

## THE EMERGENCE OF DEFINITE AND INDEFINITE ARTICLES IN THE SPEECH OF AN ENGLISH-ITALIAN BILINGUAL CHILD

*Ludovica Serratrice<sup>1</sup>*  
*University of Edinburgh*

### 1. Introduction

The focus of the present study is on the emergence of definite and indefinite articles and the associated D features (definiteness, number and gender) in the speech of Carlo, an English-Italian bilingual child, between the ages of 22 and 28 months.

The analysis of the data will highlight the importance of the role played by morphophonological regularities in the input and the reliance on rote-learned frozen forms. It will be shown how the emergence of the functional category D is very much language-specific and cannot be taken to be a prerequisite for acquisition but must rather be viewed as the result of different acquisitional strategies.

### 2. The subject and the data

Carlo is a second-born English-Italian bilingual child whose mother is a native speaker of Italian and whose father is a native speaker of American English. Carlo was born in Scotland where he lives with his parents and two elder brothers.

The data was collected separately for English and Italian during 14 weekly 30 minute-videorecorded sessions while Carlo played with a bilingual Italian-English investigator or with a monolingual English-speaking adult. The adult interlocutors always used their own native language with Carlo and language-switching was kept to a minimum.

The data was transcribed in CHAT format by the investigator and coded for morphological and syntactic analyses.

#### 2.1. What is to be acquired?

There are minimally two requirements common to the instantiation of D in both English and Italian: firstly, the realisation of the linear order “Determiner + Noun” (i.e. the realisation of the head-first parameter); secondly, the encoding of formal features such as definiteness and number, and in Italian the feature [gender] must also be included. The claim of this study is that the realisation of abstract functional D features is triggered by the

---

<sup>1</sup>University of Edinburgh, Department of Linguistics, Adam Ferguson Building, 40 George Square, Edinburgh EH8 9LL, UK; Telephone: 00 44 131 6503961, Fax: 00 44 131 6503962; E-mail: ludovica@ling.ed.ac.uk

acquisition of the lexical items filling the D position. The focus in the present work is on definite and indefinite articles and the D features they spell out.

## 2.2. The article systems in English and Italian

As is well known in English there is one definite article *the* (unspecified for number and gender), two singular indefinite articles *a/an* (*a* before consonants and *an* before vowels) and one plural indefinite: *some*. The English-acquiring child has only 4 different forms of the article system to learn. In Italian the situation is somewhat more complicated:

Table 1. The Italian definite article system

	masc.	masc.	fem.
sing.	il	lo	la
plur.	i	gli	le

Table 2. The Italian indefinite article system

	masc.	masc.	fem.
sing.	un	uno	una
plur.	dei	degli	delle

Note that there are two singular forms and two plural forms of the masculine article. All simple-consonant-initial words require *il* (*i* for the plural); words beginning with *s+* consonant, *z*, *ps* and *gn* require *lo* (*gli* for the plural). Thus the Italian-acquiring child must learn 12 different forms and their morphophonological distribution before s/he can master the article system.

## 3.1. The emergence of definiteness, gender and number in Carlo's speech

In this section we will focus on the distribution of definite and indefinite articles in Carlo's noun phrases and on the associated features of *definiteness*, *gender* and *number*. Data come from a total of 14 videorecorded free-play sessions of approximately 30 minutes each, 7 for Italian and 7 for English, between the ages of 1;10.0 and 2;4.10. In Tables 3 and 4 below an overview of production of definite and indefinite articles in the two languages is provided:

Table 3: Carlo's production of articles in English between 1;10.0 and 2;4.6

C.'s age	MLU	∅ articles	def. articles	indef. articles	total % in obligatory contexts
(1) 1;10.0	1.156	11	0	0	0
(2) 1;10.21	1.378	24	0	0	0
(3) 2;0.1	1.393	11	0	0	0
(4) 2;0.22	1.204	25	2	0	7.4
(5) 2;2.11	1.909	19	0	4	17.3
(6) 2;2.22	2.042	18	6	30	66.6
(7) 2;4.6	2.226	16	11	21	66.6

