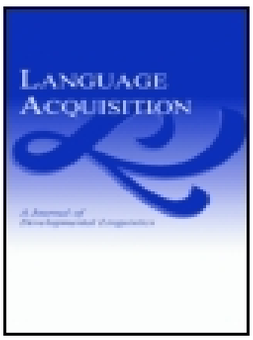

Language Acquisition

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hlac20

The acquisition of *wh*-questions: Beyond structural economy and input frequency

An D. Nguyen & Geraldine Legendre

To cite this article: An D. Nguyen & Geraldine Legendre (2021): The acquisition of *wh*-questions: Beyond structural economy and input frequency, Language Acquisition, DOI: 10.1080/10489223.2021.1968867

To link to this article: https://doi.org/10.1080/10489223.2021.1968867

Published online: 28 Sep 2021.

LANGUAGE ACQUISITION
https://doi.org/10.1080/10489223.2021.1968867

Routledge
Taylor & Francis Group

The acquisition of *wh*-questions: Beyond structural economy and input frequency

An D. Nguyen and Geraldine Legendre

Johns Hopkins University

ABSTRACT

We present in this article corpus analyses, two experiments, and a preliminary English-French comparison on children's acquisition of *wh*-in-situ. Our examination of 10,000 *wh*-questions from CHILDES reveals that the reported empirical picture of *wh*-question acquisition in English is incomplete: A type of *wh*-in-situ, probe questions (PQs), has been left out from most discussions despite its presence in child-directed speech. Unlike *wh*-in-situ echo questions (EQs), PQs are used to request new information, and parents frequently use PQs and fronted information-seeking questions in alternation. The fact that PQs share the pragmatic space with fronted *wh*-questions while involving fewer syntactic operations and exhibiting lower input frequency allows us to test both structure-based and frequency-based theories of syntax acquisition. Our comprehension task with 3;06–5;06-year-olds confirms that children accept and understand PQs as information seeking. On the other hand, results from a production task show a strong avoidance of *wh*-in-situ, which is in line with reported elicited data from French-speaking children. We reason that a structural economy-based approach alone is not sufficient to account for children's disfavor of *wh*-in-situ. Depending on the input frequency and consistency, as well as the number of variants licensed by the grammar of a given language, children may treat part of the input as uninformative and initially only learn from higher-frequent, more regularized input. Their intake is thus selective.

ARTICLE HISTORY
Received 30 July 2020
Accepted 06 August 2021

1. Introduction

The acquisition of *wh*-questions occupies a special place in studies of syntactic acquisition, most likely because *wh*-questions are central to developments in linguistic theory (Roeper & de Villiers 2011) and because high error rates in the production of *wh*-questions by English-speaking children as old as 5 have been found (Bellugi 1965), which raise important developmental questions. A variety of theories has been proposed to account for the pattern(s) found in *wh*-question acquisition, taking into account the input available to children. However, the basic empirical picture in English is still incomplete: A specific type of in-situ *wh*-questions, despite its presence in the input, has been ignored and left out in most, if not all, previous discussions: probe questions. Probe questions (PQs) can be thought of as a "fill-in-the-blank" type of question that are pragmatically similar to fronted information-seeking questions in the sense that they request new information, yet they have an in-situ surface structure similar to echo questions. For example, while reading a story to her child, a mother can stop and ask: "On Friday, he ate through five oranges, but he was still *what*?" (example from the HSLLD corpus: Dickinson & Tabors 2001).

Although English in-situ *wh*-questions are sometimes discussed in the acquisition literature (e.g., Takahashi 1991; Becker & Gotowski 2015), it is widely assumed that they are only grammatical as echo questions (EQs). EQs are generally used to ask for a repetition or a clarification of

CONTACT An D. Nguyen an.nguyen@jhu.edu Johns Hopkins University, Cognitive Science, 3400 N Charles, Baltimore, MD 21218.

a previous utterance (Authier 1993), which separates them from general information-seeking questions. Since the syntactic difference between fronted information-seeking questions and echo questions is paired with a pragmatic difference, it is unclear whether we can conclude much about the acquisition of syntax in these *wh*-questions. Thus, the presence of a syntactically different (in-situ) but pragmatically related (information-seeking-like) structure like PQs is invaluable as it can provide more insights into the acquisition of *wh*-fronting versus in-situ in particular and of *wh*-questions in general.

Investigating the acquisition of PQs is likely to shed light on theories of syntactic acquisition and contribute to an evaluation of two dominant existing models, one grounded in structural economy[1] versus another grounded in input frequency. The former has been particularly influential in discussions of the acquisition of *wh*-questions in French, which allows both fronted and in-situ strategies, among others (e.g., Jakubowicz & Strik 2008; Hamann 2006; Zuckerman & Hulk 2001). Drawing on the idea of structural economy (Chomsky 1995), Jakubowicz (2005, 2011) in particular proposes a Derivational Complexity Hypothesis, according to which derivational complexity conditions the course of syntactic acquisition. The Derivational Complexity Metric is defined as follows:

(i) A. Merging α_i n times gives rise to a less complex derivation than merging α_i $(n + 1)$ times.
B. External Merge of α gives rise to a less complex derivation than Internal Merge of $\alpha + \beta$.

In particular, the Derivational Complexity Hypothesis predicts that children will avoid structures involving more syntactic operations unless those are obligatorily required, resulting in a single syntactic option. The Derivational Complexity Hypothesis is an influential hypothesis frequently discussed in the study of *wh*-questions acquisition (e.g., Yuan 2015; Durrleman, Marinis & Franck 2016; Prévost, Strik & Tuller 2014; Hopp, Putnam & Vosburg 2019) as well as other phenomena (e,g., acquisition of Differential Object Marking: Cuza et al. 2019; acquisition of object and quantitative pronouns: van Hout, Veenstra & Berends 2011).

With respect to *wh*-questions, economy-based theories like the Derivational Complexity Hypothesis predict that children will prefer in-situ PQs over fronted *wh*-questions, given that both PQs and fronted questions can function as information-seeking questions and PQs are structurally simpler, with neither overt *wh*-fronting (movement to SpecCP) nor auxiliary inversion (T to C movement). We are leaving aside the question of whether PQs involve covert *wh*-movement, as covert (or LF) movement is said to impose less cost than overt movement in fronted questions (Chomsky 1993, 1995) regardless.

The presence of PQs in child-directed speech implies that the input for questions that children receive is more diverse than it has been previously thought. Given that diversity in input can have a number of effects on children's production and comprehension (Goodwin, Fein & Naigles 2015), we expect that including PQs will also give us more insight into the general acquisition mechanism. One important issue in learning is concerned with the amount of information children actually learn from the input: Do they learn everything they hear, or do they learn selectively, necessitating a distinction between input versus intake (e.g., Gagliardi & Lidz 2014; Omaki & Lidz 2015)? Hudson Kam & Newport (2005, 2009) find that in the context of learning an artificial language with inconsistencies, children almost always regularize the inconsistent forms and adults do not. Singleton & Newport (2004) also find a similar regularization pattern in Simon, a deaf child whose ASL input contains many inconsistencies and errors, yet his production is much more regularized and indistinguishable from children learning from native signers. Children's regularization typically happens in the direction of the more frequent option (Schwab, Lew-Williams & Goldberg 2018)—that is, when being exposed to multiple variants of a grammatical item, adults tend to match the input frequency while children tend to boost the frequency of the more frequent variants. Perkins, Feldman & Lidz (2017) and Schneider, Perkins & Feldman (2019) apply the idea of an "input filter" to computationally model the acquisition

[1]We do not consider structural accounts that rely on intervention effects such as Featural Relativized Minimality (Friedmann, Belletti & Rizzi 2009) because the present study is limited to simple object (*what, who*) and adjunct questions (*where*). See Table A1 for details.

of verb transitivity and English determiner agreement respectively. Both papers find that by allowing an assumption that part of the data is "noisy" (generated by error), the model learns better than a no-filter model. By treating the lower-frequency variants of a grammatical item as "noise" and filtering them out, the learner arrives at a more regularized and consistent grammar. The disproportionate distribution of fronted *wh*-questions in child-directed speech (>80%) in comparison with PQs (>10%) provides a good test case for such a hypothesis, which predicts that children will only regularize to high-frequency fronted *wh*-questions. As a result, input filter models make the opposite prediction from economy-based accounts: When there are two variants of a grammatical item—simple low-frequency PQs and more complex high-frequency fronted questions—children will initially prefer, and regularize to, fronted *wh*-questions in production.

