

ARTICLE

# *Wh*-in-situ in child French: Deictic triggers at the syntax-semantics interface

Katerina Palasis , Richard Faure and Fanny Meunier

Université Côte d'Azur, CNRS, BCL, France

Email: [Katerina.Palasis@univ-cotedazur.fr](mailto:Katerina.Palasis@univ-cotedazur.fr)

(Received 01 July 2022; revised 13 January 2023; accepted 02 February 2023; first published online 17 March 2023)

## Abstract

This article investigates the development of *wh*-in-situ questions in French by examining a three-year kindergarten dataset of spontaneous productions with 16 children between 2;5 and 5;11. The distribution of the *wh*-phrases is statistically examined in relation to age, verb form (Fixed *be* form *c'est* 'it is' vs. Free *be* forms vs. Free lexical verbs), and grammatical category of the *wh*-word (Pronoun vs. Adverb). Results show that *wh*-in-situ remains prevalent throughout the period despite a steady increase in *wh*-ex-situ. Verb form (Fixed vs. All free forms) is a discriminating variable for the *wh*-position in all three years, and it interacts with the category of the *wh*-word. The Fixed *be* form *c'est* favours in-situ *wh*-pronouns (*c'est qui Taz ?*), whereas the Free forms favour *wh*-ex-situ questions, and massively co-occur with *wh*-adverbs (*combien ça coûte ?*). The emergence of the ex-situ *qu'est-ce que* 'what is it that', as opposed to the in-situ *quoi* 'what', is identified as a factor accounting for the gradual increase in *wh*-ex-situ. Finally, most outliers (*wh*-in-situ with Free forms) are shown to belong to the same paradigm as *c'est* in-situ questions: non-presuppositional questions, which are visible from the frequent use of *là* 'there', like *c'est*, a deictic item.

**Keywords:** *wh*-questions; French; acquisition; copula; lexical verbs; *quoi/que* alternation; deixis

## 1. INTRODUCTION

The position of *wh*-phrases in French *wh*-questions is the centre of much attention for at least two reasons. First, Colloquial French is usually described as a 'mixed' language, showing both in-situ and ex-situ *wh*-phrases in matrix information-seeking questions (e.g., Dryer, 2013) as compared to other languages that show either only *wh*-in-situ items (e.g., Chinese; Huang, 1982) or only *wh*-ex-situ items (e.g., English; Nguyen and Legendre, 2022).<sup>1</sup> Second, French-speaking children and adults differ in their distribution of *wh*-positions. It is usually

<sup>1</sup>We focus on matrix information-seeking questions, hence leaving aside echo-questions, probe-questions and embedded structures (but see Ledegen (this issue) on the latter).



reported that children utter more *wh*-in-situ questions than adults (Zuckerman and Hulk, 2001; Strik, 2007; Strik and Pérez-Leroux, 2011; Becker and Gotowski, 2015; Thiberge, 2020). This article aims to explain this early prevalence of the *wh*-in-situ position and the difference with adult productions. It will statistically investigate the production of matrix *wh*-questions in 16 typically developing French L1 children during their three kindergarten years. The investigation will pinpoint two factors at the syntax-semantics interface that account for the prevalence of the *wh*-in-situ position in child but not adult French, namely the Fixed *be* structure *c'est* 'it is' and the adverb *là* 'there'. An attempt at the unification of these in-situ question triggers under an umbrella category (deixis) will be made as well as an examination of the counter-examples. The data will also draw our attention to a third factor: the morpho-syntactic difference between in-situ *quoi* 'what' and ex-situ *qu'est-ce que* 'what is it that', and the gradual emergence of the latter throughout the three-year kindergarten period.

### 1.1 The target language

In this contribution, we will leave Standard French aside, because we assume that preschool children are mainly exposed to and hence initially acquire Colloquial French (e.g., Hamlaoui, 2011). Colloquial French is often but not always described as a mixed language. Baunaz (2011, 2016) compared 'Non-Standard Colloquial French' (NSCF) with Chinese and described both as *wh*-in-situ languages with covert *wh*-movement. The author also proposed that NSCF displays three types of in-situ *wh*-phrases: one non-marked and two marked types (Types 1, 2 and 3, respectively). Type 1 in-situ *wh*-phrases are analysed as covert Split-DPs (i.e., variables with a phonologically null interrogative operator). They are non-presuppositional and have a rising intonation. Types 2 and 3 in-situ *wh*-phrases have specific and partitive interpretations, and fall-rise and falling intonations, respectively. Both entail covert movement of the entire *wh*-phrase. Furthermore, Baunaz (2011: 75) described in-situ and ex-situ structures in NSCF as syntactically 'similar', with identical processing loads (following Adli, 2006), and she related the difference in overt versus covert movement to 'interpretation, as witnessed by intonation'. We will elaborate on Baunaz's (2011, 2016) proposal in Section 4.

Faure and Palasis (2021) also described 'Colloquial French' as a *wh*-in-situ language with covert *wh*-movement; however, contrary to Baunaz (2011), they suggested that covert and overt movement are different. They showed that movement is less restricted in *wh*-ex-situ questions, which is evidenced by lower sensitivity to weak islands (e.g., negation; see Examples 1a vs. 1b) and Superiority (i.e., order of arguments in multiple questions; see Example 1c, its alternate c' vs. Example 1d and the answers they allow). The authors thus attributed overt movement of the *wh*-phrases to the presence of a different trigger, namely Exclusivity. A *wh*-question endowed with this feature presupposes that not all of the possible answers, say from a set of three, are correct but that two at most are (see Examples 1e vs. 1f).

- (1) *Wh*-in-situ and *wh*-ex-situ structures differ in adult data (Faure and Palasis, 2021: 64–74):  
*Context: Julien runs a stand at a car boot sale. Nadia checks out his sales at the end of the day:*
- a. \**T(u) as pas vendu quoi* ?<sup>2</sup> (no stress on *quoi*, sensitivity to weak islandhood)  
 you have not sold what  
 ‘What haven’t you sold?’
  - b. *Qu’est-ce que t(u) as pas vendu* ? (no sensitivity to weak islandhood)  
 what is it that you have not sold  
 ‘What haven’t you sold?’
  - c. *Qu’est-ce que t(u) as vendu à qui* ? (answer: *a book to Sam, a pen to John*)  
 what is it that you have sold to whom  
 ‘What have you sold to whom?’
  - c’. *À qui t(u) as vendu quoi* ? (answer: *to Sam a book, to John a pen*)  
 to whom you have sold what  
 ‘What have you sold to whom?’
  - d. *T(u) as vendu quoi à qui* ? (no stress on *à qui*, answer: *a book to Sam, a pen to John, #to Sam a book, to John a pen*)  
 you have sold what to whom  
 ‘What have you sold to whom?’  
*Context: Marie enquires which friend(s) among a group of three Jane saw last week-end:*
  - e. *T(u) as vu qui* ? (in-situ *qui* = anyone and all of them)  
 you have seen who  
 ‘Who did you see?’
  - f. *Qui t(u) as vu* ? (ex-situ *qui* = at least one person from the presupposed set cannot be the answer)<sup>3</sup>  
 who you have seen  
 ‘Who did you see?’

Baunaz’s (2011, 2016) and Faure and Palasis’ (2021) proposals both hinge upon the syntax-semantics interface and account for adult Colloquial French.<sup>4</sup> This language will be considered here as the target language for children.

## 1.2 The child system

Research has shown that French preschoolers usually produce more in-situ *wh*-questions than adults (Zuckerman and Hulk, 2001; Strik, 2007; Strik and Pérez-Leroux, 2011; Becker and Gotowski, 2015; Thiberge, 2020). Young children nevertheless also produce ex-situ *wh*-items from the onset around the

<sup>2</sup>Incomplete words are transcribed with the omitted material in parentheses.

<sup>3</sup>An anonymous reviewer wonders whether the answer ‘everyone’ is totally excluded, as predicted. According to our informants, it is not but it has a corrective flavour.

<sup>4</sup>Hypotheses on Chinese and English also assume that the default *wh*-position only is related to the *wh*-feature and that the other *wh*-position is triggered by another feature (i.e., contrastive focus for Chinese ex-situ and presupposition for English in-situ; see Cheung, 2008 and Biezma, 2020, respectively).

age of 2;0 (Crisma, 1992; Déprez and Pierce, 1993; Hamann, 2006; Strik, 2007; Prévost, 2009), and the in-situ rates vary widely from one child to another. A small review of children between 1;8 and 2;9 (Palasis, Faure and Lavigne, 2019) reported in-situ rates ranging from 1.3% (Philippe at 2;3 in Crisma, 1992) to 94.4% (Augustin at 2;04 in Rasetti, 2003) in naturalistic speech, which seems to display more in-situ *wh*-items than experimental contexts (Zuckerman and Hulk, 2001). The in-situ ratio also seems to be different with pronouns (e.g., *quoi* 'what') compared to adverbs (e.g., *où* 'where'), and there are more in-situ *wh*-phrases with the former than with the latter (Strik, 2007; Jakubowicz, 2011; Strik and Pérez-Leroux, 2011).