Table 4: Carlo's production of articles in Italian between 1;10.7 and 2;4.10

C.'s age	MLU	∅ articles	def. articles	indef. articles	total % in obligatory contexts
(1) 1;10.7	1.097	74	0	1	1.3
(2) 1;10.28	1.149	102	0	0	0
(3) 2;0.1	1.200	21	0	0	0
(4) 2;0.22	1.868	106	11	6	13.8
(5) 2;1.21	1.951	4	8	30	90.4
(6) 2;3.7	2.173	2	28	13	95
(7) 2;4.10	2.598	4	11	7	81.8

By considering the MLUs for both languages we observe that the child is in the transition phase between the one-word and the two-word stage. His production can be divided into two successive phases: the first one, between 1;10.0 and 2;2.11 for English and between 1;10.7 and 2;0.22 for Italian, characterised by the total absence of articles, followed by a second one where articles gradually appear reaching 66% in English in the last two sessions, and an average of 88.6% for Italian in the last three sessions.

In Phase I the features associated to the D system are not represented in C.'s noun phrases; in the absence of articles, the distinction between definite and indefinite reference obviously cannot be made. Number distinctions are equally absent, only singular forms appear at this stage; as for gender, its expression is strictly correlated to the presence of an agreeing modifier, if such a modifier is not present there is no way to establish if any gender categorization is operational at all in the child's system.

Only when gender and number agreement are instantiated in the form of grammatical concord between an article, or other modifier, and a noun, can we speak of syntactic realisation of the features in question.

To sum up, in Phase I none of the features (definiteness, number and gender) are represented in C.'s speech either by the use of articles or by singular and plural noun suffixes. There are neither morphological nor syntactic clues in the data that the child has discovered the syntactic position and function of the Determiner inside the noun phrase.

### 3.2. Phase II: the emergence of morphosyntactic features in the noun phrase

Around 24 months (for Italian) and 26 months (for English) Carlo starts producing his first “article + noun” combinations. The percentage in both languages is initially still fairly low (between 13 and 17%), however, while in English file 5 only the indefinite article *a* appears with 3 different nouns, in Italian file 4 we have 6 singular masculine indefinite articles (*un*), 8 singular masculine definite articles (*il*), 9 singular feminine definite articles (*la*), 2 plural masculine definite articles (*i*) and 1 plural feminine definite article (*le*) used in combination with 17 different nouns.

In the following two sections the emergence of D features in Carlo's data after the initial phase will be analysed separately for English and Italian.

#### 3.2.1. The English data

In files 6 and 7 (ages 2;2.22 and 2;4.6 respectively) C. produces a total of 17 definite articles and 51 indefinite articles in combination with 21 different nouns, only three of which are found both with the indefinite and the definite article (*ball, cat, mouse*). Except for one mass noun (*paper*) and a plural noun (*ladders*), all nominals are singular count nouns. As for their semantic distribution, with one exception (*a paper*, file 6) articles are always used appropriately in the specific reading. Given the highly deictic nature of C.'s initial usage of articles it is not surprising that we should not find any DPs interpretable generically or existentially.

The fact that indefinite articles outnumber definite ones by 34 is probably due to the nature of the contexts in which the articles are required. Let us consider the syntactic frames in which these indefinite articles appear in Table 5 below:

Table 5: Distribution of indefinite articles

a + noun	dat 0*aux a + noun	a + noun (reply to Q)
20	16	15

From an analysis of the transcripts we can establish that whenever C. uses an indefinite article it is either to name an object (a + noun; dat 0\*aux a + noun) or in reply to a question like “*What’s this/that?*”. In both cases if the object in question is new to the listener the indefinite reference is required. Only on two occasions does C. incorrectly use a definite article in lieu of an indefinite.

