination with masculine and feminine targets (for example: *u aza* instead of *una casa* 'a house (fem.)' and *u nena* instead of *una nena* 'a girl (fem.)'. In cycle two, the system experienced an important convergence with the prescriptive model (see line for C2 in Figures 2.1 and 2.2), the percentage of errors being very low.

However, for subject three, the pattern observed was the reverse. The distribution of prenoun vowels was adjusted to the correct or adult-like form mainly for feminine nouns; this girl combined vowel /a/ with 90% of the feminine nouns (see Figure 3.1), and with 29.8% of masculine ones (for example, *a gólo* instead of *el gorro* 'the hat (masc.)', *a pafé* instead of *el café*

'the coffee (masc.)'. These data together with the percentage of fuzzy vowels distributed with masculine nouns do not permit the interpretation of these vowels as allophones of articles. By contrast, the change was evident in cycle two (compare Figures 3.1 and 3.2.), where data related to the distribution and form of preN vowels with nouns of both genders would justify the consideration of such vowels as allophones of definite articles.

For the interpretation of all these data it is also necessary to consider the phonological differences between the definite and indefinite articles in Spanish. The indefinite articles (*un, una*) are stressed and the feminine form is bisyllabic. When speaking of these kinds of articles, articulatory reasons could be applicable in order to understand why children used the vowel *ú* with masculine and feminine nouns.

Finally, subject four presented a particular pattern of acquisition, very different from that of the other children. His vN productions were very scarce in cycles one and two; only 7 cases were observed in cycle one and 18 in cycle two (see data in Table 7). The only forms produced were /a/ and /e/ combined erroneously with masculine and feminine nouns (*a + Nmas. and *e + Nfem.). From the beginning this child used the forms *uno* – the masculine numeral pronoun instead of the correct indefinite article *un* – and *una*, both combined with nouns and with a numeral or quantifier function (for example: **uno coche* 'one car (masc.)', *una niña* 'a girl (fem.)'). The particular pattern followed by this child will be analysed in the Discussion section of this paper.

However, putting aside the individual differences encountered, what is worth emphasizing is that coinciding with the decrease of *0N productions and the increase of ART+N and ODET+N (see again Table 7), vN utterances gradually become allophones of ART+N, pushing the system to its convergence with the adult-like (i.e. grammatical) one.

Let us now see the data related to the distribution of articles and other kinds of determiners with N, when the former already had an adult form. Table 9 demonstrates the following:

- (1) The earliest determiners used are articles.
- (2) The first differentiation within D category concerns articles; spontaneous NP produced by subjects one and two (the less advanced) show an incipient differentiation between definite and indefinite subtypes.
- (3) The number of subtypes increases from one cycle to the next for all the children, but as Table 7 shows, their frequency is very low.

Given these results, it is necessary to account for the reasons why articles are the first determiners to be learned/used by children.

TABLE 9. *Number of different types of determiners per subject and cycle*

Subject	Cycle	Determiners	
		Number	Types ^a
Carlos	1	0	preN vowel (and some nasal) forms considered only precursors of articles.
Carlos	2	1	Only articles and minimum differentiation definite-indefinite.
Carlos	3	4	Articles, possessives, demonstratives and the quantifier <i>otro/a</i> 'other (masc./fem.)'.
Clara	1	1	Only articles.
Clara	2	4	Articles, demonstratives, possessives (3 rd pers.) and quantifier <i>otro/a</i> .
Andrea	1	2	Articles, quantifier <i>otro/a</i> .
Andrea	2	5	Articles, quantifier <i>otro/a</i> , numeral <i>dos</i> , possessive (1 st pers.), exclamative <i>qué</i> .
Arturo	1	4	Articles and quantifier <i>otro/a</i> ; possessive (1 st pers.) and demonstratives.
Arturo	2	6	Articles, quantifier <i>otro/a</i> , exclamative, possessive (1 st pers.), demonstratives and numeral <i>dos</i> .

