



# Left-dislocated subjects: A construction typical of young French-speaking children?

Christophe Parisse

## ► To cite this version:

Christophe Parisse. Left-dislocated subjects: A construction typical of young French-speaking children?. Guijarro-Fuentes, P., Larrañaga, P. and Clibbens, J. First Language Acquisition of Morphology and Syntax: Perspectives across Languages and Learners, John Benjamins, pp.13-30, 2008. halshs-00167208

**HAL Id: halshs-00167208**

**<https://shs.hal.science/halshs-00167208>**

Submitted on 16 Aug 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title: Left-dislocated subjects: A construction typical of young French-speaking children?

Author: Christophe Parisse

Affiliation: INSERM, MODYCO-CNRS-Paris 10 Nanterre University

E-mail: paris@vjf.cnrs.fr

## **Abstract**

Few constructions in child's production can be considered originals and not mere copies (partial or complete) of adult input. These rare constructions are precious to child language studies because they may reveal a developmental process at work. One of them is children's left-dislocated subjects. In French, when there is no lexical subject, the use of a subject pronoun is mandatory. When there is a lexical subject, two constructions are possible: one with the use of a personal pronoun and one without. Children tend to produce a lot of pronouns in the context of lexical subjects, which makes it a good candidate for being a specific feature of child language.

This study used two corpora of spontaneous language production, one with children aged three to four and one with children aged two to four. The data were analyzed using Tomasello's framework (usage-based theory of language acquisition). In this theory, children's language competence differs from the adults' and develops with age. The first issue was to demonstrate the existence of this specific feature. Indeed, it appeared that children produced more left-dislocated subjects than adults and did not simply reproduce a feature of adults' oral language. The presence of a developmental effect in the children's production was also strongly suggested by the results.

The second issue was to try to better understand what could be the reason for the difference in production of left-dislocated subjects between children and adults. Several explanations of the children's results were put forward, in keeping with the predictions of the usage-based theory.

Lexical and usage-based explanations of children's behaviour were shown to be unlikely. More plausible explanations were performance limitation or a consequence of the development of the use of personal pronouns, even when they are not obligatory.

It is proposed that children's left-dislocated subjects should not be considered as a copy of adults', but as a specific construction pattern that appears during the course of language development. This pattern could provide a path towards the development of more general linguistic abstractions, as proposed in the usage-based theory.

## **1. Introduction**

The theoretical approach to language acquisition presented by Tomasello (2000; 2003) argues that syntactic competence is built in a piecemeal fashion, starting with item-dependent knowledge and gradually generalizing towards the mastery of abstract syntactic structures. This approach considers that the grammars of the child and the adult are different and that the differences get smaller as the child develops her language competence.

One of the consequences of considering language acquisition as a developmental process is that it is important to pay special attention to structures produced by a child that cannot be found in adult language. The difference between the child's and the adult's productions attests that these structures are constructed by the child and not merely reproductions of adult input. In other words, they represent the child's grammar and not the adult's grammar. Pinpointing such constructions opens a window into the underlying processes of language development and allows us to go further than simple linguistic descriptions.

This technique is widely used to study the development of the lexicon and of morphosyntax, where it is easy to apply, as one needs only to find words that are not adult words

to know that they correspond to original morphological or morphosyntactic constructions (setting aside the mispronunciation of existing words). It is unfortunately not so easy when one is interested in the study of syntactic constructions composed of more than one word (from two words up to a whole utterance).

In this case, with the exception of obligatory markers and arguments that are never or seldom omitted by adults, it is not always easy to know if a construction is an original or a copy of an adult construction. Indeed, language flexibility and performance limitations are such that very few children's first syntactic combinations cannot be found in adult language (see Parisse and Le Normand 2000). This is especially true when young children's utterances are very short. The scarcity of constructions totally specific to children's language can be compensated for by looking for constructions unusual in adult's language instead of truly non-adult constructions. This can be done by comparing children's and adults' frequency of production. If there is a significant difference between children and adult productions for a given construction, then this construction becomes a good candidate for specific child construction and grammatical behaviour.

This technique is applied in the study below which deals with the development of left-dislocations of the lexical subject in French children. The choice of this syntactic structure comes from previous work (Parisse and Le Normand 2001) in which it appeared to be relatively frequent in children, although it is not considered to be the most natural construction for adults. Thus, a difference in frequency between children's and adults' use could be expected and left-dislocated subjects could represent an example of grammatical behaviour specific to children.

## **2. Usage-based theory of language acquisition**

Presentations of Tomasello's usage-based theory (UBT) of language acquisition can be found in Tomasello (2000; 2003). UBT is rooted in a construction grammar framework (see Croft 2001; Goldberg 1995). Tomasello (2002) gives a short summary of the main steps in the development of productive language:

- “1. Children hear and attempt to learn utterances, i.e. phonological forms for expressing communicative intentions; in doing so they assign a communicative function both to the utterance as a whole and to certain of its constituents (sub-functions of the utterance's overall function).
2. Children store comprehended and produced utterances and constituents, along with their functional analyses; token and type frequencies of stored pieces of language lead to the construction of (a) utterance schemas and (b) constituent schemas, both of which may contain relatively abstract 'slots' (and it is an empirical fact that children construct these abstractions only gradually and in piecemeal, item-based fashion).
3. Children produce utterances by 'cutting and pasting' from their inventory of stored utterance schemas and constituent schemas.” (Tomasello 2002, 309, 310).