The discrepancy between the ratios of in-situ and ex-situ *wh*-phrases in adult and child speech has often been ascribed to differential abilities between adults and children. Generativists posit that in-situ *wh*-questions converge by checking a *wh*-feature covertly in Logical Form, whereas ex-situ *wh*-questions converge by checking the *wh*-feature in LF and overt syntax (Bayer and Cheng, 2017 for a review). The latter is described as cognitively more costly, and derivational economy is assumed in children compared to adults (Hamann, 2006; Jakubowicz, 2011; Strik, 2012). Differing pragmatics have also been evoked in the literature, such as overuse of common ground in children entailing more *wh*-in-situ questions (Gotowski and Becker, 2016). However, hypotheses based on differential abilities make the prediction that all children, whatever their language, should initially produce more in-situ *wh*-phrases, which is not the case in target *wh*-ex-situ languages (e.g., English, Dutch, Italian, German, Swedish; Stromswold, 1995; Van Kampen, 1997; Guasti, 2000; Roeper and De Villiers, 2011; Strik, 2012).

Resorting to another type of difference between adults and children, Palasis *et al.* (2019) showed that Verb form (i.e., Fixed *be* form *c'est* 'it is'<sup>5</sup> vs. Free *be* forms and Lexical verbs) was a discriminating variable with regard to *wh*-position, and that there was a correlation between the *wh*-in-situ position and the Fixed *be* form in their child data (see Examples 2a vs. 2b and 2c). The difference between adults and children was then ascribed to different stages of the diversification of the verbal system with more *c'est* forms in children than in adults (following Guillaume, 1927; Clark, 1978; Bassano, Eme and Champaud, 2005). The hypothesis also accounted for the well-documented asymmetry between adults and children with regard to *wh*-pronouns and *wh*-adverbs: Children utter more in-situ *wh*-pronouns than adults (e.g., *quoi* 'what', *qui* 'who'), whereas both groups favour the ex-situ position for *wh*-adverbs (e.g., *où* 'where', *comment* 'how'; Zuckerman and Hulk, 2001; Strik, 2007; Strik and Pérez-Leroux, 2011; Becker and Gotowski, 2015). Palasis *et al.* (2019) reported that most pronouns were uttered with the Fixed *be* form *c'est* 'it is' that favours in-situ (see 2a); whereas most adverbs were uttered with the Free verb forms that favour the ex-situ position (see 2b and 2c).

<sup>5</sup>This *c'est* is not to be confused with the *c'est* that is used in clefts (*c'est X que* 'it is X that'). Fixed *be c'est* is a variant of the copula that appears in non-predicational copular sentences that posit an equation between two terms: *Bob c'est mon meilleur ami* 'Bob is my best friend' (Higgins, 1973; Roy, 2013 for French).

(2) Correlation between the Fixed form *c'est* 'it is' and *wh*-in-situ in child data (Palasis et al., 2019):

- a. *C'est quoi ça ?* (LI, 3;4)  
it is what that  
'What is that?'
- b. *Où elle est ta maison ?* (MT, 3;11)  
where she is your house  
'Where is your house?'
- c. *Comment on joue à ce jeu ?* (LU, 3;4)  
how one plays at this game  
'How do we play this game?'

In this study, we will examine a new dataset for 16 children in their third kindergarten year and connect these new data with the previously examined data for the same children in their first two kindergarten years (Palasis et al., 2019), hence providing a new three-year longitudinal investigation into the *wh*-questions for 16 children. More specifically, we will examine the distribution of *wh*-in-situ questions in Year 3 and compare our results with Years 1 and 2 in order to shed light on new developmental aspects of the position of *wh*-words in preschoolers. Section 2 will describe the entire dataset and the predictions we make building on previous hypotheses. Section 3 will statistically examine the data as a function of the *wh*-position (in-situ vs. ex-situ), the verb form (fixed vs. free), and the grammatical category of the *wh*-phrase (pronoun vs. adverb). Section 4 will examine the counter-examples to the correlation between the Fixed *be* form *c'est* 'it is' and the *wh*-in-situ position, provide a comparison with the target language described in Section 1.1, and elaborate on deixis. We will end with concluding remarks and future research in Section 5.

## 2. METHOD

This research will statistically investigate the production of matrix *wh*-questions in typically developing preschool French L1 children. In Section 2, we will describe the initial kindergarten dataset (2.1), the participants (2.2), the data cleaning process leading to the final dataset (2.3), and the predictions we make building on previous hypotheses (2.4).

### 2.1 The initial kindergarten dataset

The data were collected with a class of preschoolers during their three kindergarten years in the South of France (total  $n = 41,845$  utterances; see Table 1). Small groups of two to four children were gathered in a quiet room close to their usual classroom and were encouraged by the researcher to narrate their activities in and out of school, 'read' books and play games. There were child-child and child-researcher interactions, and the latter were kept as natural as possible. The children were video- and audio-recorded during 20 to 25-minute sessions, 13 times during their first year, and 10 times during their second and third kindergarten years

**Table 1.** The kindergarten corpus

Specifications	Year 1	Year 2	New Year 3 (5/10 sessions)	Total (provisional)
Dates	2006/2007	2007/2008	2008/2009	
Hours	25	20	20 (10)	
Sessions	13	10	10 (5)	
Recording types	Audio & video	Audio & video	Audio & video	
Adults	3	1	1	
Adult utterances	12,891	9,291	4,947	27,129
Children (all/ <b>L1 French</b> )	20/17	19/17	18/ <b>16</b>	
Age range	2;5-4;0	3;6-4;11	4;5-5;11	
Child utterances (all/ <b>16</b> )	15,992/13,320	14,348/12,242	11,505/ <b>10,501</b>	41,845/36,063

(which amounted to an average interval between two sessions of three to four weeks, except for holidays). The first two years were examined with regard to *wh*-questions in Palasis *et al.* (2019). This article will focus on the *wh*-in-situ questions in the newly processed data for the third year (5 sessions out of 10 are available for analysis, transcribed and coded with the CHILDES tools; MacWhinney, 2000) and on a new longitudinal analysis of the three years.

## 2.2 Participants

The participants were the native speakers of Hexagonal French in the kindergarten class described in Section 2.1. All the children in the class were asked to participate ( $n = 20$ ), but four children were excluded in the final dataset: two children did not have a native French background at home, and two children moved during the period. The set thus comprises the data for 16 children (9 female) during a three-year period ( $n = 36,063$  utterances). Figure 1 shows the distribution of the data according to each child over the three years. The children were between 2;5 (MS: youngest child at the beginning of Year 1) and 5;11 (KE: oldest child by the end of Year 3).

## 2.3 The final dataset

The children produced six types of questions that distributed in frequent and infrequent types (see Examples 3 and Figure 2). Finite matrix *wh*-questions (initial  $n = 1,084$ ) represented one of the frequent types, with an average of 34.5% of the children's questions during the three-year period.

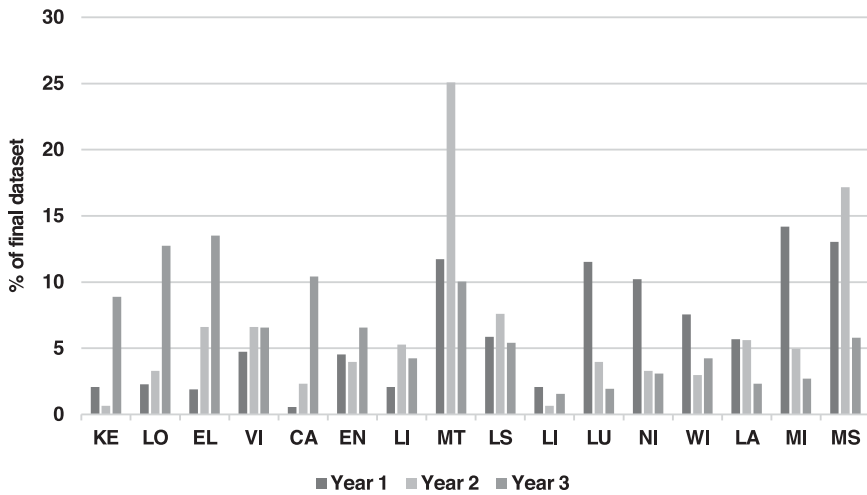


Figure 1. Participation per year per child.

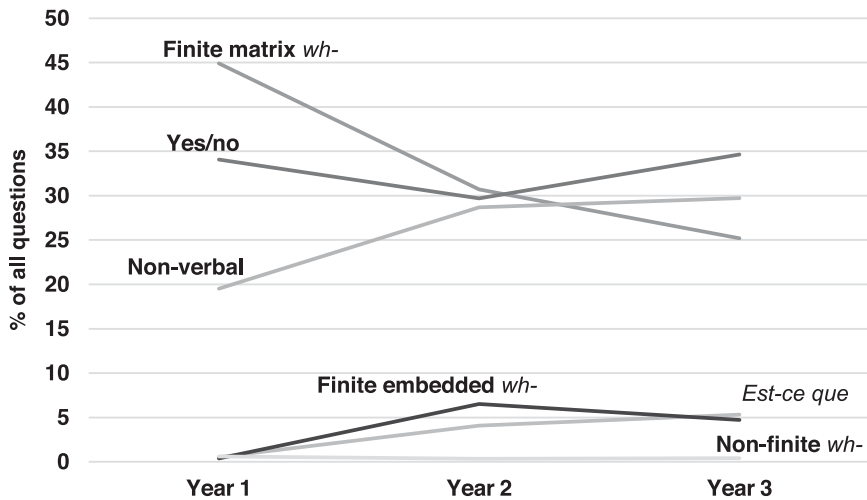


Figure 2. Distribution of question types.

(3) Six types of questions in the dataset:

- a. Finite matrix *wh*- (34.5%):  
*Où il est Kelian ?* (MS, 4;7)  
 where he is Kelian  
 'Where is Kelian?'
- b. Yes/no (32.8%):  
*On peut lire des livres ?* (MT, 4;11)  
 one can read some books  
 'Can we read books?'