^a Frequency of non-article determiner use is very low in all children, as shown in Table 7.

TABLE 10. *Productions of noun types with frequency ≥ 5 and values of the Measure of Variability*

Subject	Cycle	N ^o of N types	No of noun productions	Measure of Variability (Coefficient of Variation)
1 Carlos	1	25	358	1.2821
	2	34	471	0.3760
	3	36	444	0.1585
2 Clara	1	15	102	0.7144
	2	9	52	0.1601
3 Andrea	1	17	126	0.5892
	2	15	111	0.1058
4 Arturo	1	24	322	1.4397
	2	25	210	0.6883

Variability in nominal productions

Table 10 presents data on the computation of the Measure of Variability developed for this research. Variability was very high in the earlier phases (Carlos 1 and Clara 1), and decreased from one cycle to the next (see development of subjects one, two and three in Table 10). A way of obtaining an image of these values is by drawing a simple XY-diagram as in the

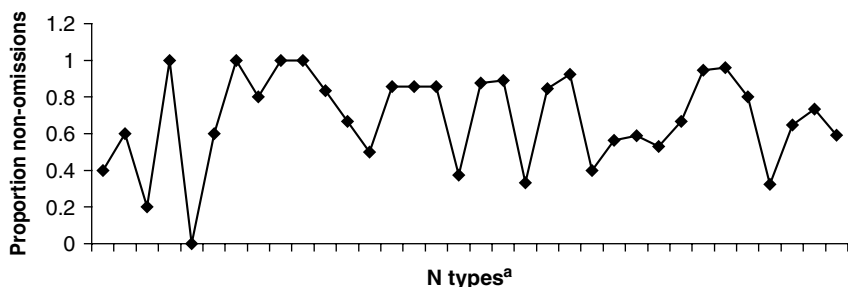


Fig. 4. Subject 1, cycle 2: Inter-type variability.

^a Noun types are ordered by frequency (ascendant order). Each point on the X-axis stands for a type. For example, the last point stands for *león*, produced 46 times by subject 1 in cycle 2. Written notation of each type does not appear due to lack of space.

example presented in Figure 4, where each nominal type (X-axis) is assigned a value which represents the proportion of times that it is produced as vN or ART+N or ODET+N. By joining these points, a variability profile is obtained, which reflects inter-type differences. The variability of the values is immediately noticeable. For reasons of space we have not included the rest of the graphs, but it is interesting to note that figures corresponding to first cycles are characteristically uneven.

This measure picks up intra-individual variability in the production of each child's repertoire of nouns. Examples of this phenomenon can be observed in the same type produced in different sessions during a cycle as in (a), in the same session (b), and even in the same utterance (c).

- (a) Carlos, cycle one, session one: *u pé* for *u(n) pez* 'a fish (masc.)', coded as vN and uttered in answer to the question 'What is this?'; session four: *pé: oto* for *pez roto* – taking a toy fish that is broken and coded as *0N
- (b) Carlos, cycle one: *pupo* – making a toy octopus eat a fish, and later on in the same session *e pupo* – putting the toy octopus into a bag.
- (c) Carlos, cycle one, session two: *pé, e pé*, pointing to a little fish.

From the first to the last cycles, the values for each N type approximated 1 (the corresponding adult value) and the percentage of non-grammatical omissions decreased at the same time as variability did (see Tables 7 and 10).

The developmental profile of subject one, the most extensively followed, seems to indicate that variability decreased continuously, but not gradually. The change from cycle one to cycle two appears to be higher than the change to cycle three.

Subject four, presented, once again, a different profile, whose main characteristics in comparison with subjects one, two and three are listed below.

TABLE II. *Elicited adjectives produced by each child*

	Subjects							
	Carlos		Clara		Andrea		Arturo	
	n ^a	errors	n	errors	n	errors	n	errors
Cycle 1	—	—	6	2	9	4	9	0
Cycle 2	9	4	12	0	8	0	12	0
Cycle 3	12	4	—	—	—	—	—	—

^a Total number of elicited productions; the task consists of twelve items.

