Ellipsis and dialogue in the early acquisition of syntax

Fátima Gómez¹, Susana López-Ornat^{1*}, Carlos Gallego¹ y María Martínez²

1 Universidad Complutense de Madrid 2 Universidad Nacional de Educación a Distancia [UNED]

Título: La elipsis y el diálogo en la adquisición temprana de la sintaxis. Resumen: Se investiga la transición de las fases de una y dos palabras a la aparición de las primeras oraciones completas. De acuerdo con una visión emergentista de la adquisición sintáctica temprana se esperaba detectar en este periodo construcciones de transición más complejas que la mera yuxtaposición de dos palabras pero menos que una oración simple. En particular, se esperaba el predominio de las emisiones de una o dos palabras, la presencia residual de oraciones correctas aunque no productivas y un aumento gradual de fragmentos sintácticos, como elipsis correctas no productivas, dado su uso en la lengua española. Con este fin, se examinó la evolución entre los 20 y 27 meses de una niña monolingüe española, registrando en video y analizando su habla espontánea en ese periodo y se codificó la estructura de las emisiones junto a sus contextos discursivos y situacionales. Los resultados confirman y especifican la hipótesis. A lo largo del tiempo se observa un incremento de construcciones pre-elípticas que mimetizan oraciones elípticas adultas. Un análisis pormenorizado muestra que las construcciones pre-elípticas evolucionan pasando de depender localmente del contexto situacional a vincularse a la estructura lingüística del turno anterior, en un proceso en el que el diálogo y sus contextos juegan un papel de-

Palabras clave: Adquisición temprana de la morfosintaxis; elipsis; diálogo.

Introduction

In the acquisition of syntax, transition from one or two words to the first simple sentence constructions raises interesting theoretical issues. According to a nativist approach, the syntactic rules are known from birth (Valian, 2009). Difficulties arise because these rules specify possible relations between lexical categories, but not between words, yet it is words that appear in the input. This forces the child first to discover how words map onto the lexical categories, with the possibility of using 'bootstrapping' such as semantics (Pinker, 1995). Alternatively, both the semi-modular (Newport, 2011), and the constructivist approach propose a statistical learning process as the basis for the first 'syntactical' abstractions (of limited scope) by the child. To this, the semi-modular perspective adds that it is cognitive conditions (variables of perception, memory, context and pragmatics) that determine whether the learner forms a statistical or a symbolic regularity (Newport & Aslin, 2012). In turn, the constructivist (or emergentist or usage-based) perspective emphasizes the developmental dynamic by which both productive syntax and the symbolic aspect of syntactic abstraction (Marcus et al., 1999) gradually result from a continuous acquisition process. That process stems from specific, local constructions, used for a particular item and gradually progresses from early statistical abstractions to reach fully abstract forms ('rules') (McClelland & Plaut, 1999).

* Dirección para correspondencia [Correspondence address]: Susana López Ornat, Dpto.Psicología Básica II (Procesos Cognitivos), Facultad de Psicología, Universidad Complutense de Madrid, 28223 Madrid (Spain). E-mail: slopezor@ucm.es **Abstract:** The transition from the one and the two-word phases to the first complete sentences was investigated. Within an emergentist scope, it was hypothesized that it would be possible to identify transitional constructions, more complex than the mere juxtaposition of two words but syntactically less mature than simple sentences. Specifically, predominance of one and two-word productions, and a marginal use of correct but nonproductive sentences were predicted, together with a gradual increase of syntactic fragments, i.e., correct and non-productive ellipses, given their frequent use in the Spanish language. Hypotheses were tested through a longitudinal study of a monolingual Spanish girl, from the age of 20 to 27 months. Weekly video sessions over seven months recorded her spontaneous utterances. Those were coded together with their situational and speech contexts. The results supported and specified the hypothesis. Over time, there was an increased use of constructions we have called preellipses because they mimic adult ellipses. Detailed analyses showed preelliptical constructions evolved from showing a local dependency on their production context, to becoming linked to the linguistic structure of the previous turn. Dialogue and its context played a fundamental role in these transitional steps into syntax.

Key words: Early syntactic acquisition; ellipsis; dialogue.

Within the emergentist approach, which is followed by the current study, several "transition to syntax steps" have been well defined, such as frame and slot for English, which express a very limited, gradually increasing pre-syntactic generalization (Pine & Lieven, 1997). Frames are constructions repeated in the same form with a position in which the lexicon is interchangeable ("give me water", "give me milk", etc.). Other forms of syntactic transition are partial regularities (López Ornat, 1994; Mariscal et al., 2010), forms that are morphosyntactically correct but only for some of its values, not for all (López Ornat, 1994; Mariscal et al., 2010). This occurs, for example, in the gender agreement acquisition process in the Spanish language, where fewer errors occur in the Det + N agreement for feminine than for masculine (López Ornat, 2003; Mariscal, 2008; Smith et al., 2003). Within this perspective, and based on the ellipsis phenomenon, it is hypothesized that there will be transitional constructions during the one or the two-word phases in the Spanish learning process. These transitional constructions would be more complex than the simple juxtaposition of two words but less complex than a simple sentence.

The investigation of transitional constructions to syntax with production data such as ours raises an important theoretical and methodological problem: with their use, the child's production is much enriched, and the researcher might get the impression that this type of utterance is already fully syntactical, productive, even when it is not (Tomasello, 2000, 2003). The origin of the current study is precisely this. We analyzed Mendía's production, a Spanish-speaking girl observed during the one and two-word phases. Our preliminary, informal, inspection of the longitudinal data collected, identified, during the one and two-word phases,

and before age 2 years, cases where it seemed the child had already constructed elliptical sentences. For example, the dialogue between the child (CHI) in our study, Mendía, at 23 months and her mother (MOT):

MOT: he is also asleep, look! (indicating a doll).

CHI: éte tído ["este dormido"] ("this one sleeping") (indicating a doll which is sleeping).

CHI: éte no ["este no"] ("this one no") (indicating a doll which is not sleeping).

MOT: not that one.

The second utterance of the child, apart from the phonology, could have been produced by an adult. The child's response appears superficially to be a correct elision of the verb: "This one is not asleep." From a nativist point of view (Drodz, 2002), it could appear that the syntactic rules are "already there" and the child is constructing correct elliptical sentences albeit suppressing part of the phonological realisation. In contrast, for example Bloom & Lahey (1978) identified early elliptical productions such as our example but considered that children produced them long before knowing the rules which governed them.

