



## Chapter 3. Young children's uses of referring expressions

Christine da Silva-Genest, Anne Salazar Orvig, Haydée Marcos, Stéphanie Caët, Julien Heurdier

### ► To cite this version:

Christine da Silva-Genest, Anne Salazar Orvig, Haydée Marcos, Stéphanie Caët, Julien Heurdier. Chapter 3. Young children's uses of referring expressions. Anne Salazar Orvig; Geneviève de Weck; Rouba Hassan; Annie Rialland. The Acquisition of Referring Expressions, 28, John Benjamins Publishing Company, pp.82-113, 2021, Trends in Language Acquisition Research, 9789027208354. 10.1075/tilar.28.03das . hal-03374815

**HAL Id: hal-03374815**

**<https://hal.univ-lorraine.fr/hal-03374815>**

Submitted on 17 Jan 2023

**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.

da Silva-Genest, C., Salazar Orvig, A., Marcos, H., Caët, S., & Heurdier, J. (2021). Young children's uses of referring expressions. In A. Salazar Orvig, G. de Weck, R. Hassan, & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogical approach* (pp. 82 - 113). John Benjamins Publishing Company.

Doi: 10.1075/tilar.28.03das

© 2021 John Benjamins Publishing Company

## CHAPTER 3

### Young children's uses of referring expressions

Christine da Silva-Genest<sup>1</sup>, Haydée Marcos<sup>2</sup>, Anne Salazar

Orvig<sup>2</sup>, Stéphanie Caët<sup>3</sup> and Julien Heurdier<sup>2</sup>

<sup>1</sup> DevAH & ATILF, Université de Lorraine

<sup>2</sup> Laboratoire CLESTHIA, Université Sorbonne Nouvelle - Paris 3

<sup>3</sup> Unité Mixte de Recherche 8163 Savoirs Textes Langage, Université de Lille

#### Abstract

This chapter investigates the repertoire and the uses of referring expressions in natural dialogues of 28 French-speaking children, aged between 1;7 and 2;6 years old. We focus on three strong forms (nouns, strong demonstrative pronouns and strong personal pronouns) and three weak forms (clitic personal pronouns, null forms and fillers). Their uses are analyzed for the following factors: the linguistic development, the type of referent (participants vs. entities), the syntactic function, and the attentional and discursive status of the referent. The results reveal contrasted uses of referring expressions. Nouns are mainly used for new or reintroduced referents, generally in syntactic functions other than subject, whereas clitic personal pronouns (as well as the other weak forms) are mainly used for given referents and always as subjects. The type of referent (entities vs. participants) also influences the use of referring expressions. The discussion addresses the issue of the intertwining of morphological, syntactic and pragmatic development.

**Keywords:** referential choice, accessibility, strong forms, weak forms, syntactic function, attentional status, discursive status, entities, self, addressee

## **1. Introduction**

At about the age of two, and jointly with grammatical development, young children acquire the different uses of language. Among other things, their ability to communicate about world objects, actions, and events develops. Their means of expressing reference multiply and diversify, reflecting their sensitivity to discourse-pragmatic factors such as referent accessibility for the addressee (for several reviews, see Serratrice & Allen, 2015). Although some studies have focused on the various ways in which children refer to themselves (Budwig, 1995; Caët, 2013; Morgenstern, 2006), only a few have looked simultaneously at the development of how children refer to world entities and to discourse participants. Do the same factors affect the acquisition of these two referential paradigms and do they do so in the same way? The aim of this chapter is to study young French-speaking children's repertoire of referring expressions and their uses, in particular the choice between strong forms and weak forms. We consider the impact and interactions of several factors, including linguistic development, the type of referent (i.e., entities and discourse participants), the syntactic function of the

referring expressions used, and the attentional and discursive status of the referent.

### *1.1 Children's first referring expressions*

The repertoire of referring expressions emerges gradually. In French, as in many other languages, the first words produced by children are usually "noun candidates" employed to refer to an entity. Also appearing early are demonstrative pronouns (*this* and *that* in English: Diessel, 2006; Rozendaal & Baker, 2010; *ça* in French: Hamann, Rizzi & Frauenfelder, 1996). As early as the one-word utterance stage, when children produce predications about a referent, we find null forms and fillers (Veneziano & Clark, 2016).

Most studies on French (Girouard, Ricard & Gouin Decarie, 1997; Hamann et al., 1996; Jakubowicz & Rigaut, 1997) date the emergence of personal pronouns at about the age of two. Children produce most of the forms by around 30 months. The pace of acquisition varies considerably, probably because of the morphological diversity of French. During a rather extended period, adult-like forms alternate with null forms and fillers (see: Yamaguchi, Salazar Orvig, Le Mené, Caët & Rialland, Chapter 2, this volume).

The status of null forms and fillers are a matter of theoretical debate. In non-null subject languages the former have been seen either as a phenomenon

specific to child grammar (Hyams, 1986; Rizzi, 1993/1994), as an indication of children's performance limitations (Bloom, 1990; for French, Jakubowicz & Rigaut, 1997) or as the result of the distributional features of the language heard by the child (Theakston, Lieven, Pine & Rowland, 2001). Fillers are accounted for in phono-prosodic terms (e.g., Demuth, 2001) and are seen as precursors of grammatical morphemes (Dressler & Kilani-Schoch, 2001; Veneziano & Sinclair, 2000), especially pronouns in preverbal position (see Yamaguchi et al., Chapter 2, this volume). The status of clitic personal pronouns is also under debate. In French, a clitic personal pronoun acting as the grammatical subject cannot be separated from the verb by expressions other than negation (*ne*) or other clitic pronouns (e.g. *le*, *en*) as in *il mange* ('he eats'), *il le mange* ('he eats it'), *\*il aussi mange* ('\*he also eats'). Some authors see clitic pronouns as inflectional morphemes (Culbertson & Legendre, 2008), whereas for others they are arguments (De Cat, 2005). Whatever the perspective, clitic personal pronouns can be likened to other weak forms that tend to occur in certain discourse-pragmatic contexts. Since we focus in this research on the pragmatic dimension, the question of whether clitic pronouns are free or inflectional morphemes is less relevant.

In the process of acquiring the pronoun paradigm, children master the different forms and their variations in usage which are dependent on both the type of referent (entities, the speaker, the addressee) and the cognitive status of that referent.

### *1.1.1 Factors affecting the use of expressions referring to entities*

The choice of a type of linguistic expression partly depends on discursive and cognitive factors. According to research in this field, regularities in the use of strong versus weak forms rest on a predominant factor, namely, the speaker's representation of the cognitive accessibility of the referent to the addressee (Ariel, 1990; Gundel, Hedberg & Zacharski, 1993): weak forms (such as null forms or clitic personal pronouns for French; Cornish, 1999) encode referents that are in focus whereas the use of strong forms (lexical forms, emphatic and strong pronouns) generally orients the listener towards referents with other cognitive statuses, whether activated, familiar, or only type identifiable (Gundel et al., 1993), or less accessible (Ariel, 1990). In other frameworks (Chafe, 1976; Givón 1983), these strong forms also mark discursive discontinuity, or contrast between entities.

Many studies on children's early productions have taken this approach. In a study with English-speaking children based on Gundel et al.'s hierarchical scale of givenness (1993) and the implications it entails, Gundel and Johnson (2013) found evidence of contrasted uses of personal and demonstrative pronouns: the former were used for already-mentioned referents, the latter for activated but not yet mentioned referents.

Moreover, in adult discourse, these uses appear to be tightly linked to syntactic choices, since syntactic structure is not independent of information management, as suggested by Du Bois (2003). This author showed that languages tend to avoid having several new (lexical) elements in the same utterance. Also, subjects of transitive verbs — which generally correspond to given referents — tend to be realized by null or weak forms, whereas new referents are encoded in the object function using lexical forms or strong morphemes. Other studies (Chafe, 1976; Givón, 1983) have shown that the subject function is often the preferred one for the topic (and thus, for given referents).

Based on Du Bois's (2003) work, Clancy (1993) and Allen (2000) proposed an analysis situated at the interface between syntax and pragmatics (see also Allen, 2006). In these studies, informativeness (a notion later replaced by "accessibility") is seen as the product of a complex set of features such as the referent's newness vs. givenness, the presence of competing referents, animacy, and references to oneself or the addressee, or reference to entities. Studies on various languages have brought out the main characteristics of referents with little or no accessibility, which include newness, contrast, and absence from the immediate context (for Korean: Clancy, 1993; for Inuktitut: Allen, 2000; Skarabela, Allen & Scott-Phillips, 2013; for Japanese: Guerriero, Oshima-Takane & Kuriyama, 2006; for Hindi: Narasimhan, Budwig & Murty, 2005; for English: Guerriero et al., 2006; Hughes & Allen,

2006; for Mandarin: Huang, 2011). Each of these factors affects the choice of referring expressions, and their effects are cumulative. Accordingly, strong forms have a higher probability of occurring in contexts that have two or three of these characteristics and weak forms are more probable otherwise (Allen, 2008; Hughes & Allen, 2015).

Similar results have been obtained in research on young French-speaking children (Salazar Orvig et al., 2004; Salazar Orvig et al., 2010a). Here also, we find a tendency to choose weak forms (clitic personal pronouns, fillers, and null forms) for given referents and strong forms (nouns, demonstrative pronouns) for first mentions. Nouns acting as grammatical subjects that refer to given referents are frequently associated with contrast or opposition, whereas clitic personal pronouns more often occur in a context of adding predications for an already established topic (Salazar Orvig et al., 2010b; Salazar Orvig & Morgenstern, 2015). Demonstrative pronouns (Salazar Orvig & da Silva, 2014) are preferred for activated referents – not mentioned in the discourse but nevertheless included in the activity – and also, like nouns, in contexts of contrast. However, the question of how these factors interact has not been addressed directly for French, which is what we propose here.