Acquisition sequences: Intersubject differences

For subjects one, two and three, the above-mentioned data reveal the following sequence of acquisition of NP: $*0N=>vN=>ART+N \text{ } \& \text{ } ODET+N$.

By contrast, subject four scarcely used vN structures. The most characteristic trait of his language was the high percentage of determiner omissions; in cycle two he did not reach Brown's criterion. The sequence of acquisition for this child was: $*0N=>ARTidf. (*uno/una)+N \text{ } \& \text{ } ODET+N$.

In cycle one he produced only some indefinite articles; however, he produced **uno* (quantifier pronoun) instead of *un* (indefinite article). It is remarkable that in cycle two all ART+N productions ($n=79$) had this form ($*uno+N$). From cycle one to cycle two new determiners were incorporated, although their use was not too generalized.

(2) Morphological gender agreement: Spontaneous and elicited adjectives

In this section we will present mainly qualitative and global analyses of morphological patterns observed in the production of spontaneous and elicited adjectives.

The number of adjectives produced spontaneously was very low when compared to the high frequencies of nominal types in NP structures. Even in the elicitation task, it was not always possible to obtain the full set of adjectives. Table 11 shows data related to this task for the four children.

Taking into account both spontaneous and elicited tokens, the patterns obtained are:

- (1) A common sequence of acquisition is revealed: non-analyzed-units $=>$ minimal productivity $=>$ generalization and errors $=>$ absence of errors.
- (2) Errors were not very frequent (see Table 11), but the majority consisted in the production of the masculine gender morpheme *-o* instead

TABLE 12. *Summary: NP structure and gender agreement with adjectives*

	Brown's criterion for DET + N structure (only 10% omissions)	N° of determiner types ^b	Morphological agreement [adjectives]
CARLOS 1	NO	0	Non-analyzed-units
CARLOS 2	NO	1	Minimal productivity
CARLOS 3	NEAR ^a	4	Productivity + errors
CLARA 1	NO	1	Minimal productivity
CLARA 2	YES	4	No errors
ANDREA 1	NO	2	Productivity + errors
ANDREA 2	YES	5	No errors
ARTURO 1	NO	2	Productivity + 1 error
ARTURO 2	NO	6	No errors

^a Percentage of omissions of D is 12.4%.

^b For all subjects frequency of Det + N productions was lower than 10%, as shown in Table 7.

of the feminine one *-a*. The following are some examples: (a) **malo* 'bad (masc.)' applied to *bruja* 'witch (fem.)' (Carlos, cycle one); (b) **(r)oto* 'broken (masc.)' for *caja* 'box (fem.)' (Andrea, cycle one; Arturo, cycle one). It is necessary to state that in the first cycles, particular adjectives tended to appear only with particular nouns too; that is to say, children tended to produce adjectives tied to only one or very few nouns, learned in a very specific context. For example, Clara restricted the production of *fría* 'cold (fem.)' only to the noun *agua* 'water (fem.)', learned in bath routines.

- (3) Errors affected both prototypical and non-prototypical nouns (for example: *pelota *amarillo* 'yellow (masc.) ball (fem.)', *flor *rojo* '*red (masc.) flower (fem.)'. Also, non-prototypical but very frequent nouns, such as *mano* 'hand (fem.)', were produced correctly from an early stage (Arturo and Clara, cycle one, produced *mano roja* 'red (fem.) hand (fem.)' and *flor roja* 'red (fem.) flower (fem.)').
- (4) Intra-type variability was observed in spontaneous productions: one single noun could be distributed with both masculine and feminine adjectives. For example: Clara, cycle one, produced the adjectives *bueno/buena* 'good (masc./fem.)' for the noun *enanito* 'dwarf (masc.)', and Arturo (cycle one) said *campana rota* 'broken (fem.) bell (fem.)' and *campana *roto* '*broken (masc.) bell (fem.)'.
- (5) Gender agreement errors disappeared (except for sporadic over-generalizations) when determiner omissions in NP were very low (10% or less) (see Table 12).