Studies previously carried out on the language development of this child (López Ornat, Nieva & Martínez, 2008; Nieva, 2009, 2013) show that between 20 and 27 months, the duration of our study, the child first went through the "one-word phase" (20-23 months) and then into the "twoword phase" (25-27 months). The above example belongs to the two-word phase. Although not forming an elliptical sentence, those two words do present morphosyntactic markers which relate them to each other and also to the child's previous utterance. Taken "vertically", those two successive utterances form a complete "sentence", e.g.: "This one is not [asleep]". What is important, is that the grammatical constraints which characterize the early "ellipses" make these constructions unambiguous, by which they rank higher than "two-word" productions in linguistic development and become closer to sentences. In this sense, they constitute first steps into syntactic construction, regardless of their consisting in one, two or three words. Another reason for the interest in early "ellipses" is that they highlight the developmental importance of vertical construction. The function of this type of construction has already been identified for both the two-word phase (Behrens & Gut, 2005) and the earlier transition from one word phase to two word phase (Carranza et al., 1991). That is to say, the early vertical "ellipses" are a cooperative mode of linguistic construction and can be an observable and measurable index of social determination in the process of language acquisition, since verticality is a shared text, dialogue.

To our knowledge, no systematic investigation has been carried out in any language on the early production of vertical ellipses during the "two-word phase", and their part in the transition to syntactic construction. In our view this study is perfectly consistent with others seeking to identify

exactly what are the constructions that enable the child to make the transition to early syntax by way of learning processes. As implied in the research already reviewed, it is assumed these processes are facilitated by a pragmatic context (e.g.: vertical construction) where linguistic advances are obtained (e.g.: the pre-ellipses) avoiding the memory load demanded by a horizontal construction of those same ellipses. At the same time, the proper use of vertical pre-elliptical constructions can result from short scope statistical abstractions, done on the relations of formal variants of words (e.g.: morphology) to their semantic function within weakly generalized contexts.

Ellipsis

The function of early "ellipsis" in syntactic acquisition does not appear to be known, as we see, neither is the importance of ellipsis in the daily language of adults. In the words of Lise Menn:

People say lots of things [...] that are treated as complete in spite of being only phrases, [...]. The idea that people do or should talk to each other in complete sentences is just silly; the question Where's my mittens? can be answered equally well by They're on the bed or just On the bed. Even bed is possible, although grumpy-sounding. Only some language textbooks for foreigners and some programs for children with language difficulties insist on complete sentences all the time (Menn, 2011, p.46).

As is known, ellipsis is a type of construction where one or more words that would be necessary for the complete sentence are omitted without detracting from the meaning, for example: "John has read the same book as Pedro [has read]" (RAE, 2001). The RAE account of Spanish grammar indicates that the elision of a constituent is subject to certain structural restrictions known as "recoverability conditions". Moreover, ellipsis may be situational, discursive or both: the elided constituent may be in the material context (situational ellipsis), in the previous text (discursive ellipsis) or in both. Discursive ellipsis is so named because the structure of the production includes the elided grammatical categories, e.g.: the grammatical features of person, number, gender ought to be explicit in the same sentence that includes the ellipsis (Brucart, 1999). In the example we gave earlier we have a situational ellipsis in the second turn by the child ('éte no' (["este no"], "this one no"), indicating the other doll which was not sleeping). It is an ellipsis because (1) the elided verbal phrase (is [not] asleep) is recoverable by the interlocutor (MOT) as it appears already in the two immediately prior turns (MOT and CHI) and (2) the structure of the elliptical production includes the elided grammatical categories, i.e.: "éste" connotes the masculine gender and the singular of the noun "muñeco" (doll) and "no" is a semantically appropriate adverb in this construction. Adopting the criteria described and taking a more conservative stance, we assume in this study that utterances such as our example are not true ellipses, although superficially they seem to be, but have been

learned through association with the situations in which they occur (Bloom & Lahey, 1978), i.e.: they are in essence local, not productive. For this reason, we have called them preellipses, precedents of ellipsis. It seemed possible that over time, but within the one and the two-word phases, these pre-ellipses might lose their local nature, gaining linguistic generalization and thus bringing the child closer to the adult model.

Within the field of language acquisition, ellipsis is known as an acquisition which facilitates narrative for children over 3;6 who have been constructing correct sentences for some time (Berko & Bernstein, 2009; Berman, 2009; Berman, & Slobin, 1994). For older children, of 5 to 12 years, there is the investigation by Callahan, Walenski & Love (2012) into comprehension of elliptical sentences of varying degrees of ambiguity. Recent constructivist work (Lieven, 2008) has identified elliptical productions (in English) in children aged 2;4 and 2;8, as well as in their caretakers. And works within the nativist tradition, (Allen, 2009) point to the high frequency of ellipsis in CDS (Child Directed Speech) in many languages, included English, and raise the question of how, under these conditions, children receive the data they need to define the structure of their language. Finally, within the area of linguistic dysfunction, the excessive or sometimes incorrect usage of ellipsis has been found with language impairment (Kolk, 2001) or SLI children (Pérez, 1997; Serra,

As previously mentioned, in this research an ellipsis ought to meet grammatical and lexical constraints such that the listener can retrieve the omission and hence that the ellipsis as a whole contains all the omitted categories in its production, including, for example, grammatical concordances. For a small child, it may be possible to yield accurate situational ellipses based only on the context-ellipsis association. However, in the case of a discursive pre-ellipsis, the "context" is linguistic, which implies a leap in abstraction of the representations to be processed; always, of course, excluding routine dialogue.

On the other hand, private language excluded, all language is social. The children's output in language production studies is extracted from dialogue situations. Many researchers, from pioneers such as Bloom, to more recent, such as Veneziano (1999, 2010), have taken this into account and have analyzed early language within the setting of the dialogue. They have considered the structure of the child's utterance not only by itself (horizontal perspective) but also the structure (semantic, pragmatic, grammatical) that results from taking into account the immediately prior turn in the dialogue (vertical perspective). This has enabled the discovery, for example, of advances, during the "one" to the "two word phase" that are expressed at first vertically, until the child is able to use the same construction as a stand alone, horizontally (see also Carranza et al., 1991). Therefore, in our study we are investigating the next developmental advance, the transition from one or two words to the first simple sentences, within child-adult dialogue. All constructions lacking syntactic organization that correspond to utterances of one, two, or three words, are considered as non-sentences (NOR).