- c. Non-verbal (25.5%):  
*Et le blanc ?* (EL, 5;4)  
 and the white  
 'And (what about) the white one?'
- d. *Est-ce que* (3.7%):  
*Est-ce que t(u) as encore un tablier ?* (MS, 5;0)  
 is it that you have still an apron  
 'Do you still have an apron?'
- e. Finite embedded *wh*- (3.1%):  
*Je sais pas comment i(ls) s'appellent.* (LU, 4;9)  
 I know not how they <sup>REFL</sup> call  
 'I don't know what their names are'
- f. Non-finite *wh*- (0.4%):  
*À construire quoi ?* (VI, 5;3)  
 at build what  
 '(they help) To build what?'

The longitudinal dataset in Figure 2 shows that the ratio of finite matrix *wh*-questions decreases during the period in favour of yes/no and non-verbal questions. To the contrary, the ratio of children's utterances compared to adults' utterances in the entire dataset increases during the period (Year 1: 50.8%, Year 2: 56.9%, Year 3: 68%). The progression shows that there are more child-child and fewer child-adult interactions in Year 3 than in Year 1. We hypothesize that the decrease in child *wh*-questions pertains to this change in addressee, children using more non-verbal and yes/no questions than matrix *wh*-questions when addressing another child than when addressing an adult.

Table 2 shows the initial dataset of 1,084 matrix *wh*-questions distributed according to *wh*-word (*quoi/que* 'what', *où* 'where', *qui* 'who', *pourquoi* 'why', *comment* 'how', *à qui* 'to whom', *à quoi* 'to what', *quand* 'when', *quel* 'which', *lequel* 'which one', *combien* 'how many') and according to syntactic structure. The syntactic structures stem from the Verb form hypothesis that suggested that verb form (i.e., Fixed *be* form *c'est* 'it is' vs. Free *be* and lexical forms) is a discriminating variable for *wh*-position in child data (Palasis *et al.*, 2019). Examples (4) illustrate the 11 different structures as a function of the position of the *wh*-item and the verb form in the dataset. Structures 1a, 1b, 4 and 6 display in-situ *wh*-questions and Structures 2a, 2b, 3a, 3b, 5, 7 and 8 display ex-situ *wh*-questions.

(4) Structures according to *wh*-position and verb form:

- a. Structure 1a (S Vlex *wh*-):  
*Je fais quoi ?* (MT, 4;11)  
 I do what  
 'What do I do?'
- b. Structure 1b (S Vbe *wh*-):  
*Je suis où ?* (LO, 5;9)  
 I am where  
 'Where am I?'



**Table 2.** Distribution of *wh*-questions according to *wh*-item and syntactic structure (initial 3-year dataset)

*	<i>quoi/</i>		<i>où</i>	<i>qui</i>	<i>pourquoi</i>	<i>comment</i>	<i>à</i>		<i>quand</i>	<i>quel</i>	<i>lequel</i>	<i>combien</i>	Total
	<i>que</i>						<i>qui</i>	<i>quoi</i>					
1a	80	14	12	0		8	0	2	0	6	2	3	127
1b	0	59	0	0		1	1	0	0	1	0	0	62
2a	0	8	45	94		62	1	0	0	2	2	1	215
2b	0	91	3	7		0	1	0	0	0	0	0	102
3a	1	0	1	0		1	0	0	0	0	0	0	3
3b	0	8	0	0		0	0	0	0	0	0	0	8
4	206	1	34	0		1	56	1	1	1	9	0	310
5	12	0	1	10		0	5	0	0	0	0	0	28
6	0	2	55	0		0	0	0	18	0	0	0	75
7	2	0	37	0		0	0	0	2	0	0	0	41
8	102	0	10	0		0	0	0	1	0	0	0	113
	403	183	198	111		73	64	3	22	10	13	4	1084

\*Structures: 1a: S Vlex *wh*-, 1b: S Vbe *wh*-, 2a: *wh*- S Vlex, 2b: *wh*- S Vbe, 3a: *wh*- Vlex S, 3b: *wh*- Vbe S, 4: *c'est wh*-, 5: *wh-c'est*, 6: *c'est wh- que* S V, 7: *wh- c'est que* S V, 8: *wh-est-ce que* S V.

- c. Structure 2a (*wh*- S Vlex):  
*Combien ça coûte ?* (LA, 5;2)  
how much that costs  
'How much does that cost?'
- d. Structure 2b (*wh*- S Vbe):  
*Où il est Kelian ?* (MS, 4;7)  
where he is Kelian  
'Where is Kelian?'
- e. Structure 3a (*wh*- Vlex S):  
*Que se passe-t-il ?* (VI, 3;5)  
what REFL happens it  
'What is happening?'
- f. Structure 3b (*wh*- Vbe S):  
*Elmer où es-tu ?* (LS, 4;3)  
Elmer where are you  
'Elmer where are you?'
- g. Structure 4 (*c'est wh*-):  
*C'est qui Taz ?* (LS, 5;1)  
it is who Taz  
'Who is Taz?'
- h. Structure 5 (*wh- c'est*):  
*À qui c'est ?* (LI, 5;1)  
to whom it is  
'To whom does it belong?'

- i. Structure 6 (*c'est wh- que S V*):  
*Mais c'est quand que tu montres la maison ?* (MT, 3;11)  
 but it is when that you show the house  
 'But when is it that you show the house?'
- j. Structure 7 (*wh- c'est que S V*):  
*Mais quand c'est qu'on va à l'école ?* (MS, 3;9)  
 but when it is that one goes to the school  
 'But when is it that we go to school?'
- k. Structure 8 (*wh- est-ce que S V*):  
*Qu'est-c(e) qu'il a dit ?* (VI, 5;4)  
 what is it that he has said  
 'What did he say?'

Previous research has put forward that some *wh*-words, contexts and syntactic structures favour either ex-situ or in-situ *wh*-questions. In order to investigate the factors that possibly interact with the position of *wh*-items, we needed to examine only the *wh*-items, contexts and syntactic structures that allow the alternative in-situ/ex-situ *wh*-positions. The following items ( $n = 366$ ; mainly in shaded cells in Table 3) were hence discarded from the initial 1,084 set:

- The *wh*-item *pourquoi* 'why' ( $n = 111$ ) appears only ex-situ. This *wh*-word is assumed to be generated in the left periphery, higher than the other *wh*-words (Rizzi, 2001; Hamann, 2006; Myers and Pellet, 2014).
- The *wh*-item *qui* 'who' always appears to the left of the finite verb when it is a subject ( $n = 58$ ).
- Cleft structures (Structures 6 and 7;  $n = 116$ ), because we do not address the specific matter of *wh*-fronting in clefts in this article (but see Oiry, 2011).
- The D-linked *wh*-items (mainly *quel* 'which', *lequel* 'which one', and variants thereof;  $n = 59$ , 23 *quel/lequel* + 36 with other *wh*-words), because of the possible interaction between D-linking and the in-situ position in French (Coveney, 1989; Obenauer, 1994; Chang, 1997; Boeckx, 2000; Cheng and Rooryck, 2000).
- Questions with subject-verb inversion (Structures 3a and 3b;  $n = 11$ ), because they allow ex-situ *wh*-items only, and these questions belong to Standard French, which we do not examine in this article.
- Questions with any other characteristic of Standard French as opposed to Colloquial French (i.e., non-elided nominative clitics before consonants and discontinuous negation; Palasis, 2013): only one *wh*-question was discarded due to full *il* in front of a consonant (i.e., *et il [/] il voit quoi ?* 'and what does he see').
- Immediate (self-)repetitions ( $n = 10$ ).

Table 3 shows the final dataset with 718 *wh*-questions produced by 16 children during their three kindergarten years. Year 3 ( $n = 158$ ) is examined for the first time and is added to Year 1 ( $n = 368$ ) and Year 2 ( $n = 192$ ) that have been investigated in previous work.

**Table 3.** Distribution of *wh*-questions according to *wh*-item and syntactic structure (final 3-year dataset)

*	<i>quoi/</i> <i>que</i>	<i>où</i>	<i>qui</i>	<i>pourquoi</i>	<i>comment</i>	<i>à</i> <i>qui</i>	<i>à</i> <i>quoi</i>	<i>quand</i>	<i>quel</i>	<i>lequel</i>	<i>combien</i>	Total
1a	69	14	12	0	7	0	1	0	0	0	2	105
1b	0	57	0	0	1	1	0	0	0	0	0	59
2a	0	7	0	0	62	0	0	0	0	0	1	70
2b	0	86	0	0	0	0	0	0	0	0	0	86
3a	0	0	0	0	0	0	0	0	0	0	0	0
3b	0	0	0	0	0	0	0	0	0	0	0	0
4	200	1	34	0	1	48	1	1	0	0	0	286
5	12	0	1	0	0	5	0	0	0	0	0	18
6	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8	93	0	0	0	0	0	0	1	0	0	0	94
	374	165	47	0	71	54	2	2	0	0	3	718

\*Structures: 1a: S Vlex *wh*-, 1b: S Vbe *wh*-, 2a: *wh*- S Vlex, 2b: *wh*- S Vbe, 3a: *wh*- Vlex S, 3b: *wh*- Vbe S, 4: *c'est wh*-, 5: *wh-c'est*, 6: *c'est wh- que* S V, 7: *wh- c'est que* S V, 8: *wh-est-ce que* S V.