In sum C. correctly supplies an indefinite article 51 times out of 53, 96.2% of the time, an extremely high percentage which could well be taken as an indication of a precocious mastery of the indefinite/definite reference distinction. However, how reliable are the data? When using an indefinite article followed by a noun, the child is not so much engaged in *identifying* a referent as in *naming* one; i.e. his preoccupation is not to introduce a new referent to a context of discourse which he can subsequently refer back to, but it is simply to name the object *per se*. This lack of interest in introducing new referent for discourse purposes is further demonstrated by the fact that, aside from 8 cases in which C. repeats the same noun phrase for a number of consecutive times, there are no instances in which C. uses a definite article to refer to a previously identified object. In sum, considering the self-repetitions and the lack of further mention of an object once it has been identified, it seems that, at least at this stage, naming objects is C.’s primary mode of linguistic interaction in English.

Definite articles do however appear in C.’s speech and they account for 25% of his article usage. Apart from the two incorrect instances mentioned above, all of C.’s other definite articles are used appropriately inasmuch as they refer to objects which are inferable from the context, or which have been previously mentioned by the adult interlocutor.

In sum, although indefinite and definite articles are basically used correctly in C.’s speech, this does not automatically imply that he has mastered the identifying role of indefinites and the anaphoric function of definites, in fact there are indications to the contrary.

As for number specification, apart from one plural noun, all the rest are singular nouns and they correctly refer to a singular referent. The absence of plural nouns could either be an artifact of the sampling method, or it could reflect a genuine avoidance of plurals by the child. In languages of the world the singular is typically regarded as the unmarked form (e.g. Greenberg, 1966), a fact which is reflected in the literature on acquisition where emergence of plural forms normally follows that of singular forms. The appearance of plurals is strictly related to the notion of singularity vs. plurality, a concept that may take some time to develop. Moreover, the phonological structure of the English morpheme is complex and relatively

difficult to master. The English plural morpheme has three allomorphs: [s, z, ez]; in plural formation a voiced fricative is added to the root (e.g. “claws”), the plural marker can involve adding a syllable (e.g. “buses”), or creating a final cluster (e.g. “cats”). The literature on phonological acquisition indicates that children have difficulties with these phonological forms and that fricatives and final clusters are relatively late to emerge (Ingram, 1981). The phonological complexity of the English plural may well be one of the contributing factors to the later emergence of plural forms compared to singular forms, cross-linguistic evidence has shown earlier acquisition of the plural morpheme in languages where the marking system is simpler (e.g. a vowel change as in Italian) (Slobin, 1973; Leonard, Sabbadini, Leonard & Volterra, 1987). Unfortunately because no plural referents are identified by C. in the files under consideration it is impossible to establish what if any knowledge of number specification and number agreement the child has.

### 3.2.2. The Italian data

As will be shown below, C.’s usage of articles in Italian is both quantitatively and qualitatively different from his usage in English. To start with, between 2;1.21 and 2.4.10 suppliance of articles in obligatory contexts in Italian stabilizes over 80 % with a peak of 95% and an average of 88%, a figure very close to the cut-off acquisition point of 90% correct morphemes in obligatory contexts set by Brown (1973). In English, on the other hand, articles start appearing later (2;2.11 as opposed to 2;0.22 in Italian) and reach a mere 66.6% in obligatory contexts by 2;4.6.

The figure for indefinite articles in the four Italian files examined (files 4, 5, 6, 7) is 76, while definite articles total up 58, as opposed to 55 for indefinites and 17 for definites in English. The 134 instances of indefinite and definite articles appear with 78 different nouns (37 feminine nouns and 41 masculine nouns); 9 masculine nouns are used both with the definite and the indefinite article and only 2 feminine nouns appear with both articles. Of the 41 masculine nouns 6 are plural nouns, 4 of which also appear in their singular form (*il cane/i cani, il pesce/i pesci, il serpente/i serpenti, and l’albero/\*l’alberi*)<sup>2</sup>. As for the feminine gender we find 6 plural nouns, only one of which appears in the singular as well (*la mela/le mele*). Altogether only two nouns, the masculine noun *pesce* (“fish”) and the feminine noun *mela* (“apple”) appear with three different forms of article: *una/la mela- le mele, un/il pesce- i pesci*.