### *1.1.2 Factors affecting the uses of expressions referring to the discourse participants*

To talk about the participants of the interaction taking place (themselves or/and the addressee), children may alternatively employ different types of referring expressions. In most studies on self-reference, the use of different self-referring expressions has generally been explained in terms of pragmatic factors (Budwig, 1995; Budwig & Wiley, 1995; Caët, 2013; Morgenstern, 2006; Nelson, 1989). Null forms, fillers, and *I* are often found when children refer to their internal states or desires. Children may also use different first-person pronouns for the subject function (*I*, *me*, or *my*), in accordance with the type of agency they are expressing (depending, controlling, or contrastive), and they may use their name when expressing opposition with other referents or when telling stories. Studies on references to the addressee in English are scarce, probably because the morphology of second-person pronouns varies less than first-person pronouns, null forms in assertions can be difficult to distinguish from imperatives, and fillers are rare (see Caët, 2013; for French, Yamaguchi et al., Chapter 2, this volume). Budwig and Wiley (1995) and Imbens-Bailey and Pan (1998) found that children use null forms and *you* to make requests or proposals. Caët (2013) observed that *Mommy/Daddy* could be employed by children to express opposition with other referents, as they do when referring to themselves but could also be produced in contexts where the interlocutor's parental status was at play. Not

all studies have found the same pragmatic functions associated to the same forms, perhaps due to differences in the method used to collect data, or to interindividual variability.

## 1.2 *The aim of this study*

In order to gain further insight into the factors that affect children's use of referring expressions in French, we took a broad range of factors into account and analyzed the way in which they interact. We contrasted three weak forms<sup>11</sup> (clitic personal pronouns, which are in the process of being acquired, null forms, and fillers) and three strong forms (nouns, strong demonstrative pronouns, and strong personal pronouns). Naturally, an important factor is the existence of the various lexical and grammatical forms in the child's repertoire, which depends on his/her developmental level. In addition to this aspect, the following factors enter into play in the here-and-now of the dialogue: the type of referent (entities or participants), the syntactic function, and the referent's attentional and discursive status (givenness) in the dialogue.

---

<sup>1</sup> Other forms, such as dislocations and the clitic demonstrative pronouns in the construction *c'est+X* ('it's+X'), are treated respectively by Klein, Jullien & Fox, Chapter 6, this volume), de Weck, Hassan, Heurdier, Klein & Salagnac (Chapter 9, this volume) and Vinel, Salazar Orvig, de Weck, Nashawati & Rahmati (Chapter 10, this volume).

In considering these different factors, we focused on three main aspects: the types of referring expressions produced and their referential and discursive context of production, the way in which the factors described above interact in promoting the use of one of the possible forms, and the links between the grammatical and pragmatic levels in the development of referring expressions in young children.

## **2. Method**

After briefly presenting the study participants and the data collected (see 2.1), we describe our method of analysis (see 2.2).

### *2.1 Participants and data collection*

The participants were 28 children between the ages of 1;7 and 2;6.<sup>2</sup> Some children were videotaped twice at the same age (2;3); five children were

---

<sup>2</sup> The data were drawn from several corpora compiled in previous studies: Anaé and Madeleine, from the Paris Corpus; the Childcare Corpus; the Standardized play corpus; the Nashawati corpus; the Salazar Orvig corpus; the Yamaguchi corpus. See Appendix I for further details.

videotaped several times at different ages. This made for a total of 60 filmed sessions. The children were observed during family interactions. The utterances of all participants were transcribed. The contextual information needed to interpret the verbal productions was also noted.

Given that the age period (around the age of 2 years) is strongly marked by interindividual variability, we used mean length of utterance (MLU) rather than age as a measure of development. Three MLU groups were set up. The data for each group are presented in Table 1.

MLU group	MLU range	No. of children	No. of sessions	Age range
MLU GR.1	1.3 – 1.92	12	20	1;7 - 2;3
MLU GR. 2	2.04 – 2.5	9	18	1;10 - 2;4
MLU GR. 3	2.52 – 2.96	12	22	1;10 - 2;6

Table 1. Participants, by MLU group, age and number of sessions

## 2.2 *Coding system*

We sorted the referring expressions by grammatical category and then analyzed the occurrences in terms of the following factors: MLU, syntactic function, type of referent, and attentional and discursive status. Due to lack of space, this chapter does not include reference to second order entities or expressions employed in a non-referential fashion (naming, expletive and

vocative uses, etc.). Section 2.2.1 presents the types of referring expressions. Section 2.2.2 describes the statistical analyses and the method used to assess intercoder agreement. The categories retained for analyzing the five factors are presented separately in Section 3.

### 2.2.1 *Types of referring expressions*

The referring expressions used were classified according to their grammatical category in adult language.

#### *Strong linguistic expressions*

- **Nouns** (common, proper, and adjectival) such as *la peau* in [*mwa ʒvø la po*] (*‘moi j(e) veux la **peau**’, ‘me, I want the **peel**’*).
- **Strong demonstrative pronouns** such as *ça* (*‘this’* or *‘that’*), *celui-ci* (*‘this one’*), etc., as in [*sa pik*] (*‘**ça** pique’, ‘**that**’s scratchy’*).
- **Strong personal pronouns** as in [*mwa fe fato*] (*‘**moi** fais château’, ‘**me** make castle’*).

#### *Weak linguistic expressions*

- **Clitic personal pronouns**, as in [*ɛl ɛ dədã*] (*‘**elle** est dedans’, ‘**she**’s inside’*).

- **Preverbal fillers**, as in [*ema?*] (*F*<sup>3</sup> *mange?*, '*F eats*').
- **Null forms**, as in [*va pa oli*] (*va pas au lit*, '*doesn't go to bed*').

We also counted all other forms and grouped them under the heading **All Other Forms**. These included dislocations, clitic demonstrative pronouns, and relative, indefinite, numeral, adverbial, possessive, and interrogative pronouns.

### 2.2.2 *Intercoder agreement and statistics*

Coding was done using a coding guide. After the coding of each dialogue by one of the coders, a different coder recoded 10% of each dialogue independently. The intercoder agreement rates were as follows: linguistic-unit category: 95.81% (Cohen's kappa  $k=0.95$ ), syntactic function: 91.72% (Cohen's kappa  $k=0.888$ ), referent type: 94.52% (Cohen's kappa  $k=0.914$ ), attentional and discursive status of referring expressions: 86.97% (Cohen's kappa  $k=0.796$ ).

Several techniques were used to analyze the data. First, we studied the distribution of referring expressions for the various factors by means of visual inspection of each table. Then, the potential influence of five fixed-effect factors (i.e., MLU group, syntactic function, type of referent, and attentional

---

<sup>3</sup> In the interpretations and translations of children's utterances, 'F' stands for a filler syllable.

and discursive status) on the use of the six linguistic units selected was explored with mixed-effect binomial regressions (using the `glmer` function of the `lme4` R package; Bates, Mächler, Bolker & Walker, 2015), including two random-effect variables to control for the participant (the child) and the session. Mixed models were chosen because the number of sessions per child varied, and mixed models are able to handle missing data and repeated measures. Partition trees (Hothorn & Zeileis, 2015) were drawn with the `ctree` function included in the "party" R package to further explore interactions between the factors.

### **3. Factors affecting the use of referring expressions**

First, we look at the distribution of forms from the developmental standpoint. The question raised is whether, between the beginning and end of the period under consideration, we will find a linear progression in the children's appropriation of referring expressions (see 3.1). Then, we consider the frequency of the different referring expressions, according to the factors considered: syntactic function (see 3.2), type of referent (see 3.3), and attentional and discursive status (see 3.4). In Section 3.5, we present an

analysis of the interactions between the factors and a qualitative analysis of some uses of expressions for referring to entities and participants.

### 3.1 *Impact of linguistic development on the use of referring expressions*

Table 2 presents the distribution of linguistic units for the whole corpus (all groups pooled) and for each MLU group taken separately. Overall, the percentage of nouns was much higher than the percentages of the other forms. Null forms were the second most frequent type of expression in our corpus.

	MLU1 %	MLU2 %	MLU3 %	OVERALL %	N=
Nouns	40.38	36.87	29.62	34.70	1602
SDem*	10.43	8.08	4.58	7.19	332
SPP	3.42	4.86	3.18	3.81	176
CPP	7.28	7.39	16.87	11.31	522
Fil	8.63	6.12	6.77	7.00	323
NF	15.92	14.90	12.18	14.01	647
All other forms	13.94	21.78	26.61	21.98	1015
N=	1112	1584	1921	4617	4617

\* SDem: Strong demonstrative pronouns, SPP: Strong third-person pronouns, CPP: Clitic personal pronouns, Fil: Fillers, NF: Null Forms

Table 2. Distribution (in percentage) of the referring expressions for each MLU group

The percentage of nouns and strong demonstrative pronouns decreased with MLU. Conversely, in MLU Group 3 the frequency of clitic personal pronouns increased. For strong personal pronouns there was no clear difference



between the MLU groups. This lack of variation also existed for null forms and fillers, which could be expected to decrease with MLU.<sup>4</sup>

### 3.2 *Impact of syntactic function on the use of referring expressions*

As noted in the introduction, past research has shown that syntactic function affects the choice of a referential term. This section is devoted to verifying this relation for each of the six types of referring expressions. In the analyses, we made the distinction between the following syntactic functions.

