Chapter 5

Matrix Wh-questions in French

1.1   Introduction

This chapter focuses on the topic of matrix Wh-questions in French and their acquisition by L1 learners of French. One of the most interesting characteristics of French matrix Wh-questions is the fact that they allow for a rather large variation with respect to the constructions used and, more specifically, with respect to movement operations within those constructions. In particular, while many languages have an obligatory rule for the derivation of a Wh-question, (e.g., move Wh-element to spec-CP, invert the auxiliary and the subject, etc.), French allows matrix questions to be derived via what seems to be an optional rule (i.e., the Wh-element can be moved to spec-CP or remain in its base position, the auxiliary/verb can move to a pre-subject position, or remain in a post-subject position). As a result, speakers of French have a variety of options when they produce a matrix Wh-question, and children learning French are exposed to a variety of Wh-structures in their input. Although there are constraints on the use of the various structures (both syntactic and semantic/pragmatic, see (Coveney 1996), several structures are allowed in the same syntactic and semantic/pragmatic context. In many cases, therefore, children are exposed to Wh-questions in their input that differ in structure but not in their interpretation.

This optionality in children's input is the subject of the current investigation. The main question asked in this chapter is the following: what are children's preferences with regard to the various options that French allows for Wh-questions, and more importantly, how do these relate to the preferences of adult French speakers?

The acquisition of Wh-questions in French has been the subject of much investigation in recent years (e.g.: Crisma 1992, Weissenborn 1991, Hulk 1996, Hamann 2000, Plunkett 1999a,b), which has shed some light on the topic providing several findings relevant to the current discussion. First, all the children investigated (with the exception of Philippe; CHILDES corpus, investigated by Crisma, Hulk and Hamann) seem to prefer the Wh-in-situ structure (in which the Wh-element remains in its base position) in early stages; none of them produced subject-

auxiliary/verb inversion[23]. It has also been claimed that such inversion structures are not produced by caregivers and thus have no effect on the child's early grammar. All investigations of this subject have been based on an analysis of spontaneous speech rather than on experimental data. The purpose of the current investigation --the first experimental investigation of this subject-- was to produce experimental results that would allow a quantitative and qualitative comparison between the preferences of children and the preferences of their parents, as representative tokens of the input to which children are exposed.

1.2   Variation in the target grammar

French seems to allow for optionality with regard to the structure of Wh-questions, as can be seen from the examples in (1)–(4) below.

(1)   Inversion (I)

a.   Comment as-tu   fait   ça ?   (Clitic inversion)
        how   have-you   done that
'How did you do that?'

b.   Où   est (allé) ton père?   (Stylistic inversion)
        where is   (gone) your father
'Where is your father (gone)?'

c.   A quelle heure le train est-il parti ? (Complex inversion)
        at what time the train has he left
 'At what time did the train leave?'

(2)   Wh+ESK
        Comment est-ce-que tu   as   fait   ça?   (Wh+ESK)
        how   ESK   you have done that
'How did you do that ?'

---

[23] Except for the Belgian child Léa from the York corpus who is reported to have produced inversion wh-questions (Plunkett to appear, 1999).

(3)   Wh-fronted, no Inversion   (F)

a.   Comment tu   as   fait   ça ?   (Fronting)
     how   you have done that
  'How did you do that?'

b.   Qu'est-ce que tu   as   fait?   (KESK)
      KESK =what you have done
'What did you do?'

(4)   In-Situ (IS)
      Tu   as   fait   ça   comment ?
      you have done that how
'How did you do that?'

For the purposes of the current investigation, the various options are divided into 4 groups, described briefly in the following sub-sections, corresponding to examples (1)-(4).

1.2.1   Inversion (I).

In inverted questions, both the Wh-element and the verb are moved to a pre-subject position. This group includes 3 different types of inversion: 'Clitic inversion' (a 'Germanic' V2 type inversion in which the subject-clitic follows the finite verb, auxiliary or lexical); phrasal subjects are excluded in this structure); 'stylistic inversion' (a 'Romance' type inversion in which the nominal subject follows both the auxiliary and the past participle. (Clitic subjects are excluded in this structure); and 'Complex inversion' (V2 type with a phrasal subject of the form subject phrasal – verb finite – subject clitic.)

1.2.2   Wh+est-ce que (Wh+ESK).

Wh+ESK structures are characterized by fronting of a wh-word (other than que 'what') followed by the marker *est-ce que*. Following Rooryck (1994), *est-ce que* is considered here to be a complex Q-morpheme, an unanalysed chunk that is base-generated (merged directly) in C^0. This complex interrogative complementizer ESK is restricted to matrix interrogatives in standard French. In colloquial varieties, it can also appear in embedded interrogatives.

We reject the analysis proposed by Plunkett (1998, 1999), who analyses *est-ce que* as a string containing a clause in which inversion of the verb *est* around subject *ce* has applied, and a complementizer *que*. Plunkett calls questions containing a fronted Wh-word followed by the string *est-ce que* "periphrastic questions". The main argument in favor of this analysis is based on the prediction it allows one to make about the acquisition of periphrastic questions. Plunkett compares her own analysis to one in which Wh+ESK would be analyzed as routine (cf. Rooryck 1994). She claims that the competing analysis predicts – contrary to fact – that periphrastic questions are acquired (i.e. produced) early, at the same age as (simple) fronted Wh-questions. Her own analysis correctly predicts that periphrastic questions are acquired late, at the same age as other complex sentence structures. However, analyzing Wh+ESK as routine is not the same as analyzing ESK as a Q-morpheme. Hulk (1996) shows that the analysis of Wh+ESK with ESK as a Q-morpheme correctly predicts that the acquisition of this type of question should coincide with the acquisition of a full-fledged CP projection. Wh+ESK structures appear at the same time as other CP constructions, such as clefts (a matrix question might be analysed as adjunction or movement to IP and does not necessarily contain a full-fledged CP projection). Consequently both Plunkett's analysis of *est-ce que* and the one proposed by Hulk/Rooryck make the same predictions with respect to acquisition.

