THE LATE ACQUISITION OF PREFERRED ARGUMENT STRUCTURE IN VENEZUELAN SPOKEN SPANISH

Paola Bentivoglio

Universidad Central de Venezuela

Preferred Argument Structure (PAS), originally proposed by Du Bois 1987 in his study on an ergative Mayan language, has been shown generally to obtain in many other languages, ergative and non-ergative as well, thus lending credence to the suggestion that PAS may be a universal. In this article I analyze the acquisition of PAS in a corpus of narratives by children speaking Venezuelan Spanish, aged 7;0 through 10;7, with two purposes: i) comparing the characteristics of PAS of these children with those of younger children; and ii) proposing a tentative schedule of the acquisition process.

Keywords: Language acquisition, Venezuelan Spanish, Preferred Argument Structure.

1. INTRODUCTION[*]

The present study aims at ascertaining the presence of *Preferred Argument Structure* (PAS) in Spanish spoken by children pertaining to two different age groups (7;0 through 7;9) and (10;0 through 10;7) and comparing the results whith those obtained in a similar research on a corpus of younger children speaking the same dialect.

PAS, a hypothesis developed by Du Bois (1987), pertains to the form and discourse role of the "core" arguments of the verb, that is, the subject and the direct object. At this point, it is necessary to say that, in the following analysis, and in order to distinguish the subject of an intransitive verb from the subject of a transitive verb, I follow the practice initiated by Dixon (1979), whereby the

subject of a one-argument verb is denoted by an **S**, the subject of a two-argument verb by an **A**, and the direct object by an **O**.

PAS has both a grammatical and a pragmatic dimension, as we may observe in Table 1:

Table 1: Dimensions and constraints of Preferred Argument Structure

	Grammar	*Pragmatics*
Quantity	One Lexical Argument Constraint	One New Argument Constraint
Role	Non-lexical A Constraint	Given A Constraint

Source: adapted from Du Bois (1987:829)

The grammatical dimension of PAS is expressed by two "constraints"[1] relating to the presence or absence of full, lexical NPs, in the same clause. The first constraint, called the "One Lexical Argument Constraint", reflects the preponderance of clauses in which only one of the core arguments is represented by a full NP. The second constraint, termed the "Non-lexical **A** Constraint", refers to the fact that the single lexical NP tends to occur either in the **S** or in the **O** role, but not in the **A** role. In its related pragmatic constraints, depicted in the right column of Table 1, PAS can likewise be expressed in terms of two constraints. First, the "One New Argument Constraint" establishes that clauses tend to contain no more than one piece of new information; and second, the single piece of new information tends to be linguistically encoded in the **O** or in the **S** role, but rarely in the **A** role, hence the "Given **A** Constraint".

PAS has been attested in several languages typologically very different; it has been also found in Venezuelan Spanish spoken by adults as well as in written Old Castilian (XI and XIV century) and in written contemporary Venezuelan Spanish. A more recent research (Bentivoglio 1996) has shown that PAS is already present in the speech of 3 to 5 year-old children, but not with all the characteristics of adult PAS. Other studies on child language, from the same viewpoint, are Clancy (1993) on Korean, and Allen and Schröder (forthcoming) on Inuktitut.

The present article is dedicated to describing the progressive acquisition of PAS in children of 7;0 to 10;7 years of age, trying to determine whether or not children of the older group have fully reached the adult target.

2. METHODOLOGY

The data consist of twenty narratives obtained by showing to children of two age groups a short film, *Picnic* (Woods 1993). The film, with only background sound but no words, shows a family of rats who go on a ride to have a picnic in the country. The film was shown to each child separately; after having seen it, each child was taken to a different non adjacent room and requested to tell what s/he had seen to the researcher who was conducting the experiment. The narratives were tape-recorded and later transcribed.

---

[1]The proposed *constraints* do not constitute absolute prohibitions or requirements, but rather strong patterning preferences that can be measured in natural discourse.