In the next section these patterns will be discussed in relation to the main hypothesis being tested by this work.

DISCUSSION

The results of this study provide evidence which is consistent with constructivist approaches to the acquisition of grammatical knowledge by children (López-Ornat, 1994; Pine & Lieven, 1997; Tomasello, 2000; Mariscal, 2001), specifically to the acquisition of NP structure and gender agreement in Spanish. At the same time, the focus on intra-individual variability has provided interesting data, and also a new way of looking at this developmental phenomenon, which we consider crucial in order to understand the dynamics of the language acquisition process.

NP structure and development: Adult-like from the beginning?

Firstly, as predicted by our hypothesis number (1), the general pattern of NP acquisition found in the subjects of our sample is consistent with the one found in the single-case classical longitudinal study by Mariscal (1996), and with other studies in Spanish (Aguirre, 1995; López-Ornat, 1997; Lleó, 1997; López-Ornat, 2003). A gradual decrease of agrammatical determiner omissions and an early presence of productions that include a prenoun vowel were observed.

During the last cycles, the reduced percentage of omissions and the correct use of articles and some other (although not very frequent) subtypes of determiners – with nouns of both masculine and feminine gender – allow us to attribute certain syntactic knowledge to the children. However, we have found different pieces of empirical evidence related to the first cycles of this study that raise doubts about the validity of innatist arguments (see, for example, Aguirre (1995) for Spanish data) which attribute syntactic categories to young children from the beginning or very early on. On the one hand, prenoun forms do not distribute according to the adult-like model. Some early productions, as for example *e pé* – referring to *el pez* ‘the fish (masc.)’ – could give the impression of an adult-like kind of utterance. However, results from the detailed analysis of form and distribution of vN productions with masculine and feminine nouns provide arguments against such an idea. Subtle individual differences or preferences for certain preN vowel forms, and imbalances dependent on the gender of the noun shown do not seem compatible with a categorical definition of these linguistic units. Inter-subject differences can be expected in any learning process and are dependent on the particular acquisition process of every child. These differences can be explained by alluding to probabilistic or partial-kind knowledge about formal and distributional properties of the particular noun tokens that are acquired by each subject (see also Pizzuto & Caselli, 1992).

During the last cycles, when omissions decrease and the correct production of articles (or allophones) and other determiners increases, these

differences and imbalances disappear. Only then does the system seem to converge with the norms of the language.

On the other hand, although the percentage of determiners used grows from cycle to cycle for all subjects in our study, only articles were used from the beginning. Some other subtypes of D are used in the last cycles but, as our results show, very infrequently.

The early use of articles in child language can be explained by means of frequency factors. However, even accepting this restriction, if children had a syntactic category, the use of other subtypes of determiners could be expected from earlier on. That is, formal knowledge about an individual subtype of D should generalize to at least some other members of the category. Why do young children not use possessive or demonstrative determiners while at the same time these lexical forms, used as pronouns, are found in their language (for example, *ete*, *este* 'this (masc.)' pointing to something, or *oto*, *otro* 'another (masc.)' asking for more candies)? It is precisely the generalization expected by the syntactic-like type of knowledge of this category that is not found in the data. Even in the last cycles only a limited overlap between articles and other determiners in NP children's productions was found (see also Pine & Lieven (1997) for a comparison between definite and indefinite articles used in English).

Moreover, precise vowels produced before nouns in the first cycles are closer to definite than to indefinite articles. And as we mentioned before, definite articles are the most frequent determiners in Spanish.