There is, however, another argument that favors the analysis adopted here to the one proposed by Plunkett. Rooryck shows that the formal properties of *est-ce que* as a Q-morpheme in C^0 and *est-ce que* as an intervening sentence are not the same. As an intervening sentence it bears descending intonation, while as a Q-morpheme it does not bear any intonation. As a sentence, it is interpreted as, "does it mean that...?", but as a Q-morpheme it is interpreted as, "is it true that…?". In a sentence *est* can be used in the past (or future) tense, but this is impossible when *est-ce que* has the status of a Q-morpheme.

1.2.3   Fronting (F).

In the 'Fronting'-structures (the term 'Fronting' is used in this chapter in the technical sense i.e., to refer to 'fronting without inversion'), the Wh element is fronted but the verb remains in a post-subject position. This category also includes the 'KESK' structure (fronting of the "word" *qu'est-ce que*, analyzed here as one word: KESK 'what'; for a different analysis of this structure see Plunkett 1999a). The

analysis of Wh+ESK as a routine was rejected above, since this analysis predicts that Wh ESK questions would be acquired at the same time as simple fronted wh-questions, which is not the case.

It appears, however, that questions with *qu'est-ce que* differ from the other Wh+ESK questions in this respect: children seem to begin producing *qu'est-ce que* questions and fronted Wh- questions at the same stage. This can be seen the most clearly in Philippe's acquisition data (Hulk 1996). Philippe produces both his first *qu'est-ce que* question and his first fronted Wh-question (with *où*) at age 2.1.19. He continues to produce both question types very frequently. It is only at age 3.03.12 that he produces his first Wh+ESK questions: *où est-ce qu'il roule?* 'Where ESK he drives?' Around that time he also produces his first cleft sentences as well as other constructions with a full fledged CP-projection. Hulk (1996) therefore proposes analysing *qu'est-ce que* not as the Wh-word que + ESK, but as an unanalysed chunk that behaves like other simplex Wh-words, such as for example *où*, and *comment*. The distribution of KESK is more restricted than the distribution of other simplex-Wh-words: KESK never appears in-situ or with inversion, facts that also hold for other Wh-words such as *pourquoi.* [24].

In the acquisition data of other children mentioned in the literature, the time-gap between the first *qu'est-ce que* question and the first (other) Wh+ESK question is not always as large as Phillipe's, but for all children production of *qu'est-ce que* questions precedes the production of other "periphrastic questions".

In the present work, we follow Hulk (1996) in classifying *qu'est-ce que* questions as fronted Wh-questions rather than as Wh+ESK questions.

1.2.4   In-situ (IS).

In in-situ questions no inversion takes place; the Wh-element remains in-situ.

1.3   The Nature of the Variation

Regarding the nature of the variation described above and the question of whether these structures are truly optional, we make the assumption that the optionality here is not genuine (following Bolinger 1977, Coveney 1996 and others); that is, that the structures above are not totally interchangeable and are subject to syntactic, morphological, semantic, pragmatic, and possibly other restrictions. Examples of

---

[24] This argument can be strengthened by the observation that Phillipe uses not KESK but KES (*KES c'est*). This supports the view of KESK as an unanalyzed form.

the influence of such factors on the grammaticality of the various options are listed below.

Syntactic factors: The Wh-element – the element *pourquoi* 'why' does not allow the Wh-in-situ option (**il est parti pourquoi ?* 'He has left why?')   nor the stylistic inversion option: (**pourquoi sont partis les enfants ?* 'why have left the children?').   The element *que* 'what' does not allow the 'fronting' option (**que tu fais?* 'What you do?') nor the in-situ option (**tu fais que?* 'You do what?')   (See also Coveney 1996 for more on the influence of the Wh-element on speakers' preferences).   Clitic inversion and stylistic inversion are dependent on whether the subject is clitic or phrasal.   Negative questions do not allow the in-situ option (* *tu n'as pas réparé la voiture comment?* 'You neg have not repaired the car how'); embedded questions allow for only the fronting option (**je me demande comment as-tu fait cela* 'I me wonder how have you done that', etc.).   The type of question (i.e. argument vs. adjunct) is also said (e.g. in Plunkett 1999b) to be a factor in speakers' preferences.   Our results will also show a syntactic influence of the wh-element on preferences, rather than grammaticality differences, among the various options.

Morphological factors: the element *que* 'what' can appear only in the inversion structure while in in-situ 'what'-questions the form *quoi* is obligatory (whether the morphology dictates the position or the position dictates the morphology is a question that will not be discussed in this chapter; see Bouchard & Hirschbühler 1986, Hirschbühler 1979, Obenauer 1976, Pollock 1992, Pollock et al. 1999).

Semantic/pragmatic factors: Convey 1996 observes, on the basis of a corpus analysis of adult speakers of French, that context and speaker's intentions, as well as other discourse related factors (e.g. whether or not the interrogative is read as an emphatic question, rhetorical question, request for information, etc.), might have a strong influence on the choice of structure.   Cheng & Rooryck (1999) argue, partly based on Chang 1997, that Wh-in-situ questions are associated with a strongly presupposed context.   For example, Chang (1997) notes that negative answers are not legitimate answers to in-situ questions. (For example, question: *Marie a acheté quoi?* 'Mary has bought what?'; answer: ??? *rien* 'nothing') although they are fine as answers to a 'fronting' question.   Moreover, they show that Wh-in-situ questions have a special intonation pattern that is similar to the one found in yes/no questions.

The examples above demonstrate that the various options for Wh-questions in French are not totally interchangeable and are subject to the influence of different factors[25]. Nevertheless, in many cases it seems that all options are allowed, without influencing the interpretation or the grammaticality status of the question. In the present work, the focus of investigation is on such structures; thus all the questions that are elicited from the subjects in the experiment are ones that allow for all options, with the exception of *pourquoi* 'why' questions, which cannot appear in an in-situ structure).