## 2.4 Predictions

Building on previous hypotheses detailed hereunder, we make the following predictions:

- i. Because there are no major changes in the grammar of the children between Years 1, 2 and 3 (as established in Section 2.3 with the near absence of markers of Standard French in Year 3), we predict that there should be no significant difference in the distribution of the in-situ versus ex-situ rates between Year 3 and the previously studied Years 1 and 2. We will address this matter in Section 3.1.
- ii. The Verb form hypothesis (Palasis et al., 2019) makes the prediction that the distribution of in-situ and ex-situ *wh*-questions in Year 3 of the dataset should be related to Verb form (Fixed vs. Free forms). We will test this prediction in Section 3.2.
- iii. Child datasets usually report an asymmetry in the position of the *wh*-items according to age (children vs. adults) and to grammatical category of the *wh*-item (pronouns vs. adverbs; Zuckerman and Hulk, 2001; Strik, 2007; Jakubowicz, 2011; Becker and Gotowski, 2015). Two hypotheses exist on this asymmetry in child questions: a relationship between the position of objects in the verbal phrase and the in-situ position of the corresponding *wh*-pronouns, and a relationship between the verb form in the sentence and the position of the *wh*-word whatever its grammatical category (Strik, 2007 and Palasis et al., 2019, respectively). The first hypothesis makes the prediction of a clear-cut asymmetry according to grammatical category

(i.e., in-situ pronouns vs. ex-situ adverbs); the second hypothesis predicts that *wh*-pronouns and *wh*-adverbs can appear in-situ or ex-situ depending on the verb form. We will test both predictions in Section 3.3 and ponder their respective developmental scopes.

### 3. RESULTS AND PRELIMINARY DISCUSSION

The final dataset in Table 3 ( $n = 718$  *wh*-questions) was submitted to the Cochran-Mantel-Haenszel Test (henceforth CMH; Cochran, 1954; Mantel and Haenszel, 1959) for repeated 2x2 tests of independence. We investigated the relationship between the verb form (Fixed *be* form *c'est* 'it is' vs. Free *be* forms vs. Lexical forms) and the position of the *wh*-item (In-situ vs. Ex-situ) in each utterance. This was done by specifically programming an Excel spreadsheet. The data were submitted in as many 2x2 tables as there were participants with available data for each test in order to account for the quantitative and qualitative differences between participants (McDonald, 2011; see Figure 1 for the distribution per child). A significant result at the  $X^2$  CMH test means that the variables are dependent (one has an effect on the other). In the following subsections, we will test the predictions stated in Section 2.4 on the evolution of *wh*-in-situ compared to *wh*-ex-situ (3.1), on the impact of the different verb forms (3.2), and on the asymmetry between *wh*-pronouns and *wh*-adverbs (3.3).

#### 3.1 Distribution of the *wh*-questions according to the position of the *wh*-item (in-situ vs. ex-situ)

We examined the distribution of the *wh*-questions in Year 3 and longitudinally from Year 1 to Year 3 in order to describe the overall progression of the in-situ position in the dataset. The filtering of the data (described in Section 2.3) pointed towards an absence of radical change between Years 1 and 2 taken together and Year 3 in terms of grammatical characteristics (i.e., nominative clitics and negation). We therefore predicted no significant difference of distribution between both *wh*-positions in Year 3 compared to the first two years.

Figure 3 shows that *wh*-in-situ questions remain prevalent in Year 3 (55.7%), as they were in Year 1 and Year 2 (67.4% and 59.4%). A repeated-measures ANOVA including Year (1, 2, 3) and Position (ex-situ, in-situ) as factors confirmed that there are overall significantly more in-situ questions than ex-situ ones ( $F(1,15) = 26.649$ ;  $p < .001$ ; eta squared = .640) and no significant differences in the distribution between the three years ( $F(2,30) = 1$ ; n.s.). Prediction (i) is thus borne out. Nevertheless, Figure 3 also shows a steady progression in the *wh*-ex-situ position during the period. Specific post-hoc comparisons using Bonferroni corrections showed that there are more in-situ questions than ex-situ ones in Year 1 ( $t(15) = -4.418$ ,  $p < .01$ ), while their proportions become equivalent in Years 2 and 3 (respectively:  $t(15) = -1.575$ , n.s.;  $t(15) = -1.826$ , n.s.).

We also considered the *wh*-words individually in order to uncover possible specific developmental patterns.<sup>6</sup> There were no significant differences in the in-

<sup>6</sup>We thank an anonymous reviewer for this very interesting suggestion.

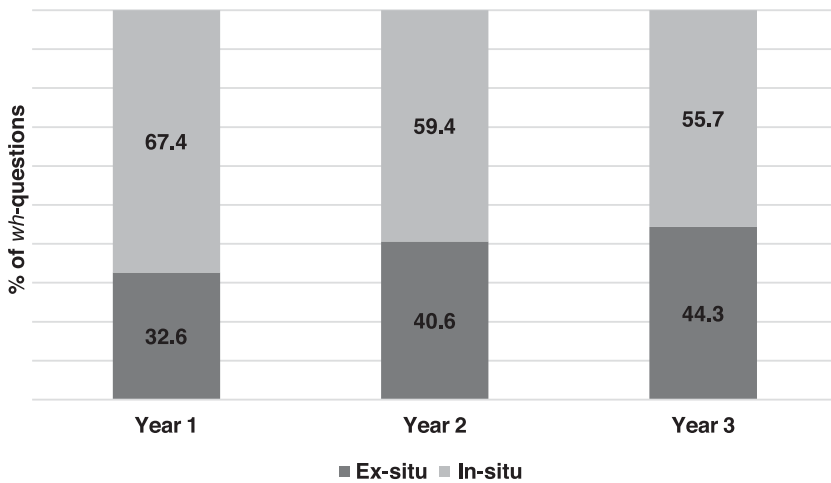


Figure 3. Longitudinal distribution of matrix *wh*-questions according to the position of the *wh*-item.

situ/ex-situ distributions between the three years, except for the *wh*-word *quoi/que* 'what' that showed a significant decrease in the in-situ position between Year 2 and Year 3 (73.5%–59.0%; Fisher's Exact Test for Count Data:  $p < .05$ ). In Section 3.2, we will investigate the detail of the in-situ/ex-situ distribution according to the verb form, and we will elaborate on the significant decrease in the in-situ *quoi* 'what' at the end of Section 3.3.

### 3.2 Distribution of the *wh*-questions according to the verb form (Fixed *be* form vs. Free *be* forms vs. Lexical verbs)

We examined the distribution of the *wh*-questions according to the verb form (Fixed *be* form vs. Free *be* forms vs. Lexical verbs) in Year 3 (Table 6) and longitudinally with Years 1 and 2 (Tables 4 and 5). We repeated the series of analyses carried out on the first two years on the third year: All *be* forms vs. Lexical verbs, Free *be* forms vs. Fixed *be* form, Free *be* forms vs. Lexical verbs, Fixed *be* form vs. Lexical verbs, and Fixed *be* form vs. All Free forms (Free *be* and Lexical).

The initial analyses of the first two years opposing All *be* forms to Lexical verbs (following the binary Lexical hypothesis on adult data; Coveney, 1995) had shown that the position of the *wh*-word in child data was also related to the type of verb, linking the in-situ position with *be* forms (without distinguishing Free *be* forms and Fixed *be* form) and the ex-situ position with Lexical verbs (Palasis et al., 2019: 225). In order to test if this is still the case in Year 3, we performed, as it had been done in Years 1 and 2, a CMH Test for repeated 2x2 tests of independence. The analysis of the independence between *wh*-position and verb type –that is Lexical Verbs vs. All *be* forms (free and fixed) – gave a  $X^2$  CMH = 31.28, 1 *df*,  $p < .0001$ . Again, in the third year, the in-situ *wh*-position is significantly more frequent with *be* forms and the ex-situ *wh*-position is significantly more frequent with Lexical Verbs. Cramér's

**Table 4.** Distribution of *wh*-questions according to verb form and *wh*-position (Year 1)

	Ex-situ		In-situ		Total	
	n	%	n	%	n	%
Free <i>be</i>	58	60.4	38	39.6	96	26.1
Fixed <i>be</i>	14	8.0	162	92.0	176	47.8
Lexical verbs	48	50.0	48	50.0	96	26.1
Total	120	32.6	248	67.4	368	100.0

**Table 5.** Distribution of *wh*-questions according to verb form and *wh*-position (Year 2)

	Ex-situ		In-situ		Total	
	n	%	n	%	n	%
Free <i>be</i>	20	62.5	12	37.5	32	16.7
Fixed <i>be</i>	2	2.7	71	97.3	73	38.0
Lexical verbs	56	64.4	31	35.6	87	45.3
Total	78	40.6	114	59.4	192	100.0

V, which determines the degree of association between variables, indicates a large effect size ( $V = 0.44$ ; Cramér, 1946).

To refine this result in Year 3, we further tested the impact of each *be* form on the position of the *wh*-word separately. Since one child (NI) had no data for this test, data were submitted in only 15 2x2 tables. Analyzing the relationship between the *wh*-position and the Free *be* forms vs. the Fixed *be* form using CMH, the result was  $X^2 \text{ CMH} = 12.97$ , 1 *df*,  $p < .001$ . This showed that, as it was the case in the first two years, the *wh*-position is significantly more likely to occur in-situ with the Fixed *be* form and ex-situ with the Free *be* forms. Cramér's V indicates a large effect size ( $V = 0.533$ ). This result allows us to separate the *be* forms into two significantly different categories, hence distinguishing Free *be* forms and the Fixed *be* form *c'est* 'it is', as previously observed in Years 1 and 2.