Beside the influence of frequency on the learning process, vN productions show that children make distributional and phonological analyses of input (Veneziano & Sinclair (2000) used the term 'superficial analysis'). Several studies (e.g. Gerken, Wilson & Lewis, 2005) have shown very early abilities of distributional and formal analysis prior to production. Children, therefore, can detect, store and analyze units in the input without assigning clear function or meaning, especially if they are combined with the most ubiquitous of linguistic units, that is, nouns. Clearly, vN units take some of the properties of the future D category, but, as has been shown by our data, only some. We consider the term 'proto-articles' (Lleó, 1997, 2001) appropriate in order to describe them, as long as it does not imply any kind of system-wide syntactic category. These vN utterances, when neither their functions nor their forms are completely adjusted to the normative model, could be explained, as some authors have proposed, by alluding to phonological bootstrapping processes (Peters & Menn, 1993; Gerken, 1996; López-Ornat, 1997, 2003).

Variability in nominal productions

The focus of this research on intra-individual variability at the level of each child's particular NP productions has provided a set of data that deserves

theoretical consideration. On the one hand, the kind of longitudinal design used for the gathering of data has enabled us to compile a corpus of NP productions in a single developmental period, large enough to calculate an index of variability. These corpora show the co-existence of a variety of NP utterances, from *0N to ODET+N, and the value of the variability measure used in this study decreases as the acquisition process advances.

Variability, as has been measured in this study, indicates a differential 'treatment' of the N types produced by each child. That is to say, in contexts of compulsory determiner use, some exemplars have higher probabilities of being produced as 'more grammatical' (as vN, ART+N or ODET+N) structure than others. If we take this variability as an expression or index of partial grammatical (morphosyntactic) knowledge, a decrease of inter-type variability is the expected phenomenon – as stated in our hypothesis number (2) – and that is what is reflected by our data as the grammar acquisition process advances.

Some researchers who have focused on the study of language acquisition from a generativist point of view (Valian, 1991; Aguirre, 1995) attribute this lack of generalization in determiner use to performance factors. However, such affirmations have not been empirically supported, but rather seem to reflect 'the decision of analyzing data at the level of the hypothesized syntactic categories' (Pine, Lieven & Rowland, 1998: 38). Performance errors can always be adduced as an argument to account for variability data, thereby maintaining the hypothesis of a hidden competence or preformed knowledge underlying children's utterances.

What kind of performance factors could be responsible for inter-type variability? It could be hypothesized that when young subjects face the production of multiword utterances, prosodically weak units might be omitted. There is an abundant corpus of evidence indicating that children who omit functional words, such as articles, do process such units at certain formal levels (i.e. positionally or distributionally) (Gerken & McIntosh, 1993), although frequency and familiarity of nouns also influence the probability of article production or omission (Boyle & Gerken, 1997).

This study does not cover the analysis of the interaction between the length of utterances in which each NP is inserted and the production/omission of determiners in the contexts of obligatory use. However, it is important to highlight that during the first cycles, MLUs of our subjects did not reach the value of 2; that is, the majority of our children's utterances were not long enough to allude to this kind of performance limitations. In Mariscal (1997) we analyzed the relationship between the number of syllables in each N type and the form of the NP utterance in which it was inserted. The results were rather unclear. For subject one, there was a significant relationship between these variables; monosyllabic nouns were produced more often as vN or ART+N than multisyllabic

ones. But for subjects two, three and four this relationship was non-significant.

However, early productions of vN structures as well as first productions of indefinite articles seem to be linked to particular items, irrespective of their number of syllables. This evidence of lexical specificity constitutes a kind of local effect that does not seem compatible with the limited performance account.

Inter-type variability could possibly be accounted for by a combination of lexical specificity effects, a fuzzy or partial knowledge (distributional and/or prosodic) on what WILL constitute the D category, and (only in some cases) by articulation effects due to the length of utterance. In order to disentangle this set of variables, experimental designs would be necessary, although difficult to carry out with young children. What we have tried to emphasize throughout this paper is the necessity of considering intra-individual variability in language production as a crucial phenomenon. Variability, as other researchers have maintained (van Geert & van Dijk, 2002), is not necessarily an index of some kind of error. Even though the measure of this phenomenon in our study has been quantitatively simple, it was intended to show that variability can be considered an indicator of development. That is to say, the observed decrease of variability could be a sign of the convergence of the system with conventional or grammatical forms.