The purpose of the experiment discussed in this chapter is to check the general preferences of children and adults with regard to the options above, and more specifically, whether the preferences of children differ from those of adults.

1.4   Forming a prediction

The general proposal of this dissertation, presented in chapter 3, predicts that given "alternatives" in the input, children will opt for the more economical. The current case of Wh-questions in French answers our tentative requirement for an apparently optional structure, but it differs from other cases discussed elsewhere in this dissertation, because it offers more than two "alternatives". This means that instead of looking for a preferred or a less-preferred option, we can establish an economy-based hierarchy that will use relative economical value to order the four main options presented above for Wh-questions in French. In principle, under the assumption that movement is less economical than non-movement (or than covert-movement), the following hierarchy can be established: the least economical structure is the Inversion structure, because it involves two-movement operations [+Wh-movement, +SAI]. The 'Fronting' option is more economical, since it involves only one movement operation [+Wh movement, -SAI]. The in-situ option is the most economical, because it does not involve any movement [-Wh movement, -SAI]. If ESK is an unanalyzed item inserted in C, the Wh+ESK structure falls in this hierarchy, between Inversion and Fronting, because it involves one Move operation and one Merge operation. A Merge operation is viewed in Chomsky (1995) as less costly than movement but more costly than non-movement.

---

[25] As shown in chapter 1, several authors (e.g., Bolinger 1977, Clark 1987) have claimed that there might be a principle of grammar that forbids such interchangeability in any language for any structure, i.e., a principle of a 'one form – one meaning' relationship between word-order and interpretation.

This hierarchy from most to least costly (Inversion, Wh-ESK, Fronting, In-situ) represents, therefore, the order of preferences expected under the current proposal. Recall, however, that we assumed in chapter 3 that children's preferences for the economical options are expected to be influenced by the preferences of adult speakers as they appear in the input. The structures that appear frequently in the input, even if they are less economical, might influence children to hypothesize intermediate rules that allow them to parse and produce these structures. As a result, the prediction of the current proposal is not necessarily for a total absence of the less economical structures and a 100% production of the most economical one, but rather for preferences that favor the economical structure, in comparison to adult-preferences. The underlying hypothesis is that children begin with a general preference for the most economical option, and based on the frequency of the less economical option in the input, they produce the less economical option more and more frequently until it reaches the frequency in adult speech. The costliness-hierarchy proposed above is based on the existence or absence of two movement operations, Wh-movement and Subject-Auxiliary/Verb Inversion (SAI). The predicted results of a comparison between children and adult productions can be sub-divided into two separate predictions. These are presented in (5), below.

(5)   Predictions
When comparing children's Wh-questions with those of their parents (as representatives of the input), we should find that:

      a. Children use fewer Inversion and Wh+ESK questions than their parents do.
      b. Children use more Wh-in situ questions than their parents do.

2.   The experiment

Methods:

Subjects: 33 French-speaking children (ages:4.0 - 5.9, mean age:4.7) and 22 adults (of which 15 parents, and 7 other adults) participated in the experiment. The subjects were from 4 different areas (all in the area of Paris: Stains (5 children, 3 adults), St. Denis (6, 6), Paris (19, 8) and different areas labeled as 'other' (3, 5).

Procedure: The method used was question-elicitation. The subject was presented with a puppet that is said to be shy and not willing to talk to the experimenter. The experimenter explained that she wished to know more about the puppet and asked the child to help since the puppet wouldn't answer the experimenter's questions. The experimenter then presented 32 Wh-questions, one by one, in the form of embedded questions (for example, "I want to know where he has gone"). Recall that the embedded form in French allows for only one option. The child was asked to address the questions to the puppet, and was thus required to turn them into direct questions ("Where has he gone?"). The child was free[26] to choose any of the orders[27] available in his grammar. Possible answers are presented below.

(6) Experimenter: Je veux savoir où   il est allé.
                     I   want to-know where he is gone
'I want to know where he went'
Possible responses:
a.   Où est-il allé ?   (I)
b.   Où est-ce qu` il est allé ?   (Wh+ESK)
c.   Où il est allé ?   (F)
d.   Il est allé où ?   (IS)

In testing the adults, no puppet was used. In order to achieve willing participation from the adults, the experimenter explained that we wished to compare the language of children to that of adults and explained the procedure used with the children. In order to achieve as much similarity as possible between the context in which children and adults were tested, the adults were asked to produce the questions as if they were addressed to the puppet. To avoid a formal register, the adults were

---

[26] The use of embedded questions (which are always in the 'fronting' form) as stimuli might create a bias in the subjects' responses, leading them to prefer the 'fronting' option. However, as in responses to a repetition task, conclusions can be drawn from responses that deviate from this possible bias. Note that the prediction that children will prefer in-situ structures goes against this bias. As compared to a repetition method, the current task was found to be superior because it encourages the child to create 'her own' structure (in many questions the pronoun had to be changed) rather than simply repeat, where the bias is even greater. We return to the possible bias effect in the discussion.

[27] Note that, as shown previously, the range of all 4 options (I,ESK,F,IS) is not always available even in matrix questions. Out of the 32 questions elicited from each subject (see appendix), 27 can be formulated in all four options. The questions with a more limited range of possibilities are the four *pourquoi* 'why' questions, which can be formulated only in I and F structures, and question #19 for which the ESK and F options are less natural.

asked to try to use the level of language they used when talking to children[28].   The adults knew neither the purpose of the experiment, nor that it focused on different question-structures.

The 32 questions presented included seven Wh-elements: *qu* 'what' (7 tokens), *où* 'where' (7 tokens), *comment* 'how' (6 tokens), *quel* 'which' (5 tokens), *pourquoi* 'why' (4 tokens), *quand* 'when' (2 tokens), and *combien* 'how many' (1 token).
The responses of the subjects were recorded with an audio tape-recorder and later transcribed and analysed.