- **Subjects.** Grammatical subjects, as in [ʒvø pasɔ] (*j(e) veut poisson* ‘*I want fish*’).
- **Complements.** Arguments other than subjects. These included direct and indirect objects – as in [ʒəvwa loli] (*je vois Loli* ‘*I see Loli*’), complements in presentative constructions – as in [sa se amwa] (*ça c’est à moi* ‘*that, it’s mine*’), circumstantial complements, and noun determiners.
- **Expressions in verbless utterances.** Pre-syntactic utterances and noun phrases in answers to the interlocutor's questions, as in [ɛbebe] (*un bébé* ‘*a baby*’) in reply to the mother's question *Y avait quoi dedans?* ‘*What was in there?*’.

---

<sup>4</sup> Table 2 also shows that the percentage of “all other forms” rises to almost 22% of the referring expressions. As indicated in section 2.2.1, this grouping lumps together a set of referring expressions (dislocations, clitic demonstrative pronouns, and relative, indefinite, numeral, adverbial, possessive, and interrogative pronouns) that are heterogenous both in their formal and functional aspects. In all of the tables in the chapter, their total percentage is noted only to show the relative importance of the six types of referring expressions under focus in this chapter.

- **Uncertain.** In some cases, the syntactic function could not be determined because the context of the referring expression was itself uninterpretable.

For null forms and fillers, the function was retrieved on the basis of linguistic as well as contextual clues such as the ongoing activity and the adult's discourse. In Example 1, the referent's function in Olg151 was determined by identifying the actors, the object being handled, and the form of the mother's utterance (Mot152).

(1) Olga, 2;3, MLU 2.65<sup>5</sup>

Mot152	celui-là / attends celui-là on va se le mettre là-dedans	'that one / wait that one let's put it in there'
Olg151	[mɛtlada] 'met(tre) là (de)dans'	'put in there'

Table 3 gives the distribution of linguistic expressions for each syntactic function.

	Subject %	Complements %	Verbless %	Unct %	N=
Nouns	4.83	42.15	73.52	70.33	1602
SDem*	2.17	4.86	18.66	6.59	332
SPP	2.37	5.50	4.43	4.40	176
CPP	22.04	5.82	0.00	1.10	522
Fil	14.94	1.20	0.00	2.75	323
NF	17.41	21.91	0.09	9.89	647
All Other Forms	36.24	18.57	3.30	4.95	1015
N=	2028	1255	1152	182	4617

\* SDem: Strong demonstrative pronouns, SPP: Strong third-person pronouns, CPP: Clitic personal pronouns, Fil: Fillers, NF: Null Forms

Table 3. Distribution (in percentage) of referring expressions for each syntactic function

<sup>5</sup> Example captions indicate the name of the child, his/her age (years; months) and the Mean Length of Utterance (MLU) for the cited session. The first three letters of the child's first name are given (e.g. Cle for Clément), Mot stands for mother, Fat for father and Obs for observer. When the children's utterances are transcribed phonetically (between square brackets [ ]), the interpretation in French is given below in inverted commas. An approximate English translation is also given between inverted commas. Braces indicate uncertain transcriptions or alternative interpretations. {X} stands for uninterpretable or inaudible segments. In the interpretations and translations, 'F' stands for a filler syllable. '/' stands for a pause. For more details on transcription conventions, see Appendix II.

Table 3 points out a tendency for different types of expressions to appear in a preferred syntactic function. Subjects were realized mostly by weak forms (clitic personal pronouns, fillers, and null forms), whereas nouns were rare for the subject function. Complements were realized mainly by nouns, and nouns were the preferred form in verbless utterances.

### 3.3 *Type of referent*

We considered two main types of referents.

- **References to entities.** These included references to elements of the objective world, called first-order entities by Lyons, who stated, "First-order entities are such that they may be referred to, and properties may be ascribed to them" (Lyons, 1977:443). An example is *dame* ('*lady*') in [*pati ladam*] ('*pa(r)tie la dame*', '*gone the lady*').
- **References to participants.** Participants were defined by their role in the interlocution and not as entities in the world (Benveniste, 1966). Among the references to participants, we distinguished referring to

the speaker (SELF), referring to the addressee (ADD), and referring to both the speaker and the addressee (WE<sup>6</sup>).

Do these two types of referents involve different uses of referring expressions?

	Entities	Participants			
	%	SELF %	ADD %	WE %	All participants %
Nouns	41.09	6.39	30.38	0.00	9.93
SDem*	9.05	0.00	0.00	0.00	0.00
SPP	0.16	20.97	12.03	0.00	17.95
CPP	8.80	13.75	36.08	62.32	21.01
Fil	5.29	15.97	3.16	13.04	13.62
NF	10.98	29.44	10.76	21.74	25.77
All other forms	24.63	13.48	7.59	2.90	11.72
N=	3670	720	158	69	947

\* SDem: Strong demonstrative pronouns, SPP: Strong third-person pronouns, CPP: Clitic personal pronouns, Fil: Fillers, NF: Null Forms

Table 4. Distribution (in percentage) of referring expressions for each type of referent

Generally speaking, there were many more references to entities than to participants. For referring to entities, nouns were more frequent than weak forms. For referring to participants, the distribution of null forms, clitic personal pronouns, fillers, and strong personal pronouns was more homogeneous. Demonstrative pronouns were never used. In addition, we found variations that depended on whether the children were referring to themselves (SELF), their addressee (ADD), or both (WE). For self-references, the most prevalent expressions were null forms, strong pronouns, and fillers,

<sup>6</sup> “WE” is used in this chapter as the label for the category of reference to self and the addressee.

whereas for referring to the addressee, nouns and personal pronouns were used the most. For WE, the children clearly preferred clitic personal pronouns and null forms.

These results indicate contrasting uses of referring expressions – particularly clitic personal pronouns – that were dependent upon the type of referent.

### 3.4 *Attentional and discursive status of the referent*

We took into account the participants' joint attention to the referent and its previous mention in the dialogue. We considered four categories: new, activated, discourse-given, and reintroduced<sup>7</sup>.

- **New.** The expression introduces a new referent under the addressee's focus of attention.

(2) Daniel, 2;3, MLU 2.5

*Daniel is trying to get a car into a toy house. There is an obstacle.*

Dan201 [abonum] ((il sort un bonhomme ((he gets a man out of the house))  
de la maison))

‘**F bonhomme**’

‘**F man**’

Obs61 ah! il y a un gros bonhomme là. ‘oh! there's a big man there.’

- **Activated.** The referring expression is the first mention of a referent already under the attention of the interlocutors.

---

<sup>7</sup> Our definitions of the categories for the attentional and discursive status of the referent do not necessarily correspond to those used in other studies (Chafe, 1976; Gundel et al., 1993).

(3) Loli, 2;3, MLU 2.94

*Loli is eating a banana. The referent 'peau de la banane' ('banana peel') has not yet been mentioned and is under the joint attention of the two speakers.*

Lol3	[mwəʒəvəlapo]	
	'moi je veux <b>la peau</b> '	'me, I want <b>the peel</b> '
Mot3	ça c'est la peau oui.	'that, it's the peel yes.'

- **Discourse-given**<sup>8</sup>. The referent has already been mentioned by the child or his/her interlocutor within the same thematic sequence.

(4) Arnaud, 2;3, MLU 2.96

*The observer has just left the room after positioning the camera.*

Arn6	[{uwe} / klɛmari: ?]	
	'où (elle est/ est) / Claire-Marie?' 'where is / Claire-Marie?'	
Mot7	ben je sais pas.	'uh I don't know'
Mot8	Alors	'so'
Arn7	[ðva owalɔ̃ klɛmari:]	
	'F va au salon <b>Claire-Marie</b> '	'F {goes/go} into the living room, <b>Claire-Marie</b> '

Here, in Arn7 the child mentions "Claire-Marie" for the second time (see Arn6).

- **Reintroduced**. The child reintroduces a referent that had been mentioned in a previous thematic sequence, with a focus on other topics in-between, as in Example 5 (Arn20).

(5) Arnaud, 2;3, MLU 2.96

*This sequence came after Example 4, following a topic change.*

Arn20	[wɛle klɛmari ?]	
	'où <b>elle</b> est <b>Claire-Marie</b> ?' 'where is <b>she</b> , <b>Claire-Marie</b> ?'	
Mot25	ben je sais pas où elle est passée/ elle va revenir	'uh I don't know where she went/ she's going to come back'

---

<sup>8</sup> We will also refer to this category as "given".

This analysis was conducted both for entities and for participants. By definition, participants never had a new status, but they could have any one of the other three statuses. Table 5 gives the distribution of referring expressions according to the referent's attentional and discursive status.

	New %	Activated %	Discourse-Given %	Reintroduced %	N=
Nouns	53.68	32.72	31.18	47.95	1602
SDem*	8.68	20.58	5.01	7.71	332
SPP	0.26	0.62	4.98	1.45	176
CPP	1.32	5.76	13.73	7.47	522
Fil	2.11	4.53	8.12	5.30	323
NF	3.95	9.67	16.67	6.99	647
All Other Forms	30	26.13	20.32	23.13	1015
N=	380	486	3336	415	4617

\* SDem: Strong demonstrative pronouns, SPP: Strong third-person pronouns, CPP: Clitic personal pronouns, Fil: Fillers, NF: Null Forms

Table 5. Distribution (in percentage) of referring expressions for each attentional and discursive status

Most of the expressions encoded given referents (which reflects dialogue cohesion). In spite of this predominance, different distributions of referring expressions were found for each status. Table 6 shows that nouns were the most frequent type of expression for referring to new or reintroduced referents. On the other hand, nouns were used proportionately less for given or activated referents. Clitic personal pronouns obtained the opposite pattern: they were mainly associated with given referents and were hardly ever chosen in new contexts. They were not used much in activated and reintroduced contexts either, but were not as rare as when the context was new. Fillers and null forms exhibited the same pattern, and strong demonstrative pronouns were found the most for activated referents.

### 3.5 *Interactions between factors*

Interactions between factors were analyzed for the following binomial variables.