A more detailed account of the theory from a syntactic point of view is found in Tomasello (2006). The issues of grammar learnability and the nature of the adult's grammatical knowledge are discussed in Parisse (2005).

## **3. Left-dislocated subjects**

The analysis of left-dislocated subjects using Tomasello's framework is rather infrequent, especially for French. So it seemed that the presentation of two recent works on the subject,

Blasco-Dulbecco (1999) and De Cat (2002), although unrelated from a theoretical perspective, would provide more insightful analyses of the language data and would help the reader understand the language structures analysed in this paper.

French is a non pro-drop language, which means that subject pronouns are mandatory when there is no lexical subject. For example, in ‘il rit (he is laughing)’, the use of ‘il (he)’ is mandatory. When there is a lexical subject, it takes the place of the personal pronoun. However, it is possible to insert a personal pronoun between the lexical subject and the verb: the construction ‘le bébé rit (the baby is laughing)’, and the construction ‘le bébé il rit (the baby he is laughing)’ are both correct. Constructions of the second type are called left-dislocated subjects. They belong to the larger category of dislocations, a wide class of syntactic constructions that has already been studied, especially in French, where this grammatical construction is frequently used. Left-dislocated subjects represent the most frequent type of dislocation in French (Blasco-Dulbecco 1999).

These constructions are often called double-markings of subject (i.e., the first marking is the lexical noun, the second marking is the subject pronoun). They are attested in children’s production as well as in adults’ production and can be used in affirmative utterances, for example, ‘*papa il aime pas trop* (dad he does not like [it] too much)’, as well as in interrogative utterances, for example, ‘*et ton papa il est pas là?* (and your dad he is not here?)’.

There is some controversy as to whether the subject pronoun is really optional. For this to be true, the two constructions, ‘NP pro:subj V’ and ‘NP V’ should have the same meaning and function. This possibility is clearly presented in Grevisse and Goose (1993): ‘the pronoun is redundant with the subject that is in its usual position – in oral language, the pronoun is considered as necessarily belonging to the verb ... no pause is made between nominal subject and

pronoun'. However, dislocations are useful as a way of introducing emphasis (to put extra-force on a element of the utterance in order to make it more important). In this case, there is usually a pause between the nominal subject and the pronoun, or at least a specific intonation contour (De Cat 2002). The use of dislocations also allows the NP to be moved around and material to be inserted between the NP and the verb. This also contributes to emphasis and focalisation effects (Wilmet 2003). Nonetheless, it remains true that many occurrences of dislocations are not emphatic. For Ashby (1977), less than half of all dislocations are clearly emphatic. Even more important is the fact that emphasis can be produced without the presence of a dislocation, so the existence of a personal pronoun between a lexical subject and a verb cannot be explained by emphasis only.

Blasco-Dulbecco (1999) contains a historical presentation of dislocations. It appears that the phenomenon already existed in Latin and in Old French. Dislocations were limited to subjects and adjuncts without prepositions in that period. Prescription against its use is attested in Chifflet (1659), which shows that it was quite common at that time, and already considered as a marker of orality as it is today. Thus, the use of dislocations is not an evolutionary feature of French. Neither is it a phenomenon linked to a grammaticalisation process, contrary to some authors' opinions, for example Ashby (1977), who suggested that the pronoun in front of a verb was becoming an obligatory construction in French. It remains true, however, that ancient and modern dislocations have different formats, and that double-marking did not then exist, so that a diachronic trend is not outright impossible. There is also no historical material available concerning oral language, so it is possible that constructions used in daily life were slightly different. What is known for certain is that it was often used in the 17<sup>th</sup> and 18<sup>th</sup> centuries with long subject noun phrases (Grevisse and Goose 1993). The explanations of the phenomenon

given then were also quite similar to those given today, especially that of dislocations being a means of introducing or emphasising language material.

Blasco-Dulbecco's (1999) original work is based on a large set of attested examples (1500 oral language examples and 500 written language examples). Her first result is that dislocations account for 10.19% of all subjects oral French, and 2.83% of dislocated subjects in written French. Other results are that 69.25% of dislocations in oral language are left-dislocations, and that a majority of these in oral French (76.56%) concern subjects. The large difference between the figures for oral and written language shows the importance of using oral language data to assess the importance of a linguistic construction.

Blasco-Dulbecco makes a thorough linguistic analysis of all dislocation types, using the 'pronominal approach' of Blanche-Benveniste et al. (1984). Dislocated elements may correspond to various syntactic components: subject, object, oblique. Blasco-Dulbecco starts her analysis with the oblique components. It appears that they are typical illustrations of dislocation usage, as they allow a lot of variation, some being used with either subject or object, some before the verb and some after the verb. When there is no preposition in the dislocated element, it is difficult to know exactly what is the relationship between the dislocated element and the verb, and the genuine cases of double-marking may be less frequent than they appear at first. This explains why the agreement in gender or number between the dislocated element and the verb (or the pronoun) is often missing. Her linguistic analyses led her to conclude that dislocations may cover a wide variety of situations. They may introduce new language material, but this is not always the case, especially when there is no exact reiteration of the dislocated element in the sentence. They are also a means of organising repetitions and creating enumeration effects. Finally, some elements are systematically used with a dislocated construction. There is a large difference in the nature of the function of dislocations before and after the verb. Before the verb, they have no



syntactic function, but create a frame that helps organise the utterance that follows. After the verb, they represent a means of extending the verbal group, but have no pragmatic function.