A study on *wh*-questions acquisition that includes PQs therefore has the potential to generate insights both for the field of child language acquisition in general and for the field of syntactic acquisition specifically. In this article, we will first briefly summarize the characteristics of PQs (section 2). We then provide new corpus evidence that PQs are commonly used in child-directed speech but rarely spontaneously produced by children (section 3). In section 4, we present results from an experimental study to (i) reinforce the claim that children interpret in-situ PQs not as EQs but as information-seeking questions and (ii) assess whether children are willing to produce such constructions given a pragmatically appropriate setting. Section 5 provides a brief comparison of our results with results from existing studies on the acquisition of *wh*-questions in French-speaking children. We revisit economy-based accounts versus the input filter hypothesis in our general discussion of all results in section 6. Section 7 concludes the article.

2. A characterization of PQs

PQs and EQs are both in-situ *wh*-questions on the surface:

(1) a. (PQ)
A: That's a what?
B: A computer.
b. (EQ)
A: That's a computer.
B: (That's) a WHAT?[2]
A: A computer.

However, PQs and EQs are different across many linguistic dimensions, including pragmatics, phonology, and syntax. Pragmatically, EQs obey a strict linguistic context requirement. As Banfield (1982) has observed, EQs can only occur as a reaction to a prior utterance. The strict context requirement, plus the specific purpose of EQs as a request for clarification or repetition, leads to a strong presupposition that the addressee knows the answer and can provide it when asked. For example, (2b) and (2c) are infelicitous responses to the EQ in (2a). The response to an EQ must be the original utterance, a synonym of the original utterance, or at the very least, a description that is coreferential with the original utterance (Blakemore 1994).

(2) a. A: Jimmy just bought an accordion yesterday.
B: Jimmy just bought a WHAT?
b. A: #I don't know.
c. A: #maybe an accordion?

Instead of asking for repetition or clarification of a previous utterance, PQs can be thought of as a "fill-in-the-blank" type of question: The addresser prompts the addressee for a piece of information by providing the base structure of the answer with a blank slot for the addressee to fill in. Due to this pragmatic characteristic, PQs are frequently used in particular discourse contexts, such as quiz shows (Comyn 2013), courtrooms, classroom, and child-directed speech—our target of interest. These are

[2]We will systematically represent EQs with a *wh*-phrase in caps for ease of identification.

situations in which the addresser is more interested in assessing the addressee's knowledge than in the answer itself. In the examples in (3) and (4), it is clear that the addressers already knew the answer before they asked the PQ(s).

(3) (ABC Nightlin e show, 1994, from COCA[3] corpus)
Teacher: I need to know about displacements. They have a what?
2nd student: Distance.
Teacher: They have a fixed distance and fixed what?
3rd student: Direction.
Teacher: And fixed direction. Fixed distance and fixed direction. Kim, number three. Tell us what you have, Kim. A displacement of how many?

(4) (excerpt from Adam, Brown corpus)
Mother: and he had a sister named what?
Child: Tony.
Mother: no, Tony was the little baby. His sister's name was what?
Child: [. . .] I don't know.
Mother: Her name is Sheila.

Since the addresser is requesting new information that has not been previously mentioned in discourse, it is perfectly acceptable if the addressee does not know the answer, unlike in the case of EQs. In (4), the first answer the child provides is wrong, and the final answer is "I don't know." There is no restriction on the possible set of answers that the child can consider, as long as the answers do not digress from the main question. This is similar to information-seeking questions. Thus, based on answers alone, PQs are functionally more similar to fronted information-seeking questions, as answers to both types convey new information.

With regard to prosody, while EQs have a distinctive intonational pattern, consisting of a rising pitch accent and heavy stress on the *wh*-phrase (Authier 1993), PQs have a flat or even falling pitch accent, similarly to information-seeking *wh*-questions (Reis 2012). We examined the duration and F0 characteristics of the *wh*-word's[4] vowel in 50 EQs and PQs extracted from three CHILDES audio corpora in English: HSLLD (Dickinson & Tabors 2001), Van Houten (1986), and Weist (Weist & Zevenbergen 2008), to confirm this result. The questions were forced-aligned using the Montreal Forced Aligner (McAuliffe et al. 2017) and analyzed using the PRAAT software (Boersma & Weenink 2019). As shown in Figure 1, the *wh*-word pitch contours of EQs and PQs follow opposite directions.

In brief, although PQs are more similar to EQs in surface structure, their pragmatic use is more similar to that of fronted information-seeking questions. PQs are distinct from EQs and should not be grouped under or together with EQs. Rather, they are a subtype of information-seeking questions. In child-directed speech, as we will see next, they are a second syntactic option available to adult caregivers.

3. Corpus studies

3.1. *Child-directed speech*

The presence of *wh*-in-situ questions in child-directed speech is occasionally noted in the child language-acquisition literature: For example, Becker & Gotowski (2015) report that 16% of all *wh*-questions produced by adults in Eve's data (Brown 1973) are *wh*-in-situ, and Gotowski (2017)

[3]The Corpus of Contemporary American English (COCA) is a widely used corpus of American English containing 1.1 billion words.
[4]Since the *wh*-words in almost all of the samples were in the final position of the sentence, it is impossible to tell whether the differences between EQs and PQs arise from the stress on the *wh*-word or from the sentence-final prosody. Having PQs or EQs with sentence-medial *wh*-word or multiple *wh*-words would address this problem; however, such sentences are rare in corpora. As most studies on EQs have analyzed their unique intonation in terms of their stressed *wh*-word, we suggest that the differences between PQs and EQs emerge from the *wh*-word instead of sentence-final prosody.

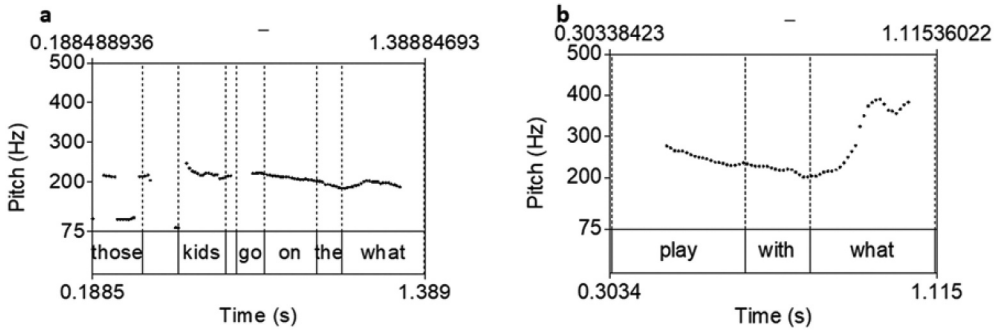

Figure 1. (1a) Example pitch track of a PQ; (1b) Example pitch track of an EQ.

similarly reports 22% of *wh*-in-situ questions in Adam's input data. However, to the best of our knowledge, there has been no attempt to differentiate between in-situ EQs and in-situ PQs. The in-situ structures are typically all assumed to be EQs. Given that PQs and EQs are independent of each other and should not be grouped into the same category, it is crucial to evaluate their respective presence in child-directed speech. We conducted our own corpus analysis to offer a more accurate picture of *wh*-in -situ English questions in this context.

To get an estimation of the percentage of *wh*-in-situ questions in the input, one data file for each month between 2;00 and 4;00 years of age of 10 children was randomly selected, including Sarah and Adam: Brown corpus (Brown 1973), Shem: Clark corpus (Clark 1978), Trevor: Demetras corpus (Demetras 1989), Abe: Kuczaj corpus (Kuczaj 1977), Lily and William: Providence corpus (Demuth, Culbertson & Alter 2006), Naomi: Sachs corpus (Sachs 1983), Roman: Weist corpus (Weist & Zevenbergen 2008), and Laura: Braunwald corpus (Braunwald 1985).

Using the CLAN tool (MacWhinney 2000), we first extracted all *wh*-questions from child-directed speech in the selected files. Since we are only interested in cases in which an in-situ utterance is possible, we further excluded subject *who/what* questions (whose analysis is ambiguous between in-situ and vacuous fronting), embedded *wh*-questions, "what-if," "how come," and "what about" questions and limited our search to only main clause questions, resulting in a total of 9,039 questions. Questions that do not carry any other piece of information besides the *wh*-phrase (e.g., "*now what?*," "*because why?*," "*so what?*," "*for what?*") or expressions that are not actually used as genuine sentential questions such as "*you know what?*" were also excluded. From the extracted data, we recovered a total of 1,361 in-situ questions, which take up 15.1% of all main clause questions in this sample of child-directed speech. The percentage of in-situ questions in the input varied among children. Out of 10 children, two received more than 20% of in-situ questions, five received between 10% and 20%, and three received less than 10%.