In Table 6 and 7 below a breakdown of C.’ Italian articles is provided:

---

<sup>2</sup> In the string “\*L’alberi” the noun suffix is correct while the article should have been *gli* in stead of *l’*.

Table 6: C.'s Italian definite articles

sing. masc.	sing. masc.	sing.fem.	plur. masc.	plur. masc.	plur. fem.
il	lo	la	i	gli	le
36	0	27	6	0	7

Table 7: C.'s Italian indefinite articles

sing. masc.	sing. masc.	sing. fem.	plur.masc.	plur. masc.	plur. fem.
un	uno	una	dei	degli	delle
37	4	15	0	0	0

A fairly clear-cut picture emerges from the data above: firstly, singular forms by far outnumber plural ones, a fact which has repeatedly been observed in the literature (Leonard, Sabbadini, Leonard & Volterra, 1987; Pizzuto & Caselli, 1992; Leonard, Bortolini, Caselli & Sabbadini, 1993; Caselli, Leonard, Volterra & Campagnoli, 1993). In fact, if plural definite articles account for a mere 17% of all definite articles, plural indefinites are not attested at all. It must be noted however that plural indefinites are not very common in the speech of C.'s adult interlocutors either:

Table 8: Number of plural indefinite articles in the speech of C.'s adult interlocutors:

dei	degli	delle
4	1	5

The limited exposure to the use of plural indefinites in the input could therefore be another factor contributing to the absence of such determiners in the child's speech.

Another characteristic of C.'s corpus in line with previous studies on monolingual acquisition of Italian morphology (Pizzuto & Caselli, 1992; Caselli et al., 1993; Leonard et al., 1993) is the difficulty encountered with articles with restricted phonetic contexts such as *uno*, *lo*, *gli* and *degli*. As mentioned in the introduction, these masculine articles are used when the subsequent word begins with *ps*, *z*, *gn*, or *s* cluster; otherwise *un*, *il*, *i* and *dei* are used. Because only a phonetically restricted subset of masculine words will require the *uno*, *lo*, *gli* and *degli* forms, the prediction is that the child will start by hypothesizing that *un*, *il*, *i* and *dei* are the prototypical masculine forms, and only later will he realize that some phonetic contexts require the use of the *uno*, *lo*, *gli* and *degli* forms. By this rationale we would expect

to find a number of commission errors where the *lo/gli, uno/degli* forms are replaced by the more common and already established *il/i, un/dei* articles. The data in C.'s corpus only partially bear out this prediction: omission errors are in fact still predominant over commission errors. Altogether there are 30 omissions of either *uno/lo* or *gli/degli* and only 2 commission errors (*l'alberi, l'elefanti*). Admittedly, out of the 32 omissions 11 belong to files 2 and 3 (ages 1;10.28 and 2;0.1 respectively) a stage in which C. does not produce any articles at all and therefore they should be regarded as omission of the determiner in general rather than avoidance of a particularly difficult form. Of the remaining 21 omissions 17 are found in file 4 (age 2;0.22) a stage in which articles are only just beginning to appear, and the other 4 in files 5, 6 and 7 (ages 2;1.21, 2;3.7, 2;4.10 respectively). As for the substitution errors (*l'alberi, l'elefanti*) they are most likely segmentation errors in which the correct forms *l'albero* and *l'elefante* are mis-analysed as a whole rather than a Det + N string (cf. the 11 occurrences of *quello l'orsacchiotto* in file 7). The only correct phonetically restricted form appearing in C.'s data is the singular indefinite article *uno* (4 occurrences in file 6).

In sum, if on the one hand it is true that phonetically restricted articles are not significantly represented in C.'s data, it is equally true that nouns requiring such articles do not account for a significant proportion of all obligatory contexts, neither in C.'s speech, nor in that of C.'s adult interlocutors. The low frequency cue strength of the phonetically restricted forms is in all likelihood the single factor determining their delayed emergence and mastery.