De Cat (2002) targets French child language as well as French adult language. Her work is based on the generativist framework, which makes it difficult to compare with Blasco-Dulbecco's work and with the present work. However her description of linguistic data is very informative and some of her observations are certainly very useful for the analysis of the present work. De Cat's work is based on a corpus of child-adult interactions, on a corpus of elicited judgements, and on data elicited for the study of intonation. The main argument of the author is that constructions such as 'NP pro:subj V' are different from constructions such as 'NP V' because the pronoun in 'NP pro:subj V' construction is not a simple inflectional morpheme but a syntactic argument that bears a theta-role (see also De Cat 2005). This means that the first construction is not a case of double-marking. The author demonstrates that it is possible to differentiate the prosody of a construction with dislocation and that of a construction without dislocation, and that 'NP pro:subj V' children's constructions have a true dislocation prosody more frequently than a non-dislocation prosody. She finds that in declarative sentences, 31% of root clauses contain a dislocated element. In wh-questions, this figure rises to 46%, and falls to 18% in yes/no questions. Although it is difficult to compare, the percentage of dislocations is certainly higher than in Blasco-Dulbecco's study. Another interesting result is that when there is a nominal expressing the subject, 46% of subjects appear after the verb with subject clitic on the verb, 39% appear before the verb with subject clitic on the verb, and 16% appear before the verb without subject clitic on the verb. De Cat gives these percentages for three dialects of French. There are some differences between the dialects, the main one being that in French French, there are more dislocated subjects after the verb and less true lexical subject (heavy subjects) than in Belgian or

Canadian French. Children use dislocations at a very early stage of language development, at the beginning of multi-word constructions, and demonstrate an early mastery of the pragmatic use of dislocations. Her main conclusion is that dislocations are discourse-motivated and constrained by the information structure; they are generated by adjunction, using minimal syntax, and they have no argumental status.

#### **4. Goal of the study**

The goal of the present paper is to find out whether the production of left-dislocated subjects are compatible with UBT. This implies obtaining some specific results.

First and most important, the developmental hypothesis has to be confirmed. There should be differences between adults' and children's grammars and a developmental process in children's grammar. These differences in grammars should manifest themselves in children's and adults' productions. Two consequent characteristics of children's and adults' productions will be checked below:

- 1) Whether a difference between adult and child production of left-dislocated subjects is confirmed or not;
- 2) Whether the children's use of them has a developmental trend or not.

Second, if the first test indicates that the existence of a usage-based account is possible, various hypotheses of UBT or related to UBT about the causes of a developmental trend can be tested. These hypotheses are:

- 1) Is there a lexical preference? Are there some specific verbs where dislocations are used by all children, whereas they are not for other verbs?

- 2) Could it be the result of item-based constructions? Are there some specific verbs where, on a child by child basis, dislocations are used, whereas they are not for other verbs? In this case, the lexical preference would change from one child to another.
- 3) Is there a relationship between the use of LS v. DS constructions and performance limitations? This could be checked by looking at the length of children's utterances. If one or the other type of construction is used more systematically in short or long utterances (taking into account the supplementary pronoun in DS constructions), this could mean that it is more complex than the other. Knowledge of this fact would help to understand the underlying characteristics of both types of constructions.
- 4) Is dislocation a consequence of the development in the use of personal pronouns? As language develops, production of obligatory subject pronoun becomes more and more systematic and easy for the child. If the development of dislocations is correlated with the development of subject pronouns, the production of dislocations could be a side effect of the development of pronouns before children learn not to use them when unnecessary.

As numerous corpus analyses are presented in this paper, it would be awkward to start by presenting all seven groups of results first and then a lengthy discussion about each group. So instead a short discussion will be presented with each series of results, which will also explain why the analyses were performed and organised as they were. After this, the totality of the results will be summarized, followed by a general discussion about what underlies the production of subject dislocation by children.

## 5. Methods

The following work is based on analyses performed on two corpora of spontaneous language production. These corpora are described below and will be referred to as the ‘Bath corpus’ and the ‘House corpus’. The two corpora correspond to different children and were collected in different settings. They were not collected for the purpose of this study, but to serve as reference for spontaneous language production (Chevrie-Muller et al. 1997; Le Normand 1986).

### 5.1. *Bath corpus*

The procedure used was the ‘*Bain des poupées*’ from a normalized language-testing tool (Chevrie-Muller et al. 1997). Children were recruited in kindergarten and recordings were made on site. All children were seen separately, in a quiet room within the school. During this procedure, a professional speech therapist plays with the child using a standardized set of toys in a standardized situation: giving a bath to two dolls. The speech therapist asks the children questions about the process involved (taking a bath) and about what the child does at home in a similar situation. The therapist interacts with the child until the question is answered and sufficient language material is produced. The interview goes on until the whole set of questions is answered.

The corpus contains 108 recordings of children of three different age groups: 3;0, 3;6 and 4;0 for a total of 21,667 utterances and 121,042 words. The recording characteristics for each age group are summed up in Table 1.