To classify whether each in-situ question is a probing or an echo question, we examined the surrounding utterances in close detail. For example, although it is not clear enough to tell if the in-situ question in (5a) "*It's a what*?" is an EQ based on the prior utterances alone (as the structure of the question does not match exactly with what the child said previously), we can rely on the mother's response to classify this *wh*-question as an EQ. When there was not enough information to uniquely infer whether a question should be an instance of EQ or PQ, it was regarded as ambiguous. (5b) is an instance of an ambiguous in-situ question. The adult and the child were looking at a book. The adult's in-situ question could be an EQ given that its structure matches with the previous utterance of the child. However, the child's answer was not a repetition or clarification; moreover, the fact that the adult was asking the child about details in a book also suggests that the adult knew the answer, and this could be a follow-up question to lead the child to describe the scenario in the book.

(5) a. (excerpt from Adam, Brown corpus)
Cousin:[5] is it a square?
Child: no square, is clown.
Cousin: *it's a WHAT?*
Child: Mommy
Mother: it's a clown, he said.
b. (excerpt from Roman, Weist corpus)
Adult: Look at this bunnies. What do you think is happening?
Child: They are gonna catch [. . .].
Adult: They're gonna catch what?
Child: They're gonna hide in creek.

In general, EQs appear more frequently than PQs in phrasal questions (e.g., "*did what?*" or "*little what?*"). However, when considering only full sentential questions (e.g., "*it is a what?*"), PQs appear more frequently (54.6%) than EQs (33.2%). The results are summarized in Table 1. In general, children gave appropriate answers to these questions, indicating that they understood these structures. In particular, children responded appropriately to EQs about 90% of the time, though occasionally (about 10% of the time) they ignored the question and gave no answer.

It is important to note that caregivers frequently alternate between PQs and fronted questions and use them in child-directed speech as if they are interchangeable, as shown in the following example from the Weist corpus. In (6) the father keeps restating the question over and over again to get the answer he wants ("the baby is going to cry"). This further supports the claim that PQs and fronted information-seeking questions are closely related, pragmatically speaking, at least in child-directed speech.

(6) Father: hey Roman, if the dinosaur roars what's the baby gonna do?
Child: it gonna roar and it's gonna say like this (roar).
Father: yeah but if the dinosaur roars the baby is gonna be what?
Child: scared.
Father: no the baby's scared what's it gonna do?
Child: it gonna eat the thing.
Father: no no the baby, the baby's gonna what?

3.2. *Child production*

The search process for *wh*-in-situ questions in child production was similar. Data files for each month between 2;00 and 4;00 years of age of the same 10 children were randomly selected, using the same inclusion and exclusion criteria. Our search returned a total of 10,241 *wh*-questions, 407 (3.9 %) of which were in-situ *wh*-questions. However, most of these in-situ utterances were not genuine questions but corresponded to either a fixed expression such as "*for what?*" (31 counts, 7.6%) or an expression that was not actually intended as an information-seeking or echo question such as "*you know what?*" or "*guess*

Table 1. The distribution of in-situ *wh*-questions in child-directed speech.

	FULL SENTENTIAL	PHRASAL
EQ	227 (33.1%)	290 (43%)
PQ	369 (53.7%)	242 (35.9%)
AMBIGUOUS	91 (13.2%)	142 (21.1%)
TOTAL	687	674

[5]The "cousin" is an older child only producing adult-like utterances throughout the corpus.

what?" (266 counts, 65.4%). Children sometimes also asked questions and answered them themselves (e.g., "*they buy some more scrambled eggs. Three what? Three scrambled eggs*"). Such utterances were excluded from the analysis, leaving 75 in-situ questions (i.e., less than 1% of all *wh*-question production). The final result is summarized in Table 2. It is worth noting that the majority (52 counts, 70%) of these utterances came from a single child, Adam, possibly due to his unique variant. The rest of the children produced rather few in-situ questions.

Many of the in-situ questions were ambiguous (e.g., Mom: "*that's ocean*"; Child: "*ocean what?*"), as it is not clear whether the child was simply repeating the last word his mom said (despite adding a *wh*-phrase) or he was genuinely asking for clarification (the mother usually did not give a clarifying answer). Some of the utterances were EQs:

(7) Mother: He was talking about President Kennedy.
Child: Talking about WHAT?

There were also occurrences of PQs.[6] For example, in the following occurrence, to ask his mother about the new object that wasn't mentioned in the text previously, the child used an in-situ question:

(8) Child: Mommy, this is a what?
Child: It's a what?
Mother: Paper punch.

Summing up, out of all genuine in-situ questions, children produced more PQs than EQs. However, the number of such utterances is very small—less than 1% of all *wh*-question utterances and 70% of them produced by a single child. Overall, our results confirm the claim that children rarely produce in-situ *wh*-questions (Valian & Casey 2003; Becker & Gotowski 2015).

4. Experimental studies

Our corpus analyses show that children rarely produce PQs spontaneously; however, they are able to respond appropriately when adults use PQs. We conducted two experiments to confirm the basic results from the corpus studies—that (i) children accept and understand PQs as information-seeking questions (comprehension study), and (ii) children do not produce PQs by themselves, even in an appropriate pragmatic setting (production study).

4.1. *Experiment 1—Comprehension study*

Given the assumption in the language acquisition literature that English-speaking children only understand in-situ *wh*-questions as EQs (e.g., Takahashi 1991; Becker & Gotowski 2015), our comprehension task sought to investigate whether children would be able to differentiate these two types on *wh*-in-situ (i.e., give repetition to EQs and new information to PQs). We leave out

Table 2. The distribution of in-situ questions in child production.

	FULL SENTENTIAL	PHRASAL
EQ	2 (16.7%)	14 (22.2%)
PQ	6 (50%)	21 (33.3%)
AMBIGUOUS	4 (33.3%)	28 (44.5%)
TOTAL	12	63

[6]Adam frequently produced a PQ immediately after a fronted *wh*-question, most likely mirroring the adult behavior mentioned in (6). However, adults typically rephrase the original question into an in-situ PQ only if the child fails to answer the fronted one. Adam, on the other hand, did not wait for a response. This suggests that Adam may have used PQs in a different way compared to adults, as he asked these questions without knowing the answer.

(i) Child: What is that?
Child: It's a what?
Researcher: I don't know what it is, do you?

fronted *wh*-questions in this task because their inclusion would lead to a mismatch in conditions (three types of *wh*-questions but only two types of answers, repetition versus new information), which can induce a response bias. Note that young children as young as 20 months of age can already demonstrate above chance accuracy in simple *wh*-questions comprehending tasks (Seidl, Hollich & Jusczyk 2003).

4.1.1. *Method*

4.1.1.1. *Participants.*

Twenty children were recruited for the study. All of them were native English speakers. The mean age of the children was 4;01 (range: 3;06–5;06; 7 boys, 13 girls). Of them, one child was excluded due to an unusually high number of irrelevant answers, and two children were excluded due to failure to follow instructions.

Fourteen adults were additionally recruited to serve as a control group. Of them, two were excluded because they were outliers, i.e., their scores were three standard deviations away from the mean score of the sample. This left us with 12 adult participants (age range: 19–24; four males). All of them were university students.

4.1.1.2. *Materials.*

4.1.1.2.1. *Task design.* The context of 12 scenarios making up the experiment was explicitly specified as a classroom-like setting involving a participant, a storyteller (research assistant), and an alien classmate (the experimenter). The role of the alien was to comment on the stories as the storyteller told them. Each scenario led to a target question. In total, there were 12 *wh*-in-situ questions (six PQs and six EQs). Of them, four were "*what*"-questions, another four were "*who*"-questions, and four were "*where*"-questions; each *wh*-word appeared twice in PQs and EQs. No subject *wh*-question was included in the experiment, given the ambiguity of analysis (in situ/vacuous fronting). Each question had three possible answers, including a target, a nontarget, and a wrong/irrelevant answer. In each scenario, the two characters Bill and Jill would pass by an event but did not get to observe the full development of it. They talked to each other about the event, with one character saying: "*I wonder [what happened]*." The alien classmate then turned to the participant and whispered what he thought had happened. The alien's opinions, however, always violated Grice's Maxim of Quantity by being underinformative. The alien would give a description that matched both the target answer and a nontarget answer, essentially narrowing down the choices from three to two, but not enough to uniquely identify the target answer. Half of the time, the storyteller was able to hear the alien and acknowledged his answer, and he would turn to the participant to ask for his/her own answer using a *wh*-in-situ structure (PQ condition). Half of the time, the storyteller noticed the alien was saying something but could not hear it clearly, and he asked for the participant's help for clarification using a *wh*-in-situ echo question (EQ condition).

A sample scenario with illustrations is provided in Figure 2. In an echo trial, the target answer would be "*the white building*," the nontarget answer would be "*the hospital*," and the irrelevant answer would be "*the apartment/the library*." In a probing trial, the target answer would be "*the hospital*," the (underinformative) nontarget answer would be "*the white building*," and the irrelevant answer would be "*the apartment/the library*."