What this set of results shows seems to be compatible with a characterization of the acquisition process of NP and gender agreement as a gradual and progressive construction, yielded by the interaction between formal (and language-dependent) properties of the input and the processing abilities of the learners. In agreement with other authors who have studied the process of acquisition of NP and gender agreement from other perspectives (see, for example, Sicuro Corrêa & Name (2003) for Portuguese), we consider it necessary to hypothesize the use of early and already well-attested children's processing abilities and their application to the 'task' of distinguishing morphophonological classes within the category of determiners. This task is faced at a very early stage by Spanish children, as is shown by present and previous data. But what the present study does not share with innatist or generativist approaches is the need to call for a special computational system that would operate upon the formal features of a functional category such as D, considered as a unitary entity. Contrary to this last position, what the results of this study seem to highlight is the 'progressivity' in the construction of NP and gender agreement and its dependence on a process of learning.

Inter-subject differences

Another finding derived from this research concerns the inter-subject differences in the general pattern of NP acquisition. We have already

mentioned subtle individual differences related to the form of vN in the cases of subjects one, two and three. For subject four, however, the data showed a different pattern of acquisition. This child did not use the intermediate structure vN but used as his first determiners the indefinite articles (*un/una*) and the pronominal form **uno* – used erroneously instead of the masculine indefinite article *un*. That is to say, this child did not use non-salient and opaque forms as definite articles or its predecessor forms, but selected stressed units with a more clear-cut function or meaning, as the numeral *uno*, used in D position. It is possible that he could perceive articles in NP structures, but, for unknown reasons, was not producing them. This occurred even in the last cycle, when there were no gender agreement errors in his productions.

We do not have a precise explanation for this particular pattern of development, but inter-subject differences are consistent with a dynamic system view of language acquisition, which mainly emphasizes the process, and it is more compatible with constructivist models than with innatist ones.

Evidence from gender morphology

Subject four's error **uno* (used instead of the indefinite article *un*) is clearly related to morphological gender marks in Spanish. This error is coherent with the pattern observed in elicited adjectives, where all our subjects tended to produce masculine adjectives combined with some feminine nouns during first cycles.

In contrast with evidence provided by other authors (Pérez Pereira, 1991), we do not consider it necessary to conceive the masculine form as the unmarked one since, in spontaneous language, the opposite error pattern is also observed; for example, productions such as *nene *mala* 'child (masc.) *bad (fem.)'. Children tend to learn their first adjectives linked to particular nouns – in the example, *mala* used for stepmother was learned in the context of the Snow White tale. During initial phases these forms are only used as non-analyzed units, as happens with other kinds of bound morphemes in Romance languages. The reason for the absence of feminine instead of masculine errors in the elicitation task could be explained by the form of the eliciting question used: *De qué color es +NP?* 'What colour is +NP?'. This question could be perfectly answered by always using a masculine adjective if the listener attends to the word *color*, which is masculine, and not to the noun in the NP.

Another piece of evidence concerned with gender morphology worth mentioning is the absence of errors in certain non-prototypical nouns. For example, *mano*, which is feminine, was produced correctly in combination with feminine adjectives by all of our subjects, whereas learners of Spanish

as L2 tend to produce an overgeneralization error in this case (for example, *mano* **pequeño* 'hand (fem.) *small (masc.)). In order to explain such evidence it is necessary to allude again to frequency and lexical-specific factors, which tend to prevent errors in L1 acquisition, as has been attested by different researchers.