Table 1: Characteristics of the ‘Bath corpus’

|     |                  | Children         |          |      | Adults           |          |      |
|-----|------------------|------------------|----------|------|------------------|----------|------|
| Age | n.<br>recordings | n.<br>utterances | n. words | MLU  | n.<br>utterances | n. words | MLU  |
| 3;0 | 37               | 169              | 502      | 2.96 | 229              | 1345     | 5.87 |
| 3;6 | 29               | 127              | 439      | 3.38 | 166              | 905      | 5.51 |
| 4;0 | 42               | 131              | 477      | 3.63 | 177              | 950      | 5.37 |

Note: All figures, except the number of recordings itself, correspond to average figures per recording

## 5.2. *House corpus*

Children from this corpus were recruited in nurseries and kindergartens. Some recordings were made in a special room in the laboratory but most were made in a separate room in the child's institution. All children were seen separately with a caretaker or a person familiar with the child. In this procedure (Le Normand 1986), children are in a play situation. They are not interrupted when they start speaking and may go on for as long as they wish. The familiar adult observer (usually one of the parents) plays with the child. The role of the adult observer is to ask questions if the child does not produce much language on her own. The length of the recording is standardized to 20 minutes. This procedure does not elicit a lot of language material from the observer and is not used to assess adults' language.

The corpus contains 316 recordings going from age 2;0 to age 4;0 by increments of 3 months. The average number of recordings per age is 35, the total number of utterances is 32,643 and the total number of words is 108,675. The recording characteristics for each age group are summed up in Table 2.

Table 2: Characteristics of the ‘House corpus’

| Age | Nb.<br>recordings | Nb.<br>utterances | Nb.<br>words | MLU  |
|-----|-------------------|-------------------|--------------|------|
| 2;0 | 40                | 66                | 108          | 1.51 |
| 2;3 | 31                | 83                | 185          | 2.13 |
| 2;6 | 36                | 89                | 225          | 2.45 |
| 2;9 | 36                | 110               | 364          | 3.22 |
| 3;0 | 40                | 113               | 397          | 3.41 |
| 3;3 | 34                | 130               | 457          | 3.51 |
| 3;6 | 34                | 118               | 464          | 3.70 |
| 3;9 | 33                | 117               | 475          | 3.92 |
| 4;0 | 32                | 109               | 454          | 4.02 |

Note: Nb. utterances and nb. words correspond to average number of utterances and words per recording

## 6. Corpora analyses

### *6.1. Are children’s dislocations mere reproduction of adult language?*

Is it possible that the use of dislocations by children simply reproduces adults’ use? This would mean that dislocations are not a very interesting feature of children’s language to study, as they would then provide no particular insight into the development of children’s grammars. The Bath corpus can be used to test this hypothesis, as it contains both children and adults spontaneous productions.

Left-dislocated subjects (DS) have the following structure: *lexical noun group* (noun group where the head is a lexical noun) followed by *subject personal pronoun* followed by *verb* (any tense or type). In all these constructions, the subject personal pronoun is in the third person. Variations concerned gender (feminine or masculine) and number (singular or plural). Standard lexical subjects (LS) are constructions which are not subject dislocations, and consist of *lexical noun group* followed by *verb*. LS constructions are similar to DS constructions minus the personal subject pronoun. Table 3 gives the average number of such constructions found for the children and the adults in each age group in the Bath corpus. None of these two constructions was frequent, and some recordings had none at all. This confirmed previous observations made in English, German, and Russian that showed that lexical subjects are not frequent in oral language (Miller and Weinert 1998; Biber 2003).

Table 3: Total number of lexical subjects and subject dislocations in the Bath corpus and ratio between DS and LS structures

|     | Children |    |       | Adults |    |       |
|-----|----------|----|-------|--------|----|-------|
| Age | LS       | DS | DS/LS | LS     | DS | DS/LS |
| 3;0 | 13       | 17 | 1.31  | 23     | 16 | 0.69  |
| 3;6 | 7        | 12 | 1.71  | 9      | 5  | 0.55  |
| 4;0 | 1        | 9  | 10.50 | 7      | 12 | 1.71  |

Children's LS constructions were less frequent than DS constructions for all age groups (see Table 3). For adults, this was not true, except for the 4;0 age group. The number of DS constructions produced by children did not appear to be much different from the number of DS

constructions produced by adults  $t(107) = .51$ ,  $p = .66$ , whereas the number of LS constructions produced by children was significantly lower than the number of LS constructions produced by adults  $t(107) = 2.04$ ,  $p = .021$ .

For each recording, the difference between the number of DS and LS constructions was computed. A statistically significant difference was found between children and adults,  $t(107) = 1.94$ ,  $p = .027$ . The difference remained significant,  $t(17) = 2.44$ ,  $p = .013$  when the recordings with no DS nor LS construction (none for the child and none for the adult) were omitted from statistical computation. This suggests that the low frequency of DS and LS constructions did not influence the significance of the statistical results.

It thus appears that children do not produce more dislocations than adults. The figures are roughly similar. Taking into account the fact that adults produced a slightly larger number of occurrences and of words, the children's percentage of use of DS constructions was slightly higher than the adults'. However, it was the difference between the number of LS constructions and DS constructions that truly separated children from adults (see Table 3). In similar situations, when it was necessary to provide the lexical subject in an utterance, children tended to use a dislocated construction where adults often used an '*a priori* simpler' lexical subject construction with no subject pronoun.