- **MLU group.** MLU2 vs. MLU1, and MLU3 vs. MLU1.
- **Type of referent.** Entities vs. Participants

Among the participants, we selected the referent for which the expression was the most frequent (see Table 4), that is, addressee (ADD) vs. others for nouns and clitic personal pronouns, and SELF vs. others for null forms and fillers. This factor was not tested for strong demonstrative pronouns because they never referred to participants.

- **Attentional and discursive status.** We tested discourse-given vs. all others and new vs. all others, except for strong demonstrative pronouns, for which activated was the most frequent status.
- **Syntactic function.** We tested subject vs. all other functions. This factor was not tested for fillers, which, by definition, are in preverbal position. In this case, the function overrides the other variables.



The model with the best fit for each referring expression was chosen by backwards elimination of nonsignificant factors. Tables 6 and 7 present the results of the logistic regressions. In order to further assess the interactions between these factors, partition trees (Figures 1 to 6) were constructed with the factors that yielded significant values in the logistic regressions. We considered the type of referent to be a single factor with four variables: SELF, ADD, WE, and entities.

Fixed Effects	Est*.	S.E.	Z	p	Random Effects	Var.	S.D.	C-Value **
Nouns								0.88
Intercept	-1.592	0.187	-8.513	<.0001	Session	0.381	0.617	
Attentional and discursive status (New)***	0.359	0.135	2.665	.0077				
Type of referent (Entities)	2.105	0.174	12.099	<.0001				
Syntactic function (Subject)	-3.281	0.117	-28.072	<.0001				
Participant (ADD)	2.189	0.286	7.658	<.0001				
Strong Demonstrative Pronouns								0.85
Intercept	-2.291	0.192	-15.57	<.0001	Session	0.706	0.841	
Attentional and discursive status (Activated)	1.634	0.150	10.92	<.0001	Child	0.260	0.510	
Syntactic function (Subject)	-1.721	0.171	-10.09	<.0001				
Strong Personal Pronouns								0.96
Intercept	-1.3603	0.3277	-4.151	<.0001	Session	1.059	1.029	
Type of referent (Entities)	-5.069	0.511	-9.930	<.0001				
Syntactic function (Subject)	-2.448	0.233	-10.532	<.0001				
Participant (SELF)	1.171	0.308	3.799	0.0001				

\*Est.: Estimate; S.E.: Standard Error; Var.: Variance; S.D.: Standard Deviation.

Number of observations: 4617, Grouping factors: Sessions: 60, Children: 28

\*\* The models were assessed using C statistics, which indicate whether the predicted binomial outcome is better than chance. Hosmer, Lemeshow and Sturdivant (2013) considered that C-values below .5 indicated a model that is not better than chance whereas C-values above .7 are reasonable and above .8 are strong.

\*\*\* The category in parentheses on fixed-effect rows indicates the category value compared with the intercept

(reference level). For binomial factors, the reference level is the other factor.

Table 6. Regression tables for strong referring expressions (nouns, strong demonstrative pronouns, strong personal pronouns)

Fixed Effects	Est*	S.E.	z	P	Random Effects	Var.	S.D.	C-Value**
Clitic Personal Pronoun								0.87
Intercept	-4.563	0.327	-13.945	<.0001	Session	0.358	0.598	
MLU (Group 2)***	-0.0209	0.377	-0.055	.9558	Child	0.522	0.722	
MLU (Group 3)	0.811	0.345	2.351	.0187				
Attentional and discursive status (Discourse-given)	0.688	0.162	4.255	<.0001				
Attentional and discursive status (New)	-1.350	0.479	-2.823	.0048				
Syntactic function (Subject)	2.195	0.136	16.102	<.0001				
Participant (ADD)	1.362	0.224	6.079	<.0001				
Fillers								0.76
Intercept	-3.490	0.226	-15.427	<.0001	Session	0.342	0.585	
Attentional and discursive status (Discourse-given)	0.372	0.173	2.149	.0317	Child	0.322	0.568	
Attentional and discursive status (New)	-0.801	0.387	-2.073	.0382				
Participant (SELF)	1.187	0.134	8.848	<.0001				
Null Forms								0.75
Intercept	-2.703	0.175	-15.466	<.0001	Session	0.297	0.545	
Attentional and discursive status (Discourse-given)	0.591	0.134	4.395	<.0001	Child	0.193	0.439	
Attentional and discursive status (New)	0.756	0.292	-2.594	.0095				
Syntactic function (Subject)	0.220	0.093	2.369	.0178				
Participant (SELF)	0.917	0.107	8.563	<.0001				

\*Est.: Estimate; S.E.: Standard Error; Var.: Variance; S.D.: Standard Deviation.

Number of observations: 4617, Grouping factors: Sessions: 60, Children: 28

\*\* The models were assessed using C statistics, which indicate whether the predicted binomial outcome is better than chance. Hosmer, Lemeshow and Sturdivant (2013) considered that C-values below .5 indicated a model that is not better than chance whereas C-values above .7 are reasonable and above .8 are strong.

\*\*\* The category in parentheses on fixed-effect rows indicates the category value compared with the intercept (reference level). For binomial factors, the reference level is the other factor; for multilevel factors (i.e. MLU) the reference level is the level not appearing in any row of the table.

Table 7. Regression tables for weak referring expressions (clitic personal pronouns, fillers, and null forms)

### 3.5.1 Nouns

As Table 6 shows, the use of a noun was affected by four factors. The probability that a child would choose a noun was significantly higher when the referent had a syntactic function other than subject, when it was new to the interaction, when it referred to an entity rather than to a participant, and, if it did refer to a participant, when the participant was the addressee.

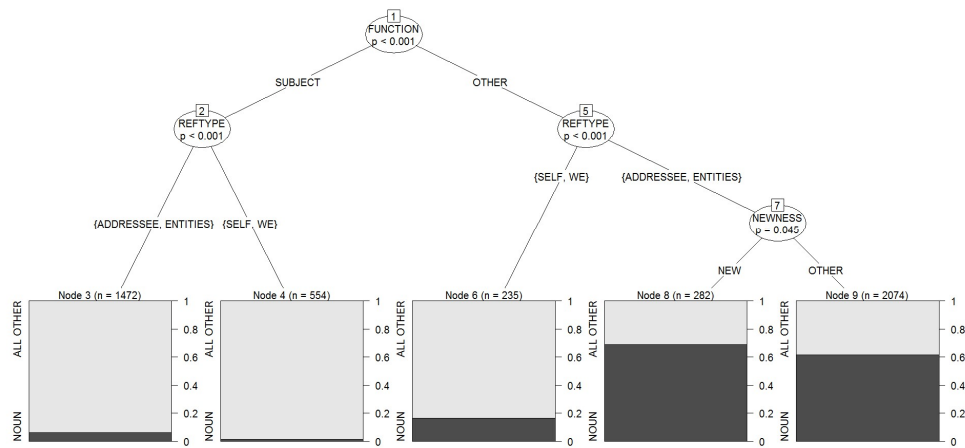


Figure 1. Binary partition tree for referring uses of nouns

Figure 1 shows that syntactic function was the main factor that accounted for the use of nouns, which were infrequent in the subject function. For both subjects and non-subjects, the type of referent was the second factor for nouns: nouns were used more frequently to refer to entities or to the addressee than to the SELF or WE.

For the non-subject functions, we can see that the partition tree lumps together references to entities and references to the addressee, although the former

were much more frequent than the latter (see Table 4). In this analysis, the proportion of nouns for new referents (necessarily entities) was significantly greater than the proportion for non-new referents (activated, discourse-given, and reintroduced taken together). However, the number of nouns for non-new referents was far from negligible, in such a way that function and referent type overrode attentional status. The effect of this was that nouns were more frequent among given referents (31.18 % in Table 5). The analysis of a few examples helps account for this unexpected pattern of noun use.

In Example 6, the child uses the noun *camion* ('truck') for a given referent.

(6) Elodie, 2;2, MLU 2.04

*Elodie and her mother are playing with Legos. She just took apart a garbage truck.*

Mot22	ben il est où le camion?	'so where is it, the truck?'
Elo85	[ipati]	
	'i(l) pa(r)ti'	'it's gone'
Mot123	il est parti le camion?	'is it gone the truck?'
Mot124	qu'est-ce qu'elle a fait Elodie au camion? / c'est tout ce qui reste	'what did she Elodie do to the truck? /that's all that's left'
Elo86	[a! / ajajapy / japyjamðpubel {xx}]	
	'ah! / a (il)ya (il)ya p(l)u(s). / ya p(l)u(s) camion poubelle'	'oh! / no more, no more / no more garbage truck.'
Mot125	y a plus / il est parti	'no more / it's gone'
Elo87	[pakamjðpubel?/ padkamjðpubel?]	
	'pas camion poubelle / pas d(e) camion poubelle?'	'no garbage truck / no garbage truck?'
Mot126	ben oui	'well yes'

In this excerpt, the child uses two types of referring expressions: the clitic pronoun *il* ('it') acting as the subject (Elo85), and the noun *camion* ('truck') in functions other than subject (object of a presentative in Elo86, or alone in a verbless utterance in Elo87). Note also that nouns occurred in "negative" utterances.

Example 7 suggests that when a child is speaking about a participant, nouns tend to appear in contexts implying a comparison or an opposition between (potential) referents.