Two question fillers were included to keep children engaged. In each filler, three possible choices were also introduced; however, there was no right answer and no visual or auditory cues about which option should be chosen. For example, in a filler, three types of drinks were shown on the screen as the storyteller said to the participant: "*Bill is very thirsty on this hot summer day. Can you choose a drink for Bill?*"

4.1.1.2.2. *Pragmatic considerations.* In a PQ trial, we expect that participants will generally prefer to be informative and precise (Grice 1975) and thus would give the target answer over the alien's underinformative answer when being quizzed by the storyteller. At the same time, in an EQ trial,

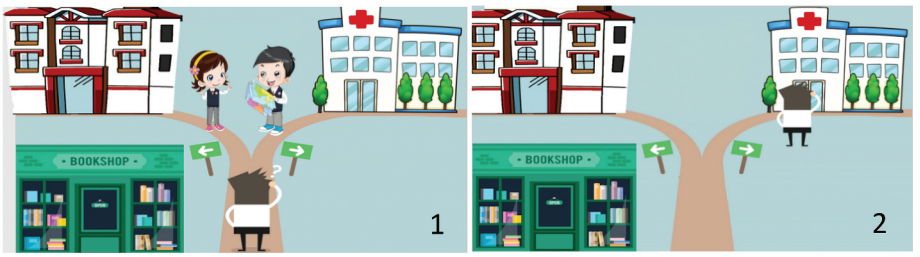

Figure 2. Trial example in the comprehension task.
Billy and Jilly are standing in the middle of the road. To their right is a hospital, to their left is an apartment, and going down is a library[7]. While waiting for Billy to read the map, Jilly sees a man appear. He looks at all the buildings carefully, as if he is trying to decide which one to go to. However, by now Billy has figured out the map: "Let's go up that way," he says, and the two kids walk away. After they have gone, the man finally walks into one of the buildings. But Billy and Jilly do not see this. On their way, they talk to each other about the man.

Jilly says: "I wonder where the man went".
Alien puppet: "I think the man went to the white building".
Storyteller: "The man went where/WHERE?"

participants should know that the storyteller is interested in what the alien has said instead of their own opinion; thus they can only repeat the alien's answer even though they may perceive it as being underinformative.

Typically, the original speaker (i.e., the alien) should be the one to respond to the storyteller's EQ instead of a third party (the participant). Therefore, we opted to maximize pragmatic plausibility by adding more details to the setup. In the first practice trial for EQ, the storyteller reminded the participants that the alien liked to voice his (unsolicited) opinions, but only to the participant. The participant was told that the alien was afraid of the storyteller and only talked to the participants, refusing to answer the storyteller's questions directly. Sometimes the storyteller could hear the alien and sometimes not, but the storyteller was interested in everyone's opinion and specifically wanted to know what the alien said. The alien would hide behind the participant when he heard the storyteller ask the EQ. The storyteller then asked the participants to help the alien out whenever he got shy. This EQ scenario was designed to reflect a typical experience in a classroom setting—the teacher tells a story and asks a question, and someone in the back responds but not loud enough for the teacher to fully hear the answer. The teacher then asks an EQ and someone else in the front who heard the answer can repeat it for the teacher. Most child participants responded correctly in their first attempt, with some responding correctly to the EQ even before the storyteller started explaining further about the alien, indicating that they understood the objective of an EQ.

Pragmatic cues, including hand gestures (gesturing toward the participant in a probing trial or putting a hand to ear in an echo trial) and cue words appropriate to a classroom setting ("*Class*, [PQ]" or "*Hmm*, [EQ]"), were included in certain fixed trials to increase the pragmatic plausibility of the task and make the questions more natural sounding. The cues for each type of questions were controlled so that they matched in number (e.g., hand gesture was used for one PQ and one EQ) and properties (e.g., the cue words had the same length). We will return to the use of such pragmatic cues in the Discussion section.

4.1.1.2.3. *Prosody of the questions.* The questions were not recorded and instead were asked directly by the research assistant playing the storyteller role for pragmatic reasons: It would be pragmatically implausible if the whole story is told by the storyteller in his natural voice, but whenever he asks a question, the question is in a recorded voice played through the computer. Recording the whole story

[7]In the experiment, we used animation effects to display the building one by one so that it was clear to the child which building was which. Assuming that 4-year-olds cannot read yet, a "bookstore" drawing was chosen to represent the "library" to maximize illustration as "library" drawings are typically a generic building with no books shown.

would solve the problem of potential inconsistency, but it would have made the story less engaging and less interesting to young children. Thus, we had both the story and the questions in the storyteller's natural voice. Since prosody served as an important factor that participants could rely on to distinguish the two types of questions (an EQ has an exaggerated pitch accent on the *wh*-phrase while a PQ does not), we examined the possibility that the storyteller might not reliably produce exactly the same prosody for the same question in every trial and experiment. We conducted a post hoc acoustical analysis using the Praat software (Boersma & Weenink 2019), excluding from data analysis any trials in which the prosody of the question is significantly different from the rest of the sample. Two data points (out of 170) of EQs in the child experiment were thus excluded from the final analysis due to inconsistent prosody. Since each participant provided 10 data points, we ended up with a total of 168 data points for children and 120 data points for adults.

4.1.1.3. ***Procedure.*** The whole task took approximately 20 minutes. Participants were explicitly told that they were in a classroom where they would be listening to a story along with an alien "classmate" named Terry. The storyteller was played by a research assistant, and the alien classmate puppet was played by the main experimenter. Participants were told at the beginning that the alien was scared of the storyteller, and he would only talk to the participants but not the storyteller. Illustrations of the story were shown on a big TV screen in the testing room. Participants were directed to pay attention to the illustrations.

At the end of each scenario, one cartoon character would raise a problem. The alien would whisper to the participants his thought about the problem. Following that, either a *wh*-in-situ EQ or a *wh*-in-situ PQ would be asked by the storyteller. When the participants had finished answering the question, the story continued to the next scenario. To keep the child participants engaged in the task, the storyteller would reward them with a sticker after every four questions.

The first two scenarios of the story were used as practice trials (one with a PQ and one with an EQ). In both practice trials, if the participants responded incorrectly, the storyteller would try to guide them to the target answer by providing hints or suggestions, though never explicitly corrected them by giving out the target answer. Such feedback was only given in the practice trials. The data from the two practice trials were not included in the analysis, resulting in 10 answers per participant.

4.1.2. *Results and discussion*

Participants' responses were divided into three categories: target, nontarget, and irrelevant/wrong answers. A target answer means that the participants gave the right information to the right type of question (e.g., repeating the alien's answer in an EQ trial). A nontarget answer means that the participants gave the right information to the wrong type of questions (e.g., repeating the alien's answer in a PQ trial). An irrelevant/wrong answer means that the participants gave wrong information (e.g., saying the man went to the library when he actually went to the hospital). The distribution of the answers by category is shown in Table 3. Children only gave irrelevant answers 4.8% of the time, which indicates that they were able to understand and pay attention to the story. They also provided almost twice as many target answers (61.4%) as nontarget answers (33.7%).

However, children were not as good as adult controls in interpreting the intention of the two types of questions. Although the overall adult performance was not perfect (91.7%), more than half (7 out of 12) of the adult participants achieved perfect accuracy. The other five adult participants made errors but only toward the end of the task, which could be a result of a loss of attention due to the task being childish and overly easy for them. In contrast, the performance of the child participants ranged from 50% to 77.78% (or 50% to 80% *target* answers when excluding wrong answers), with none of them ever achieving perfect accuracy. The data are represented in Figure 3.

Excluding irrelevant answers, children otherwise correctly produced target answers over nontarget ones 64.5% of the time. None of the child participants consistently produced only one type of answer to all 10 questions throughout the experiment. In other words, every child used both types of answers

(echo-appropriate and probing-appropriate) at least once. Given the relatively small sample size, nonparametric Wilcoxon tests were conducted. The test showed that children correctly produced the target answer significantly above 50% chance level (p = .001, effect size = 0.80). However, the results varied within each type of questions. Children performed significantly better with PQs than with EQs (M_{PQ} = 76.6% and M_{EQ} = 51.3%, p < .001, effect size = 0.97) (Table 4). While their accuracy was significantly above 50% chance for PQs (p < .0001, effect size = 1.2), it was only at chance level for EQs (p = .67, effect size = 0.1) This disparity was not observed in the adults' performance. Adults were equally good at inferring the intention of PQs and EQs (Ms = 91.7%). Adults also outperformed children both overall (p < .0001, effect size = 1.08) and within each type of questions (ps < .01).

Although children seemed to struggle slightly more with object "who"-questions, a Kruskal-Wallis test suggests that there was no significant difference in the performance within each subtype of *wh*-questions of both adults, H(2) = 2.56, p = .28, and children, H(2) = 1.47, p = .48. Figure 4 illustrates this.