Bearing in mind that Lexical verbs are also free verb forms, we then examined Lexical Verbs in relation to Free *be* forms. The result was not significant ( $X^2 \text{ CMH} = 0.135$ , 1 *df*,  $p = .7$ ). This result shows that, as in the first two years, the *wh*-position is not significantly different when comparing Free *be* forms and Lexical Verbs, which are also free forms. Therefore, we confirm that the type of free verb (Lexical vs. *be*) is not a discriminating variable for the position of the *wh*-words in neither year of the dataset.

To test if the form of the verb correlates with the position of the *wh*-word, we looked at the impact of the Fixed *be* form on the position of the *wh*-word and compared the Fixed *be* form with Lexical Verbs. The result was significant ( $X^2 \text{ CMH} = 42.74$ , 1 *df*,  $p < .0001$ ). As in the first two years, the *wh*-position is significantly more in-situ with the Fixed *be* form and more ex-situ with Lexical verbs. Cramér's V indicates a large effect size ( $V = 0.649$ ).

**Table 6.** Distribution of *wh*-questions according to verb form and *wh*-position (Year 3)

	Ex-situ		In-situ		Total	
	n	%	n	%	n	%
Free <i>be</i>	8	47.1	9	52.9	17	10.8
Fixed <i>be</i>	2	3.6	53	96.4	55	34.8
Lexical verbs	60	69.8	26	30.2	86	54.4
Total	70	44.3	88	55.7	158	100.0

Finally, to confirm the importance of the Verb form (Fixed vs. Free) as a discriminating variable with respect to *wh*-position, we looked at the impact of the Fixed *be* form on the position of the *wh*-words by testing the Fixed *be* form against All Free forms, regardless of verb type. The result was significant ( $X^2$  CMH = 46.516, 1 *df*,  $p < .0001$ ) as it was for the first two years. The *wh*-position is significantly more in-situ with the Fixed *be* form and more ex-situ with All Free forms (Free *be* forms and Lexical verbs). Cramér's *V* indicates a large effect size ( $V = 0.598$ ).

Altogether, this pattern of results is in line with what was observed for the first two years. In the third year too, the Verb form (Fixed vs. Free) is a discriminating variable for the *wh*-position, regardless of Verb type (All *be* vs. Lexical verbs), and the effect size is larger when investigating Verb form (Fixed vs. Free:  $V = 0.598$ ) compared to Verb type (All *be* vs. Lexical verbs:  $V = 0.44$ ). Finally, the different effect sizes between the three categories (Fixed *be* vs. Free *be*:  $V = 0.533$  and Fixed *be* vs. Lexical:  $V = 0.649$ ) illustrate the greater syntactic and semantic differences between Lexical verbs and the Fixed *be* form, which differ with regard to two features (i.e., lexicality and fixity), than between Fixed and Free *be* forms, which differ in one feature only (i.e., fixity).

The tests undertaken in Section 3.2 on Year 3 of the dataset allow us to confirm the Verb form hypothesis: the in-situ position of the *wh*-phrases is related to the Fixed *be* form *c'est* 'it is' against all Free forms. Prediction (ii) is thus borne out. Though significant, this link does not hold for the entire dataset. We will elaborate on the interface between syntax (the position of the *wh*-phrase) and semantics (the question type) in Section 4, when we discuss the outliers to the in-situ/Fixed *be* correlation (i.e., the *wh*-in-situ questions with Free verb forms).

### 3.3 Distribution of the *wh*-questions according to the grammatical category of the *wh*-word (pronouns vs. adverbs)

In Section 2.4 (Prediction iii), we reported that children usually utter more in-situ *wh*-pronouns than adults, whereas both groups favour the ex-situ position for *wh*-adverbs (Zuckerman and Hulk, 2001; Strik, 2007; Strik and Pérez-Leroux, 2011; Becker and Gotowski, 2015). We mentioned two hypotheses on the prevalence of in-situ *wh*-pronouns in children compared to adults. Strik (2007) suggested an effect of the embedded position of arguments in the verbal phrase compared

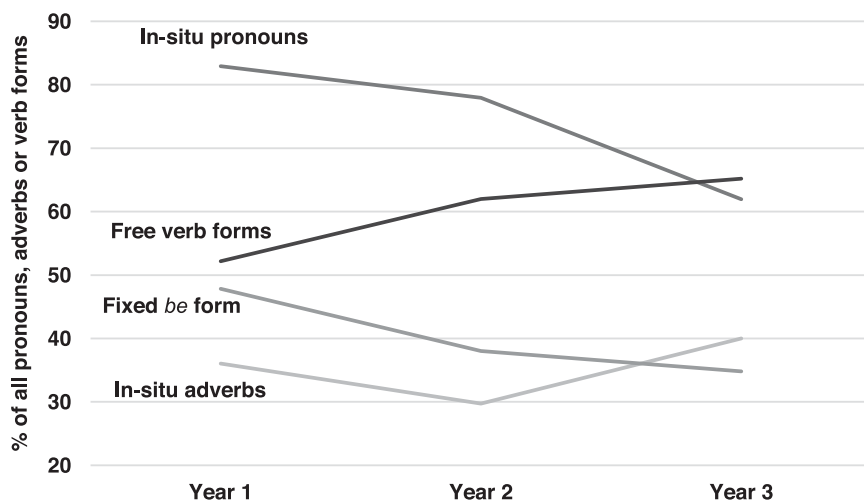


Figure 4. Evolution of ratios for *wh*-in-situ pronouns/adverbs and Fixed *be*/free verb forms.

to adjuncts that are merged higher in the structure. The VP internal position then renders pronouns more costly to move than adverbs. Palasis *et al.* (2019) reported a possible effect of the verb form on the position of the *wh*-phrase whatever the grammatical category of the *wh*-word. The ratio of the *wh*-in-situ phrases (pronominal and adverbial) is then expected to follow the ratio of the Fixed *be* form *c'est* 'it is', because the latter favours *wh*-in-situ questions. The first hypothesis makes the prediction that the ratio of *wh*-in-situ pronouns should decrease in favour of more *wh*-ex-situ pronouns as a function of the children's cognitive development. The hypothesis does not make any particular predictions on the evolution of the position of *wh*-adverbs. The second hypothesis makes the prediction that the ratio of *wh*-in-situ altogether (pronouns and adverbs) should decrease if the Fixed *be* form *c'est* 'it is' decreases, because verb form and *wh*-position are related. Conversely, *wh*-words co-occurring with Free verb forms are expected to be ex-situ. The developmental aspect of this hypothesis hinges upon the diversification of the verbal system with more *c'est* forms in children than in adults (Guillaume, 1927; Clark, 1978; Bassano *et al.*, 2005).

We examined the evolution of the in-situ/ex-situ ratios for *wh*-pronouns and *wh*-adverbs separately, and the evolution of the ratio of fixed and free verb forms during the three-year period. A repeated-measures ANOVA including Year (1, 2, 3) and Grammatical category (Pronouns, Adverbs) as factors indicates an effect of the grammatical category of the *wh*-item ( $F(1,15) = 50.143$ ;  $p < .001$ ; eta squared = .770), and that the distribution between the three years is not significantly different ( $F(2,30) = 1.513$ ; n.s.).

Figure 4 shows the difference between *wh*-in-situ pronouns and adverbs. The ratio of *wh*-in-situ pronouns decreases constantly (82.9%–78.0%–61.9%) in favour of more ex-situ pronouns, as expected considering that adult Colloquial French is usually described as displaying more ex-situ *wh*-questions than child



**Table 7.** Distribution of *wh*-adverbs according to syntactic structure (Year 3)

		n	%		
1a	S Vlex <i>wh</i> -	7	15.6	35.6	95.6
1b	S Vbe <i>wh</i> -	9	20.0		
2a	<i>wh</i> - S Vlex	19	42.2	60.0	
2b	<i>wh</i> - S Vbe	8	17.8		
8	<i>wh</i> - ESK S V	0	0.0		
4	c'est <i>wh</i> -	2	4.4	4.4	4.4
5	<i>wh</i> - c'est	0	0.0	0.0	
	Total	45	100.0	100.0	100.0

French. The decrease could then well stem from the cognitive development suggested in Strik (2007). An intriguing fact then is the evolution of the ratio of in-situ *wh*-adverbs. They first decrease and then increase during the period (36.1%–29.7%–40.0%). Since the last year is not significantly different from the first year, there does not seem to be a developmental/cognitive constraint underpinning the ratio of in-situ/ex-situ *wh*-adverbs. We also note that despite the decrease in *wh*-in-situ pronouns in Figure 4, the observed decline is not significant either.

We then compared the evolution of the *wh*-positions – pronouns and adverbs – with the evolution of the verb forms – Fixed *be* and Free forms. It seems that the decrease in *wh*-in-situ pronouns parallels the decrease in Fixed *be* form, and that the decrease-increase in in-situ *wh*-adverbs mirrors the Free verb forms pattern, which means that the ex-situ *wh*-adverb pattern parallels it. The evolution of the in-situ/ex-situ ratios in pronoun and adverb *wh*-positions could then all be accounted for within the Verb form hypothesis that opposes the Fixed *be* form to All Free forms (Section 3.2).