Hypotheses

To recap, the following hypotheses will be tested: (1) when analyzing the child's output in its dialogic context, transitional constructions will appear in the one or in the two-word phases exhibiting greater linguistic organization than the mere juxtaposition of "two words" but less than sentences. In particular, between 20 and 27 months, a predominance of one and two-word utterances (NOR), the presence of syntactical transition constructions such as correct but non productive pre-ellipses (EPS), and the marginal presence of sentences, correct but not productive (OR), are expected. The second hypothesis (2) expects that the seven month period to be analyzed, all of it during the one and two-word phases, will not be homogeneous, rather the gradual nature of this transition will be expressed in changes in the proportions and characteristics of these constructions (NOR, EPS, OR) over time (T). The third hypothesis (3) expects the linguistic complexity of EPS (transitional constructions) to gradually increase. Specifically, it is expected that over time (T), both the distribution of situational, discursive and 'both' EPS and their semantic specification will significantly change. Finally, the distance in turns of the discursive EPS to its linguistic referent will be analyzed, from a purely exploratory point of view, given the lack of prior information on this characteristic.

Method

Participants

This is a longitudinal study (n = 1), carried out on a monolingual Spanish girl (Mendía), with normal development, without any neurosensory alterations, living in Madrid. The family are of middle socio-economic status.

Procedure

The study is based on longitudinal data recorded and transcribed by Nieva (2013) from video recordings focused on the child's play and daily activities at weekly intervals for a total of seven months. Mendía was filmed from the age of 1 year 8 months (1;8) to 2 years 3 months (2;3) in familiar settings (home, park, holiday home) in interaction with her close relatives: mother, father, grandmother. The recordings capture every detail of the unfolding scenes, with objects and situations easily observable. Each recording lasts between 30 and 47 minutes, with one exception of 16 minutes (see Table 1).

Temporal Sampling

The original recordings (Nieva, 2013) comprised 26 filmed sessions. Given the developmental nature of the hypotheses, these have been divided into five time periods (T1, T2, T3, T4 and T5) instead of analyzing only the initial T1 compared with the final T5. The study begins with the T1 slice, when the child is 20 months old. The five time periods cover seven months of the child's development. The child's follow-up during these months tries to capture all her development until she produces the first complete sentences. Each time slice corresponds to a different age (1;8, 1;9-1;10, 1;11, 2;1 and 2;3). With the constraint that each time slice should be at least a month separate from the following slice, we have selected those that contain the best sound quality. Each time slice contained three sessions except for T4 and T5, given the large amount of material recorded. In particular, the last time slice (T5) consists of a single, lengthy session, very abundant in linguistic production. Thus, there were 12 sessions in total (see Table 1).

Table 1. Temporal Sampling

Table 1. Temporal Samp		Age	Duration	Partner	
1	Mendía 01	1;8.03	30′ 29′′	Mother	
1	Mendía 02	1;8.09	38′ 32′′	Mother	
1	Mendía 03	1;8.18	33′ 10′′	Mother	
2	Mendía 08	1;9.22	16′49′′	Mother	
2	Mendía 09	1;9.27	35′ 26′′	Mother	
2	Mendía 10	1;10.3	49′30′′	Mother	
3	Mendía 16	1;11.15	45′ 13′′	Mother	
3	Mendía 17	1;11.22	45′ 19′′	Mother	
3	Mendía 18	1;11.29	44′ 54′′	Father	
4	Mendía 24	2;1.13	47′ 52′′	Mother	
4	Mendía 25	2;1.21	46′ 18′′	Mother	
5	Mendía 31	2;3.3	44′27′′	Mother	

Data Transcription

For this study, the transcription of the output was completely revised by the authors. Its original version (Nieva, 2013) used CHILDES' CHAT (MacWhinney, 2000) format. In this format, the exact utterances of each participant are transcribed orthographically, clarifying within brackets the equivalent word when necessary. The additives or filler elements (fillers) are transcribed with @fs. Each transcript line is linked to both audio and video, making it easy to access the original in the complete situation. This linkage was indispensable for the re-analysis of the output for this study. Agreement between transcriptions for this study was high and disagreements (less than 5%) were jointly resolved. Moreover, a third judge, blind to the hypotheses, coded phonetically all constructions in which there were additional sounds difficult to interpret, mainly but not exclusively, pre-

fixes to nouns and verbs. This acoustic analysis was twofold: simple listening and with the help of Praat, a tool for the computer analysis of phonetics (Boersma & Weenink, 2001). In addition, the phonetic transcriptions produced by this third judge were compared with this study's transcription and agreement in over 90% of the cases was found.

Coding

Five types of constructions were coded: one, two or three-word utterances (NOR), "sentences" (OR), "sentences" with an error (OR-E), pre-ellipses (EPS) and pre-ellipses with an error (EPS-E). These are explained as follows:

- NOR (an *utterance with one, two or three words*): utterances consisting of at least one recognizable word and typically two juxtaposed words; they may include three, and there is no indication of any morphosyntactic organization. Examples:
 - 1- Mendía (CHI) is staking cubes on top of each other. While she picks up a green cube, she says:
 - *CHI: atí petíllo. ["Aquí amarillo"] ("Here yellow").
 - 2- The Mother (MOT) and Mendía (CHI) are reading a storybook, in which there is a child picking apples.
 - *MOT: what is the child doing?
 - *MOT: Look # Mendía.
 - *MOT: what is he doing?
 - *CHI: nene má:na ["Nene manzana"], ("Child apple").
- OR ("sentence"): utterances having the surface form of a sentence, with at least one word which in the adult language would be a verb and with all the mandatory grammatical morphemes. It equates to a construction which is complete morphologically, semantically and pragmatically. It may be immature phonologically, provided that this does not affect a mandatory morphological marker and is not so immature that it cannot be understood by persons outside the family. Examples:
 - 1- The child picks up a cushion to sit on the floor and play. She asks her mother:
 - *CHI: éste é de Menía? ["¿Este es de Mendía?"] ("Is this Mendía's?").
 - 2- The child is drawing on a magnetic board with her "pen". While looking at another instrument she says:
 - *CHI: ése no pi:ta. ["Ese no pinta"] ("That one doesn't paint").
- OR-E ("sentence" with an error): has the form of a sentence with a single error. The error may be (1) phonological,

- (2) of omission or (3) of commission. Complementarily to the previous criterion, it is coded as a phonological error when the version produced by the child can only be understood by her close family. An error of omission is where the child omits a mandatory structure, a function word, an inflective or derivational morpheme. An error of commission is where the child chooses the wrong function word, morphological marker, or syntactic order. Examples:
- 1- Looking at a jigsaw puzzle, the mother asks Mendía to look for the raincoat piece; Mendía picks it up and says:

*CHI: ése es e@fs chubacá. ["Ese es el chubasquero"] ("That is the raincoat").