Scoring: Subjects' responses were divided into 4 categories (see the examples in (1)-(4) and in (6) above): (I) Inversion, (wh+ESK), (F) fronting without inversion and (IS) in-situ. Questions that did not get a response were not counted. The difference between adults and children was calculated with a one-way analysis of variance. The correlation among the different Wh-elements was based on the proportional results --since each Wh-element had a different number of items-- and was calculated with a Pearson correlation.

3.   Results and Discussion

Table 3 below presents the general results for the 22 adults and 33 children in raw numbers and in percentages[29].

---

[28] This request is ambiguous. It can lead to a more 'childish' level but also to a more formal (prescriptive) level. Further research must include more sophisticated settings in order to limit adults' use of a formal register.
[29] The numbers in table 3 represent the subjects' relevant responses to all 32 questions. As pointed out in note 7, five of the questions cannot be formulated in all four options. These are nevertheless included in table 3 (and all other tables except table 9 and figure 2) as exclusion of them does not effect the results in any relevant way. The numbers and percentages for the remaining 27 questions are: children – I=16(2%), F=662(90%), IS=57(8%), adults – I=353(62%), ESK=30(5%), F=151(27%), IS=31 (6%).

Table 3 - Total results for adults and children

	Wh-fronted +Inversion	Wh+ESK	Wh-fronted without inversion (+KESK)	Wh-in situ	Total
Children (n=33)	40 (5%)	0	784 (89%)	57 (6%)	881
adults (n=22)	412 (62%)	36 (5%)	188 (28%)	32 (5%)	668

3.1   Inversion :

Table 3 shows a clear difference between the children and adults with regard to the use of inversion. The adults use inversion in 62% of the cases (this high percentage is unexpected based on what is known about the use of inversion in the spoken French), while children virtually never use it[30]. A one-way ANOVA confirms the significance of this difference: t= - 10.524, df = 53, p< 0.05. It might be possible to explain the lack of inversion in the children's responses through the possible bias toward the 'Fronting' structure, since this was used in the (embedded) stimuli (see note 6).   However, the fact that children produce zero responses of the clitic-inversion structure, while producing 57 responses of the in-situ structure, shows that the deviation from such (possible) bias is one-sided; it is clearly toward in-situ and not toward inversion.   Furthermore, the total lack of inversion in children's production is in accordance with previous results of spontaneous-speech analysis.

A more specific analysis of the inversion types produced by the adults is given in table 4 (the structures in table 4 correspond to the examples in (1)a-c):

Table 4   Types of Inversion responses used by the adults.

Clitic inversion	351	(85%)
Stylistic inversion	50	(12%)
Complex inversion	11	(3%)
Total	412	

It is usually claimed (Coveney 1996, Plunkett 1999b), that the three inversion structures in table 4 appear only rarely in the spoken language. In this respect, the current results are different from spontaneous speech analysis data since inversion

---

[30] The 5 % inversion questions the children produced represent two sentences and are both of the Stylistic Inversion type. Moreover they are "predicative" questions with the verb être 'be' in the simple tense form: 15x "Où est sa maman" 'where is his mummy?', 23x "Quel jour est ton anniversaire" 'what day is your

in general, and the clitic-inversion in particular, are the most prominent structures in the experimental data. The complex inversion -- produced 11 times by the adult subjects--is said to not appear at all in the spoken language. The reason for this difference, and for the high percentage of inversion in general in our data, might be that the adults used a more formal register than they would have used in normal circumstances, despite the instructions to speak as though they were addressing children (see note 8), because they knew they were talking to a language researcher. Nevertheless, the current results will indicate that children do hear these inversion structures in their input and are influenced by them.

3.2   Wh+ESK:

The Wh+ESK category receives only 5% of the adult responses and 0 responses from the children. This result is also surprising with respect to the adults, since these structures are of the most frequent in the adults' speech in Plunkett's (1999) investigation. This fact might also be attributed to the use of a more formal register by the adult subjects, leading them to prefer inversion structures over Fronting and Wh+ESK structures. The complete lack of such structures in children's production is in accordance with results of previous analyses as well as with the predictions made in this chapter. As both Wh+ESK and inversion responses are lacking in the production of children in the current study, a difference with respect to the order of their acquisition cannot be detected; recall that the economy-based approach adopted in this chapter predicts that inversion is acquired later. Future research with either older children or with stimulus questions designed especially for this purpose might reveal differences between these structures.

3.3   Fronting:

With respect to the third category the results are clear and surprising: children used more fronted questions than the adults. Note, however that in the predictions made in (5) above, no specific prediction was made with respect to the 'Fronting' structures. This effect may be the result of experimental artifact and no reliable conclusion can be drawn from it. As mentioned before, the use in the stimuli of embedded form, which allows only for the 'Fronting' option, might have caused the children to prefer this structure in their responses. Furthermore, the use of a more

---

birthday?'. One could characterize them as semi-formulaic; they cannot be taken to show knowledge of inversion.

formal register by the adults decreased the adult's usage of 'Fronting', which is less formal. These two possible artifacts (the Fronting-bias of the children and the formal register of the adults) are possibly responsible for the finding that children produce more 'Fronting' structures than adults. The KESK structure, analyzed here as an unanalyzed form in Spec-CP (and thus categorized under Fronting), is the preferred option (in 'what'-questions) for both children and adults (82% and 51% respectively). There is a clear difference for both populations between the use of KESK and Wh+ESK; the first is much more frequent than the second (for the adults: KESK = 51%, Wh+ESK= 5%, and for the children KESK=82%, Wh+ESK= 0%). This difference supports the view of Wh+ESK and KESK as two different structures, as proposed in Hulk (1996) and adopted here. Note that the use of KESK structures by the children cannot be attributed to the possible bias toward Fronting mentioned above, as it is different from the embedded (*ce qu'*) form.