### 3.2.3. Definite and indefinite reference in Italian

In Italian, unlike in English where naming seems to be C.'s primary mode of interaction, naming and deictic reference are both well accounted for in the data. Moreover, the emergence of the earliest Verbs and Adjuncts is paralleled by an increase in the number of definite articles. However not even in Italian there is any clear indication that C. has mastered the anaphoric function of definite determiners and the contrastive use of definite and indefinite articles.

### 3.2.4. Gender and number agreement

C.'s performance as far as gender and number agreement are concerned is extremely good. There are only 3 instances in which he uses an inappropriate article: *l'alberi* (twice) and *l'elefanti* (file 6). In both cases the number specification expressed by the plural noun suffix is correct, but the plural masculine article *gli* would be required. Recall that the article *gli*, together with *lo, uno* and *degli* belongs to a number of article forms occurring in phonetically restricted contexts which are typically problematic in the acquisition of the Italian article

system. It is therefore not surprising that C. should commit an error precisely with this kind of article. At an earlier age C.'s strategy was to omit the article altogether: however, of the 32 omissions of *uno*, *lo*, *gli* or *degli* only 4 are found in files 5, 6 and 7, the remaining 28 are spread over files 2 and 3 (11), a stage in which C. does not produce any articles, and file 4 (17) where articles are just beginning to appear.

C.'s excellent performance in the mastering of gender and number agreement in Italian raises a fundamental question: to what extent can the child be credited with actual knowledge of morphosyntactic agreement? Adult-like mastery of gender and number features in Italian entails the realisation of agreement between prenominal determiners and other modifiers (e.g. adjectives) and the noun. In the period examined there are no prenominal adjectives in the data, therefore our analysis will be limited to the agreement between articles and nouns.

Both the masculine and the feminine gender are well represented with a total of 37 feminine nouns and 41 masculine nouns, however there is a discrepancy in the number of singular "article + noun" combinations with 77 masculine noun phrases over only 42 feminine ones, an imbalance towards the masculine gender previously reported in the literature on the acquisition of gender (Hernández-Piña, 1984). In the plural the figures for masculine and feminine noun phrases are almost identical: 6 plural masculine noun phrases and 7 feminine ones. The low number of plural noun phrases in the period under consideration is most certainly related to the later emergence of the plural observed elsewhere in the literature on language acquisition.

C. seems to have mastered the feature gender as agreement is 100% correct for both masculine and feminine "article + noun" combinations. As for number, the low percentage of plural forms (9.8 % of the total number of noun phrases) does not provide sufficient evidence to suggest that C. has completely figured out the morphosyntactic correlates of the notion of singularity vs. plurality, although all the plural forms in the corpus are in fact correct.

C.'s adult-like performance lends itself to a number of possible interpretations that I will consider in more detail below. Firstly, it could be argued that in fact C. does not know anything about gender and number agreement, and that he has simply rote-learned the "article + noun" combinations storing the correct article together with the noun rather than identifying the noun and the article separately and matching them when required according to the relevant features.

Secondly, we could claim that the child has actually mastered the notion of gender and number agreement and that therefore his excellent performance is the logical consequence of this realisation.

Thirdly, we could hypothesize a more gradual process whereby initially a number of “article + noun” combinations which are very frequent and salient in the input are rote-learned by C. and only subsequently does he begin to recognize the pattern and extend it to other nouns and articles across the board.

At least two objections can be raised against the first hypothesis: first of all if C. actually learned nouns together with their articles as one lexical entry we would not expect a phase in which he produces bare nouns, but we do find bare nouns in the data between 1;10.7 and 2;0.22. Secondly, the number of different nouns preceded by an article totals up 78, far too large a figure to think that C. may have memorized them all.