Coded as phonological error.

2- While the Mother is drawing colored balloons on a sheet of paper, Mendía points at one and says:

*CHI: éte # é bó:bo. ["Este es globo"]. ("This is balloon").

Coded as error of omission. The correct sentence would have been "this is **the/a** balloon". The child omitted the mandatory article "the/a".

- 3- While putting a nappy on a doll, Mendía says:
 - * CHI: ía caca a@fs peté:co. ["Tenía caca a muñeco"] ("The & fem. doll had poopoo").

Coded as an error of commission because the *filler* is given the feminine gender where it should have been masculine. The correct sentence would have been "ía caca e@fs peté:co". ("The & masc. doll had poopoo")

- EPS (pre-ellipsis): the pre-ellipsis has the surface form of an adult ellipsis: a sentence which is incomplete but correct semantically, morphosyntactically and pragmatically. The elided constituent can be found in (1) the context-situation, (2) prior discourse or (3) both at the same time. It is assumed that the child's ellipses are local, non-productive. Therefore, they are coded pre-ellipses, precedents of ellipses. Examples:
 - 1- The child touches a blemish on her mother's arm as she says:
 - * CHI: é: pú:a ["Es pupa"]. ("Is owie").

Mendía omits the subject of the sentence, which is physically present in the situation and moreover she touches it while speaking. 2- Mother and child are doing a puzzle pairing things that go together; there is only one pair left to complete and mother asks:

```
*MOT: Are you going to do it?
```

*MOT: or will I?

*CHI: no: yo ["No, yo"] ("No, me").

The child elides the verb, which is present in the previous turn.

Excluded are routine constructions, for example: While the mother tidies the storybooks they were coloring, the child comments:

```
*CHI: mía, mía ["Mira, mira"] ("look, look").
```

Also excluded are constructions modeled by the mother in the dialogue, for example: While they are using toy saucers to draw circles, the mother asks:

```
*MOT: what are plates for?
```

*MOT: for \pm /?

*CHI: pa comé ["Para comer"] ("For eating").

- EPS-E (single-error pre-ellipsis): a pre-ellipsis with a single error is coded EPS-E (if two or more errors occur, it is coded NOR). An error in an EPS may be of omission or of commission. Examples:
- 1- Holding a ball the mother comments that it could be something for their dogs to play with and asks:

```
*MOT: are you going to let them play with it?
*CHI: no: é: Mendía ["No, es [de] Mendía"]. ("No, is Mendía['s]").
```

In this example, the correct pre-ellipsis would be "No, [the ball] is Mendía's"; however, the child has omitted "de" ['s]. It is coded pre-ellipsis with an omission error.

2.- Mother and child are looking at picture cards and naming them one by one.

```
*MOT: let's see # what is that?
*CHI: o@fs cóse [: Coche] (:Car).
```

Here, the error occurs in the gender of the *filler*. The correct pre-ellipsis would be: "Un coche" [a car]. It is coded as pre-ellipsis with an error of comission. Single words and routine expressions such as proper nouns, determinants, greetings, exclamations, adverbs, and "más" [more], "todo" [all], "ya" [already/now], "sí" [yes], "no" and "caca" [poo], are not included in the coding, nor are any doubtful utterances.

For the analysis related to the third hypothesis (the advance in complexity of the EPS over time), the following codes were used:

- EPS *situational, discursive, both*: according to whether the elided constituent formed part of the situational context, the prior discourse or both. Regarding the discursive or 'both' EPS, it is important to remember that in all cases where the child responds with a routine utterance were removed from the data. Therefore, in these discursive or "both" EPS the child is continuing a sentence started by others. The following are some examples of the three types of the *locus* of the elided constituent, as defined by the Royal Academy of the Spanish Language (Brucart, 1999; RAE, 2010) and adopted here:
 - The child is trying to unscrew a doll which contains another:

*CHI: no se@fs á:be ["No se abre"] ("It doesn't open").

[She omits: the doll]. This is coded situational EPS.

2.- The child is drawing in the presence of her father

*FAT: Mendía # what are you doing? *CHI: pitáno ["Pintando"] ("Drawing").

[She omits: "estoy" ("I'm")] This is coded discursive EPS

Mother and child are tidying away some balls they have been playing with.

*MOT: let's see ... more.

*MOT: more balls.

*CHI: no hay ("There aren't.")

[She omits: more balls] This is coded 'both' EPS.

- Distance in turns: How many turns of speech separate the adult utterance and the child's pre-ellipses is coded for all EPS discursive or 'both'.
- Finally, coding is applied to the semantic change in an EPS, followed across the time slices and considering all its to-kens. Note that the contexts in which it is used were expected to be wider and more differentiated each time.

Data analysis

In order to analyze the different types of constructions and their evolution, firstly the relative frequencies of each type (NOR, OR, OR-E, EPS and EPS-E) were obtained using CHILDES' CLAN program. Afterwards, the significance of the changes in frequency of the constructions across the five time slices was tested using a chi-square independence test. Adjusted residuals analyses were performed using SPSS Statistics 19. In all cases, calculations and analyses were separated for types and for tokens, since their frequencies can reflect different internal processes. For example, the presyntactic type "ese no" [not that one] is used by Mendía on 34 occasions (34 tokens) during record 24 (age 2;1) which appears in T4. She uses it to refer to a doll she does not want her Mother to take, to a puzzle piece that does not fit, a toy that is not the one she wants, etc... Thus 1 type, and 34 tokens. Given the repetitive nature of the games and interactions at these early ages, this differentiation in the analysis of the data is essential. There are types with high frequency and others practically unique, and their developmental implications are clearly distinct. Analyzing linguistic structures diachronically, different frequencies of tokens reflect different uses and different frequencies of types indicate productivity. The chi square test of independence was also applied to the number of words in NOR constructions, the locus of the elided constituent and the distance in turns of the EPS over the time slices.