3.4   Wh-in-situ:

Table 3 above does not reveal a difference between adults and children with respect to the use of Wh-in-situ (t=0.312, df=53, P>0.05). Nevertheless, looking at the in-situ responses of the individual adults (Table 5, below), one can see that 2 of the 22 adults are responsible for 27(84%) of the 32 in-situ responses, causing the distribution of adults' responses to be extremely skewed (skewness=3,342).

Table 5   Individual in-situ responses of the 22 adult subjects.

	Parents															Other Adults						
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
IS	0	0	0	0	1	2	0	0	1	0	0	0	0	0	0	0	1	0	17	0	0	10

The fact that the adults' in-situ responses are so skewed calls for a reevaluation of the data. Recall that the primary purpose of testing adult speakers in the current study was to create a sample that represents the data to which children are exposed in their early years (the input). The clear difference between subjects # 19 and #22 and the rest of the subjects (see table 5 above) should lead us to ask if the data from these two subjects can be taken as representative of the input. Notice that these two subjects were not parents; thus the total in-situ responses of parents differ from that of the other adults. Two strategies are available here to overcome this problem.

The first strategy is to recalculate the data for children and their parents alone, ignoring the responses of adults who are not parents (n=7, 15 remaining) and of children who are not matched with a parent (n=15, 17 remaining[31]). This strategy of ignoring the non-parents' responses is based on the logical assumption that parents are better representatives of child input than non-parents.

A second strategy is a statistical analysis that ignores all *outliers* (both adults and children) and recalculates the results for subjects who scored fewer than two standard deviations away from the group's mean. This strategy is based on the assumption that outliers came from a population different from the rest of the subjects and consequently do not represent the population the study intended to sample. For example, it might be the case that these outliers' responses were influenced by a dialect, by a second language or by some other factor not relevant to the present study. Under this analysis, two subjects are excluded from the adults group and 5 subjects are excluded from the children group.

The recalculated results for children and their parents are presented in table 6 below:

Table 6   Comparison between children and their parents

	Wh-fronted +Inversion	Wh+ESK	Wh-fronted without inversion	Wh-in situ
Children (n=17)	21 (4%)	0	424 (88%)	37 (8%)
Parents (n=15)	306 (68%)	15(3%)	128 (28%)	4 (1%)

A t-test for matched pairs for the in-situ responses of children and their parents revealed a significant difference (t=2.678, df=16, p<0.05) and led to the conclusion that children use more in-situ structures than their parents. Notice that with respect to the other structures (Inversion, Fronting, and Wh+ESK), the reanalysis of the data for children and their parents alone does not change the results.

The results of recalculation using the second strategy (omission of *outliers*) is presented in Table 7:

Table 7   Responses of children and adults after omission of outliers.

	I	Wh+ESK	F	IS	Total
Children n=28	34 (5%)	0	697 (92%)	25 (3%)	756
Adults n=20	397 (66%)	31 (5%)	172 (28%)	5 (1%)	605

[31] One parent was the mother of three of the children, hence the unequal numbers of children and parents.

Table 7 shows that the outliers strategy yields results similar to the exclusion of non-parents. A t-test comparing in-situ responses of children and adults after omission of the outliers also revealed a significant difference between the two groups (t=22.86, df=46, p<0.05), which led to the conclusion that children produce more in-situ responses than adults, while all other responses remained similar to the responses of all children and adults (c.f. Table 3, above), as well as to responses of children and their parents alone.

The fact that these two different strategies (exclusion of non-parents and exclusion of outliers) led to similar results supports the conclusion that the original results (table 3) represent a sampling bias in the input and that a difference between children and adults does exist with respect to the use of Wh-in-situ structures.

The general conclusion of this section is thus that children produce more in-situ responses than adults. Note that if there is a bias toward 'Fronting' in children's responses (as mentioned above), the preference of children for in-situ, shown in this section, might actually be greater in practice.

3.5   Comparison between the different Wh-words:

Coveney (1996) observes that different Wh-elements (e.g., *what*, *where*, *when*, etc.) are used with different structures. Specifically, the Wh-element serves as a factor influencing speakers' choices of the various options. Examination of the individual responses of both adults and children in the current data reveals similar findings.

What-questions in French are particularly interesting, because the form of the word corresponding to English *what* varies according to the question-type (as throughout, we limit ourselves to matrix questions):

- in situ questions: quoi
- fronted questions without inversion: KESK recall that this is considered here as an unanalysed element
- fronted questions with inversion: que

Interestingly, the children did not make form-errors: they used only *quoi* in in-situ questions and only KESK in fronted questions without inversion; they did not use inversion with *que*. On the other hand, the adults almost never used in-situ "what" questions: they produced both fronted questions with KESK without inversion (51%) and fronted questions with *que* with inversion (45%). For both parents and children, the majority of "what" questions are fronted KESK questions without inversion.

Another interesting observation can be made concerning the wh-word *où* 'where'. The children used it in situ (11 times) but less often than the wh-word *quoi* (34 times) and only in two sentences that contained the verb *être* 'be':

(7)   C'est où Paris?
      it is where paris

(8)   Sa maman est où?
      his mummy is where

The sentence in (7) above, with *où* in situ, is also found two times in the adult data. However, the adults also produced a number of questions (18) in which fronted *où* was followed by the interrogative element/marker ESK: "*où ESK il aime jouer avec eux*?" 'where ESK he likes to play with them'. They rarely used this structure with other Wh-words. The children never produced such questions.

The Wh-word *pourquoi* "why" can never be used in-situ in French, and is also ungrammatical in so-called 'Stylistic Inversion' constructions. The children never used it in these two constructions. The adults, however, used it rather frequently in constructions with only Fronting and no inversion, even when the subject was a clitic that could have appeared in a V2-type inversion: "*Pourquoi tu veux parler avec moi*?" 'Why you want speak with me?'.