In order to test the parallels observed in Figure 4, we examined the position of the adverb and pronoun *wh*-words as a function of the Verb form (Fixed *be* form vs. Free *be* forms vs. Lexical verbs). Table 7 shows that adverb *wh*-words ( $n = 45$ ; *où* ‘where’, *comment* ‘how’, *quand* ‘when’, *combien* ‘how many’) massively co-occur with Free verb forms in the third year (95.6%; Structures 1a, 1b, 2a and 2b), similarly to the first two years (99.6%; same structures in Palasis et al., 2019). In the third year, two adverbs only appear with the Fixed *be* form (*c’est comment* ? ‘how is it’ VI, 5;2 and *c’est quand* ? ‘when is it’ CA, 5;3). The Verb form hypothesis predicts that we should not find a significant difference in the position of the *wh*-word when comparing Structures 1b and 2b (Free *be* forms) with Structures 1a and 2a (Lexical verbs), since the latter are also free verb forms.

We repeated the series of analyses carried out on the first two years on the third year. We examined the independence between the *wh*-position and the verb form with *wh*-adverbs. Only 11 children had sufficient data for this test ( $n = 41$ ). The CMH test was carried on Lexical verbs vs. Free *be* forms, which represent 95.6%

**Table 8.** Distribution of *wh*-pronouns according to syntactic structure (Year 3)

		n	%		
1a	S Vlex <i>wh</i> -	19	16.8	16.8	53.1
1b	S Vbe <i>wh</i> -	0	0.0		
2a	<i>wh</i> - S Vlex	0	0.0	36.3	
2b	<i>wh</i> - S Vbe	0	0.0		
8	<i>wh</i> - ESK S V	41	36.3		
4	<i>c'est wh</i> -	51	45.1	45.1	46.9
5	<i>wh</i> - <i>c'est</i>	2	1.8	1.8	
	Total	113	100.0	100.0	100.0

of the *wh*-adverb questions in Year 3. As expected, the result is not significant ( $X^2$  CMH = 0.52, 1 *df*,  $p = .471$ ). This result confirms that, as in the first two years and as in Section 2.3 for the third year, the *wh*-position is not significantly different when comparing Free *be* forms and Lexical verbs (also free forms) in this subgroup of adverbial *wh*-words. We conclude that adverbs are mainly ex-situ throughout the period (Year 1: 63.9% and Year 2: 70.3% in Palasis *et al.* 2019; Year 3: 60.0% in Table 7), because they massively and steadily co-occur with Free verb forms (Free *be* forms and Lexical verbs) in all three years.

Table 8 shows that *wh*-pronouns co-occur with Free and Fixed verb forms (53.1% and 46.9%, respectively), but that nearly half the pronouns (45.1%) appear in one configuration: in-situ with the Fixed *be* form (Structure 4). The rest of the pronouns mainly co-occur with Lexical verbs, either ex-situ with the *est-ce que* interrogative marker (Structure 8: 36.3%) or in-situ (Structure 1a: 16.8%). The *wh*-pronoun *quoi/que* 'what' is overwhelming ( $n = 105$ ), and there are a few *qui* 'who' ( $n = 5$ ), *à quoi* 'to what' ( $n = 2$ ), and *à qui* 'to whom' ( $n = 1$ ). Since *wh*-pronouns never appear with Free *be* forms, we examined the relationship between the position of the *wh*-pronouns with the Fixed *be* form as opposed to Lexical verbs only. The CMH test was performed on the data in 15 children ( $n = 112$ ) and showed a significant result ( $X^2$  CMH = 27.598, 1 *df*,  $p < .00001$ ). This result confirms that the *wh*-items are more in-situ with the Fixed *be* form than with Lexical verbs in this subgroup of pronominal *wh*-words.

Taken together, these results seem to show that the grammatical category (adverb vs. pronoun) is not a discriminating variable for the position of the *wh*-phrase. We suggest that the in-situ/ex-situ ratios of *wh*-pronouns and *wh*-adverbs and the well-documented discrepancy with adult data pertain to the ongoing development of the verbal system of the children, who gradually flesh their system out by replacing *c'est* forms with lexical verbs.

In addition to this gradual verb diversification, the new data examined in this contribution draw our attention to a second developmental phenomenon. Indeed, we saw in Section 3.1 that there were no significant differences in the in-situ/ex-situ distributions between the three years, except for the *wh*-word *quoi/que* 'what' that showed a significant decrease in the in-situ position

**Table 9.** Outliers in the dataset according to *wh*-word and syntactic structure

		<i>quoi/que</i>	<i>où</i>	<i>qui</i>	<i>comment</i>	<i>à qui</i>	<i>à quoi</i>	<i>combien</i>	Total
1a	S Vlex <i>wh</i> -	69	14	12	7	0	1	2	105
1b	S Vbe <i>wh</i> -	0	57	0	1	1	0	0	59
5	<i>wh</i> - <i>c'est</i>	12	0	1	0	5	0	0	18
	Total	81	71	13	8	6	1	2	182

between Year 2 and Year 3 (from 73.5% to 59.0%; Fisher's Exact Test for Count Data:  $p < .05$ ). Interestingly, the 'what' pronoun in French is the only *wh*-word with different morphological forms according to the position (i.e., in-situ *quoi* vs. ex-situ *que*). Moreover, the form of the pronoun varies between Colloquial French, which displays *qu'est-ce que* 'what is it that' without subject inversion in the question (e.g., *qu'est-ce qu'il a dit ?* 'what did he say'), and Standard French, which displays *que* 'what' with subject inversion (e.g., *qu'a-t-il dit ?*). The data thus illustrate that the children are fleshing out their *wh*-word inventory by gradually integrating the alternation between in-situ *quoi* and ex-situ *qu'est-ce que* 'what is it that', which results in the latter being more frequent over time. This development together with the verb diversification seem to account nicely for the decrease in the overall *wh*-in-situ ratio (Figure 3), the decrease in the pronoun *wh*-in-situ ratio (Figure 4), and the absence of decrease in the adverb *wh*-in-situ category (Figure 4).

#### 4. OUTLIERS TO THE CORRELATION BETWEEN FIXED *BE* FORM AND *WH*-IN-SITU AND DEICTIC HYPOTHESIS

Finally, we listed the outliers to the correlation between the Fixed *be* form *c'est* 'it is' and the *wh*-in-situ position, that is the questions that display Free verb forms and *wh*-in-situ (Structures 1a and 1b) and the Fixed *be* form *c'est* 'it is' with the *wh*-ex-situ position (Structure 5). The total number of outliers amounted to 182 (25.3% of the total dataset, see Table 9).

In this section, we will examine the *wh*-in-situ questions only (Structures 1a and 1b,  $n = 164$ , 22.8% of the total dataset) in order to seek variables that could account for the unexpected *wh*-in-situ position in these structures. We explore two hypotheses:

- Outliers share with the Fixed *be* structure (Structure 4) a feature that triggers the usage of a *wh*-in-situ instead of a *wh*-ex-situ.
- Outliers exhibit types of *wh*-in-situ that are different from those found in the Fixed *be* questions.

Taking a closer look at the corpus, we observed that it is replete with in-situ questions such as (5) and (6).

**Table 10.** Distribution of *là* in *wh*-questions

	In-situ	Ex-situ	Total
Fixed <i>be</i>	13	0	13
Free <i>be</i>	11	3	14
Lexical verbs	18	13	31
Total	42	16	58

- (5) *Et là i(l) dit quoi?* (LA, 3;11)  
 and there he says what  
 ‘And there what does he say?’  
 (6) *Là y a quoi?* (KE, 5;7)  
 there there is what  
 ‘What is there there?’

Note the presence of the adverb *là* ‘there’ in both cases. A survey of our corpus showed that *là* is surprisingly frequent among the outliers. We examined the presence of *là* in detail. Items like *celui-là* ‘that one’ or *ce chat-là* ‘that cat’ were excluded, since *là* is part of the discontinuous demonstrative morpheme *ce...là* and is hence not the adverb *là* (see 7a). Similarly, *là-dedans* ‘there inside’ and embedded versions like (7b) were not included.<sup>7</sup>

- (7) Presence of *là*:  
 a. *Celui-là* ‘that one’:  
*Et celui-là ça va où lui?* (NI, 3;0)  
 and that one that goes where him  
 ‘And where does that one go?’  
 b. Embedded *là*:  
*Il est où le truc qui est là à sa vache euh?* (MS, 3;0)  
 it is where the thing that is there to his cow uh  
 ‘Where is the thing that is there (and that belongs) to his cow uh?’

Based on the occurrences in Table 10, we ran a Fisher’s Exact Test of independence. The result is significant ( $p < .01$ ), which shows that the distribution of *wh*-in-situ and *wh*-ex-situ with respect to the presence of *là* is above chance and that there is a correlation between the presence of *là* and the presence of *wh*-in-situ.

*Là* is a complex marker in French. Although it is originally a locative adverb meaning ‘here’ or ‘there’, it developed a range of other deictic uses, including temporal, narrative, discursive and interactional ones. In its temporal and narrative uses, it serves as a linker between two parts of a story. This use is not

<sup>7</sup>Likewise, Kleiber (1995: 10) hypothesizes that these items have a different reference mode from lone *là* and treats them separately.

featured in our questions. As a discursive and interactional marker, *là* is used by the speaker to draw the addressee's attention to her/his discourse and her/his utterance. According to Smith (1995), *là* is the least specified of the French locative adverbs (in contrast with *ici* 'here' and *là-bas* 'over there'). It is characterized by the features [-far], [-speaker], which account for its local as well as its discourse usages (by default, it serves to involve the addressee). In contrast, Kleiber (1995: 23) suggests that *là* refers to something that is given, manifest or accessible. When it refers to a location, spatial referentiality is already active in the context. Note however that the two approaches are not incompatible. To draw the attention of the addressee to an object, this object must be salient or at least already accessible.