Finally, we submitted the child data to a logistic mixed-effect model using the lme4 package in R (Bates et al. 2015). The dependent variable was the Accuracy of each question. Age, Question Type (probing vs. echo), Pragmatic Cues (hand gesture, cue words, or none) were included as fixed factors. Participant and Question Item were included as random factors. The R syntax for this was: Accuracy ~ Age + Type + Cue + (1 | ID) + (1 | Item).

There was a significant effect of Question Type (β = 1.29, SE = 0.37, p < .001) but no effect for Age (β –0.06, SE = 0.23, p = .80), hand gesture (β = 0.27, SE = 0.6, p = .66), or cue words (β = 0.46, SE = 0.68, p = .5). This suggests that the random inclusion of pragmatic cues did not make any trials significantly easier than others.

Table 3. Distribution of the answers by category.

	Target answers	Nontarget answers	Wrong answers
Children (3;06–5;06)	61.4%	33.7%	4.8%
Adults	91.7%	8.3%	0%

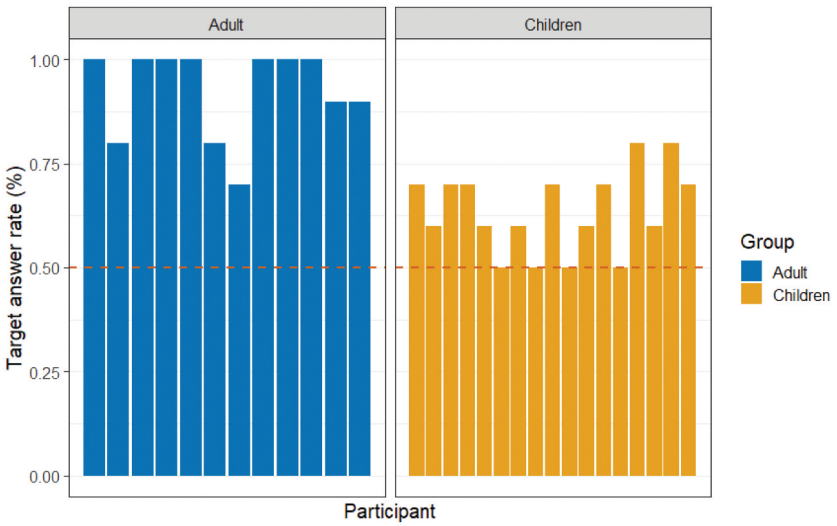

Figure 3. Target answer rate of children and adults.

Table 4. Percentage of choosing target over nontarget answers.

	PQs	EQs	Overall
Children (3;06–5;06)	76.6%*	51.3%	64.5%*
Adults	91.7%*	91.7%*	91.7%*

*Significance above chance level.

The results suggest that children are able to make use of prosodic information in language comprehension to differentiate among the two types of in-situ *wh*-questions. Although the extra pragmatic cues that were included to increase the naturalness of the questions may have aided the differentiation process, the prosodic difference between PQs and EQs was the only factor that was consistently present in every trial. The extra pragmatic cues, on the other hand, were not as reliable: Some trials consisted of only hand gestures, some consisted of only cue words, some consisted of both, and some consisted of none. Results from the mixed-effect analysis show that there was no effect of pragmatic cues; in other words, questions with extra pragmatic cues were as challenging as those without any such cues. If children couldn't reliably use pragmatic cues to differentiate the two types of questions, it must be that they employed prosodic cues. However, the fact that children overall performed worse than adults suggests that they may not be as sensitive to the prosody of questions as adults are. This is in line with previous studies, which claim that although children are able to use prosodic information in sentence processing, they use such information less effectively than adults do to infer the intended meaning (e.g., Snedeker 2008; Ito et al. 2012; Sekerina & Trueswell 2012; Hupp & Jungers 2013).

4.1.3. *Summary of Experiment 1*

The results show that in the comprehension task, children performed above chance level. They were able to provide almost twice as many target as nontarget answers. If children (wrongly) assumed that PQs and EQs have a similar intention, we would expect the percentage of target answers to be the same as nontarget answers. The significant difference in percentage shows that children, at the very least, recognized that two different types of in-situ *wh*-questions were asked, and the fact that there was a strong preference for target answers over nontarget answers shows that they were able to assign the right intention to the right type of question with moderate accuracy.

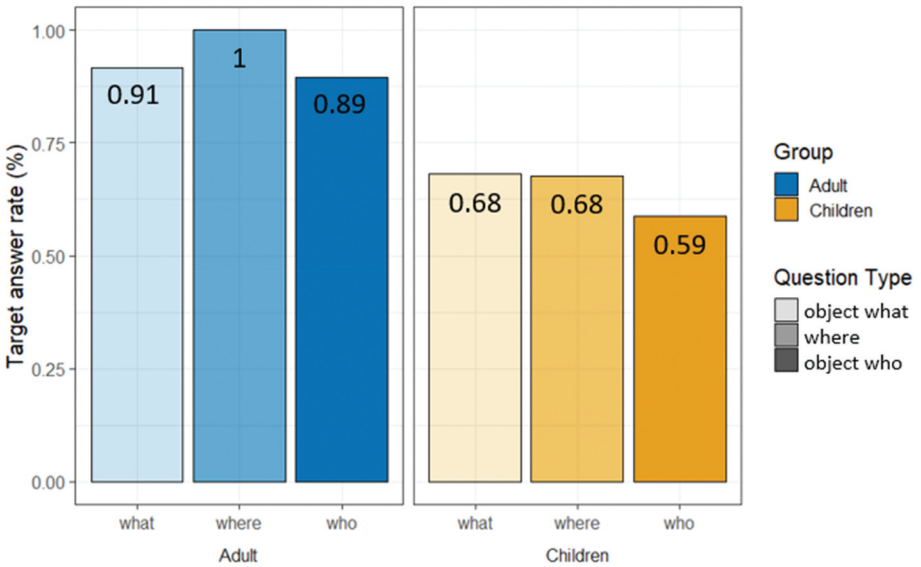

Figure 4. Target answer rate by subtype of *wh*-question.

4.2. *Experiment 2—Production study*

Although the corpus analysis in section 3 shows that children do not produce *wh*-in-situ, this could be because children are rarely in a pragmatically appropriate context for PQs. In this task, we tested children's production of *wh*-questions in contexts where PQs and fronted information-seeking questions are both acceptable and can be used interchangeably.

4.2.1. *Method*

4.2.1.1. ***Participants.*** After completing experiment 1, both children and adult participants were given a 10-minute break in an adjacent room before getting invited back to the testing room for experiment 2. Note that the comprehension task (experiment 1) was always completed prior to the production task to ensure uniform prior exposure to *wh*-in-situ.

4.2.1.2. ***Materials.*** Participants were introduced to an alien character, Beeple. Beeple came to planet Earth to learn about the Earth and its culture. Before Beeple left to return to his planet of origin, we wanted to make sure that he had learned enough about Earth. Thus, the participants' task was to ask Beeple multiple questions to quiz his knowledge.

There were 12 trials in total, which included four object "*what*"-, four "*where*"-, and four object "*who*"- questions. In each trial, participants were shown an illustration of Beeple standing next to certain objects or characters. The experimenter would prompt the participants by saying: "*Let's ask Beeple if he knows about [general description of the object].*" Beeple's responses were prerecorded. After the participants asked the question, the experimenter would play the recorded audio file. The recorded answers were acoustically modified to sound alien-like. An example of a practice trial with an in-situ question is presented in Figure 5.

The scenarios were designed to match with the context in which PQs are used in child-directed speech. Typically, PQs are used when (i) the addresser already has an answer in mind and (ii) the addresser is more interested in assessing the addressee's knowledge than the answer itself. These conditions are satisfied in the production task. First of all, we ensured that children knew the answer to the *wh*-question by only using simple target objects or characters (e.g., *apple, pizza*, or *mom*, etc.). Secondly, the task was set up so that participants were interested in the alien's ability to answer, as they needed to decide whether the alien had learned enough to return to his planet. Thirdly, the participants were placed into an "authority" role: They knew more about English than an alien who was learning human language, and they were encouraged to give the alien feedback ("Good job, Beeple" if the response was correct, or "That's wrong" followed by a correction if it was incorrect). It is important to note that these are contexts in which PQs are often used but are not meant to strongly favor PQs. Given that we were interested in testing economy-based accounts, such contexts in which PQs and fronted questions are both acceptable can be used interchangeably are ideal, as we can tease apart whether a PQ production is primarily motivated by economy preference or by pragmatic constraints.

4.2.1.3. ***Procedure.*** In each scenario, the experimenter instructed the participants to ask Beeple about an object or a person in his surrounding environment. The first two scenarios were used as practice trials. To avoid a strong priming effect on PQs and demonstrate that participants had the freedom to choose the type of questions they wanted to use, we included both an in-situ and a fronted *wh*-question in the practice trials (randomly introduced as questions 1 and 2). After the practice trials, participants were encouraged to produce the questions by themselves, with no feedback or correction given.