## *6.2. Does the use of dislocations follow a developmental path?*

The use of dislocations appears to be a feature of children's language, but this does not characterize or explain the use of dislocations by children. One important question that needs to be answered is whether the use of dislocations follows a developmental trend, or whether it is the same for all children and will change only through maturation at a later age. This can be measured with the House corpus, which follows children from age 2;0 to age 4;0. Results in



Table 4 present the number of DS and LS constructions occurring in the Bath corpus. Contrary to what was observed in the Bath corpus, the number of DS constructions tended to increase with age to a much greater extent than the number of LS constructions. This is shown in the increased ratio of the number of DS constructions to the number of LS constructions (see Table 4, column 4).

Table 4: Total number of lexical subjects and subject dislocations in the House corpus and ratio of DS to LS structures

| Age | LS | DS  | Ratio DS/LS |
|-----|----|-----|-------------|
| 2;0 | 21 | 1   | 0.06        |
| 2;3 | 22 | 10  | 0.45        |
| 2;6 | 33 | 23  | 0.70        |
| 2;9 | 38 | 54  | 1.42        |
| 3;0 | 76 | 75  | 0.99        |
| 3;3 | 26 | 101 | 3.91        |
| 3;6 | 36 | 80  | 2.22        |
| 3;9 | 45 | 102 | 2.27        |
| 4;0 | 60 | 103 | 1.71        |

The results obtained above could be confirmed using the Léveillé corpus (Suppes, Smith, and Léveillé 1973) from the CHILDES database. This corpus follows closely, two or three times a month, the development of language interaction between a child and his mother from age 2;2 to 3;3. The recordings were made at home, which contrasts with the corpus presented above. The

same computations as above were done on the L  veill   corpus. The development in the production of LS and DS constructions by the child and his mother are presented in Figure 1. At first, the child produced many more LS than DS constructions. The number of LS constructions then tended to diminish whereas the number of DS constructions remained the same. Finally, the number of DS constructions increased similarly to the number of LS constructions. The most interesting result was that the development of the number of LS and DS constructions produced by the child did not in any way seem related to the use of these constructions by his mother. The mother used a much larger number of LS constructions than DS constructions for a long period (from age 2;2 to age 2;11) and, though she tended to use a reduced number of both constructions in the last recordings (from age 3;0 to age 3;3), this did not prevent a sudden growth of the use of such constructions by the child.

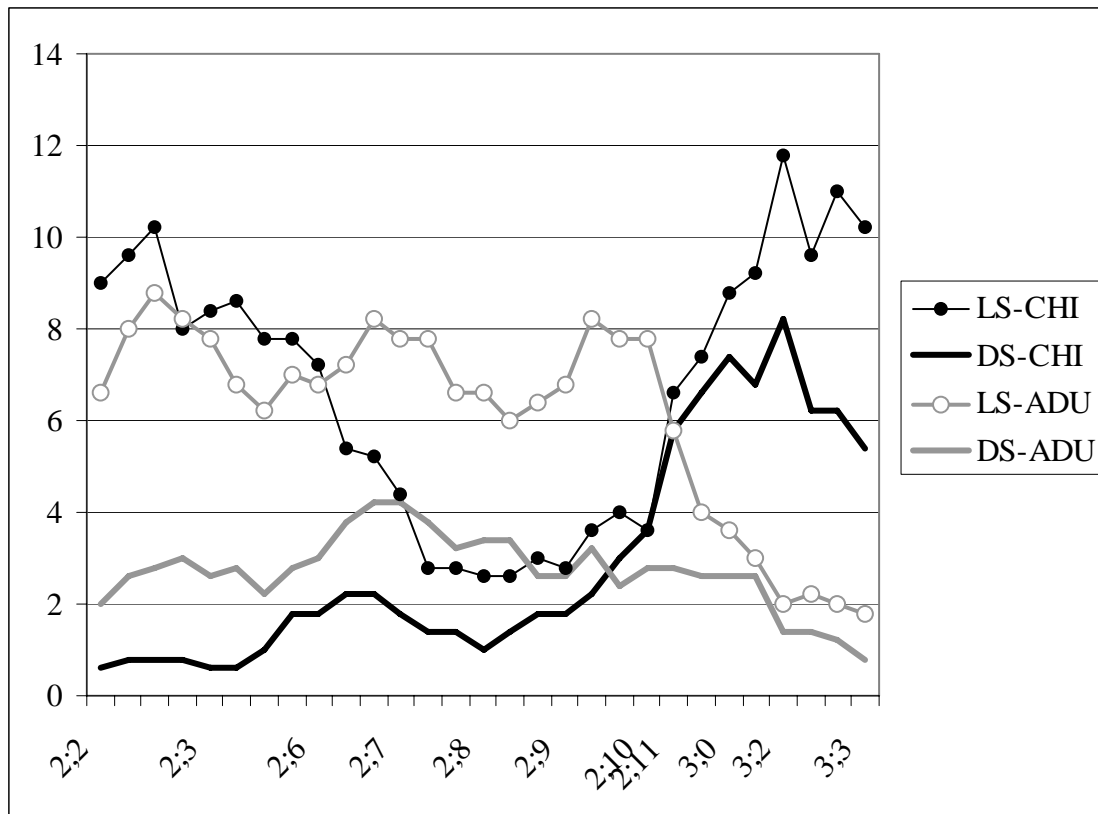


Figure 1: Developmental trajectory for a child from age 2;2 to 3;3 (Leveille corpus)

It appeared that children frequently produce dislocated constructions. There was no obvious explanation as to why this construction would be chosen instead of the simple lexical subject construction, which children also used. The fact that they used both constructions prevents any explanation in terms of missing syntactic knowledge. The fact that they did not simply copy the adults' production suggests that regulating mechanisms specific to the children were at work.