The children were attending two different schools, both located in Caracas. All of their parents are Venezuelan, but not all of them are Caracas-born. The children's age ranges from 7;0 to 10;7. For the purpose of this study they were divided into two groups: i) from 7;0 to 7;9 (henceforth the 7 years group); and ii) from 10;0 to 10;7 (henceforth the 10 years group). There are five boys and five girls in each group, respectively. The exact age of each child and more details can be found in Appendix A.

In order to conduct the PAS analysis, all NPs (full, stressed and unstressed pronouns, zeros) fulfilling the core roles of **A, S,** and **O** were coded according to the following set of linguistic variables: i) clause type (main/non-main); ii) syntactic role (**A, S, O**); iii) form (**Ø, P**[ronoun], **N**[oun], **C**[lause]); animacy (animate/inanimate); identifiability (identifiable/non-identifiable); and information status (new/non-new). All items were also coded according to the age group (7 and 10) and sex (femenine/masculine) of the children. The coded data were then cross-tabulated, submitted to a χ^2 test and finally analyzed with the help of Goldvarb 2.0, a computer program designed for Macintosh by Rand & Sankoff (1990).

Let us observe in example (1) the narrative produced by a boy, who belongs to the 7 years group. The two letters, separated by a hyphen, in the left part of the text, correspond to the syntactic role (**A, S, O**) and to the form (**Ø, P, N**) of the NP which appears in italics in the text:

(1)   *Picnic*: a narrative by Erick, a 7;2 boy.

1.		... que unos ratoncitos estaban ... estaban ...
2.	**A-N**	y *el papá* estaba arreglando
3.	**O-N**	*un carro*
4.		para irse de ... de camping,
5.		y después Ø arreglaron ...
6.	**A-Ø**	*Ø* lo *arreglaron*
7.	**S-Ø**	y *Ø se fueron* de camping y ...
8.		Y había una niñita, una ratoncita,
9.	**S-P**	*que* se cayó por el camino,
10.	**S-N**	y *los ratones* siguieron,
11.	**O-P**	no Ø *la* vieron,
12.		y ... los ... los otros ratones fueron a jugar
13.		y Ø se divirtieron.
14.		Después Ø se dieron cuenta,
15.	**A-Ø**	cuando *Ø* le *iban a dar*
16.	**O-N**	*la leche* ...

1.	'the... that some little rats were ... were ...
2.	and **their daddy** was fixing
3.	**a car**
4.	to go ... camping,
5.	and then [they] fixed ...
6.	**[they]** fixed it
7.	and went camping and ...
8.	And there was a little girl, a little rat,
9.	**who** fell during the trip,
10.	and the rats went on,

11.	[they] didn't see **her,**
12.	and ... the ... the other rats went playing
13.	and had a nice time.
14.	Later [they] realized [that she fell],
15.	when **[they]** were about to give [them]
16.	**the milk ...'**

Examples of NPs fulfilling the **A** role are: the full NP *el papá* 'the daddy' in line 2 of (1), and the zero form *Ø lo arreglaron* '[they] fixed it' in line 6. No examples of pronouns in the **A** role were found in this particular fragment. The **O** role is exemplified by the full NP *un carro* 'a car' in line 3 and by the unstressed pronoun *la* 'her' in line 11. These examples should be sufficient to illustrate the core arguments of a transitive verb. As for S, the sole argument of an intransitive verb, all three possibilities are present: *los ratones* 'the rats', a full NP in line 10, a relative pronoun in *que se cayó por el camino* 'who fell during the trip' in line 9, and finally a zero form in *Ø se fueron de camping* 'Ø went camping' in line 7.

All tokens were coded for the semantic variable of animacy: *el papá* 'the daddy ' in line 2 was coded as animate, whereas *un carro* 'a car', in line 3, and *la leche* 'the milk', in line 16, were coded as inanimate. It may be noted that in this particular corpus there are no animate referents who are also human, all of them being either animals or inanimate things.