(7) Olga, 2;4, MLU 2.65

Olg198	[seamwa] 'c'est à moi'	'it's mine'
Mot199	c'est à toi oui	'yes it's yours'
Olg199	[wa:] {xx}	{xx}
Mot200	sais bien que c'est à toi	'know very well it's yours'
Olg200	[epapukmamã] 'est pas pour <i>maman</i> '	'is not for <i>Mommy</i> '
Mot201	c'est pas pour maman	'it's not for Mommy'

In Example 7 the child refers to her mother using *maman* ('Mommy') in a negative utterance [epapukmamã] 'est pas pour *maman*' ('is not for Mommy') after she identifies herself as the possessor of the object of attention (Olg198). At the syntactic level, nouns tended to be used in utterances containing negative adverbs and/or cleft constructions. At the discursive level, previous and/or following utterances of the same thematic sequence contained similar indicators of opposition (negative adverbs or cleft structures) or clarification questions from the adult on the child's project with *pourquoi?* ('why?').

### 3.5.2 Strong demonstrative pronouns

Table 6 shows that for strong demonstrative pronouns, only two factors had an impact: attentional and discursive status and syntactic function. These forms were more probable for activated referents than for given or new ones, and were less probable as subjects.

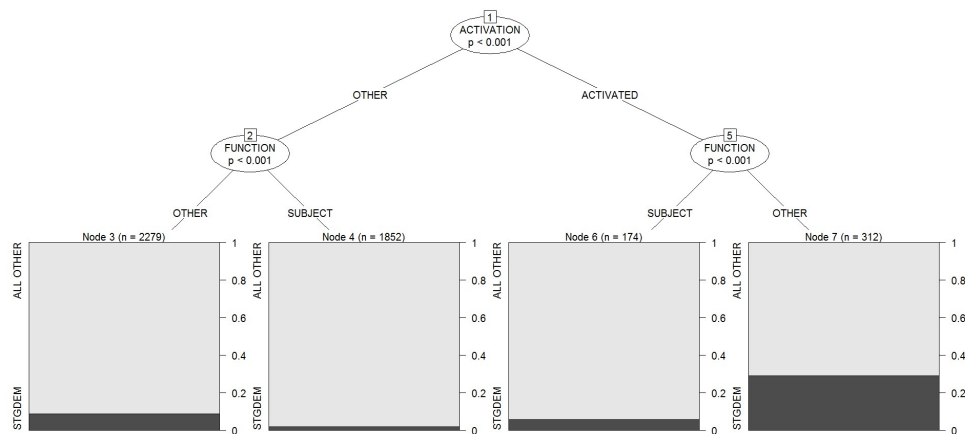


Figure 2. Binary partition tree for strong demonstrative pronouns

Figure 2 shows that the main factor was activation, followed by function. The effect of function was more prevalent for activated referents than for non-activated ones. As a whole, strong demonstrative pronouns were preferentially used in an activated context and a non-subject function, as in Example 8.

(8) Margaux, 2;3, MLU 2.62

*Margaux has just said she doesn't like water*

Mot74	oh si pourtant tu adores ça	‘oh but yes, you love that’
Mar65	[e ! / <i>sa</i> ] ((pointe le dessin sur le paquet))	((pointing to a drawing on the package))
	‘eh ! / <i>ça</i> ’	‘hey! / <i>that</i> ’
Mot75	oui c'est un singe	‘yes, it's a monkey’

In Mar65, the child uses the demonstrative pronoun *ça* (‘*that*’) to introduce a referent that is part of the same activity even though it has not yet been mentioned.

But *ça* also appeared when the referent had already been mentioned (discourse-given). It was used in contrasts too, as in Example 9 (utterance Clc24):

(9) Clémence, 2;3, MLU 2.56

*Clémence and her mother are playing lotto.*

Mot48	essaye ça pour voir si ça marche	‘try that to see if it works’
Clc24	[no se <i>pas</i> a]	
	‘non c'est pas <i>ça</i> ’	‘no it's not <i>that</i> ’
Mot49	ah bon	‘okay’

### 3.5.3 Strong personal pronouns

In Table 6, we saw that the children were more likely to use strong personal pronouns when referring to themselves, and less likely when the referent was an entity in the subject function.

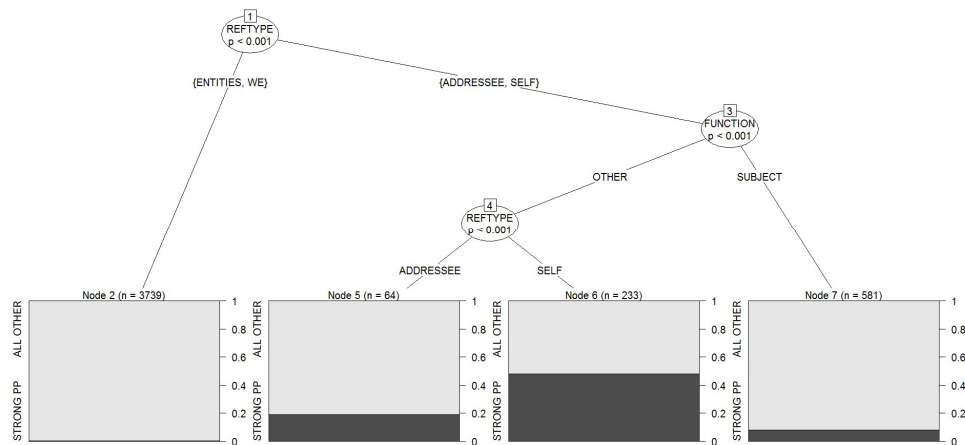


Figure 3. Binary partition tree for strong personal pronouns

Figure 3 shows that the most prevalent factor was the type of referent. For referring to entities or WE, strong pronouns were rare. On the other hand, they were more frequent for references to the self and to the addressee, especially *toi* ('you') and *moi* ('me'). In these cases, the highest percentages were obtained when the referent was not the subject – usually in conjunction with the presentative *c'est* ('it's') – or was in a simple or cleft construction that marked, or even contrasted, possession or agency (Example 10).

(10) Cécile, 1;10, MLU 2.83

Fat60	bon moi je vais mettre la petite vache dans le dans le-	'so me/ I'm going to put the little cow in the in the-'
Céc51	[nɔ̃ sɛmwa] 'non c'est <i>moi</i> '	'no it's <i>me</i> '
Fat 61	c'est toi?	'it's you?'
Céc52	[wi] 'oui'	'yes'
Fat62	ah bon	'oh'
Céch53	[sepɑtwa] 'c'est pas <i>toi</i> '	'it's not <i>you</i> '



### 3.5.4 Clitic personal pronouns

Table 7 shows that the use of clitic personal pronouns was significantly affected not only by MLU (higher frequency in MLU3), but also by attentional and discursive status (lower probability of using a clitic pronoun for new referents and higher probability for given referents) and type of referent (higher probability of using a pronoun for the addressee). There was also a significant impact of syntactic function, with clitic personal pronouns being more probable for subjects.

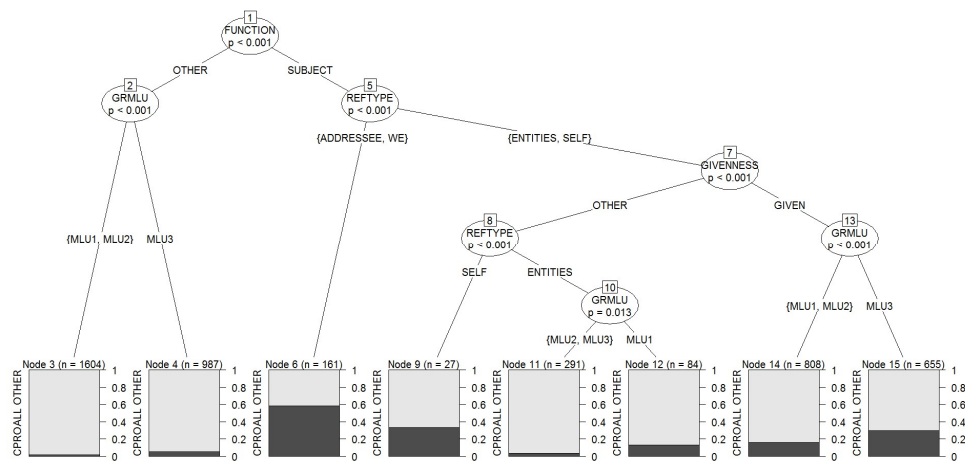


Figure 4. Binary partition tree for clitic personal pronouns

Figure 4 shows that syntactic function (subject vs. all other functions) was the main factor, since clitic personal pronouns rarely appeared in non-subject functions. Accordingly, 75 clitic personal pronouns were used for a function other than subject; they were slightly more frequent in MLU Group 3 than in

the other two groups. Among those 75 clitic personal pronouns, 57 were direct or indirect objects,<sup>9</sup> of which 47 had the given status and 7 were in joint-attention contexts. For the subject function, the type of referent influenced the number of pronouns used. For the addressee or WE, the children produced up to 60 pronouns. For the SELF and entities, the use of clitic personal pronouns was influenced by givenness and MLU. When the referent was not discourse given, only self-references triggered a certain number of pronouns. And when the referent was given, there were more pronouns for MLU3, irrespective of the referent type. This occurred, for instance, in Elo85 of Example 6, *i pati* ('it's gone'), which is a typical case of plain continuity, where taking up a referent corresponds to continuation of a predication (here, a response about a shared topic). When the child did not produce a pronoun in such contexts, he/she most often employed null forms or fillers, as in Example 11 (Clc 63 and 64):

(11) Clémence, 2;2, MLU 2.85

Mot73	il dort ?	'he's sleeping?'
Clc57	[idɔʁ]	
	'i(l) dort'	'he's sleeping'
(...)		
Mot83	oh oh oh Petit Ours Brun regarde	'oh oh oh Little Brown Bear! look!'
Clc63	[pa fɛ dodo]	
	'pas fait dodo'	'not sleeping'
Mot84	qu'est-ce qu'il fait Petit Ours Brun ?	'what is Little Brown Bear doing?'
Clc64	[afɛ {solatal}]	
	'F fait {sur la ta(b)le}'	'F 's doing {on the table}'

---

<sup>9</sup> The other cases corresponded to the adverbial pronoun *en* ('of it', 'of them'), as in [jānapy?] ('y en a p(l)us?', 'there is/are no more of it/them', Olga).