Figure 5. Practice trial example in the production task.
Experimenter: "Let's ask Beeple if he knows the word for the food the boy is eating. Let me show you how to do that: Hey Beeple, the boy is eating what?"

4.2.2. *Results and discussion*

4.2.2.1. ***Practice trial performance.*** The practice trials were similar to a repetition task. The experimenter demonstrated first how to ask a question from Beeple and asked the participants to repeat after him. Although adults had no trouble following the instruction and repeating the question the experimenter asked (either in-situ or fronted), we observed several interesting behaviors in children.

First, although children had no trouble understanding the repetition request for a fronted question, it took them longer to repeat an in-situ question. Even though the experimenter asked them to simply repeat the question, six children (35%) immediately gave an answer. Second, seven of them "auto-corrected" the in-situ question into a fronted one (e.g., "*where are the kids going?*" even though the experimenter said "*the kids are going where?*") or made a "failure to delete" error when fronting the *wh*-phrase, resulting in two copies ("*where are the kids going where?*") (see Crain and Nakayama (1987) and Roeper and de Villiers (2011) for more discussion on double marking errors in children production). These observations point to children's preference for fronted *wh*-questions but also their awareness of in-situ strategies.

4.2.2.2. ***Main trials performance.*** One 3;08-year-old child refused to produce questions by herself, and thus her data were not included in the analysis. Not every child participant successfully completed all 10 main trials; in particular, one child completed nine and one child only completed eight, resulting in a total of 157 data points. Among child participants, only one child produced PQs (with no mistake) and only for two trials (2/157 = 1.27%). The remaining child participants consistently used fronted *wh*-questions. They did make some grammatical errors in their fronted *wh*-questions. Common mistakes included auxiliary omission and absence of subject-verb agreement. In addition, one child produced only "*what*"-questions regardless of the scenario (e.g., "*what is the boys playing at?*" instead of "*where*" and "*what is the fairy talking to?*" instead of "*who*"). Overall, the percentage of well-formed fronted *wh*-questions was 67%; 33% of the utterances included at least one type of grammatical error. Since the goal of the production task was to test whether children were willing to produce PQs given an appropriate context, we will not discuss further the grammatical errors found in fronted *wh*-questions.

One adult participant produced PQs for eight out of 10 trials and spontaneously commented that he found PQs "easier to produce." However, the remaining 11 adults also consistently preferred fronted *wh*-questions throughout the production task. In total, there were 10 in-situ questions

produced by each adult participant. Although this number is small, a chi-squared test with Yates correction suggests there is a difference in performance between adults and children ($X^2 = 6.56$, $df = 1$, $p = .01$). The data are summarized in Table 5.

4.2.3. *Summary of Experiment 2*

Children did not show a preference for the structurally simpler in-situ structure. In fact, they strongly preferred producing fronted questions despite making grammatical errors in the process. This result is in line with other elicitation studies in Brazilian Portuguese (Vieira & Grolla 2020)[8] and French (e.g., Gotowski 2017), which we will discuss next.

5. Comparison of English and French data

Before turning to a general discussion of the experimental results, we present a (preliminary) comparison between the *wh*-in-situ usage of English-speaking and French-speaking children—based on preexisting literature for the latter—to provide a broader picture of the acquisition of alternative *wh*-question strategies. Both French and English allow fronted and in-situ *wh*-questions to be used as information-seeking questions, though French also allows for fronted *wh*-questions without inversion (among other main possibilities illustrated in (10), based on Shlonsky 2012). Fronting is also possible in conjunction with an invariant Q-marker *est-ce-que* (10d). In general, the presence versus absence of subject-auxiliary/verb inversion (10b) versus (10c) marks a register difference (with inversion being characteristic of the formal register).

(10) a. *Tu as quitté qui?* (in situ object *wh*)
you have left who
'Who have you left?'

b. *Qui tu as quitté?* (fronted object *wh* without inversion)
who you left

c. *Qui as-tu quitté?* (fronted object *wh* with inversion)
who have you left

d. *Qui est-ce que tu as quitté?* (fronted object *wh* with Q-marker)
who Q you have left

e. *C'est qui que tu as quitté?* (clefted wh)
It is who that you have left

f. *Qui c'est que tu as quitté?* (movement of the cleft pivot)
who it is that you have left

In many cases, French *wh*-in-situ can be used as an alternative to *wh*-fronting. As shown in (11), it is acceptable to use French *wh*-in-situ as the first sequence of an exchange, similarly to the PQ example in (3).

(11) *Pardon, il est quelle heure?* (Adli 2006:184)
sorry it is what time
'Sorry, what time is it?'

Table 5. Adults' and children's elicited production of *wh*-questions.

	PQs	EQs
Children	2/157 (1.27%)	155/157 (98.73%)
Adults	10/120 (8.3%)	110/120 (91.7%)

[8]The lack of reported data on child-directed speech precludes a fuller comparison with Brazilian Portuguese, but the elicitation results are informative on their own. In general, the large set of available *wh*-strategies is similar (but not identical) to French, and in-situ *wh* can be used for information-seeking questions. Children ages 4;06–5;06 are shown to prefer fronted *wh* (~80%) to in-situ *wh* in two experimental conditions, one establishing a Common Ground (hence an enriched presuppositional context) and the other not. Adult controls show no preference in the prominent Common Ground condition (50.5% fronted *wh* vs. 49.5% in-situ *wh*).

Although Chang (1997), Cheng & Rooryck (2000), and others have claimed that French in-situ *wh*-questions are associated with a strong presupposed context not found in fronted questions—as shown in the negative answer contrast in (12)–(13)—such a claim is controversial. In at least the colloquial register of the language, there is strong evidence that the pragmatic constraint attributed to *wh*-in-situ does not hold. Mathieu (2004) provides the following in-situ examples that elicit perfectly acceptable negative answers (the colloquial register is identifiable from the absence of the negative scope marker *ne* in the answer (*j'ai pas faim* 'I am not hungry') in (15)). This is the register relevant to child language acquisition, where *ne* is also characteristically absent from many child-directed utterances (Culbertson 2010).

(12) Q: *Marie a acheté quoi?* A: **Rien*.
Mary has bought what Nothing
'What has Mary bought?'

(13) Q: *Qu'est-ce que Marie a acheté?* A: *Rien*.
what Q Mary has bought Nothing

(14) Q: *Tu fais quoi dans la vie?* (Mathieu 2004:18)
You do what in the life
'What do you do for a living?'
A: *Rien. Je suis au chômage.*
Nothing. I am unemployed.

(15) Q: *Tu veux manger quoi ce soir?* (Mathieu 2004:18)
You want to eat what tonight
'What do you want to eat tonight?'
A: *Rien. J'ai pas faim.*
'Nothing. I am not hungry.'

In terms of prosody, Cheng & Rooryck (2000) propose that French *wh*-in-situ is licensed by an intonation morpheme. However, other studies have found that rising intonation is not required of in-situ-*wh* in French (Adli 2004; Déprez, Syrett & Kawahara 2013).

Overall, in both French and English, *wh*-in-situ is a viable alternative option for fronted information-seeking questions. More work will need to be done to verify any remaining pragmatic differences between the two languages. As described in section 2, English PQs are typically used in certain contexts in which the addressers are interested in the addressee's ability to answer, such as quiz shows and child-directed speech. On the basis of spontaneous adult-to-adult corpora, Myers (2007) makes a similar claim that French *wh*-in-situ is used for highly expected questions and more answerable questions. Boucher (2010) also suggests that French *wh*-in-situ is used in highly constrained social situations such as ordering and interrogation. Thus, while it is possible that the specific usage contexts differ between the two languages, it is unclear whether one type is overall more constrained than the other.

It also remains unclear whether there are other pertinent differences between English PQs and French in-situ questions. Lambrecht (1994, 2000)[9] claims that there is an information-structural constraint in Spoken French, but not found in English, that prevents focal elements from occurring in preverbal position. This constraint predicts a preference for *wh*-in-situ or clefted-*wh* over fronted questions. However, such preference for *wh*-in-situ has not been found in adult-to-adult corpus studies (29.2%, 25%, and 17.5% as reported in Myers 2007, Boucher 2010, and Coveney 2002 respectively) as well as child-directed-speech studies—Becker & Gotowski (2015) found only 13.2% in-situ questions compared to 86.8% of fronted questions. Strik & Pérez-Leroux (2011) found 35% of in-situ but no instances of *wh*-cleft questions; fronted *wh*-questions with the Q marker *est-ce que* were the most frequent in parental input.

[9]We thank an anonymous reviewer for bringing this to our attention.