All NPs were also coded according to the pragmatic variables of identifiability and information status, following Chafe 1994 as well as Du Bois and Thompson 1991. As zero forms and pronouns are used to code linguistically information which is both identifiable and non-new, only full NPs will be relevant for the ensuing discussion. In (1) *un carro* 'a car' , in line 3, is an example of a non-identifiable entity which represents new information, whereas *el papá* 'the daddy', in line 2, also represents new information, from both contextual and situational viewpoints, but is identifiable by means of its uniqueness (cf. Du Bois 1980). Consequently, the first NP is indefinite (*un carro*), and the second one is definite (*el papá*).

Non-new or given information should be always identifiable and coded as a definite NP.[2] *Los ratones* 'the rats' in line 10 is a clear example of an entity which is both identifiable and non-new by means of a previous mention (*unos ratoncitos* 'some little rats' in line 1) and is correctly expressed by a definite NP. The same process is shown by the zero forms *Ø arreglaron* '[they] fixed', *Ø la vieron* '[they] saw her', in lines 6 and 11, respectively.

3. THE RESEARCH PROBLEM

The question I would like to raise is the following: if children aged 3 through 4 have already partially acquired PAS (see section 1) but do not get to master all its aspects even at 5 or 6 years of age, when do they reach the full-fledged adult structure? The analysis of the 7 and 10_years groups was conducted in order to determine whether or not PAS acquisition would be completed during this age period.

---

[2]Sometimes a NP carrying non-new information is the exact repetition of an already mentioned NP, that was non-identifiable and indefinite in its first mention. No such examples, however, are present in fragment (1).

Table 2 shows the distribution of the data according to clause type in the two age groups:

Table 2: Distribution of clauses according to age group (N = 20)

	Main		Non-main		Total	%
	N	*%*	*N*	*%*		
7 years	259	68	120	32	379	37
10 years	459	70	193	30	652	63
Total	718	70	313	30	1.031	100

In Table 2 we observe that both groups show approximately the same percentage of clause types: main clauses represent 68% in the 7 years group and 70% in the 10 years group. The relevant difference resides in the total number of clauses: out of 1,031 clauses, 379 (37%) belong to the younger group, and 652 (63%) to the older one. The 1,031 clauses constitute the corpus to be analyzed.

4. THE ANALYSIS

4.1. *The grammatical dimension of PAS*

Let us concentrate on the grammatical dimension of PAS (cf. the left column of Table 1) and its related tendencies, according to which, if in a clause there is a full NP, it will probably fulfill the role of S (if the clause is transitive) or **O** (if the clause is intransitive), but not that of **A**..This last argument is much more likely to appear as a zero or a pronoun than **O**.[3] Moreover, the cooccurrence of two lexical NPs fulfilling the **A** and **O** roles within the same clause —as *el papá estaba arreglando un carro* 'their daddy was fixing a car' in lines 2 and 3 of (1)— is very rare, as all previous studies on PAS have proven. The results of the analysis in the 7 years group and in the 10 years group are presented in Table 3a and in Table 3b, respectively. Due to the scarcity of pronouns in the corpus, zero (Ø) and pronouns (P) have been collapsed into one category under the label Ø+ P.

Table 3a: Distribution of **A** and **O** according to their form (N, Ø+ P, C) in the 7 years group (N = 10)

	A-Ø + P		A-N		Total	%
	N	*%*	*N*	*%*		
O-P	51	96	2	4	53	38
O-N	47	80	12	20	59	42
O-C	18	67	9	33	27	20
Total	116	83	23	17	139	100

---

[3] In Spanish the NP in the **O** role can only be a pronoun or a lexical noun, but not a zero.