There are more than one possible explanation to the results presented above. Some of the possible explanations can be tested using the material presented above (Bath and House corpora). The first explanation is that the choice of a given construction could be related to the verb used in the construction. If some constructions are verb specific for all children, then this would be a lexical choice. A second possibility is that some constructions are verb specific for specific children, which would mean they are usage-based constructions (Tomasello 2003). A third hypothesis is that the use of a pronoun allows for creating longer utterances. A fourth and final hypothesis is that the production of DS constructions could be a logical consequence in the development of the use of subject pronouns. All these four hypotheses are tested below.

### *6.3. Lexical choice*

The use of dislocations by children could be explained by many linguistic or psycholinguistic features. One of these is that this use is governed by a lexical choice. This means that some verbs will obligatorily use dislocations structure and some verbs will not, but that this is not a syntactic choice.

To test the existence of a lexical obligatory choice, it is possible to use all the children's corpora (Bath and House corpora). If there is a lexical tendency to use some types of verb with a

type of construction, then there should be a partition between verbs that govern an LS construction and verbs that govern a DS construction. This did not seem to be the case. For all the children from the House and the Bath corpora, there were 64 verbs used in one or other construction. Twenty-eight of these verbs were used only once. Six verbs were produced more than once but used only in one context, either LS or DS. These six verbs represented only 18 of the total of 677 occurrences of verbs that occurred more than once in an LS or a DS construction. The 10 most frequently used verbs in such constructions are presented in Table 5 and no verb appears significantly more often in one type of construction than another.

Table 5: Most frequent verbs used in LS or DS constructions (House and Bath corpora)

| Verb                  | N of recordings | Lexical subject | Subject dislocation |
|-----------------------|-----------------|-----------------|---------------------|
| aller (to go)         | 184             | 117             | 200                 |
| être (to be)          | 163             | 109             | 116                 |
| avoir (to have)       | 92              | 57              | 55                  |
| dormir (to sleep)     | 35              | 9               | 45                  |
| manger (to eat)       | 21              | 4               | 17                  |
| vouloir (to want)     | 16              | 5               | 13                  |
| venir (to go)         | 14              | 9               | 11                  |
| faire (to do)         | 14              | 5               | 12                  |
| regarder (to look at) | 11              | 7               | 5                   |
| rentrer (to go into)  | 11              | 3               | 8                   |

Note: the first column gives the number of recordings containing at least one occurrence of the verb. More than one occurrence may occur per recording.

#### *6.4. Item-based constructions*

There was no lexical tendency to use a verb in a specific construction when one considers the corpus of the children's productions as a whole. However, it is possible that specific lexical constructions exist, but for specific children. Some children would use some verbs only in a certain type of construction, but other children would use the same verbs in different constructions, depending on how the verbs were modelled in their input. This corresponds to a usage-based hypothesis (Tomasello 2003).

The existence of usage-based constructions should be tested. This will be done using the House corpus. The Bath corpus cannot be used here because it is too small to provide enough contexts to test this hypothesis. To test this, it is necessary to compute the individual use of LS and DS constructions for a given verb in the same recording session. This is presented in Table 6 for the House corpus. Once again, it appeared that LS constructions disappeared as children became older (verbs that were used only with LS constructions disappeared at age 4) and DS constructions became more and more frequent. It also appeared that, although some verbs were used in only one construction type and some children used only one construction type, most children used both types of construction and often the same verbs for both types (10 children out of 25 at age 4). Considering the relatively short length of the recordings, which did not provide the children with many opportunities to produce that many different constructions, it appeared unlikely that the choice between LS and DS constructions was usage-based, or that some children used the LS constructions more easily where other children seemed to prefer the DS construction, although this could still be the case for a few children.

Table 6: Use of specific constructions for the same child (House corpus)

| Age | All LS | All DS | Both LS & DS,<br>different verbs | Both LS & DS,<br>same verbs |
|-----|--------|--------|----------------------------------|-----------------------------|
| 2;0 | 12     | 1      | 0                                | 0                           |
| 2;3 | 8      | 2      | 2                                | 3                           |
| 2;6 | 9      | 4      | 6                                | 3                           |
| 2;9 | 8      | 4      | 4                                | 8                           |
| 3;0 | 7      | 8      | 5                                | 12                          |
| 3;3 | 4      | 14     | 3                                | 6                           |
| 3;6 | 4      | 10     | 7                                | 4                           |
| 3;9 | 2      | 6      | 7                                | 14                          |
| 4;0 | 0      | 6      | 9                                | 10                          |

### 6.5. Length effect

It is well known that young children are sensitive to processing limitations (Valian, Hoeffner, and Aubry 1996). It is also possible that, during the process of language development, they tend to use simpler constructions before learning to use complex constructions (predicted by UBT). Both these reasons emphasize the fact that children will more readily use a less costly and simpler construction than a construction which is more costly and more complex. If processing limitations have an effect on children's utterances, this would mean that the simplest structures

will be produced more easily and could be used as a way to circumvent these limitations when producing long utterances.