Results

A total of 2411 valid utterances from Mendía were obtained. The number of valid utterances for each time slice varied between 271 and 644, as can be seen in Table 2.

Table 2. Sample of observed tokens and types and their ratio across time

	Time slices					
	T1	T2	Т3	Т4	T5	Total
Tokens	584	271	644	599	313	2411
Types	259	183	479	386	257	1564
Type/Token Ratio	0.44	0.67	0.74	0.64	0.82	0.65

In all time slices except T1, the type/token ratio is greater than 0.5, indicating an enrichment of vocabulary starting from T2. The evolution of the five different classes of construction for each time slice are distributed as shown in Table 3 for types and in Table 4 for tokens.

Table 3. Frequency of each construction, in types, per time slice (adjusted residuals in brackets).

		Time slices					
Classes	T1	T2	Т3	T4	Т5	Total	(%)
NOR	200 (2.5*)	146 (2.9*)	363 (2.9*)	257 (-2.02*)	139 (-6.3*)	1105	70.65
EPS	21 (-2.9*)	22 (7)	54 (-2.00*)	60 (1.1)	60 (4.8*)	217	13.87
EPS-E	19 (4)	7 (-2.2*)	41 (.5)	40 (1.9)	19 (43)	126	8.05
OR-E	15 (1.9)	2 (-2*)	5 (-3.7*)	15 (.2)	21 (4.1*)	58	3.7
OR	4 (-2.02*)	6 (3)	16 (5)	14 (1)	18 (3.06*)	58	3.7

The chi-square test revealed the existence of significant changes in the frequencies of different classes of construction throughout the five time slices (χ^2 (16, N = 1564) = 90.358, p = .000). This shows that time influences the relative frequency of the production of the different constructions. Analysis of adjusted residuals shows that this effect is due to a decrease of NOR constructions in favour of EPS, and to a lesser extent of OR-E and OR, all of which show a significant increase in T5. Over the 7 months, the most frequent construction is always NOR. However, across time, the frequency of this construction decreases and, in its place, the frequencies of the other constructions, all of which have some syntactic structure, increase. The pre-syntactic construction that increases most is the pre-ellipsis. "Sentences", with a single error (OR-E) and correct (OR), are always marginal class but do increase in T5. This increase indicates the beginning of the acquisition of OR and sets the developmental "ceiling" of our study.

The analysis of tokens produces an equally significant result (χ^2 (16, N=2411) = 130.236, p=.000) and reproduces almost exactly what has been seen for types. However, there is a difference in the case of single-error ellipses (EPS-E), as this construction shows a significant increase in T4 (see Table 4).

Table 4. Frequency of each construction, in tokens, per time slice (adjusted residuals in brackets).

		ne slice					
Tokens	T-1	T-2	T-3	T-4	T-5	Total	(%)
NOR	470(3,9*)	224(3,3*)	493(1.5)	429(-1.7)	174(-8.1*)	1790	74.24
EPS	73(-1.2)	31(-1.3)	79(-1.5)	82(2)	73(5.08*)	338	14.01
EPS-E	19(-3.3*)	8(-2.3*)	45(1.05)	54(3.3*)	22(.7)	148	6.14
OR-E	18(.8)	2(-2.05*)	5(-3.4*)	16(.1)	22(5.2*)	63	2.61
OR	4(-3.7*)	6(8)	22(.7)	18(.03)	22(4.5*)	72	2.99

Given the equivalence of the results obtained for types and tokens, the remaining analyses considered tokens only. Regarding the NOR, their number of words in all time slices was analyzed. The categories were: 1W (one word), 2W (two words) and 3 + W (three or more words). The chi-square test of independence showed that the categories differed significantly over time (χ^2 (8, N=1790) = 265.685, p=0.00). The adjusted residuals analysis showed that there is a relative decrease in frequency of one-word in favour of two words which show a significant increment at T4. Meanwhile, three or more-word utterances do not show a significant presence until T5. This trend can be properly observed by considering the proportions of each NOR sort in the total NOR utterances, per time slice (Figure 1).

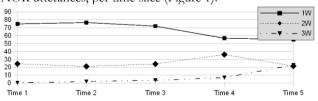


Figure 1. Development of NOR classes in percentages across five time slices (tokens).

Regarding the linguistic complexity of the EPS, these were analyzed for the *locus* of the elided constituent: if it could be found in the situation, the prior discourse, or both. The test for independence was carried out on tokens, as explained earlier. In this case, meeting the constraint that a type could be used equally as situational, discursive or both. The test (χ^2 (8, N=324) = 62.970, p=.00) showed that the relative frequency of *locus* varies significantly over time. The adjusted residuals analysis shows that there is an almost absolute predominance of situational pre-ellipsis in T1, and that it significantly decreases in T3 and T5. In addition, the final increment of the 'both' EPS in T5 is also significant. Finally, discursive EPS is marginal across the five time slices. The development of the different *locus* types of ellipses is shown in Figure 2.

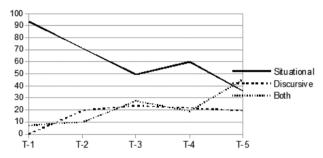


Figure 2. Development of situational, discursive and 'both' pre-ellipses (percentages).

Examining the distance in conversational turns of the omitted element in the discursive pre-ellipsis, data show that, at 80.8% of the total, the omitted element is overwhelmingly in the immediate prior turn. Only occasionally is the missing element tied to two, three or up to six turns earlier. The omitted element in situational and 'both' EPS is predominantly located in the immediately prior turn of the conversation (see Figure 3).

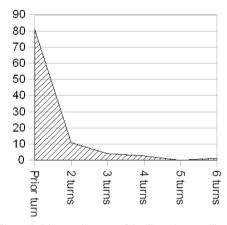


Figure 3. Distance, in turns, of the discursive pre-ellipses.