The second hypothesis, mastery of gender and number agreement, may appear too ambitious at first glance. The rote-learning of a few items and successively the gradual extension across the board of the third hypothesis seems to be the most plausible, but exactly how plausible? A survey of the data reveals that there are 4 nouns which occur throughout the period in which articles are attested (*gatto, cane, topolino papà*) and which could be considered as core rote-learned “article + noun” combinations that C. used as stepping stones to the generalization of gender and number agreement across the board. Note however that all four noun phrases are singular and masculine, the features feminine and plural are not represented at all. It is difficult to see how extension of these features to other noun phrases could have taken place if they were not part of the core memorized items in the first place. Moreover, and more importantly, the “article + noun” combinations containing these four nouns account for a mere 21% of all noun phrases, not a sufficiently high figure to consider these items as the originators of the discovery of gender and number agreement. Consider also that the four noun phrases in question co-occur with a large number of other noun phrases correctly inflected for number and gender. A somewhat unexpected fact if the four noun phrases were the starting point of C.’s mastery of agreement.

The above considerations have led me to believe that the second hypothesis is in fact the more tenable of the three. Nevertheless, the conclusion that C. has mastered the notion of gender and number agreement may be slightly premature. If, on the one hand it is true that the child’s performance is virtually perfect, it is equally true that plurals are under-represented and so are articles occurring in phonetically restricted contexts. Furthermore, the degree of overlap of definite and indefinite articles is extremely low, only 2 nouns appear both with the indefinite and the definite article (*una mela/la mela; un pesce/il pesce*). In conclusion, while C.’s performance is virtually adult-like in the noun phrases he produces, his repertoire is still incomplete and there is a general tendency to avoid plurals.

#### 4. A few concluding remarks

The major difference in the emergence of D features in the two languages lies in the different acquisitional strategies that C. has employed in these initial phases of development. While the emergence of Italian articles and the related number and gender features appears to be a phenomenon across the board as soon as articles start appearing in C.'s speech, in English the child seems to rely heavily on a small number of rote-learned forms which account for more than 50% of his article production between 26 and 28 months. Four words (*ball, dog, cat, mouse*) account for 58.8% of C.'s "article + noun" combinations, thus suggesting a very stereotyped use of articles in rote-learned strings. On the contrary, recall that in Italian the only four noun phrases recurring throughout C.'s files account for a mere 21% of the total number of articles.

While in English he does seem to start out by using a handful of memorized frozen forms without extending the use of articles to his repertoire of nouns more at large, in Italian, on the other hand, we witness a much more flexible use of articles and a realisation of the morphophonological and syntactic regularities that govern nominal agreement.

#### References

- Brown, R. (1973), *A first language: the early stages*, Cambridge, MA, Harvard University Press.
- Caselli, M.C., L. Leonard, V. Volterra, M.G. Campagnoli (1993), "Towards mastery of Italian morphology; a cross-sectional study", *Journal of Child Language* 20, 377-93.
- Greenberg, J. (1966), "Some universals of grammar with particular reference to the order of meaningful elements", in J. Greenberg (ed.), *Universals of language*, Cambridge, MA, Cambridge University Press.
- Hernández-Piña, F. (1984), *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*, Madrid, Siglo XXI.
- Ingram, D. (1981), *Procedures for the phonological analysis of children's language*, Baltimore, University Park Press.
- Leonard, L.B., L. Sabbadini, J. Leonard, V. Volterra (1987), "Specific language impairment in children: a crosslinguistic study", *Brain and Language* 32, 233-52.
- Leonard, L.B., U. Bortolini, M.C. Caselli, L. Sabbadini (1993), "The use of articles by Italian-speaking children with specific language impairment", *Clinical Linguistics and Phonetics* 7(1), 19-27.
- Pizzuto, E., M.C. Caselli (1992), "The acquisition of Italian morphology: implications for models of language development", *Journal of Child Language* 19, 491-557.
- Slobin, D. (1973), "Cognitive prerequisite for the development of grammar", in C.A. Ferguson, D. Slobin (eds.), *The crosslinguistic study of language acquisition. Volume 2: Theoretical Issues*, Hillsdale, NJ, Lawrence Erlbaum Associates.