The results of this study also indicate that the decrease in determiner omissions (or, its reverse, the generalization of X+N structures) precedes the acquisition of gender morphology. From the micro-analysis undertaken in this study, we could conclude that such an order derives from a gradualist kind of logic: the acquisition pattern observed during this process is explained by the particular development of the system. Therefore, the analysis of positional/distributional patterns, implicit in vN productions, is one of the first operations within the system's reach. The discovery of the relationships between lexical items which can occupy the same position is necessarily a subsequent process, and it requires a more detailed analysis of the form and functions associated with them. Eventually, as a product of these later analyses, the use and knowledge of gender morphology begins to emerge.

Summarizing and linking our results to the comparison between L1 and L2 learners of Spanish that was mentioned in the Introduction, children's language acquisition could very well be described as a gradual but uneven process, which advances as the children integrate different pieces of evidence (phonological, distributional, functional ...) and establish more and more relationships and regularities between them. As we proposed elsewhere (Mariscal, 1997, 2001), it is the very developmental dimension of the process which contributes to the explanation of it (Elman, 1993). Due to the temporal unfolding of the acquisition process, to the condition of 'starting small', the system-which-acquires-the-language is able to gradually tackle the learning of different aspects related to the knowledge of grammar. First, there will come all the 'superficial' aspects (Veneziano & Sinclair, 2000), then the more detailed ones (phonetic and morphological features). Development imposes such an order that the process of grammaticalization is made possible and more accessible without assuming innate or specific knowledge. The final result of this process, but not necessarily its starting point, is the emergence of more and more abstract representations which support the correct use of new items inserted in particular grammatical structures. As this kind of development is not followed by L2 learners, we could hypothesize that this is one of the reasons for the differences found in the learning processes and for the persistence of errors. We consider that the above-mentioned dynamic view of the process, which emphasizes its changing nature and its variability, not only provides a good metaphor for the explanation of language development, but it also promotes a more exploratory approach to the data.

REFERENCES

- Aguirre, C. (1995). La adquisición de las categorías gramaticales en español. Unpublished doctoral dissertation, Universidad Autónoma de Madrid.
- Alameda, J. R. & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Universidad de Oviedo.
- Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bottari, P., Cipriani, P. & Chilosi, A. M. (1993). Protosyntactic devices in the acquisition of Italian free morphology. *Language Acquisition: A journal of Developmental Linguistics* **3**, 327–69.
- Boyle, M. K. & Gerken, L. A. (1997). The influence of lexical familiarity on children's function morpheme omissions. A non-metrical effect? *Journal of Memory and Language* **36**, 117–28.
- Brown, R. (1973). *A first language. The early stages*. Cambridge, MA: Harvard University Press.
- Capdevila, M. & Serra, M. (1996). La adquisición de las categorías funcionales en niños catalanes. In M. Pérez Pereira (ed.), *Estudios sobre la adquisición del Castellano, Catalán, Euskera y Gallego*. Universidade de Santiago de Compostela, Servicio de Publicaciones.
- Dore, J., Franklin, M. B., Miller, R. T. & Ramer, A. H. (1976). Transitional phenomena in early language acquisition. *Journal of Child Language* **3**, 13–19.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition* **48**, 71–99.
- Fernández-García, M. (1999). Patterns in gender agreement in the speech of second language learners. In J. Gutiérrez-Rexach & F. Martínez-Gill (eds), *Advances in Hispanic linguistics*, 25–39. Somerville, MA: Cascadilla Press.
- Gerken, L. A. (1996). Phonological and distributional information in syntax acquisition. In J. L. Morgan & K. Demuth (eds), *Signal to syntax. Bootstrapping from speech to grammar in early acquisition*, 42–61. Mahwah, NJ: LEA.
- Gerken, L. A. & McIntosh, B. J. (1993). The interplay of function morphemes and prosody in early language. *Developmental Psychology* **29**, 448–57.
- Gerken, L. A., Wilson, R. & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language* **32**, 249–68.
- Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry* **22**(1), 27–62.
- Heinen, C. & Kadow, H. (1990). Acquisition of French. In J. M. Meisel (ed.), *Two first languages: Early grammatical development in bilingual children*, 48–71. Dordrecht: Foris Publications.
- Hernández Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Madrid: Siglo XXI.
- Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.
- Lieven, E., Behrens, H., Spears, J. & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language* **30**, 333–70.
- Lleó, C. (1997). Filler syllables, proto-articles and early prosodic constraints in Spanish and German. In A. Soarce, C. Heycock & R. Shillcock (eds), *Proceedings of GALA'97 Conference on Language Acquisition*, 251–56. Edinburgh: The University of Edinburgh.
- Lleó, C. (2001). Early fillers: Undoubtedly more than phonological stuffing. *Journal of Child Language* **28**, 262–65.
- López-Ornat, S. (1994). La adquisición gramatical: Un esquema. In S. López-Ornat, Fernández, P. Gallo & S. Mariscal (eds), *La adquisición de la lengua española*, 101–126. Madrid: Siglo XXI.
- López-Ornat, S. (1997). What lies in-between a pre-grammatical and a grammatical representation? In W. R. Glass & A. T. Pérez-Leroux (eds), *Contemporary perspectives on the acquisition of Spanish*, 3–20. Somerville, MA: Cascadilla Press.