Be that as it may, the main questions are to discover what in *là* favours *wh*-in-situ, and if *là* and Fixed *be* structures have anything in common. To discuss the matter, it is useful to go back to what was introduced in Section 1. Recall that the target (adult) language is complex with respect to the *wh*-in-situ phenomenon, since it features three types of *wh*-questions (Baunaz, 2011, 2016). Type 1 is a neutral information-seeking question, without presupposition. Type 2 is presuppositional-partitive (with rise-fall intonation on the *wh*-phrase). Type 3 is presuppositional-specific (the *wh*-phrase forms its own prosodic phrase; Vergnaud and Zubizarreta, 2005).<sup>8</sup> In adult speech, Faure and Palasis (2021) showed that *wh*-ex-situ is favoured by the presence of a contrastive feature whose specific semantics is exclusive, meaning that the question conveys the presupposition that at least one of the contextually available answers is a priori excluded.

Let us now compare our findings on child speech with this target situation. In Sections 1 and 3, we mentioned and confirmed one factor favouring *wh*-in-situ in child French: the presence of the Fixed *be* form *c'est* 'it is'. Palasis et al. (2019: 230) additionally suggested that the use of *wh*-ex-situ was favoured when the children's questions displayed a verb form with more semantic content (i.e., Free *be* forms and Lexical verbs). In the target adult language, more semantic content creates more context, in which a presupposed set can be identified. Such a situation is unlikely in a sentence with *c'est qui/quoi*, because it asks for identification or definition, and most of the time any answer is possible. Once again comparing with adult speech, this category of *wh*-in-situ sentences corresponds to Type 1 in Baunaz's (2011, 2016) classification.

What *c'est* and *là* share is that they display a deictic element (*c'* and *là*) linked to the *hic et nunc*. They pick out of the immediate context an element or a situation that was not yet in the scope of the attention of the addressee and thus initiate the identificational or definitional process that underpins the interrogative speech act. So, no relation is established with the previous context or a presupposed set, a necessary condition to have Types 2 or 3 *wh*-in-situ or *est-ce que*-less *wh*-ex-situ

<sup>8</sup>Here are Baunaz's (2016: 134 and 137) definitions (The reader is referred to the original article for details): 'A partitive *wh*-phrase is an object, which belongs to a presupposed set containing more objects. Each of the objects of the set can potentially be referents to the answer of the *wh*-phrase, i.e., all are alternatives.'

'Specificity narrows down the context to familiar individuals, excluding alternatives. A constituent question involving specificity entails an answer referring to a familiar individual that the interlocutor has in mind.'

(at this stage, children do not produce *wh-est-ce que*, apart from *qu'est-ce que* 'what is it that', as mentioned in Section 3.3).

Other data go in the same direction. In (8), the child begins the utterance with *là* and goes on with a *wh-ex-situ* question, but hesitates, stops and corrects the sentence to finally produce a *wh-in-situ* question, as expected under our analysis.

- (8) *Et là [/] et là où il est [/] il est où?* (MT, 4;4)  
and there and there where he is he is where  
'And there where is he?'

Consequently, this new dataset prominently features Type 1 *wh-in-situ*. Semantically, it does not carry a presupposition; syntactically, it does not feature fronting/movement of the *wh*-phrase. Although it is mostly materialised with *c'est* and *là* questions, unmarked Type 1 questions also occur in other guises, such as (9).

- (9) Absence of *là*:  
*Lui i(l) mange quoi?* (MS, 3;9)  
him he eats what  
'What does he eat?'

Type 1 is most likely to appear first in child speech, since it is the least marked question type in the target language. Nevertheless, the prevalence of Type 1 is not sufficient to explain all the data, and the examples in (10) and (11) do not abide by the description of Type 1 but belong to Type 2 (so far, we have found no examples of Type 3). In each case, a presupposed set is given from which the answer must be picked out.

- (10) The children must pick an item out of a set of books and games:  
*Maintenant je prends quoi alors?* (WI, 2;10)  
now I take what so?  
'So what do I take now?'
- (11) The children are discussing the different activities in a playground they are all familiar with:  
*T(u) as fait quoi?* (meaning 'which activity'; LU, 5;4)  
you have done what  
'What have you done?'

Another interesting fact is that, while the proportion of *wh-in-situ* decreases from Year 1 to Year 3 (Figure 3), the proportion of outliers increases, particularly with lexical verbs (Table 11).

Summarizing our findings in this section, we found that *wh-in-situ* structures come in two types in child speech, corresponding to Type 1 and Type 2 questions in Baunaz's (2011, 2016) description of French adult speech. Type 1 is prevalent. It is a type of mere information-seeking questions, which take the form of Fixed *be* and *là* structures in child speech. In fact, these deictic structures are the most suitable to encode the operation of identification that

**Table 11.** Distribution of outliers among *wh*-in-situ (%)

	Lexical verbs	Free <i>be</i>	Total
Year 1	19.4	15.3	34.7
Year 2	26.3	10.5	36.8
Year 3	29.6	10.2	39.8
Total	23.1	13.1	36.2

underlies the interrogative process. Nevertheless, Type 2 (partitive questions) is also present. It may be tied to the other observation that we made, namely that the number of outliers with lexical verbs increases through time. A possibility is that children gradually develop more complex types of *wh*-in-situ, such as Type 2, which leads to a growing usage of this type and to the increase of the number of alleged outliers. This hypothesis is consistent with the fact that children gradually use structures with more material, a factor favouring the linking to the context and the construction of presupposed sets. We leave the exploration of this hypothesis for future research.

## 5. CONCLUSIONS

This article examined the production of matrix information-seeking *wh*-questions in 16 typically developing French children during their three kindergarten years ( $n = 718$ , Table 3). We aimed to explain the prevalence of the in-situ *wh*-position in child data compared to adult data and to disentangle the various factors that could explain the in-situ/ex-situ ratio and its evolution during the three-year period.

In Section 3.1, we observed that *wh*-in-situ questions remain prevalent during the period despite a steady increase in *wh*-ex-situ (Figure 3). There are no significant differences in the in-situ/ex-situ distributions between the three years. We further noticed that there are no significant differences in the in-situ/ex-situ distributions of each *wh*-item considered separately between the three years, except for the *wh*-word *quoi/que* ‘what’ that shows a significant decrease in the in-situ position between Year 2 and Year 3.

In Section 3.2, we statistically examined the position of the *wh*-phrases as a function of the Verb form in Year 3 (Fixed *be* forms vs. Free *be* forms vs. Lexical verbs, which are also free forms, following Palasis et al., 2019 in Year 1 and 2). The Verb form (Fixed vs. Free) is also a discriminating variable for the *wh*-position in Year 3: The Fixed *be* form favours the in-situ position (e.g., *c’est qui Taz ?* ‘who is Taz’), whereas the Free forms favour the ex-situ position (e.g., *combien ça coûte ?* ‘how much does it cost’). Taken together, these results allow us to make the prediction that the ex-situ rate will steadily increase as a function of the diversification of the child verbal system. We leave this matter for future research.

In Section 3.3, we investigated the well-documented asymmetry between in-situ *wh*-pronouns and ex-situ *wh*-adverbs. We confirmed that the asymmetry also

pertains to the Verb form. *Wh*-pronouns are more in-situ with the Fixed *be* form *c'est* 'it is' than with the Lexical verbs. *Wh*-adverbs are mainly ex-situ, because they massively co-occur with Free verb forms. In line with the prediction on the diversification of the child verbal system, we expect that *wh*-pronouns will be more ex-situ, when children produce fewer Fixed *be* forms and more Lexical verbs. Furthermore, the evolution of the data between Year 2 and Year 3 illustrates that the children are fleshing out their *wh*-word inventory by gradually integrating the alternation between in-situ *quoi* and ex-situ *qu'est-ce que* 'what is it that'. This development together with the verb diversification seem to account for the decrease in the overall *wh*-in-situ ratio (Figure 3), the decrease in the pronoun *wh*-in-situ ratio and the absence of decrease in the adverb *wh*-in-situ category (Figure 4).

Finally, we investigated the outliers to the Verb form hypothesis (i.e., the *wh*-in-situ questions with Free verb forms; 22.8% of the dataset) in Section 4. We observed the presence of the adverb *là* 'there' in the outliers and elaborated on the fact that *là* and the Fixed form *c'est* share deictic properties. Generalising, Fixed *be*, *là* and other (unmarked) questions all belong to Baunaz's (2011, 2016) Type 1 questions in the target adult language (i.e., mere information-seeking questions). We also observed that, while *wh*-in-situ questions tend to decrease throughout the three years, *wh*-in-situ outliers increase, which we tentatively attributed to the emergence of Type 2 (and maybe Type 3) *wh*-questions, which rest on a set of presupposed possible answers. This hypothesis is left for future research.

Another important field that requires future attention is the behaviour of ex-situ *wh*-questions in child speech. The verb diversification and the increase in the *wh*-pronoun *qu'est-ce que* 'what' pertain to this domain of investigation. Some Fixed *be* and *là* structures that unexpectedly exhibited ex-situ *wh*-items, though rarely ( $n = 20$ , 2% of the corpus), also require attention. More generally, the behaviour of D-linking and *wh*-ex-situ in child speech largely remains to be explored in future research.

**Competing interests.** The authors declare none.

**Acknowledgements.** We thank the editors and three anonymous reviewers for their very constructive comments on previous versions of this contribution. We also thank the audience at the Workshop 'L'interrogative *in situ* : aspects formels, pragmatiques et variationnels' organised by Caterina Bonan, Laurie Dekhissi and Alexander Guryev in June 2022.