Finally, it was expected that over time, the pre-ellipses would evolve towards greater semantic complexity and wid-

er and more differentiated contexts of use. To look at this, the sequence of pre-ellipses was tracked. However, this yielded only one case of the same pre-ellipsis which appeared over three consecutive time slices. This was the pre-ellipsis "a guardar" ["put away"], very frequent in the early records and with a tendency to decrease after Time 3. When dealing with filming in natural settings, there is a wide variety of lexical items: games, elements and sequences change as the interests of the child over time, hence the difficulty of such tracking.

It is known that the construction "a guardar" ["put away"] is massively used by children acquiring Spanish. This child does show a very high use frequency. The analyses have been carried out on the grounds of the strong representation of this particular EPS. Structurally speaking, "a guardar" is already an elliptical construction, which omits the conjugated verb. Moreover, mother and child (it also occurs in the rhymes they sing) use it omitting the direct object (toys, for example) as well. That is to say, it is used as an imperative. The first tokens of "a guardar" were uttered by the child at 20 months of age during the act of storing objects in a specific container, i.e., they were clearly local uses. After a while (at 21 months) this pre-ellipsis was being produced in wider and more discriminative uses (meanings), before starting to tidy up, or to express preferences "put away no", or to indicate into which container it should be placed. That is to say it had a wider semantic function than its first use in a local context (Table 5).

Table 5. Longitudinal follow-up of a pre-ellipsis.

TYPE	FUNCTION (Semantic) TIME/AGE
A single type	A single semantic use (putting ob- T-1 / 20 mo
(put away) $n=83$	jects into a container)
	Several* semantic uses (putting
A single type	objects into a container, starting T-2 and T-3 /
(put away) $n=45$	to tidy up, to express preferences 21-23 mo
	-"put away no"-, to indicate in
	which container it will go)

^{*}minimum of three occurrences with the mentioned semantic function.

Discussion

Regarding the occurrence of transitional constructions within the "two-word phase" over the observed time slot (Hypothesis 1), results show that although NOR utterances are predominant, the period is not homogenous, since the frequency of this construction decreases with time and in its place the relative frequency of transitional constructions, such as EPS, increases. At the end of the period studied (2;3 years), full sentences significantly start to appear, forming the developmental ceiling of this research. A radical nativist reading (Drozd, 2002; Allen, 2009) might suggest these results are determined by the maturation of some syntactical knowledge genetically transmitted, enabling the very early construction of productive elliptical sentences. However, other results, from testing hypotheses 2 and 3 are incompatible with this idea. Thus (Hypothesis 2), it is noteworthy that

the analyzed period is not homogenous. During T1, T2 and T3 (20 to 23 months) one-word utterances predominate, marginally accompanied by all other types of constructions. In T4 (25 months) two-word utterances significantly increase, with one-word utterances decreasing in parallel and with a significant use of single-error EPS. Later on, in T5 (27 months), a significant increase in three-word and in 'both' EPS is evident, together with an emergence of full sentences. The analyses regarding the linguistic complexity of EPS (Hypothesis 3) show it changes during this period. EPS production in T1 is almost entirely situational, its relative importance gradually decreases and, by T5, 'both' EPS significantly increases. The discursive EPS is marginal throughout this development. Therefore, it seems that the development of pre-syntactic constructions -EPS- is controlled by their gradual independence from their production context. The implication is that the child is successfully associating a specific EPS utterance with a specific situation, which would correspond with a process of local learning. Seven months later (T5), after the development of the two-word stage (T4), the 'both' EPS constructions become important. The 'both' EPS are vertical; they are integrated into the linguistic structure of the dialogue. They are also situational but they comply with the linguistic constraints expressed in the previous turn. The production of the 'both' EPS becomes significant at the same moment as that of the three-word NOR utterances (T5). It could be speculated that the working memory amplitude needed to say two sequenced words is not enough to build the 'both' EPS. Perhaps adjusting to combinatorial linguistic constraints when grammatical knowledge is scarce and very weakly automated (27 months), generates too much extra processing load. It also seems that the cognitive resources needed to build three-word NOR (T5) are partly the same as those required to build 'both' EPS, and the first sentences. Also, in order to increase syntactic productivity the child has also to carry out analyses that yield at least partial morphosyntactic regularities. In fact, at 27 months, this child can both juxtapose three-word NOR constructions and start combining syntactically in a limited manner, using 'both' EPS as well as a minimum of sentences. It is known, though it is not reflected in this data, that development will continue by means of combinatorial productivity and will abandon the NOR juxtaposition path. These results show a gradual and modest development of linguistic knowledge, which does not reach discursive EPS, not even vertically. Tracking of the only EPS that could be followed (Table 5) shows, long before T5, that the child is using in T1 an identical form to what she uses in T2 or T3 ("put away"). Underlying that form, the function is developing, and it is doing so in the same sense we have seen before, that is, becoming independent from its first local semantic referent (T1).

All in all these results do not support a nativist perspective. If Mendía's EPS "already were" productive elliptical constructions, then they wouldn't have to develop from situational to 'both', neither would they have to occur only vertically, nor should the discursive EPS not increase. Moreo-

ver, the fact that this limited development takes seven months to occur would be difficult to explain, and the intermediate steps detected such as the development of semantic discrimination, the increase in the number of words in NOR utterances or the increased use in T4 of single-error EPS, would not make any sense. The vertical 'both' EPS involve the same kind of pre-syntactic progress found by other authors and/or for other languages, within a constructivist approach. For example, positional patterns or pivot grammar (Braine, 1976), "defective rules" (López Ornat, 1994), "frame & slot" (Pine & Lieven, 1997) or "productive patterns "(Cortés, 2003). The vertical 'both' EPS construction found in these data, ensures the cohesion of early dialogues. The linguistic constructions of dialogue have been widely studied (Berko-Gleason & Ratner, 2009), both as part of early pragmatic development (for a review, see Bryant, 2009) and as a determining variable in the process of early language acquisition (Veneziano, 2010). Our results make specific this relationship for the Spanish language. The development of the vertical 'both' EPS is a firm candidate for a specific mechanism of early syntactic transition in Spanish.