The wh-word *quand* "when" is never used in-situ by the children. Moreover several children replaced *quand* by *comment* 'how' (fronted without inversion) in answering question #2 (see appendix) and either did not answer or used a different construction altogether in reaction to the other question with *quand* (#28, see appendix). One child used the cleft construction *c'est quand que* … . The adults' percentage of fronted Wh without inversion is lowest for the Wh-word *quand*, which has the highest percentage for inversion.

Table 8 below presents the division of the responses across the various Wh-elements:

Table 8   Comparison between the various Wh elements (figures in %)*

| wh-word		#	I**	F	IS	I	F	IS	Wh+ESK
			Children (n=33)			Adults (n=22)			
qu'/quoi	What	7	0	82	18	45	51	3	1
où	Where	7	0	94	6	68	12	8	12
quel	Which	6	0	96	4	75	17	5	3
Comment	How	5	0	95	5	64	29	6	1
pourquoi	why	4	1	99	0	53	43	0	4
quand	When	2	0	100	0	70	9	5	16
combien	How many	1	0	97	3	67	19	9	5

* Figures are given in % since each wh-element was represented by a different number of items (see the third column of table 8 for numbers of items for each element). Note that the percentages here are calculated within each wh-element, and thus do not add up to the percentages presented in other tables.

* * The stylistic inversion fixed patterns (see ft.30) of the children were not calculated in this analysis.

Comparing the adults' and children's responses across the various Wh-elements reveals a very interesting finding. Under a traditional assumption that children learn the rules of their native language from external input alone, one would expect children to show preferences similar to their parents' with respect to the different Wh-elements. In other words, if 'what'-questions in the input are most characterized by a Fronting structure, one would expect that children would show the same preference and use more Fronting (and therefore fewer in-situ structures) with this element. In statistical terms, one would expect a correlation between children's and adults' responses when these are divided across the various elements, namely that adults' Fronting responses (column 8 in table 8) would positively correlate with children's Fronting[32] responses (column 5 in table 8) and that adults' in-situ responses would positively correlate with children's in-situ responses. However, a Pearson correlation analysis reveals that these responses are *negatively* correlated: Wh-elements that received more Fronting responses from adults, received less Fronting from children, and Wh-elements that received more in-situ responses from adults received fewer in-situ from children. The correlation analysis further shows that adults' Inversion responses are positively correlated with children's Fronting responses (Pearson= 0.912 p<0.05) and adults' Fronting with children's in-situ (Pearson= 0.912 p<0.05).

These correlations are depicted in figure 1 below.   The responses to the four *pourquoi* 'why' questions are not included in the correlation analysis in figure 1, as these do not allow the wh-in-situ option (see fts. 27 and 29).

Figure 1: Correlation between adults' I and children's F and between adults' F and children's IS
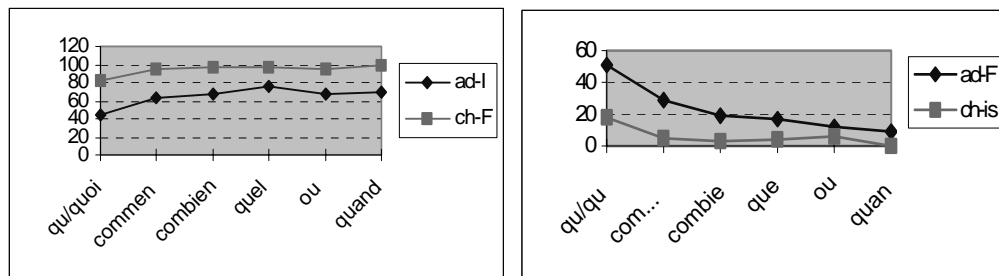
Three conclusions can be drawn from this finding. The first is that children acquire Wh-questions for each Wh-element separately, as argued by De Villiers (1991), and do not generalize from one element to the whole group. The second is that children *do* hear inversion structures in their input and take them into account when acquiring Wh-movement. The third is that children do not acquire their preferences based on external factors alone (the input), but rather integrate the input with an internal tendency toward economy. This integration of factors causes them to stay 'one step down' from the structures they hear in the input, in the following sense: an element that is observed to be identified with Fronting without inversion will be used by the children with the in-situ structure as an early strategy, while elements that are identified with inversion will cause children to move forward more quickly in their acquisition of the more marked structures. They will abandon the early in-situ structure more quickly and use the Fronting without inversion structure more frequently.

---

[32] Note that there might be a ceiling effect in the children's 'fronting' responses. However, as the point here is to compare the responses of children and adults, and as there is no ceiling effect in the adults' data, the conclusions made based on this table seem to be valid.

The general conclusion to be drawn from this correlation is therefore that children's acquisition process is characterized by an interaction of the input with their tendency toward economy.

3.6   Argument vs. adjunct questions

Another aspect that should be checked here is the division of questions into argument, adjunct, and copula/predicative questions. Plunkett (1999b) claims that this division is one of the major predictors of the choices adults and children make with regard to the various Wh-structures. Table 9 below presents the results based on this division.   Questions that did not fall under any of the 3 categories, or whose status was not clear, are labeled 'other'.

Table 9   Responses divided across question types (figures in %)

Q-type	Adults				Children		
	I	Wh+ESK	F	IS	I	F	IS
Argument (n=7)	44	1	52	3	0	84	16
Adjunct (n=18)	64	7	25	4	0	98	2
Cop./Pred. (n=4)	70	7	16	7	0	85	15
Other (n=3)	80	5	9	6	0	89	11

At first glance, looking at the data in the categories in Table 9 above seems informative, in that the type of the question does indeed seem to relate to the preferences of both children and adults.   Fronting is most common in argument questions for adults and in adjunct questionsfor children. Children's in-situ occurs most often in argument questions. As mentioned above, the only questions in which stylistic inversion was used by children (these do not appear in table 8 since they were considered to be semi-formulaic) are predicative questions.