Table 3b: Distribution of **A** and **O** according to their form (N, Ø+P, C)
              in the 10 years group (N = 10)

	A-Ø + P		A-N		Total	%
	N	*%*	*N*	*%*		
O-P	63	84	12	16	75	38
O-N	67	77	**20**	23	87	44
O-C	29	78	8	22	37	19
Total	159	80	40	20	199	100

The results of the two groups of children are very similar: both groups highly favor constructions where **A** arguments are represented by zeros or pronouns, the overall percentage being 83% in the younger group and 80% in the older one. The overall percentage of full NPs is equally low in both age groups: 17% in the 7 years and 20% in the 10 years group. The most important finding relates to the reduced number of clauses that, in both groups, have a full NP in the **A** as well as in the **O** role: 12 out of 140 cases (9%) in the 7 years group and 20 out of 199 (10%) in the 10 years group. These results, on one hand, differ strikingly from those obtained in the analysis of children aged 3 through 6, who show a much higher presence of full NPs; but, on the other, are identical to those obtained in the analysis of a similar adult corpus.[4] Table 4 allows the comparison of the number of clauses with two full NPs in the **A** and **O** roles found in the mentioned corpora:

Table 4: Cooccurrence of full **NPs** in **A** and **O** roles in children and adults

	N	*%*
3-4 children	9/43	21
4-5 children	15/95	16
5-6 children	11/64	17
7-8 children	12/140	9
10-11 children	20/199	10
Adults	87/911	10

In Table 4 we may appreciate that the percentage of cooccurrence of two full NPs in the core arguments in the same clause is very high in the 3-4 years group (21%), but decreases consistently in the following two groups (4-5 and 5-6) reaching its lowest point (9%) in the 7-8 group, and finally equals the adult level in the group of children aged 10 through 11 (10%). The comparison resumes the findings related to the first contraint proposed by Du Bois 1987 for the grammatical dimension of PAS: the tendency regarding the presence of no more than one lexical argument per clause is present in the children since a very early age but its frequency varies decreasing from 21% at the age of 3-4 to 9-10% at the age of 7-8, when the adult target is reached.[5]

The second tendency of the grammatical dimension of PAS (cf. Table 1) consists in the high proportions of zero and pronouns in the **A** role and, conversely, the high proportion of lexical

---

[4] The adult corpus I am referring to consists of 30 narratives obtained by showing to a group of Venezuelan adult speakers the same film (*The pear story*) shown to children aged 3 through 6.
[5] This result confirm nicely the prediction made in Bentivoglio 1996: "It seems reasonable to believe that the percentage of full NPs in the core roles of a transitive clause will decrease with age".

nouns in the other two roles: **S** and **O**. The distribution of zeros and pronouns, on one hand, and full NPs,[6] on the other, in the **A, S,** and **O** roles is depicted in Table 5a for the 7 years group and in Table 5b for the 10 years group.

Table 5a: Distribution of **A, S, O** according to their form (N, Ø+P) in the 7 years group (N = 10)

	Ø+P		N		Total	%
	N	*%*	*N*	*%*		
A	116	83	23	17	139	27
S	193	80	47	20	240	46
O	53	38	86	62	139	27
Total	362	70	156	30	518	100
	$\chi^2 = 91.41$	2 d.f.	p < .00000			

Table 5b: Distribution of **A, S, O** according to their form (N, Ø+P) in the 10 years group (N = 10)

	Ø+P		N		Total	%
	N	*%*	*N*	*%*		
A	159	80	40	20	199	23
S	353	78	100	22	453	53
O	75	38	124	62	199	23
Total	587	69	264	31	851	100
	$\chi^2 = 119.1$	2 d.f.	p < .00000			

Tables 5 a and 5b show great similarities of results in all roles and forms analyzed: both groups use high percentages of zeros and pronouns and low percentages of full nouns in the **A** and **S** roles. The contrary holds for the **O** role, which, in both groups of children, reaches the extremely high proportion of 62% of full nouns *viz.* a reduced amount of pronouns (38%). The comparison of these results with those obtained in the analysis of both the 3-6 children and the adult corpora are presented in Table 6.