Finally, the distribution of in-situ versus *wh*-questions in child-directed speech is also similar in the two languages, with *wh*-in-situ accounting for approximately 10%–20% of all types of *wh*-questions. For example, Becker and Gotowski (2015) report a *wh*-in-situ rate of 16% and 16.6% for the English and French child-directed speech data, Thus, any attempt to explain the performance of English-speaking children should also be able to explain the performance of the French-speaking ones, especially in light of the discussion showing the similarity between English PQs and French *wh*-in-situ in the colloquial register that is characteristic of child-directed speech.

Overall, our experimental results on English-speaking children are both similar to, and different from, reported data in the literature on French-speaking children. In elicited production tasks (Table 6), on the one hand, both populations of learners show a preference for fronted over in-situ *wh*-questions. On the other hand, French children still produce some *wh*-in-situ utterances; English-speaking children almost totally avoid in-situ structures. Note that although the fronted-*wh* category in French includes both fronted questions with inversion and those without inversion, the majority of the utterances produced by children in both settings (spontaneously and in elicitation tasks) were fronted questions without inversion. Hence, the fronted *wh*-questions produced by French-speaking children tend to be at least one instance of syntactic movement simpler (no subject-auxiliary/verb inversion[10]) than those produced by English-speaking children. However, if simplicity were the main motivation, we would expect French-speaking children to produce more *wh*-in-situ, which are also one instance of syntactic movement simpler than fronted questions without inversion. Thus, the fact that French fronted questions can be simpler than their English counterpart is not sufficient to explain why French-speaking children produce relatively more *wh*-in-situ in an elicitation task than English-speaking children.[11]

Corpus data of English-speaking children again show a strong preference for fronted *wh*-questions in spontaneous production, but reported[12] results from French corpus-based studies overall are mixed (Table 7). Crisma (1992) finds that Phillippe (from the Leveillé corpus) does not produce any *wh*-in-situ until 2;06 and has a much higher rate of fronted *wh*-question production compared to *wh*-in-situ. Palasis, Faure & Lavigne (2019) report a comparable rate of *wh*-in-situ and fronted *wh*-questions (not counting clefted questions). On the other hand, Hamann (2006) finds a strong preference for *wh*-in-situ. Although the children in these three

Table 6. Elicited production of French vs. English *wh*-questions.

	Language	Wh-in-situ	Fronted Wh	Age range
Gotowski (2017)	French	12%	51%	3;09–5;08
Cronel-Ohayon (2004)	French	24.2%	64.6%	4;00–6;00
Prévost et al. (2017)	French	25.9%	74.1%	4;00–4;05
Strik (2007)	French	~ 20%	no report	4;00–4;06
Our study	English	1.27%	98.73%	3;06–5;06

[10]French has verb raising all the way to C in questions; hence fronted questions in the absence of an auxiliary involve longer head movement, which translates into additional instances of Merge under the Derivational Complexity Metric of Jakubowicz (2011). The economy of movement hypothesis predicts more fronting with auxiliaries (shorter chains) than with lexical verbs in French, which has yet to be tested.

[11]Zuckerman & Hulk (2001) report a lower production rate of French in situ (6%, after omission of outliers [n = 5], the rate drops to 3%). The very low level of *wh*-in-situ may at least partly be the consequence of the method of elicitation in which an indirect question with a clause initial *wh*-phrase and no inversion was used as a prompt (*Je veux savoir où il est allé* 'I want to know where he went') despite the fact that one possible answer is *Où il est allé?* 'Where did he go?,' which is likely to have inflated the proportion of fronted *wh* without inversion (89%).

[12]The elicited production result (~20% *wh*-in-situ) is culled from several sources (Strik 2007, 2008) and is an estimate based on limited text descriptions of the results. A pilot study (as reported in Strik, 2008) is not reported because the elicitation method was similar to that used by Zuckerman & Hulk (2001) and discussed in fn. 11.

Table 7. Spontaneous production of French vs. English *wh*-questions (corpus studies).

	Language	Wh-in-situ	Fronted Wh	Age range
Crisma (1992)	French	26.19%	73.81%	2;01–2;07
Hamann (2006)	French	80.04%	19.96%	1;08–2;09
Palasis et al. (2019)	French	43.8%	42.7%	2;06–4;11
Our study	English	0.07%	99.3%	2;00–4;00

French corpus studies include younger ones than our own analyses of English corpora, the relatively close age range yet opposite results in Crisma's and Hamann's analyses suggest that age may not be the main factor determining a preferred *wh*-variant.[13]

The asymmetry between corpus-based and elicited results could be due to different production contexts. For example, a lab experiment may remind children of a formal classroom setting where fronted questions are regularly used, thus leading to a bias for fronted questions. Another possibility is that the higher rate for *wh*-in-situ in French is mostly due to the presence of simple *C'est wh* questions (e.g., *C'est qui Paul?* 'Who is Paul?'). Such questions are not a possible option in the elicitation tasks reported in Table 6 but account for the majority of spontaneously produced *wh*-in-situ (~70%) reported in Palasis, Faure & Lavigne (2019).

Given that there is significant variation between the spontaneous production studies and that there are unresolved controversies about possible pragmatic differences between PQs and French *wh*-in-situ, we will limit our further discussion to elicitation results, in which the production contexts are controlled so that fronted and in-situ questions are both acceptable. In section 6, we will discuss the overall preference for fronted questions over *wh*-in-situ in elicitation tasks in French- and English-speaking children as well as their difference in production rates in light of the two models we introduced at the outset.

6. General discussion

In this article, we have shown that besides EQs, in-situ *wh*-questions asking for new information (PQs) are also present in CDS. This finding challenges previous studies that rely on the assumption that in-situ *wh*-questions in English can only function as EQs or that children never hear nonecho *wh*-in-situ questions in English (Yip & Matthews 2000, 2007). Our comprehension task further demonstrates that even children as young as 4 years of age are able to differentiate between the two types of questions. Their moderately high accuracy with PQs in the comprehension task (76.6%) suggests that children understand and accept in-situ PQs as information-seeking questions. This directly contradicts a claim commonly found in the child language acquisition literature that children only recognize *wh*-in-situ as EQs (e.g., Takahashi 1991; Becker & Gotowski 2015; Park-Johnson 2017).

Why did children perform significantly worse with EQs than with PQs? One potential answer suggested by a reviewer is that children in our study were biased to answer according to their belief instead of the alien's, given the use of the attitude verb *think* and the infelicity in the alien's response. Previous studies have shown that children tend to evaluate *think* sentences based on reality or their own beliefs instead of others' belief, leading to their poor performance on false belief tasks (e.g., de Villiers 1995; Papafragou, Cassidy & Gleitman 2007). However, we rule out this hypothesis because our task is not a belief evaluation task but a repetition task in the case of EQs. De Villiers & Pyers's longitudinal study (2002) shows that children within the same age range as those in our study do not have difficulty repeating *think* sentences that report false beliefs. By the third round of the study, when their participants' age range was between 3;07 and 4;05, their performance on such a task was above

[13]It is worth noting that the Leveillé corpus analyzed in Crisma (1992) represents French speech from the early 1970s, which is likely to be more formal than more recent spoken speech. Interestingly, Larrivée (2019) provides adult data from the Orleans corpus at two different time points on *comment* 'how' questions, documenting a change in use of in situ from the early 1970s to 2014. Limiting the analysis to information seeing vs. echo use, he demonstrates that they are increasingly used as requests for new information (as opposed to requests for repetitions/clarifications). This is a clear pragmatic change within a short time interval of 40 years, and it suggests that other pragmatic changes may have taken or are taking place.

90%. It is thus unlikely that the use of the verb *think* in the task is the main reason behind the lower accuracy with EQs. Instead, given the dominance of information-seeking questions in the input, we hypothesize that either children have a bias for more precise answers, or their default interpretation of questions is information-seeking. In the first case, since answers to PQs are more informative, children's failure to respond correctly to EQs may be due to cognitive factors tied to executive control, e.g., children's inability to suppress the (more obvious) answer that they already had in mind (Gualmini et al. 2008). In the latter case, since the majority of questions children are exposed to are questions asking for new information, it is possible that they have a default or bias toward an information-seeking interpretation. To get the nondefault interpretation, children would need to rely on additional cues such as prosody. EQs would be more challenging because EQs require children to notice and interpret the prosody correctly, and studies have shown that they do so less effectively than adults (e.g., Snedeker 2008). Note that children have no issue responding to EQs in spontaneous settings when there is no competing option, as found in our corpus work. Hence the low performance on EQs in our task is not an indication of children's inability to process EQs but more likely a problem with accessing the right repetition answer when there is a new-information competitor.