However, the division across the various wh-questions as presented in table 8 seems to be even more informative with respect to the acquisition process than that presented in Table 9. For example, the question with the element *combien* 'how many', which was an argument question, received fewer in-situ responses by the children than the other argument questions, all of which contained the element *que/quoi* 'what'. Even more importantly, the correlation found between adults' I and children's F, and between adults' F and children's IS, does not appear in the current argument/adjunct division. This suggests that children generalize the input they

receive within the same Wh-element but not necessarily within the same question type.

3.7   Relation to age

An underlying assumption of the current experimental study is that although the children tested here were of a limited age group (ages 4-5;9), they represent a stage in a continual process. In other words, the children's preferences that were revealed in this study are seen as a sample of a changing process that progresses from the initial stage to the final- adult stage. For example, the fact that children show a greater preference than their parents for the in-situ structure is taken to show that this preference exists in younger ages as well, and furthermore, that this preference is greater in younger ages and decreases as the child progresses toward the adults' preferences. A prediction that follows from this assumption is that in the children tested in the current experiment, an association exists between the child's age and the number of in-situ structures she produces: the younger the child, the larger the number of in-situ structures. One must bear in mind, though, that the age of a child is not very reliable as a predictor of a child's linguistic knowledge in such a comparative study – the language of a specific 3-year old might exceed the language of another 4-year old. However,. other possible predictors such as mean-length-of-utterance (MLU) are not available in an experimental study such as the current one. Figure 2 below shows the association between age and the number of in-situ structures. The children are divided into two groups: below the mean age (4:7) and above the mean age.

Figure 2: Mean in-situ response divided into age groups
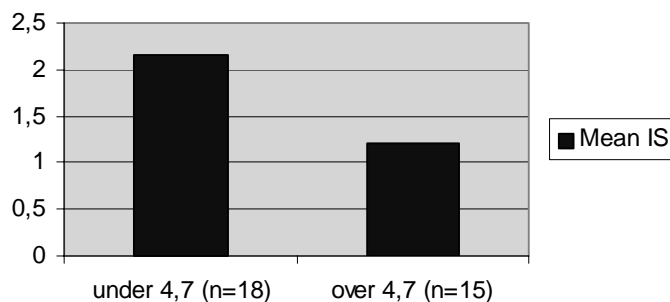
Although a correlation between age and in-situ responses does not yield a significant result (n=33,pearson=0,122, p>0.05), figure 2 shows that the younger children produced more in-situ responses than the older children. Based on the limited power of age as a predictor of progress, one can only hypothesize that such a comparison based on MLU rather than age would have yielded a significant result[33].

3.8   Comparison of the results with existing data

Spontaneous speech analysis of seven French-speaking children is reported in the literature: Philippe, (CHILDES, analyzed in Crisma 1991, Hulk 1996, and Hamann 2000), Augustin, Maria (Geneva project, see Hamann 2000), Lea, Max, Anne (York project, see Plunkett 1999b) and Fabien (Weisenborn 1991) .

Hamann (2000) discusses the acquisition of Wh-questions by Philippe (from CHILDES), Augustin and Marie (from the Geneva project). Whereas Philippe starts out with Wh-Fronting and only produces Wh-in-situ later, the two other children acquire Wh-in-situ first. This asymmetry in acquisition routes, which is most marked at the Root Infinitive stage, is taken by Hamann to indicate that different processes are involved in Wh-in-situ and in Wh-Fronting, respectively. She does not discuss the other types of Wh-questions. She only remarks that the kind of Wh-questions that Marie acquires after Wh-in-situ questions is what Hamann calls "the routine inversion" *Qu'est-ce que c'est*, at age 2;1.28. At the same time, Marie uses the first fronted Wh-question: *où il est le canard*? The fact that *qu'est-ce que* questions pattern with fronted Wh-questions is consistent with the findings of the current study (contra Plunkett 1999a).

Plunkett (to appear, 1999b) considers the acquisition of Wh-questions by three francophone children from Belgium, Canada, and France, respectively. She distinguishes four developmental stages: francophone children begin to use prepositional Wh-questions around age 2; at that time these questions are limited to questioning the predicate in a copular construction. Subjects are generally absent and the Wh-word is in situ[34]. In the second stage, subjects may continue to be absent, questions containing lexical verbs appear, and this makes available, according to Plunkett, the structure necessary for the associated appearance of

---

[33] The individual results show that one of the oldest children (5;7) produced as many as 7 in-situ responses. It is reasonable to assume that this child is less advanced than the others (i.e., has a lower MLU than her peers). Once this child is taken out of the calculation, the correlation between age and IS responses becomes significant (n=32, pearson=0.354, p<0.05).

[34] Plunkett's data show a case of one child who persistently uses null subjects in both in-situ and moved wh-questions.

object and adverbial questions. Subject questions are absent. The Wh-word is in situ. The third stage of Wh-acquisition is characterized by a general upsurge in their use. The Wh-word can both be in situ and moved: children have acquired the ability to choose. This choice is available only after firm evidence of the availability of CP structure in the child's grammar, according to Plunkett. Elliptical questions appear, as do clefts and periphrastic questions. The beginning of a fourth stage in the development of Wh-questions is marked by the productive use of periphrastic questions, according to Plunkett. The Wh+ESK structure is reported to appear in the fourth year in Plunkett's data. She states that the required structures are not all fully developed until around five years of age.

The child Lea, from Belgium, is special in that the recordings continue until a later stage than the recordings of the other two children of the York project. Moreover, she produces simple inverted questions which are absent in the other children in the L1 French acquisition literature.

In general, of these seven children only Philippe (Crisma 1991, Hulk 1996) is reported to not prefer the Wh-in-situ structure in his early stages[35]; in fact, he shows a completely opposite pattern: he begins with only fronted questions while in-situ questions emerge in a later stage and never seem to be preferred. The other six children begin with Wh-in-situ as the early strategy, while Fronting emerges later and seems to be less preferred, as it is in stages in which both Fronting and in-situ are used.