The measurement of utterance length could provide an explanation to children's use of either LS constructions or DS constructions. As LS constructions are shorter than DS constructions, they should be easier to use and produce and their syntactic structure should also be simpler. This would make them useful to compensate for performance limitations. Concurrently, when children want to produce long utterances, they should automatically use the simpler construction to compensate for increased processing demands. This means that, in longer utterances, LS structures should be used more often than DS structures, even taking the supplementary element (the personal pronoun) present in DS structures into account.

For all children of the House corpus, the average length of the DS utterances was 7.73 (6.73 minus the personal pronoun) whereas it was only 5.07 for LS utterances. This was also true when considering only the older children who often produced both types of constructions (DS: 8.10 – 7.10 minus the pronoun; LS: 5.84). This is in clear contradiction with the idea that the shorter structure should help children to produce longer utterances, even taking the supplementary pronoun of DS structures into account.

This would mean that either processing limitations bears no weight on the choice of one structure or another, or that LS structures, although they are shorter or correspond to a simpler tree structure, are not easier to process than DS structures. Further test should be done to choose between the two hypotheses. However, the difference between the length of the utterances that include LS or DS constructions needs to be explained.

Whatever the results of further tests, there is indeed a possibility that a shorter structure (LS) is more complex than a longer one (DS). Two reasons may explain this:

First, from a phonetic point of view, when a DS construction contains a pronunciation pause, it allows the child to produce a global utterance in two parts, each simpler than the counterpart LS construction. This pause can be considered as arising from limitations in performance. If the DS does not contain a pause but a rising pitch, then the phonological argument does not apply, but the second argument below does.

Second, from a grammatical point of view, the use of two loosely related constructions (DS) –a lexical dislocated subject and a verb phrase that includes a pronoun– could be simpler for a child to compute than a unique construction (LS) that includes a lexical subject and a verb. This is because both grammatical structures in DS are simpler than the unique LS construction and because the operation of creating a DS –appending two syntactically independent constructs– is a very simple and very frequent operation in child language, and thus is very easy to produce. It would then be quite similar to the production of two successive but independent clauses.

If this explanation is verified, then it could be that performance limitations and construction simplicity are behind the use of DS constructions.

#### *6.6. Consequence of the development of pronoun marking on verbs*

A last hypothesis is that DS constructions are produced because they contain verbs in their most usual form, with a personal subject pronoun before them. This would mean that personal pronouns are produced somewhat automatically, as if they were agreement markers, contrary to De Cat's hypothesis (2002). The reason behind the increase in production of DS constructions would be that the use of personal pronouns becomes more and more automatic as children grow older. If pronouns are produced automatically before a verb in most contexts, the DS constructions would merely be LS constructions where the production of an optional pronoun is not blocked by the child. If this is true, there should exist a correlation between the development



of personal pronoun marking and the development of DS constructions. This condition is not sufficient in itself to prove the hypothesis, but if not verified, would prove it wrong. To test this, the ratio of the number of subject pronouns produced before a verb by the number of verbs was computed for each recording. The correlation between this ratio and the number of DS constructions was then computed. The results showed a significant correlation value,  $df = 316$ ,  $r = .428$ ,  $z = 8.09$ ,  $p < .0001$ , for the House corpus. But no significant correlation was found in the Bath corpus, either for the children,  $df = 108$ ,  $r = .113$ ,  $z = 1.159$ ,  $p = .24$ , or for the adults,  $df = 108$ ,  $r = .039$ ,  $z = .396$ ,  $p = .69$ . The difference in the House corpus and Bath corpus results may be the consequence of the low number of DS produced in the latter corpus or from the difference in the two situations. This discrepancy suggests that more data would be necessary to confirm the hypothesis of an automatic production of the subject pronoun.

## **7. Conclusion**

The results above should be handled with care because of the very small number of DS and LS constructions effectively produced by children and adults. It is quite possible that with larger numbers, the results would change. For instance, the differences in use of LS and DS structures for the adults showed that there were unknown factors involved in the production of these structures. It should be clear that the results need to be confirmed using more reliable methods or situations. For example, as DS and LS are indeed unusual constructions in oral language, the results obtained would be easier to confirm using induced language instead of spontaneous language. Further analysis of the contexts in which LS and DS structures are used is also necessary. It would be especially interesting to know whether children's productions are similar analysis to adults' productions in their context.

Within the limits of the present results, the comparison between the number of subject dislocations produced by children and by adults showed that young children (especially children aged three to four) produced more dislocations than adults. There was a developmental trend in this use, as children aged two did not produce many dislocations and none were produced in large quantities before age three. In fact the –surprising– pattern that seemed to emerge was that children began by using more LS than DS constructions, then reversed the pattern when they were three to four years old, and would probably later go back to the first situation, which appears to be the adult norm (this last result was absent from the data as the children were not followed after age four, but they should develop towards the adult norm later).

A series of statistical measures showed that it was unlikely that the reason behind children's use of dislocations at age three and four was lexical or item-based in its nature, or the result of a specific developmental path for some children. A possible explanation for the use of dislocations was that they are a way of producing longer utterances by avoiding performance limitations or using simpler syntax (see part 6.5). This would mean that dislocations are constructed using mental operations such as concatenation which are less complex than operations involved in generating non-dislocated constructions. When children will become older, processing limitations will diminish and the number of DS would also go down (as is attested in adults' productions). Another explanation was that the production of subject pronouns before a verb had become sufficiently automatic that children could not stop their production, even when not in an obligatory context. This automatic behaviour will also change at a later age when the use of LS constructions will increase again.