In Clc57, the child takes up the clitic pronoun just used by her mother. Later, whereas the mother reintroduces the referent by way of a noun (Mot83), the child adds a predication using a null form, and then, after a second question from the mother, produces another predication using a filler [a]. These three weak forms are in plain continuity with respect to an already mentioned referent. On the other hand, strong forms were preferred in contrast contexts, as above in Example 6.

The same held true for references to participants. In the excerpt below, the child first refers to her interlocutor (the mother) using a noun; then she uses a clitic pronoun.

(12) Pauline2, 2;3, MLU 2.52

*Pauline and her mother are drawing hats.*

- |       |  |                               |
|-------|--|-------------------------------|
| Pau54 | [mẽtnã se ẽ puʁ mamã]                  |                               |
|       | ‘maintenant c'est un pour maman’       | ‘now it's one for Mommy’      |
| Mot96 | il est comment le chapeau d(e) maman ? | ‘what is Mommy's hat like?’   |
| Pau55 | [tyfɛʃapo?]                            |                               |
|       | ‘ <b>tu</b> fais chapeau?’             | ‘ <b>you</b> 're making hat?’ |
| Mot97 | oui.                                   | ‘yes.’                        |

This excerpt begins with a contrast introduced by the child between a hat already drawn and a second one not yet drawn. The subject clitic pronoun referring to the mother (Pau55) is produced by the child in an utterance without contrast that is in continuity with her own utterance.

Figure 4 shows that a few rare clitic personal pronouns were used for not-yet-mentioned entities. The qualitative analysis of these cases brought out the importance of joint attention to these referents.

(13) Maxime, 2;3, MLU 1.45

*Maxime and his mother are playing cards with pictures of rabbits on them.*

Mot66 il est pas triste là-hein . 'he's not sad there/ huh. ((third  
 ((troisième 'lapin')) là *il* est un 'rabbit')) there *he*'s a little sad,  
 peu non ((quatrième 'lapin')) no? ((fourth 'rabbit'))'  
 Max 60 [ʁegaɪplœʁ]  
 'regarde *i(l)* pleure ((cinquième 'look *he*'s crying ((fifth 'rabbit'))'  
 'lapin'))'

In this example, Maxime and his mother talk about the emotions of various rabbits, on which both of them are focusing. In Max60, the child mentions the fifth rabbit for the first time using the clitic personal pronoun *i(l)* ('*he*'). This referent is similar to the preceding ones introduced by the mother. Note that in the previous speaking turn (Mot66), the mother had also introduced a referent belonging to the same category, via a clitic personal pronoun.

### 3.5.5 Fillers

In Table 7, we can see that the children were more likely to use fillers when referring to the SELF and less likely when referring to entities not yet mentioned in the discourse.

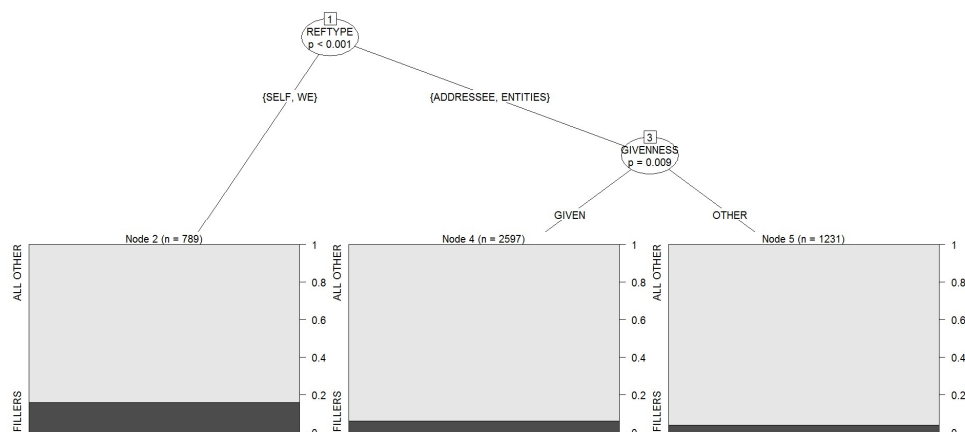


Figure 5. Binary partition tree for fillers

Figure 5 shows that fillers were more frequent in references to oneself (SELF) or to the dyad (WE). In the other cases, altogether scarce, fillers were preferentially used for given referents.

Example 14 illustrates a case of references to the self (Elo13) and to WE (Elo11) by means of a filler. In Elo13, the filler is in continuity with the preceding utterance. This preceding utterance expresses an opposition pertaining to the agent of the current process via an utterance containing not only an adverb of negation but also a cleft construction about the agent and a noun in a self-reference (Elo12).

(14) Elodie, 2;2, MLU 2.04

Mot12	là / ici / on le met là.	‘there/ here/ we put it there.’
Elo11	[ɔmela] ‘ <i>F</i> met là.’	‘ <i>F</i> put there’
Elo12	[nɔseelodikife {sa}] ‘non c’est <i>Elodie</i> qui fait ça.’	no it’s <i>Elodie</i> who does that.
Mot13	c’est Elodie qui fait ça alors mets là.	‘it’s Elodie who does that, so put there.’

Elo13      [la / amela / sela]  
               'là / *F* mets là / c'est là.'                      'there / *F* put there / it's there.'

Regarding the occurrences of fillers for non-given entities, many uses were for activated referents (19 out of 38), as in Example 15.

(15) Arnaud, 2;2, MLU 2.96

*Arnaud and his mother are putting toy beads into holes in a game board. Arnaud touches the rug.*

Arn36      [pukwaɪmwela]  
               'pourquoi *F* mouillé là.'                      'why *F* wet there.'

In this excerpt, Arnaud (Arn36) attracts his mother's attention to the rug where they are sitting. This not-yet-mentioned referent is part of the shared activity. Moreover, it is actualized by a gesture (a touch) accompanying the deictic *là* ('there').

### 3.5.6 Null forms

The children were more likely to use null forms (see Table 7) when the referent was discourse-given, acting as the subject, and referring to themselves. They used null forms less often for new referents.

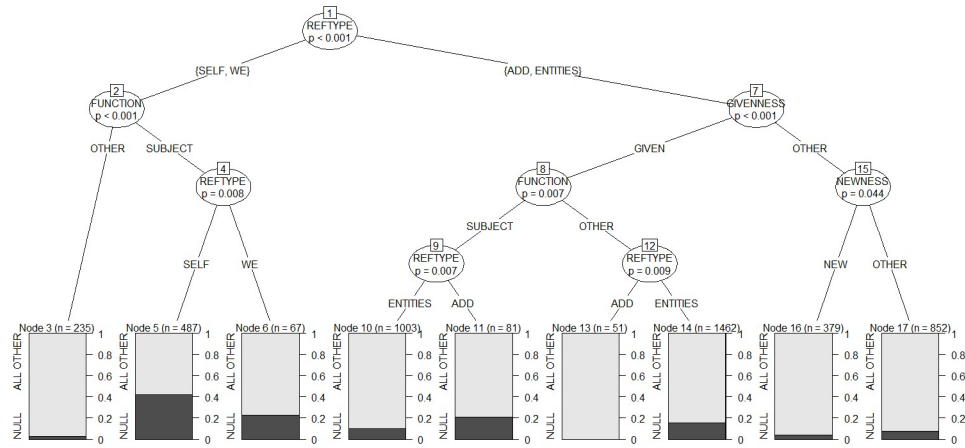


Figure 6. Binary partition tree for null forms

Figure 6 shows that these factors interacted in a complex way. The type of referent was the primary factor. The secondary factors for the SELF and WE differed from those for the addressee and entities on the other. The SELF and WE references involved more null forms when the referring expression was acting as the subject. In this case, the SELF category triggered more null forms than the WE category did.

Furthermore, the production of null forms to refer to the addressee (ADD) or to entities partly depended on givenness. Non-given referents (nodes 16 and 17 on the right) were seldom encoded using null forms, and when they were, the referents were activated or reintroduced. The greatest number of null forms was found for given referents acting as the subject. As we saw above in Example 11, null forms in the subject function alternated with fillers and clitic personal pronouns: we found more null forms for the addressee than for

entities, whereas the opposite was true for non-subject functions (see Example 1).

Example 16 illustrates the case of a null form acting as the subject and referring to the addressee. Pauline, the child, first uses the clitic subject *t'* (‘you’) to offer her mother some of her bun, saying [tãvø?] (‘*t'en veux?*’, ‘do you want some?’). When the mother refuses, the child produces an explanation while referring to her mother without mentioning her. We were able to identify this reference, thanks not only to the preceding discursive context, but also to the content of the rest of the utterance and the mother’s reformulation.

(16) Pauline2, 2;3, MLU 2.52

Pau125	[tãvø ?]	
	‘t'en veux?’	‘do you want some?’
Mot246	non merci .	‘no thanks.’
Pau126	[dezøne avøk papa?]	
	‘déjeuné avec papa ? ‘	‘ate with Daddy?’
Mot247	oui ((rit)) j'ai déjeuné avec papa	‘yes ((laughing))I ate with Daddy
	oui c(e) matin.	yes this morning’.

#### 4. Discussion

In this chapter, we attempted to account for the repertory of referring expressions and their uses among 28 French-speaking children in a natural-dialogue situation taking place in their homes. We raised the question of how



the use of referring expressions is related to linguistic development (measured by MLU) and to syntactic and discourse-pragmatic factors, and of how these different factors interact with each other. We approached this issue from the angle of the opposition between strong and weak forms.