Although the children reported in this study are older than the children of the spontaneous speech analysis (they can be classified as being in stage 3 or 4 in Plunkett's terms), and although the percentages of the various structures in the current data differ from what is reported in the literature for both children and adults, the *comparative* preferences of children and adults shown by the current data are in line with the existing data. The conclusion based on combining the existing – spontaneous speech analysis- with the current experimental results, is that children begin with Wh-in-situ and no inversion, then acquire the Fronting, Wh+ESK and inversion structures, in that order. The present work shows that children work their way from the assumed initial-in-situ structures toward the adult patterns for each

---

[35] Another child who appears on CHILDES is Grégoire, but this data is rather limited. Plunkett (1998) mentions Grégoire as a child who begins with Wh-fronting rather than in-situ. A look at the Grégoire files shows, however, that in all his files together there are fewer than 20 Wh-questions; in some of them the Wh-word is fronted and in some it is in-situ. Furthermore, most of these questions seem to be repetitions of adults' speech or fixed patterns.

Wh-element separately.   The acquisition of the two movement operations (Wh-movement and inversion) is guided by a combination of the input they receive and their tendency for economy. This conclusion confirms the predictions above.

4.   Summary and conclusion

This chapter presented results from an experimental elicitation task with French-speaking children and adults. The purpose of the experiment was to elicit Wh-questions and to check the preferences of children and adults with respect to the various strategies available in spoken French. The main purpose was to compare the responses of children and adults and see whether there is a difference between the two populations.

The results revealed the following findings:

- Children used no inversion structures while adults used them frequently
- Children use no Wh+ESK structures but frequently use KESK structures (in 'what' questions).
- Children use more in-situ structures than adults
- Children do not make form-errors with respect to the various forms available for the Wh-element 'what' and do not make structure errors with respect to the limitation on the element *pourquoi* 'why'
- The choices of both adults and children with respect to the various options is related to the nature of the Wh-element used and to the specific question presented by the experimenter
- Children's 'Fronting' structures appear with the same Wh-element that adults use with inversion, and children's in-situ responses appear with the same Wh-element that adults use with 'Fronting'

These findings led us to the following conclusions. The current data, as well as previous investigations, indicate that French-speaking children prefer the Wh-in-situ in the early stages relatively to its frequency in their input. The other structures appear in the following stages: Fronting without inversion, Wh+ESK, and inversion (cf. Weissenborn 1991, Plunkett 1999a). Although this pattern is not entirely

without exception (cf. the data on Phillipe [36]), a clear majority of the children tested conform to it, which lends support to this conclusion.

The stages proposed above for the acquisition of Wh-questions in French reflect an economy-based hierarchy in which the in-situ is the least marked structure (since it is the most economical), and inversion is the most marked. Thus, children seem to be sensitive to this hierarchy and to hold a 'rule of thumb' that directs them to not move to the next stage before sufficient evidence is gathered. A similar proposal for accounting for child language through the notion of economy of derivation is made in Van Kampen (1997). She claims that children acquire language in a hierarchical fashion, beginning with structures that require less movement and represent a smaller discrepancy between PF and LF representation. Although some of the assumptions made in this chapter are different from those in Van Kampen (1997), the data reported here support her predictions in full. As noted earlier, the notion of economy adopted in the current approach differs from the notion of economy in Van Kampen (1997). In this study, the notion of economy is seen as a syntactic principle used by children in the task of acquisition, while in Van Kampen it is described as a performance strategy rather than a principle of U.G. Furthermore, the current approach was described as 'global', as it relates only to cases of "optional" structures that are all grammatical in the target grammar, while a reduction of PF-LF discrepancy belongs to what we called the 'local' approach and also relates to structures that are obligatory in the adult grammar. It seems, however, that the predictions of the two approaches converge in the current case of Wh-questions in French.

 The fact that somewhat older children still show a greater tendency than their parents to prefer the more economical structures, implies that for younger children the choice of the 'cheaper' option is also the result of a strategy and not due to incapability as might be proposed, with respect to the present data, by several current theories (e.g., truncation: Rizzi 1995; merger: Roeper 1996, Powers 1996; underspecification: Wexler 1994; lack of functional projections: Radford 1996; etc.). With a mean age of 4.7, the children in the current study cannot be said to be in a stage characterized by a less than fully projected phrase marker, but nevertheless

---

[36] Interestingly, Phillipe, who produces Fronting structures from the early stages, uses no other structures in this period. Like the other children, that is, he uses a single structure for his early questions. Such an exception to the general pattern (namely, subjects who produce the marked option first) is predicted in Clark and Roberts (1993). These authors hypothesize that a certain percentage amount of "mutation" with respect to parameter setting must exist in order to ensure successful acquisition.

they still reflect a tendency for economy. This conclusion therefore offers an indirect argument for a continuity approach. It shows that while approaches that assume a non-full phrase marker can account for early stages alone, the current economy-based approach can account for both the early and the late stages of acquisition of Wh-questions in French.

An additional conclusion is that children treat each Wh-expression separately (cf. De Villiers 1991). When sufficient evidence is gathered for a specific Wh-element, the children will move to the next stage with respect to this element, but will not generalize to the other elements. Such lexically-based acquisition is also shown for the acquisition of V-movement in Dutch. In Wijnen (1999), children are shown to move certain verbs in the early stages, while for other verbs movement is attested only later.

The final, general conclusion of this work is that acquiring an "optional" operation such as Wh-movement and subject-verb inversion in French involves balancing two factors: the input and the tendency for economy. The contribution of the current chapter, therefore, to the general discussion of this dissertation, apart from supporting the general proposal, is the demonstration of the interaction between economy as an internal tool of acquisition on the one hand and the input on the other.