From a formal point of view, this gradual learning of pre-elliptical constructions perhaps has effects that go far beyond the acquisition of elliptical constructions as such. It is possible that the acquisition and correct manipulation of a variety of pre-elliptical constructions helps the child to achieve his/her first complete sentences by simply filling in the omitted constituent from pre-ellipses already in her repertoire. For example, in T3, Mendía says "náme" ["dame"/"give me"] asking for an object she is pointing to and / or looking at. In T5, she already produces the OR: "dáme la ná:na" ["dame la rana"/"give me the frog"] to request a flashcard. She manages to produce a sentence by adding a constituent she has usually omitted, using a process less demanding of working memory, than constructing a complete sentence from scratch. This particular relationship between the development of pre-ellipses and the acquisition of sentences has not, however, been the object of this study, for which the emergence of sentences is an end point. On the other hand, it would be expected that vertical, and certainly horizontal discursive ellipsis, would develop after and would depend on the prior acquisition of complete sentences. For example, all of Mendía's discursive EPS or 'both' EPS, over the five time slices, were vertical, i.e. they omitted a constituent expressed in the previous dialogic utterance. The late character of horizontal discursive ellipses is also implicit in the pioneering work of Bloom & Lahey (1978) who defined ellipses as an omission of redundant elements that contributes to the cohesion of speech, and who added, "for example: leaving out the words you and going in question: where are you going?, when responding 'to the store'" (p. 223), which is an example of vertical ellipsis in dialogue. Mendía never produced a horizontal discursive EPS such as "I'm going to play with plasticine, and you too". Another interesting detail that emerges from these results is that, during the 20 to 27 months period, the distance of a single turn between the adult utterance and the infant's vertical EPS seems to be the norm in adult-child dialogue (Figure 5). This analysis on the distance in turns has filled the information gap on this aspect of the structure of adult-child dialogue.

The results obtained for NOR constructions are consistent with what is currently known about such constructions (Veneziano, 1990; Carranza et al., 1991). To recap, these results show a "one-word" phase between 20 and 23 months (T1, T2, T3), in which there are also "two-word" NOR constructions. This is followed (T4, at 25 months) by a "twoword" phase with a remarkable decrease in "one-word" utterances. Finally (T5, at 27 months), a significant use of three-word NOR constructions is observed, together with a decrease in "two-word" combinations, probably because they have been treated morphosyntactically, being placed in one of the categories with some morphosyntactic organization (EPS, EPS-E, OR-E, OR). The data confirm something well known in the field: the index of linguistic development is not so much the number of words in an utterance but the presence or absence of morphosyntactic organization however partial.

Taken together, the results allow to hypothesise the gradual learning of syntactic construction in Spanish and already clearly reflect a gradation from simple to complex in the constructions, i.e.: first, "one-word" utterances and situational pre-ellipsis, afterwards, "two-word" utterances, and last, "three-word" utterances, the 'both' EPS and the start of full sentences. It is noteworthy that during the seven months studied, all within the "one or two-word" phases, there were no clear cut boundaries in this development. The most primitive constructions (NOR) coexisted throughout (though decreasing) with the transitional constructions (EPS) and at T5, all these coexisted with the most advanced (OR); additionally no type of construction appeared or disappeared abruptly. Methodologically, it should be emphasised that the data have been obtained by focusing on linguistic development within dialogue, not simply taking the output of the child or the input from the adult. The rationale is that in dialogue, language forms are "negotiated" between the child's level and the speech of the adult and it is from these that the child extracts her input (Bråten, 2009; Goldstein et al., 2010; Veneziano, Sinclair & Berthoud, 1990; Veneziano, 2010). Recall that in the first example (Introduction), the second intervention of the girl would have been coded -as output a "Det. & No" ("This one no"), and perhaps as a NOR. It is solely by considering the dialogue that one can see that it is a complete and correct (superficial) ellipsis. We are aware that it would be interesting to follow the linguistic development of this girl after 27 months to pin point how, in her case, the advances involving pre-ellipses contributed specifically to the early acquisition of her first sentences. We also will, in future work, analyze the detail of her errors, those committed in EPS (yielding EPS-E) and in OR (yielding ORE), searching for further detail on this developmental process. This investigation also raises the question of the role played in her development by the language

directed to her (CDS -Child Directed Speech). As has been seen, from an nativist perspective (Allen, 2009) one would expect that the elliptical speech of parents would impede children's acquisition of syntax by 'impoverishing the stimulus' of their linguistic input. However, data from this child suggests that the elliptical constructions in CDS can facilitate early syntactic development. A forthcoming paper examines this question using the CDS in the sessions reported here. Finally, another question which arises from this study is the function, in this development, of the nature, nominal or verbal, of the constituents elided or expressed, analyzed in both CDS and the child's output. This variable, specifically

linguistic, will also be explored as part of the investigation into the early process of construction of early syntax.

Acknowledgements.- This research has been supported by the DGI:SEJ2007-67810/PSIC project and is part of the project on early grammaticalization led by E.Veneziano: EMERGRAM (2007-2011): The emergence of grammaticality in children: Cognitive, linguistic and conversational factors. ANR, Agence Nationale de la Recherche, France. BLAN061 135249.

The philologist Irmgard Rebele carried out the phonetic encoding of the subject's output and also collaborated in calculating the coincidence-discrepancy of the results as a third judge. We wish to thank her for her outstanding contribution to this work.