- López-Ornat, S. (2003). Learning earliest grammar: Evidence of grammar variations in speech before 22 months. In S. Montrul & F. Ordóñez (eds), *Linguistic theory and language development in Hispanic language*, 254–74. Somerville, MA: Cascadia Press.
- López-Ornat, S., Fernández, A., Gallo, P. & Mariscal, S. (eds) (1994). *La adquisición de la lengua española*. Madrid. Siglo XXI.
- MacWhinney, B. (1991). *The CHILDES project: Tools for analyzing talk*. Hillsdale, NJ, Erlbaum.
- Mariscal, S. (1996). Adquisiciones morfofonosintácticas en torno al sintagma nominal: El género gramatical en español. In M. Pérez Pereira (ed.), *Estudios sobre la adquisición del Castellano, Catalán, Euskera y Gallego*, 263–72. Universidade de Santiago de Compostela, Servicio de Publicaciones.
- Mariscal, S. (1997). El proceso de gramaticalización de las categorías gramaticales en Español. Unpublished doctoral dissertation, Universidad Autónoma de Madrid.
- Mariscal, S. (2001). Es ‘a pé’ equivalente a DET+N? Sobre el conocimiento temprano de las categorías gramaticales. *Cognitiva* 13(1), 35–59.
- Pérez Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language* 18, 571–90.
- Peters, A. M. & Menn, L. (1993). False starts and filler syllables: Ways to learn grammatical morphemes. *Language* 69(4), 742–77.
- Peters, A. (2001). Filler syllables: what is their status in emerging grammar? *Journal of Child Language*, 28, 229–42.
- Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts? *Language* 53(3), 560–73.
- Pine, J. & Lieven, E. V. (1997). Slot and frame pattern and the development of the determiner category. *Applied Psycholinguistics* 18(2), 123–38.
- Pine, J., Lieven, E. V. & Rowland, C. (1998). Comparing different models of the development of the English verb category. *Linguistics* 36, 4–40.
- Pine, J. M. & Martindale, H. (1996). Syntactic categories in the speech of young children: The case of the determiner. *Journal of Child Language* 23, 369–95.
- Pizzuto, E. & Caselli, M. C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of Child Language* 19, 491–557.
- Sicuro Corrêa, L. & Name, M. C. (2003). The processing of determiner-noun agreement and the identification of gender of nouns in the early acquisition of Portuguese. *Journal of Portuguese Linguistics* 2(1), 19–43.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* 74, 209–253.
- Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition* 40, 21–81.
- van Geert, P. & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development* 25, 340–74.
- Veneziano, E. & Sinclair, H. (2000). The changing status of ‘filler syllables’ on the way to grammatical morphemes. *Journal of Child Language* 27, 461–500.