---

[6]For the sake of brevity, in Tables 5a and 5b, clauses fullfilling the **O** role appear under the **N** label. It is worth noting that in the corpus here analyzed there are no clauses fulfilling **A** and **S** role

Table 6: Distribution of full NPs as **A, S, O** in children and adults

	3-6 children		10-11 children		Adults	
	N	*%*	*N*	*%*	*N*	*%*
A	48	23	40	20	128	14
S	125	44	100	22	508	41
O	138	65	124	62	576	63

Table 6 reveals that **O** is the role which experiments least variation after early childhood: the presence of full NPs shows a minimal decrease: from 65% in children aged 3 through 6 to 63% in adults. Role **A** has a much higher proportion of full nouns in all children groups; even the children aged 10 through 11 do not seem to have reached the adult target: 20% vs. 14% of A-N, respectively. Results related to **S** are very difficult to interpret, as younger children show a use very similar to that of adults (44% vs. 41%, respectively), but children belonging to the 10-11 age group show a much lower percentage of full nouns in this role (22%). This difference could well be due to the different stories depicted in the prompt-movies: the 3-6 children saw *The pear story* film (Cf. Chafe 1980) whereas the other two groups of children saw *Picnic* (Woods 1993): the first film emphasizes the introduction of new participants, hence the higher rate of full NPs in the **S** role, whereas the second one emphasizes the events occurring to some of the participants, who have been all introduced in the first sequence. It seems that the use of **S**, the sole argument of an intransitive verb (quite frequently a motion verb) highly correlates with the kind of stories being narrated, whereas **A** and **O** appear to be less related to the nature of the text.

4.2.*The pragmatic dimension of PAS*

According to the pragmatic dimension of PAS (cf. the rightmost column of Table 1), in a single clause, the two following tendencies are present: i) only one argument can contain new information; and ii) the syntactic roles in which new information tends to be present are S (if the clause is intransitive) or **O** (if the clause is transitive), but rarely **A**. These contraints make it very unlikely to find, in spoken discourse, clauses like *el papá estaba arreglando un carro* 'their daddy was fixing a car', in lines 2 and 3 of example (1), where both **A** and **O** arguments —*el papá* and *un carro*—represent new information. The clause *cuando Ø le iban a dar la leche* 'when [they] were about to give [them] the milk', in lines 15 and 16 of (1), is a prototypical example of both pragmatic tendencies: on one hand, the **A** argument —i. e. the subject of the transitive clause— representing non-new information is linguistically encoded as a zero, and, on the other hand, the **O** argument —i. e. the object of the transitive clause— representing new information is linguistically encoded as a full NP (*la leche*). Note that new information does not necessarily coincide with undefinite marking, as shown by *el papá* and *la leche*, which, in spite of being entities introduced for the first time into the discourse and not situationally present, are identifiable and thus are linguistically encoded as definite NPs. In the analysis of the mechanisms used by children (and by all speakers) in order to introduce new information, it seems necessary to include the identifiability factor. The exclusion of new definite NPs would mean drastically reducing the quantity of new information: in the present corpus, for example, 77 entities, in the 7 years group, and 120, in the 10 years group, represent new information, but only 32 out of 77 (42 %), in the first group, and 59 out of 120 (49 %), in the second one, are non-identifiable and are consequently coded as indefinite. If

the identifiable new entities had not been taken into acccount, the amount of new information would have beeen less than half.[7]

The multivariate analysis was conducted by means of Goldvarb 2.0, in order to compare the results with those obtained in previous analyses of Spanish PAS. The results of the 7 and 10 years groups are presented in Table 7a and in Table 7b, respectively.