References

- Allen, S. (2009). Verb argument structure. En E. L. Bavin (Ed.), The Cambridge Handbook of Child Language (pp. 217-236). New York: Cambridge University Press
- Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. Current Directions in Psychological Science, 21, 170-176. doi:10.1177/0963721412436806.
- Behrens, H., & Gut, U. (2005). The relationship between prosodic and syntactic organization in early multiword speech. *Journal of Child Language*, 32, 1-34. doi:10.1017/S0305000904006592.
- Berko-Gleason, J. B., & Ratner, N. B. (2009). The development of language. Boston: Pearson/Allyn & Bacon.
- Berman, R. A. (2009). Language development in narrative contexts. En E. L. Bavin (Ed.), *The Cambridge Handbook of Child Language* (pp. 355-375). New York: Cambridge University Press
- Berman, R. A., & Slobin, D. I (1994). Relating events in narrative. A crosslinguistic developmental study. Hillsdale, NJ: Erlbaum.
- Bloom, L., & Lahey, M. (1978). Language development and language disorders. Somerset: John Wiley.
- Boersma, P., & Weenink, D. (2001). Praat: Doing phonetics by computer. Glot international, 5, 341-345.
- Braine, M. (1976). Children's first word combinations. *Monographs of the society for research in child development*, 41, scrial 164.
- Bråten, S. (2009). The intersubjective mirror in infant learning and evolution of speech. Amsterdam: John Benjamins.
- Brucart, J.M. (1999). La Elipsis. En I. Bosque, & V. Demonte (Eds), *Gramática descriptiva de la lengua española* (3 Vols.). Real Academia Española (pp 2787-2859). Madrid: Espasa.
- Bryant, J. B. (2009). Pragmatic development. En E. L. Bavin (Ed.), The Cambridge Handbook of Child Language (pp. 339-354). Cambridge: Cambridge University Press
- Callahan, S. M., Walenski, M., & Love, T. (2012). The Processing and Interpretation of Verb Phrase Ellipsis Constructions by Children at Normal and Slowed Speech Rates. *Journal of Speech, Language, and Hearing Research*, 55, 710-725 doi:10.1044/1092-4388(2011/10-0281).
- Carranza, J. A., Escudero, A., & Brito, A. (1991). De las palabras aisladas a las combinaciones de palabras. Anales de Psicología, 7, 163-180.
- Cortés, M. (2003). ¿Una gramática para lingüistas o una gramática infantil? Cognitiva, 15, 133-140.
- Drozd, K. F (2002). Negative DPs and elliptical negation in child English. Lang. Acquisition, 10, 77-122.
- Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences*, 14, 249-258.
- Kolk, H. (2001). Does agrammatic speech constitute a regression to child language? A three-way comparison between agrammatic, child, and normal ellipsis. Brain and Language, 77, 340-350. doi:10.1006/brln.2000.2406.
- Lieven, E. (2008). Learning the English auxiliary. En H. Behrens, & J. Benjamins (Eds), Corpora in language acquisition research: history, methods, perspectives (pp. 61-98). Amsterdam: John Benjamins.

- López Ornat, S. (1994). La adquisición gramatical: un esquema. En S. López Ornat, A. Fernández, P. Gallo, & S. Mariscal (Eds.), La adquisición de la lengua española (pp. 101-126). Madrid: Siglo XXI.
- López Ornat, S. (2003). Learning earliest grammar: Evidence of grammar variations in speech before 22 months. En S. Montrul, & F. Ordóñez (Eds.), Linguistic Theory and Language Development in Hispanic languages (pp 254-274). Sommerville: Cascadilla Press.
- López Ornat, S., Nieva, S., & Martínez, M. (2008). From single-word to multiword utterances in Spanish. The growth of linguistic complexity and its conversational support. British Psychological Society, Developmental Section Conference. Oxford Brookes University.
- López Ornat, S, Férnandez, A., Gallo, P., & Mariscal, S. (1994). La adquisición de la lengua española. Madrid: Siglo XXI.
- López Ornat, S., Gallego, C., Gallo, P., Karousou, A., Mariscal, S., & Martínez, M. (2005). MacArthur: Inventario de desarrollo comunicativo. Madrid: TEA Ediciones.
- MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Third Edition. Mahwah: Lawrence Erlbaum Associates.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* , 283, 77-80. doi:10.1126/science.283.5398.77.
- Mariscal, S. (2008). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language*, 35, 1-29. doi:10.1017/S0305000908008908.
- Mariscal, S., Nieva, S., & López Ornat, S. (2010). Observar y medir el desarrollo gramatical temprano en español. *Psicothema*, 22, 51-56.
- McClelland, J. L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? Trends in Cognitive Sciences, 3, 166-168.
- Menn, L. (2011). Psycholinguistics. Introduction and Applications. San Diego: Plural Publishing
- Newport, E, L. (2011). The modularity issue in language acquisition: A rapprochement? Comments on Gallistel and Chomsky. Language Learning and Development, 7, 279-286. doi:10.1080/15475441.2011.605309.
- Nieva, S. (2009). Les échanges conversationnels dans l'émergence des premières combinaisons de mots: une étude longitudinale. Colloque Jeunes Chercheurs en Acquisition du Langage. Laboratoire Dynamique du Langage, ENS-LSH, Université de Lyon.
- Nieva, S. (2013). Nieva Longitudinal CHILDES Corpus. http://childes.talkbank.org/data/Romance/Spanish/Nieva.zip.
- Pérez, É. (1997). Cohesión y coherencia en las narraciones de niños y niñas con trastorno específico del lenguaje. Revista de Logopedia, Foniatría y Audiología, 17, 103-111.
- Pine, J. & Lieven, E. (1997). Slot and frame patterns in the development of the determiner category. *Applied Psycolinguistics*, 18, 123-138.
- Pinker, S. (1995). Language Acquisition. En L. R. Gleitman, D. N. Osherson, & M. Liberman (Eds.), An invitation to cognitive science: Language (Vol. 1). Cambridge: MIT Press.
- RAE (Real Academia Española & Asociación de Academias de la Lengua Española) (2001). Diccionario de la lengua española. 22ª Edición. Madrid: Espasa.

- RAE (Real Academia Española & Asociación de Academias de la Lengua Española) (2010). Nueva gramática de la lengua española. Madrid: Espasa.
- Serra, M. (1997). Dificultades cognitivas y lingüísticas en los niños con Trastorno Específico del Lenguaje. Revista de Logopedia, Foniatria y Audiología, 17, 79-91.
- Smith, P., Nix, A., Davey, N., López-Ornat, S., & Messer, D. (2003). A connectionist account of Spanish determiner production. *Journal of Child Language*, 30, 305-331. doi:10.1017/S0305000903005622.
- Tomassello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4, 156-160.
- Tomassello, M. (2003). Constructing a Language. Boston: Harvard University Press
- Valian, V. (2009). Innateness and learnability. En E. L. Bavin (Ed.), The Cambridge Handbook of Child Language (pp. 15-34). Cambridge: Cambridge University Press.
- Veneziano, E., Sinclair, H., & Berthoud, I. (1990). From one word to two words: repetition patterns on the way to structured speech. *Journal of Child Language*, 17, 633-650. doi:10.1017/S0305000900010928.
- Veneziano, E. (1999). Early lexical, morphological and syntactic development in French: some complex relations. *International Journal of Bilingualism*, 3, 183-217. doi:10.1177/13670069990030020501.
- Veneziano, E. (2010). Conversation in language development and use: An Introduction. First Language, 30 (3-4), 241-249. doi:10.1177/0142723710380531.

(Article received:15-11-2011; reviewed: 9-10-2012; accepted: 16-2-2013)