The children ranged in age between 1;7 and 2;6, a crucial period for the development of grammatical paradigms. As noted above, nouns and strong demonstrative pronouns are the earliest referring expressions used, so they were expected to decrease proportionally with MLU. Strong and clitic personal pronouns were expected to increase. We also predicted that the increase in clitic personal pronouns would lead to a decrease in pre-syntactic forms (fillers and null forms). The results indicated a lower proportion of nouns and a higher proportion of clitic personal pronouns in MLU Group 3, whereas no significant tendencies were observed for the other four types of forms. Statistical data processing (with participants and sessions as random factors) allowed us to confirm that factors other than pure grammatical development play an important role<sup>10</sup>. Two important, interrelated findings can be mentioned here. On the one hand, the patterns of referring expressions differed across the various factors; on the other, the different factors considered did not act in the same way for the different forms:

---

<sup>10</sup> Marcos, Salazar Orvig, da Silva-Genest and Heurdier (Chapter 7), Hassan, de Weck, Rezzonico, Salazar Orvig and Vinel (Chapter 9), de Weck et al. (Chapter, 9), and Vinel et al., (Chapter 10 of this volume) address other factors such as the impact of the dialogue, the activities carried out, and the different speech genres.

- a) The contrast between references to entities and references to participants – and among the latter, between the self and the addressee – affected the choice of referring expressions in a complex fashion. The first uses of strong pronouns concerned self-references, whereas demonstrative pronouns served only for referring to entities. Nouns were mainly used for entities but were sometimes employed for participants, especially the addressee.
- b) Attentional and discursive status had an impact on the use of most of the referring expressions produced. Nouns and strong demonstrative pronouns were less probable when the referent was given; nouns were more probable for new and reintroduced referents; and strong demonstrative pronouns occurred most often when the referent was activated. By contrast, weak forms were strongly influenced by givenness, whereas strong personal pronouns were not affected by this factor.
- c) The syntactic function (subject vs. other functions) played an important role for all types of referring expressions: the three strong forms had the highest probabilities of occurring in a function other than subject, whereas weak forms were preferentially chosen for the subject function.

The various factors interacted in different ways, depending on the type of referring expression. For nouns and clitic personal pronouns, all of the factors studied allowed us to predict the use of referring expressions with a relatively high degree of certainty ( $C=0.87$  and  $0.88$ ). These two forms were symmetrical to each other in our data. While clitic personal pronouns were preferred for the subject function, nouns were not chosen as subjects. And while clitic personal pronouns were used for participants, nouns were preferred for entities. Lastly, clitic personal pronouns were associated with given referents, whereas nouns were associated with new or reintroduced referents. It thus seems indeed that children acquire clitic personal pronouns for functions that are complementary to those of nouns.

The other two weak forms, null forms and fillers, occurred in similar contexts to those of clitic personal pronouns. The two strong pronoun forms (strong demonstrative and strong personal pronouns) functioned differently at the attentional and discursive level: strong demonstrative pronouns were preferred for entities in an activated context, while strong personal pronouns were used the most for participants, regardless of their attentional and discursive status.

This functional specialization is not a grammatical constraint. Rather, it is probabilistic in nature (Allen, 2008). What we are seeing here is that children are highly likely to use weak forms in given-referent contexts and strong

forms for referents not mentioned recently. However, both strong forms and weak forms can be found in their non-preferential contexts.

By means of qualitative and quantitative analyses, then, we showed that clitic personal pronouns encoding referents not yet mentioned in the discourse were used for referents upon which the participants' attention was already focused. In line with cognitive views of anaphora (Ariel, 1990; Cornish, 1999; Gundel et al., 1993), joint attention may indeed suffice for constructing a shared representation of a situation (see also Skarabela et al., 2013, for omissions in Inuktitut).

On another side, we observed the use of strong forms for given referents, both participants and entities. Regarding references to participants, the qualitative analysis suggested that they occur in contexts of contrast or opposition. Our results are consistent on this point with Morgenstern's (2006) and Caët's (2013), although our frequencies were higher for the addressee than for the self. In our data, contrasts were usually achieved via strong pronouns in the case of self-reference.

Our qualitative analysis brought out similar functioning for references to entities: weak forms were preferred in contexts involving the simple addition of a predicate; strong forms were used to express oppositions and contrasts. These findings, in line with Chafe (1976), converge with Salazar Orvig et al.'s (2010b) results. In their study on the choice between omissions and explicit

forms, Hughes and Allen (2006) found that the contrast variable overrode the other accessibility factors (see Marcos et al., Chapter 7, this volume).

Thus, for entities, clitic personal pronouns indeed seem to be acquired with the initial function of encoding referents that are the most accessible or benefit from the interlocutors' joint attention. However, Figures 1 and 4 also showed that syntactic function was the predominant factor in the use of clitic personal pronouns and nouns, with the other factors being subordinate to it. This is a complex question.

It should be noted, first of all, that during the period of linguistic development at stake here, children have not yet mastered object pronouns (Hamann et al., 1996). Object pronouns – too scarce to carry weight in the Figure 4 partition – nonetheless seem to have been acquired well enough to be used in the same type of context as subject pronouns.

This brings us to the fundamental issue of the interface between grammar and pragmatics (and on a more general level, a key question in functional linguistics): Can syntactic function be considered independent of functional factors? Generally speaking, it seems difficult to disentangle "the purely grammatical" from "the pragmatic" in these conventional forms, since the subject function, as stated by Chafe (1976: 55), is "the item about which knowledge is added." As such, it is the function that serves to encode given referents, to the current topic, or to whatever is highly accessible (Allen, 2006; Ariel, 2000; Givón, 1983). Moreover, in languages like French, a cleft

construction or a dislocation is generally chosen to introduce or reintroduce a referent in the subject function (Lambrecht, 1994).

The contrast between strong and weak forms seems to precede the construction of syntactic relations. The fact that null forms and fillers appeared in the same types of contexts as clitic personal pronouns suggests that children acquire the functional contrast between referents under joint attention, and referents that are not, before they are able to handle pronouns. Clearly, in the earliest stages of language acquisition, during the period of one-word utterances, they seem to produce isolated nouns or deictics like *ça* ('that') or *là* ('there') to attract an adult's attention to objects, and they produce predications (not yet verbs) about highly accessible referents (already mentioned or under joint attention). Such utterances with non-verbalized referents fall into a relation of referential continuity with the preceding discourse. The same holds true during the presyntactic transition period where fillers, little by little, contribute to forming the verb category (Veneziano, 2017) and, thus, the subject function. This kind of proto-anaphoric function will serve as a medium for grafting the use of the first pronouns, which themselves are morphologically specialized for performing the subject function.

Lastly, our data brought out an asymmetry in the acquisition of the clitic-pronoun paradigm (and thus in the use of fillers and null forms) that was based on the type of referent. This asymmetry has often been brought up, not only

in developmental studies but also in work on person marking in different languages (Ariel, 2000), with the first and second persons being seen as more accessible than the third person. This was the case in our self-reference data (see also Yamaguchi et al., Chapter 2, this volume), but it was not so for references to the addressee or WE, which involved more clitic personal pronouns and fewer fillers and null forms than references to entities. We can put forth two possible hypotheses here: repetition of the French pronoun *on* (third-person singular pronoun often employed to mean ‘we’) in adult productions, and the preferential association of certain referring expressions with certain verbs or speech acts. This question (which lies beyond the scope of this chapter) remains unanswered and requires a study of its own.

## **5. Conclusion**

Our study confirmed prior results on the early development of children's referential skills while also deepening our understanding of this issue. We assessed the separate and combined effects of several factors involved in syntactic and pragmatic development. We also included, in the same statistical analysis, references to entities and references to participants. Finally, via qualitative analysis, we explored some cases of discrepancies

between the observed and expected relations between a given factor and the referring expressions used. In line with the data in the literature, we found a strong link between referent givenness and the use of referring expressions: weak forms were employed especially for given referents, strong forms for new or reintroduced referents. Our finding that syntactic function was the major factor supports this idea: the association between grammatical subjects and weak expressions corresponds to the fact that the grammatical subject of an utterance generally encodes what happens to be the interlocutors' current focus of attention. These results argue strongly in favor of the intricate relations between morphological, syntactic, and pragmatic development.

Our analyses showed, however, that other factors (linguistic development level, type of referent, etc.) also play a role, and that the impact of each factor is partly conditioned by that of the others. We have thus painted a highly complex picture of children's syntactic and pragmatic abilities at an early age. The qualitative analysis of the discrepant cases noted here provides a basis for future research.

### **Funding**

This research was made possible thanks to the grants from the ANR *Agence Nationale de la Recherche* (French National Agency for Research) for the funding ([Grants n° ANR-09-ENFT-055]).

### **Acknowledgements**



We thank the internal and external anonymous reviewers, as well as Shanley Allen, for their helpful comments, the team of the DIAREF project (*Acquisition des Expressions Référentielles en dialogue: approche multidimensionnelle*) for their contribution to the scientific and methodological discussions. The data analyses reported here were conducted by the authors with the collaboration of Janina Klein, Marine Le Mené and Salma Nashawati. We thank Stefano Rezzonico for his invaluable advice for the statistical analysis. A special thank goes to Vivian Waltz for her accurate and rigorous translation from French to English.

## References

- Allen, S. E. M. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics*, 38(3), 483-521.
- Allen, S. E. M. (2006). Formalism and functionalism working together? Exploring roles for complementary contributions in the domain of child null arguments. In R. Slabakova, S. A. Montrul & P. Prévost (Eds.), *Inquiries in linguistic development: In honor of Lydia White* (pp. 233-255). Amsterdam: John Benjamins.
- Allen, S. E. M. (2008). Interacting pragmatic influences on children's argument realization. In M. F. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 191-210). Mahwah: Lawrence Erlbaum Associates.

- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.
- Ariel, M. (2000). The development of person agreement markers: From pronouns to higher accessibility markers. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 197-260). Stanford, CA: CSLI Publications.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi: 10.18637/jss.v067.i01.
- Benveniste, E. (1966). La nature des pronoms *Problèmes de linguistique générale* (Vol. 1, pp. 251-257). Paris: Gallimard.
- Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, 21, 491-504.
- Budwig, N. (1995). *A developmental-functionalist approach to child language*. Hillsday, N.J: Erlbaum.
- Budwig, N., & Wiley, A. (1995). What language reveals about children's categories of personhood. *New Directions for Child and Adolescent Development*, 1995(69), 21-32.
- Caët, S. (2013). *Référence à soi et à l'interlocuteur chez des enfants francophones et anglophones et leurs parents* (PhD dissertation). Université Sorbonne Nouvelle Paris 3, Paris, France.

- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 27-55). New York: Academic Press.
- Clancy, P. M. (1993). Preferred argument structure in Korean acquisition. In E. V. Clark (Ed.), *Proceedings of the 25th Annual Child Language Research Forum* (pp. 307-314). Stanford, CA: CSLI Publications.
- Cornish, F. (1999). *Anaphora, discourse and understanding*. New York: Oxford University Press.
- Culbertson, J., & Legendre, G. (2008). Qu'en est-il des clitiques sujet en français oral contemporain? *Congrès Mondial de Linguistique Française-CMLF'08*.  
<http://www.linguistiquefrancaise.org/articles/cmlf/abs/2008/01/cmlf08308/cmlf08308.html> doi:DOI 10.1051/cmlf08308
- De Cat, C. (2005). French subject clitics are not agreement markers. *Lingua*, 115, 1195-1219.
- de Weck, G., Hassan, R., Heurdier, J., Klein, J., & Salagnac, N. (202X). Activities and institutional contexts: their role in the use of referring expressions. In A. Salazar Orvig, G. de Weck, R. Hassan & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogical approach* (pp. xxx-xxx). Amsterdam: John Benjamins.

- Demuth, K. (2001). A prosodic approach to filler syllables. *Journal of Child Language*, 28(1), 246-249.
- Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, 17(4), 463-489.
- Dressler, W. U., & Kilani-Schoch, M. (2001). On the possible rise and inevitable fall of fillers. *Journal of Child Language*, 28(1), 250-253.
- Du Bois, J. W. (2003). Discourse and grammar. In M. Tomasello (Ed.), *The new psychology of language. Cognitive and functional approaches to language structure* (Vol. 2, pp. 47-87). Mahwah, NJ: Lawrence Erlbaum Associates.
- Girouard, P. C., Ricard, M., & Gouin Decarie, T. (1997). The acquisition of personal pronouns in French-speaking and English-speaking children. *Journal of Child Language*, 24, 311-326.
- Givón, T. (Ed.) (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.
- Guerriero, A. M. S., Oshima-Takane, Y., & Kuriyama, Y. (2006). The development of referential choice in English and Japanese: a discourse-pragmatic perspective. *Journal of Child Language*, 33(4), 823-857.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274-307.

- Gundel, J. K., & Johnson, K. (2013). Children's use of referring expressions in spontaneous discourse: Implications for theory of mind development. *Journal of Pragmatics*, 56, 43-57.
- Hamann, C., Rizzi, L., & Frauenfelder, U. H. (1996). On the acquisition of subjects and object clitics in French. In H. Clahsen (Ed.), *Generative perspectives on language acquisition* (pp. 309-334). Amsterdam: John Benjamins.
- Hassan, R., de Weck, G., Rezzonico, S., Salazar Orvig, A., & Vinel, E. (202X). Variations in adult use of referring expressions during storytelling in different interactional settings. In A. Salazar Orvig, G. de Weck, R. Hassan & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogical approach* (pp. xxx-xxx). Amsterdam: John Benjamins.
- Hothorn, T., & Zeileis, A. (2015). Partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, 16, 3905-3909.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Assessing the Fit of the Model. In D. W. Hosmer, S. Lemeshow & R. X. Sturdivant, *Applied logistic regression* (pp. 153-226). New York, NY: John Wiley & Son. 10.1002/9781118548387.ch5

- Huang, C.-C. (2011). Referential choice in Mandarin child language: A discourse-pragmatic perspective. *Journal of Pragmatics*, 43(7), 2057-2080.
- Hugues, M. E., & Allen, S. E. M. (2006). A discourse-pragmatic analysis of subject omission in child English. In D. Bamman, D. T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 293-304). Somerville, MA: Cascadilla Press.
- Hughes, M. E., & Allen, S. E. M. (2015). The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. *Lingua*, 155, 43-61.
- Hyams, N. M. (1986). *Language acquisition and the theory of parameters*: Dordrecht: Reidel Publishing Company.
- Imbens-Bailey, A., & Pan, B. A. (1998). Pragmatics of self-and other-reference in young children. *Social Development*, 7(2), 219-233.
- Jakubowicz, C., & Rigaut, C. (1997). L'acquisition des clitiques nominatifs en français. In A. Zribi-Hertz (Ed.), *Les pronoms. Morphologie, syntaxe et typologie* (pp. 57-99). Saint-Denis: Presses Universitaires de Vincennes.
- Klein, J., Jullien, S., & Fox, G. (202X). Explorations in the relations between reference, syntactic constructions and prosody. In A. Salazar Orvig,

- G. de Weck, R. Hassan & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogic approach* (pp. xxx-xxx). Amsterdam: John Benjamins.
- Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Lyons, J. (1977). *Semantics* (Vol. 2). Cambridge: Cambridge University Press.
- Marcos, H., Salazar Orvig, A., da Silva-Genest, C., & Heurdier, J. (202X). The influence of dialogue in young children's uses of referring expressions. In A. Salazar Orvig, G. de Weck, R. Hassan & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogical approach* (pp. xxx-xxx). Amsterdam: John Benjamins.
- Morgenstern, A. (2006). *Un je en construction*. Paris: Ophrys.
- Narasimhan, B., Budwig, N., & Murty, L. (2005). Argument realization in Hindi caregiver-child discourse. *Journal of Pragmatics*, 37(4), 461-495.
- Nelson, K. (1989). Monologue as the linguistic construction of self in time. In K. Nelson (Ed.), *Narratives from the crib* (pp. 284-308). Cambridge, MA.: Harvard University Press.

- Rizzi, L. (1993/1994). Some notes on linguistic theory and language development: The case of root infinitives. *Language acquisition*, 3(4), 371-393.
- Rozendaal, M., & Baker, A. (2010). The acquisition of reference: Pragmatic aspects and the influence of language input. *Journal of Pragmatics*, 42(7), 1866-1879.
- Salazar Orvig, A., & da Silva, C. (2014). Demonstratives and personal pronouns in French: Concurrent factors for early contrasts. Paper presented at the 13th International Congress for the Study of Child Language, Amsterdam.
- Salazar Orvig, A., Fayolle, V., Hassan, R., Leber-Marin, J., Marcos, H., Morgenstern, A. (2004). Emergence des marqueurs anaphoriques avant 3 ans : le cas des pronoms de troisième personne. *Calap. La cohésion chez l'enfant*, 24, 57-82.
- Salazar Orvig, A., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marin, J., & Parès, J. (2010a). Dialogical beginnings of anaphora: the use of third person pronouns before the age of 3. *Journal of Pragmatics*, 42(7), 1842-1865.
- Salazar Orvig, A., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marin, J., & Parès, J. (2010b). Dialogical factors in toddler's use of clitic pronouns. *First Language*, 30(3-4), 374-402.



- Salazar Orvig, A., & Morgenstern, A. (2015). Acquisition and uses of pronouns in a dialogic perspective. In L. Serratrice & S. E. M. Allen (Eds.), *The acquisition of reference* (pp. 155–180). Amsterdam: John Benjamins.
- Serratrice, L., & Allen, S. E. M. (Eds.). (2015). *The acquisition of reference*. Amsterdam: John Benjamins.
- Skarabela, B., Allen, S. E. M., & Scott-Phillips, T. C. (2013). Joint attention helps explain why children omit new referents. *Journal of Pragmatics*, 56, 5-14.
- Theakston, A. L., Lieven, E., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, 28(1), 127-152.
- Veneziano, E. (2017). Noun and verb categories in acquisition: evidence from *fillers* and inflectional morphology in French-acquiring children. In V. Vapnarsky & E. Veneziano (Eds.), *Lexical polycategoriality: Cross-linguistic, cross-theoretical and language acquisition approaches* (pp. 381-411). Amsterdam: John Benjamins.
- Veneziano, E., & Clark, E. V. (2016). Early verb constructions in French: adjacency on the left edge. *Journal of Child Language*, 43(6), 1193-1230.

- Veneziano, E., & Sinclair, H. (2000). The changing status of 'filler syllables' on the way to grammatical morphemes. *Journal of Child Language*, 27, 461-500.
- Vinel, E., Salazar Orvig, A., de Weck, G., Nashawati, S., & Rahmati, S. (202X). The impact of speech genres on the use of referring expressions. In A. Salazar Orvig, G. de Weck, R. Hassan & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogical approach* (pp. xxx-xxx). Amsterdam: John Benjamins.
- Yamaguchi, N., Salazar Orvig, A., Le Mené, M., Caët, S., & Rialland, A. (202X). Filler syllables as precursors of referring expressions. In A. Salazar Orvig, G. de Weck, R. Hassan & A. Rialland (Eds.), *The acquisition of referring expressions: a dialogical approach* (pp. xxx-xxx). Amsterdam: John Benjamins.