Table 7a: Results of the multivariate analysis in the 7 years group

Factor	Count	% New	Weight
Animacy			
Inanimate	53/66	80	.692
Animate	24/61	39	.294
Identifiability			
Non-identifiable	32/37	86	.761
Identifiable	45/90	50	.383
n = 127 10 speakers input .649			

Table 7b: Results of the multivariate analysis in the 10 years group

Factor	Count	% New	Weight
Animacy			
Inanimate	88/117	75	.686
Animate	32/106	30	.297
Identifiability			
Non-identifiable	59/72	82	.732
Identifiable	61/151	40	.382
n = 223 10 speakers input .560			

Tables 7a and 7b show no relevant differences between the two age groups: thus the results will be commented upon jointly. The first factor which contributes significantly to new information is semantic: inanimate referents are more likely to represent new information than animate ones, the probability weight varying from .692 to .686.[8] The second factor is the non-identifiability of the entity analyzed with a weight varying from .761 to .732. No other factors were selected by the program. These results show that children of both age groups take into account exactly the same factors —animacy and identifiability— when they have to introduce new entities into their narrative discourse.: inanimate and non-identifiable entities

---

[7] For this reason I cannot compare the present results with those obtained by Hickmann *et al.* 1996 in their study about the marking of new information in four different languages, as the authors examine how children use 'local' marking, i. e. the indefinite article, in order to introduce new referents.

[8] In Goldvarb 2.0 a probability weight greater than .500 is considered to contribute to the variant under study.

are more likely to represent new information than animate and identifiable ones.[9] The comparison of the results obtained in the multivariate analyses of different age groups are presented in Table 7c:

Table 7c: Probability weights of relevant factors across different age groups

	Inanimate	Non-identifiable
3-4 children	---	.678
4-5 children	---	.754
5-6 children	---	.788
7-8 children	.692	.761
10-11 children	.686	.732

The comparison among the groups of children shows that the animacy variable becomes relevant only in the two groups of older children, in which it appears well established with almost identical values. On the contrary, the influence of the identifiability variable is present in all age groups: children aged 3 through 4 show the lowest value (.678), but all other age groups have very similar values, the maximum difference being .052. The variable syntactic role —hierarchically the most relevant in all adult Spanish corpora studied to date— is never selected by the Goldvarb 2.0 program in any of the children corpora, not even in the 10-11 group, the oldest group of children studied so far in regard to the acquisition of PAS. It seems possible to conclude that children aged 10 through 11 have not yet fully acquired the pragmatic dimension of PAS.

5. CONCLUSIONS

In order to determine at what age children speaking Venezuelan Spanish fully acquire Preferred Argument Structure, I have analyzed the narratives of twenty children divided into two age groups: 7-8 and 10-11. In the analysis, both dimensions of PAS —grammatical and pragmatic— were taken into account, with the purpose of discovering whether or not all four tendencies (two in each dimensions) are present at the same time in both groups of children.

The results in regard to the grammatical dimension of PAS, and more specifically in regard to the presence of full NPs in the core grammatical roles, are the following: children of both age groups know the language-specific discourse properties of **A, S** and **O**. Nevertheless, the comparison of the present results with those obtained in a similar study of an adult corpus clearly shows differences among the three roles: i) **O** (the object of a transitive clause): children's use matches almost perfectly that of the adult group; ii) **A** (the subject of a transitive clause): children's use goes in the same direction of the adult group but is quantitatively different in that children show a greater frequency of full NPs than adults do; iii) **A+O**: in spite of the greater quantity of full NPs in the **A** role, the proportion of full NPs in the same transitive clause seems to have reached the adult target already in the younger group of children; iv) **S** (the subject of an intransitive clause):

---

[9] The syntactic variable regarding the role in which new information is encoded is not yet a relevant factor , as it appears to be in the adult corpus.

the results of the present research are unclear, probably due to the nature of the prompt movie utilized to elicit the narratives. The overall conclusion for the grammatical dimension is that the tendency relating to the quantity of lexical arguments in a clause is totally acquired at the age of 7-8 years, whereas the tendency relating to the role in which full NPs should appear (S and O, but not A) is not yet fully mastered by children of the 10-11 group. This finding suggests the need of analyzing speakers aged 11 through 15, in order to determine when the acquisition of this part of the grammatical dimension of PAS may be completed.