Even if the current results are tentative, they are useful in allowing hypotheses to be drawn for future work and in pointing out interesting phenomena in child language development. A

proposal is that there are three different periods to be considered regarding the nature of LS and DS. At first, children start to produce what appears to be LS but is in fact the concatenation of a noun and a verb produced without morphosyntactic markers such as determiners or clitics. At this stage of development there is no difference between LS and DS from the children's grammar point of view (see footnote 1). Then, children start producing subject pronouns before all verbs, whether there is a lexical subject or not. In this second period, LS would mostly disappear and DS would become the norm. The process of concatenation that leads to DS is the similar with LS in the first period, but with elements that include a determiner or a pronoun. Finally, children learn to produce real LS constructions where lexical subject and verb are linked in a complex abstract structure and go on towards adult mastery. The previous production of DS structures would help them to develop more complex abstract categories, as hypothesized in UBT, and lead to the more complex LS construction. This development is helped by the development of written language. LS constructions are more frequent in written language than oral language (see Miller and Weinert 1998), to the point where the use of this construction is strongly advocated in the written language norm. This would lead children to integrate the LS construction in their oral language only after having started to learn the written language. A major consequence of this explanation is that LS structures would not be the prototype construction around which other constructions, such as DS constructions, are built, but only a type of construction among others, which are often easier to use, especially in oral language. The existence of a link between written language acquisition and the development of LS structures is made more plausible by the possibility of performance limitations in the production of LS structures. Written language is less prone than oral language to performance limitations, as limits on working memory are removed by the presence of an external support that allows language users to analyse and reanalyse the language data at will.

## 8. Acknowledgements

Grateful thanks to Sean Devitt for his help with the English version of this paper.

## 9. Footnotes

1.

This interpretation is much stronger than what De Cat (2002) proposes but it agrees with her basic interpretation that it is difficult to know what is a LS and a DS in young children: “Like adults, they [French children] frequently produce subject dislocations — and these are attested from the onset of word combination []. The difference between child and adult language in this context is restricted to the well-known but not fully understood null-subject phenomenon in language acquisition.” (De Cat 2002, 246).

## 10. References

- Ashby, William J. 1977. *Clitic inflection in French*. Amsterdam: Rodopi.
- Biber, Douglas. 2003. Variation among university spoken and written registers: a new multi-dimensional analysis. In *Corpus analysis - Language structure and language use*, edited by P. Leistyna and C. F. Meyer. Amsterdam: Rodopi.
- Blanche-Benveniste, Claire., J. Deulofeu, J. Stefanini, and K. van den Eynde. 1984. *Pronom et Syntaxe. L'approche pronominale et son application au français*. Edited by A. Selaf, CNRS. Paris.
- Blasco-Dulbecco, Mylène. 1999. *Les dislocations en français contemporain*. Paris: Honoré Champion.

- Chevrie-Muller, Claude., Anne-Marie Simon, Marie-Thérèse Le Normand, and Sophie Fournier. 1997. *Batterie d'évaluation psycholinguistique (BEPL)*. Paris: ECPA.
- Chifflet, L. 1659. *Essai d'une parfaite grammaire de la langue françoise*. Genève: Slatkine reprints, 1973 ed. Anvers: Jacques Van Meurs.
- Croft, William. 2001. *Radical construction grammar*. Oxford: Oxford University Press.
- De Cat, Cécile. 2002. French Dislocations. PhD, Language and Linguistic Science, University of York, York.
- De Cat, Cécile. 2005. French subject clitics are not agreement markers. *Lingua* 115 (9):1195-1219.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Grevisse, Maurice, and André Goose. 1993. *Le Bon Usage -- Grammaire française*. édition refondue par André Goosse ed. Paris - Louvain-la-Neuve: DeBoeck-Duculot.
- Le Normand, Marie-Thérèse. 1986. A developmental exploration of language used to accompany symbolic play in young, normal children (2 - 4 years old). *Child: Care, Health and Development* 12:121-134.
- Miller, Jim, and Regina Weinert. 1998. *Spontaneous Spoken Language*. Oxford: Clarendon Press.
- Parisse, Christophe. 2005. New perspectives on language development and innateness of grammatical knowledge. *Language Sciences* 27 (4):383-401.
- Parisse, Christophe, and Marie-Thérèse Le Normand. 2000. How children build their morphosyntax: the case of French. *Journal Of Child Language* 27 (2):267-292.
- Parisse, Christophe, and Marie-Thérèse Le Normand. 2001. Local and global characteristics in the development of morphosyntax by French children. *First Language* 21:187-203.

- Suppes, Patrick, R. Smith, and Madeleine Léveillé. 1973. The French syntax of a child's noun phrases. *Archives de Psychologie* 12, No.166:207-269.
- Tomasello, Michael. 2000. Do young children have adult syntactic competence? *Cognition* 74:209-253.
- Tomasello, Michael. 2002. The emergence of grammar in early child language. In *The Evolution of Language Out of Pre-language*, edited by T. Givon and B. F. Malle. London: John Benjamins.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge: MA: Harvard University Press.
- Tomasello, Michael. 2006. Acquiring linguistic constructions. In *Handbook of Child Psychology*, edited by D. Kuhn and R. Siegler New York: Wiley.
- Valian, Virginia, James Hoeffner, and Stephanie Aubry. 1996. Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology* 32 (1):153-164.
- Wilmet, Marc. 2003. *Grammaire critique du français*. Bruxelles: Duculot.