## References

- Adli, A. (2006). French *wh*-in-situ Questions and Syntactic Optionality: Evidence from Three Data Types. *Zeitschrift für Sprachwissenschaft*, 25, 163–203.
- Bassano, D., Eme, P.-E. and Champaud, C. (2005). A naturalistic study of early lexical development: General processes and inter-individual variations in French children. *First Language*, 25(1), 67–101.
- Baunaz, L. (2011). *The Grammar of French Quantification*. Dordrecht: Springer.
- Baunaz, L. (2016). French "quantifiers" in questions: Interface strategies. *Acta Linguistica Hungarica*, 63(2), 125–168.
- Bayer, J. and Cheng, L. L.-S. (2017). *Wh*-in-situ. In M. Everaert and H. Van Riemsdijk (eds), *The Wiley Blackwell Companion to Syntax* (Vol. 8: U–Z, pp. 4979–5022). Hoboken, NJ: Wiley Blackwell.



- Becker, M. and Gotowski, M. (2015). Explaining children's *wh*-in situ questions: Against economy. In E. Grillo and K. Jepsen (eds), *Proceedings of the 39th annual Boston University Conference on Language Development* [BUCLD 39] (pp. 88–100). Somerville, MA: Cascadia Press.
- Biezma, M. (2020). Non-informative assertions: The case of non-optional *wh*-in-situ. *Semantics and Pragmatics*, 13(18), <https://doi.org/10.3765/sp.3713.3718>.
- Boeckx, C. (2000). Decomposing French Questions. In J. Alexander, N.-R. Han and M. Minnick (eds), *Proceedings of the 23rd Annual Penn Linguistics Colloquium* (Vol. 6, pp. 69–80). Philadelphia, PA: University of Pennsylvania Working Papers in Linguistics.
- Chang, L. (1997). *Wh-in-situ Phenomena in French*. (MA thesis), University of British Columbia.
- Cheng, L. L.-S. and Rooryck, J. (2000). Licensing *wh*-in-situ. *Syntax*, 3(1), 1–19.
- Cheung, C. C.-H. (2008). *Wh-fronting in Chinese*. (Ph.D dissertation), University of South California, Los Angeles.
- Clark, E. V. (1978). Strategies for Communicating. *Child Development*, 49(4), 953–959.
- Cochran, W. G. (1954). Some Methods for Strengthening the Common  $\chi^2$  Tests. *Biometrics*, 10(4), 417–451.
- Coveney, A. (1989). Pragmatic constraints on interrogatives in spoken French. *York Papers in Linguistics*, 13, 89–99.
- Coveney, A. (1995). The use of the QU-final interrogative structure in spoken French. *Journal of French Language Studies*, 5(2), 143–171.
- Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton: Princeton University Press.
- Crisma, P. (1992). On the Acquisition of *Wh*-Questions in French. *Geneva Generative Papers*, 0(1–2), 115–122.
- Déprez, V. and Pierce, A. (1993). Negation and Functional Projections in Early Grammar. *Linguistic Inquiry*, 24(1), 25–67.
- Dryer, M. S. (2013). Position of Interrogative Phrases in Content Questions. In M. S. Dryer and M. Haspelmath (eds), *The World Atlas of Language Structures Online* (<http://wals.info/chapter/93>). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Faure, R. and Palasis, K. (2021). Exclusivity! *Wh*-fronting is not optional *wh*-movement in Colloquial French. *Natural Language & Linguistic Theory*, 39(1), 57–95. <https://doi.org/10.1007/s11049-020-09476-w>
- Gotowski, M. and Becker, M. (2016). An information-structural account of children's *wh*-in situ questions in French. In L. Perkins, R. Dudley, J. Gerard, and K. Hitzenko (eds), *Proceedings of the 6th Conference on Generative Approaches to Language Acquisition North America [GALANA 2015]* (pp. 24–35). Somerville, MA: Cascadia Press.
- Guasti, M. T. (2000). An excursion into interrogatives in Early English and Italian. In M.-A. Friedemann and L. Rizzi (eds), *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics* (pp. 105–128). Harlow: Pearson Education.
- Guillaume, P. (1927). Le développement des éléments formels dans le langage de l'enfant. *Journal de Psychologie*, 24, 203–229.
- Hamann, C. (2006). Speculations about Early Syntax: The Production of *Wh*-Questions by Normally Developing French Children and French Children with SLI. *Catalan Journal of Linguistics*, 5, 143–189.
- Hamlaoui, F. (2011). On the role of phonology and discourse in Francilian French *wh*-questions. *Journal of Linguistics*, 47(1), 129–162.
- Higgins, F. R. (1973/1979). *The pseudo-cleft construction in English*. (Ph.D dissertation; published by Garland), Massachusetts Institute of Technology, Cambridge, MA.
- Huang, C.-T. J. (1982). Move *WH* in a language without movement. *The Linguistic Review*, 1, 369–416.
- Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua*, 121(3), 339–351.
- Kleiber, G. (1995). D'ici à là et vice versa ; pour les aborder autrement. *Le Gré des langues*, 8, 8–27.
- MacWhinney, B. (2000). *The CHILDES Project. Tools for Analyzing Talk: Transcription Format and Programs* (3<sup>rd</sup> edn, vol. 1). Mahwah, NJ: Lawrence Erlbaum Associates, <http://childes.talkbank.org/>.
- Mantel, N. and Haenszel, W. (1959). Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *Journal of the National Cancer Institute*, 22(4), 719–748.
- McDonald, J. H. (2011). *Myths of Human Genetics*. Baltimore, MD: Sparky House Publishing.

- Myers, L. L. and Pellet, S. (2014). *Pourquoi* in spoken French. In S. Katz Bourns and L. L. Myers (eds), *Perspectives on Linguistic Structure and Context* (pp. 157–182). Amsterdam: John Benjamins.
- Nguyen, A. D. and Legendre, G. (2022). The acquisition of *wh*-questions: Beyond structural economy and input frequency. *Language Acquisition*, 29(1), 79–104. <https://doi.org/10.1080/10489223.2021.1968867>
- Obenauer, H.-G. (1994). *Aspects de la syntaxe A-barre. Effets d'intervention et mouvements des quantifieurs*. (Ph.D dissertation), Université de Paris VIII.
- Oiry, M. (2011). *L'acquisition des Questions à Longue-Distance par des enfants français. Stratégie à dépendance directe versus indirecte et questions alternatives*. Editions Universitaires Européennes.
- Palasis, K. (2013). The case for diglossia: Describing the emergence of two grammars in the early acquisition of Metropolitan French. *Journal of French Language Studies*, 23(1), Special Issue 'L'hypothèse d'une diglossie en France', Massot, Benjamin and Paul Rowlett (eds), 17–35. <https://doi.org/10.1017/S0959269512000348>
- Palasis, K., Faure, R. and Lavigne, F. (2019). Explaining variation in *wh*-position in child French: A statistical analysis of new semimaturalistic data. *Language Acquisition*, 26(2), 210–234. <https://doi.org/10.1080/10489223.2018.1513004>
- Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.
- Prévost, P. (2009). *The Acquisition of French: the Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism, and L2 Acquisition*. Amsterdam: John Benjamins.
- Rasetti, L. (2003). *Optional Categories in Early French Syntax: A Developmental Study of Root Infinitives and Null Arguments*. (Ph.D dissertation), Université de Genève.
- Rizzi, L. (2001). On the position “Int(errogative)” in the left periphery of the clause. In G. Cinque and G. Salvi (eds), *Current Studies in Italian Syntax: Essays offered to Lorenzo Renzi* (pp. 287–296). Oxford: Elsevier.
- Roeper, T. and De Villiers, J. G. (2011). The Acquisition Path for *Wh*-Questions. In J. G. De Villiers and T. Roeper (eds), *Handbook of Generative Approaches to Language Acquisition* (pp. 189–246). Dordrecht: Springer.
- Roy, I. (2013). *Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface*. Oxford: Oxford University Press.
- Smith, J. C. (1995). L'évolution sémantique et pragmatique des adverbes déictiques “ici”, “là” et “là-bas”. *Langue française*, 107, 43–57.
- Strik, N. (2007). L'acquisition des phrases interrogatives chez les enfants francophones. *Psychologie française*, 52(1), 27–39.
- Strik, N. (2012). *Wh*-questions in child bilingual acquisition of French: Derivational complexity and cross-linguistic influence. *Canadian Journal of Linguistics*, 57(1), 133–151.
- Strik, N. and Pérez-Leroux, A. T. (2011). Jij doe wat girafe?: *Wh*-movement and inversion in Dutch-French bilingual children. *Linguistic Approaches to Bilingualism*, 1(2), 175–205.
- Stromswold, K. (1995). The Acquisition of Subject and Object *Wh*-Questions. *Language Acquisition*, 4(1–2), 5–48.
- Thiberge, G. (2020). *Acquisition et maîtrise des interrogatives partielles en français : La variation sociolinguistique comme outil interactionnel*. (Ph.D dissertation), Université de Paris.
- Van Kampen, J. (1997). *First steps in wh-movement*. Delft: Eburon.
- Vergnaud, J.-R. and Zubizarreta, M. L. (2005). The representation of focus and its implications: towards an alternative account of some ‘intervention effects’. In J. Koster, H. Van der Hulst and H. Van Riemsdijk (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk* (pp. 641–660). Berlin & New York: Mouton de Gruyter.
- Zuckerman, S. and Hulk, A. (2001). Acquiring optionality in French *wh*-questions: An experimental study. *Revue québécoise de linguistique*, 30(2), 71–97.