In regard to the pragmatic dimension of PAS, the results of the analysis show that there are no relevant differences due to the different age of the children. In both age groups, new entities tend to be inanimate rather than animate, and non-identifiable rather than identifiable. The syntactic role in which a new entity may be encoded is not relevant, contrary to what happens in the corpus of adults, for whom this factor is the most important. It may be concluded that children aged 10 through 11 still have to complete the full acquisition of the pragmatic dimension of PAS.

Summing up, the overall conclusion relating to the full acquisition of PAS is that children have not reached the adult target at the age of 10-11 in at least two important components: first, the children use more frequently than adults full NPs in the A role; and second, syntactic role is not yet a relevant variable for children as it is for adults. This conclusion leaves partially unanswered the question which motivated the present study and suggests the necessity of further research in order to establish the age when children's tendencies will be qualitatively and quantitatively similar to those shown by adults.

REFERENCES.

Allen, S. E. M. and H. Schröder (forthcoming). Preferred argument structure in early Inuktitut spontaneous speech data. In: *Preferred Argument Structure: Grammar as architecture for function* (J. W. Du Bois, L. E. Kumpf and W. J. Ashby, (Eds.). Amsterdam: John Benjamins.

Ashby, W. J. and P. Bentivoglio (1993). Preferred argument structure in spoken French and Spanish. *Language Variation & Change* 5, 61-76.

Bentivoglio, P. (1994). Spanish preferred argument structure across time and space. *Documentaçio de Estudos em Lingüística Teórica e Aplicada* 10, 277-293.

Bentivoglio, P. (1996). Acquisition of Preferred Argument Structure (PAS) in Venezuelan Spanish. Paper read at the VII International Congress for the Study of Child Language, Istanbul.

Chafe, W. (1994). *Discourse, consciousness and time.* Chicago, The University of Chicago Press.

Clancy, P. (1993). Preferred Argument Structure in Korean acquisition. In: *The 5th Stanford Child Language Research Forum*, 307-314. Stanford, SCLR.

Dixon, R. M. W. (1980). Ergativity. *Language* 55,59-138.

Du Bois, J. W. (1987). The discourse basis of ergativity. *Language* 63, 805-55.

Du Bois, J. W. and S. A. Thompson (1991). *Dimensions of a theory of information flow.* Unpublished manuscript. University of California Santa Barbara.

Hickmann, M., H. Hendryks, F. Roland and J. Liang (1996). The marking of new information in children's narratives: a comparison of English, French, German and Mandarin Chinese. *Journal of Child Language* **23,** 591-619.

Rand, D. and D. Sankoff (1990). *Goldvarb 2.0.* Montréal, Université de Montréal.

Appendix A

Name	Code	Date of birth	Age*	# tokens
Alfredo	7ma	26-04-89	7;0	59
Erick	7mb	17-02-89	7;2	38
Andrés	7mc	29-10-88	7;6	66
Rodrigo	7md	12-12-88	7;6	73
Eduardo	7me	21/09/88	7;9	53
Daniela	7fa	02-03-89	7;2	40
Maryam	7fb	20-01-89	7;3	45
Yanella	7fc	05-12-88	7;5	35
Andrea	7fd	10-12-88	7;5	34
Carla	7fe	17-10-88	7;8	74
César	10ma	15-01-86	10;3	80
José Luis	10mb	25-12-85	10;4	30
Esteban	10mc	27-10-85	10;6	88
Eduardo	10md	02-10-85	10;7	107
Tomás	10me	28-10-85	10;7	73
Alina	10fa	07-05-86	10;0	156
Alicia	10fb	24-01-86	10;3	60
Verónica	10fc	18-12-85	10;4	99
Gabriela	10fd	25-11-85	10;5	28
Carolina	10fe	19-12-85	10;6	130

*The age was calculated according to the day when the recording took place (between April 20 and June 26, 1996).