Moving on to elicited production, our results show that English-speaking children prefer fronted *wh*-questions over *wh*-in-situ. Although this is in line with experimental results from most studies with French-speaking children (e.g., Gotowski 2017; Cronel-Ohayon 2004; Prévost et al. 2017; Strik 2007), there is an asymmetry in production rates between the two populations, with French-speaking children producing 12%–25% *wh*-in-situ in an information-seeking question elicitation task (close to their in-situ input rate) and English-speaking children producing only 1.27% (versus 16% in their input). We thus need a theory that can explain both the disfavor of *wh*-in-situ production in children in general as well as the asymmetry between the elicited production rates of French- and English-speaking children.

Although adults did not produce a high number of PQs in the elicited production task, a chi-squared test suggests a difference in adult performance versus children. Moreover, the pragmatic contexts used in our production task were not supposed to strongly favor PQs over fronted questions. Instead, they were contexts in which PQs and fronted questions can be used interchangeably. The question we were looking at is this: In an elicitation task where it is appropriate to use either PQs or fronted questions, which type will children choose? Adults, who routinely use structures derived by syntactic movement, are not expected to be motivated to go for PQs. However, with regard to children, structural economy-based accounts (e.g., Jakubowicz's Derivational Complexity Hypothesis) predict that they would have a bias toward simpler constructions as it would alleviate the amount of cognitive resources required to form an utterance. In other words, we would expect children to resort to the structurally simpler but pragmatically equivalent in-situ PQs at a stage when they struggle with forming fronted *wh*-questions (shown by their inversion errors in English and their strong preference for the no inversion option in French). This prediction was not borne out in either the French or our English elicitation tasks.

What then prevents children from resorting to the simpler PQ construction? We rule out the hypothesis that children are unaware of such constructions. Given that the in-situ structure is available in children's grammar (as they accept and understand PQs in the comprehension task), their failure to produce the in-situ structure in an appropriate context is not due to its unavailability but more likely due to other reason(s). We borrow from computationally oriented work on learning and propose that the input filter hypothesis (Perkins, Feldman & Lidz 2017; Schneider, Perkins & Feldman 2019), while typically overlooked in discussions of syntactic acquisition, offers a promising explanation for the cross-linguistic results. This hypothesis is relevant to cases when there are multiple variants of the same grammatical item. In our case, PQs and fronted *wh*-questions are two variants of information-seeking questions. Given that children have a prior bias for a smaller number of categories (Perfors, Tenenbaum & Regier 2011), initially children may not establish that PQs and fronted questions are two different types but instead regard PQs as "noise" data points (generated by errors) of fronted questions. Assuming that children have access to a finite number of possible grammars as defined by

Universal Grammar (UG), and that learning takes place by changing the probabilistic distribution of available grammars based on their input (Yang 2002), such "noise" data points could be viewed as uninformative and not used to update the probability of the grammar associated with it, or update cautiously and slowly. As a result, children end up producing a more regularized version of the language, even when their input contains multiple variants. Though slowly, children can eventually learn to produce PQs when they are exposed to enough data to override their initial bias.

The question remains as to what makes a variant in the input more likely to be viewed as "noise." Based on previous studies as well as our current results, we identify three possible factors: frequency, consistency, and the total number of variants in the language. Consider frequency first. Our study overall provides evidence that English-speaking children's production does not match the frequency distribution of their input. Although fronted questions only account for approximately 80% of *wh*-questions in the input, their frequency is boosted to nearly 100% in both our elicited production task and corpus data. A similar pattern has been independently found in natural language studies (Pozzan & Valian 2016), artificial language learning studies (Hudson Kam & Newport 2005, 2009; Schwab, Lew-Williams & Goldberg 2018), and single-case studies (Singleton & Newport 2004). Thus, Pozzan & Valian's (2016) find that despite 38% of noninverted *yes*/*no* questions in CDS, children almost never produce non-inverted yes/no questions (only 1 in 264 occurrences or ~0.4%) in corpus data and in an experimental setting. Hudson Kam & Newport (2005, 2009) and Schwab, Lew-Williams & Goldberg (2018) find that when there are multiple variants of a grammatical item, children boost the frequency or regularize to the more frequent variant(s); adults tend to match the frequency distribution in the input. The input filter hypothesis predicts that utterances of lower-frequency variants are more likely to be viewed as noise by children, thus taking longer to learn.

However, frequency cannot be the sole determiner of what is learned and what is initially ignored, as despite having approximately similar input frequencies of *wh*-in-situ, French-speaking children do produce this structure. Different linguistic properties between the two languages may play a partial role in this result. However, as discussed in section 5, there are persistent controversies tied to claims regarding information-structural and pragmatic differences between English PQs and French *wh*-in-situ. Moreover, such pragmatic differences may speak more directly to spontaneous production than to production in elicitation tasks, where children are given specifically designed contexts to produce in-situ questions as an alternative to fronted *wh*-questions. We therefore do not rule out the possibility of a pragmatic-centered explanation for the asymmetry in production rate between the two populations. However, given that more work needs to be done to resolve existing controversies, we choose not to commit to such an explanation for the time being. Instead, we turn to the next potential factor: consistency. Hudson Kam & Newport (2009) show that regularization is not the mere result of frequency—learners also regularize to low-frequency but high-consistency variants while initially ignoring low-frequency and low-consistency ones. When comparing 50 PQs and French in-situ questions (from two CHILDES corpora: Palasis [2009] and York [Plunkett 2002] using the procedure described in the prosody analysis in section 2), we find that French in-situ content questions show relatively less across-speaker prosodic variability than English PQs (Table 8). Such variability may make PQ utterances more likely to be considered noise.

Finally, we suggest that the belief about the number of possible variants within a language is also a likely factor. Although English only allows two variants for information-seeking questions (fronted questions and PQs) as well as two for *yes*/*no* questions (inverted and noninverted with intonation),

Table 8. Prosodic variance of information-seeking *wh*-in-situ in French and English.

	Wh-word Duration variance	ΔF_0 variance
French	0.046	35.836
English	0.085	61.153

French allows for a minimum of six options for *wh*-questions (see (10); most likely they are not all equivalently present in child-directed speech) and three for *yes*/*no* questions (inverted, noninverted with intonation, and *est-ce que* fronted questions). By being exposed to many grammatical items (*wh*-phrases) that allow for multiple variants, children's prior bias for the smallest number of variants (Perfors et al. 2011) may be weakened to increase the flexibility to learn a range of data (Perfors 2012; Yang 2016).

The input filter hypothesis is relevant to production as children evaluate whether an utterance is informative or noise in order to make inference about their target grammar. Indeed, despite the regularization behavior found in learning tasks that involve multiple variants, comprehension typically remains unaffected. Hudson Kam & Newport (2005) and Schwab, Lew-Williams & Goldberg (2018), for example, find that in a two-alternative forced-choice task, children still demonstrate good comprehension of the lower-frequency or inconsistent variant, despite their strong preference for the higher-frequency item in their production task. Similarly, our results show that children regularize to fronted *wh*-questions while demonstrating an understanding of PQs. The distinction between *input* and *intake* (Gass 1997; Gagliardi & Lidz 2014; Omaki & Lidz 2015) can shed light on the asymmetry between comprehension and production. *Input* is the data available in the environment; *intake* is the data from the input that are utilized by the language acquisition mechanism to make inferences about the grammar of the target language, i.e., that part of the input that is not treated as noise. The distinction between *input* and *intake* corresponds to the distinction between input for comprehension and input for learning (Sharwood Smith 1986; Gass 1997), which arises due to the difference in the amount of cognitive load (e.g., memory and planning) required for comprehension versus production (Hendriks & Koster 2010; Humphreys 2012). The input-filtering process hence only affects the input for learning, i.e., the intake.

In sum, investigating the comprehension and production of two types of information-seeking questions (PQs vs. fronted *wh*-questions) in English-speaking children provides insights bearing on models of both acquisition of *wh*-questions and learning models more generally. Structural economy-based accounts focusing on syntax alone are not sufficient to account for the behavior of English-speaking children with respect to their acquisition of *wh*-in-situ. Instead, the new results for English, combined with existing French data, offer support for the general input filter theory, which suggests that children learn to use infrequent and inconsistent grammatical variants more slowly and cautiously.

7. Conclusion

The presence of PQs in English child-directed speech gives rise to interesting questions about children's acquisition of questions. PQs share a similar surface structure to EQs but a similar pragmatic interpretation to fronted information-seeking questions, allowing for a study of different aspects of the acquisition process. In the realm of comprehension, we have shown that children can differentiate in-situ PQs from in-situ EQs, based on prosody. This, along with the finding from Takahashi (1991) that children can differentiate EQs from information-seeking questions, indicates that 4-year-olds can use both prosodic cues and syntactic cues to aid comprehension.

Using corpus studies, we have also shown that children typically do not produce in-situ *wh*-questions in spontaneous speech despite the presence of PQs in the input. In our experimental production study, children not only failed to produce PQs during testing, they were also reluctant to repeat PQ examples during practice trials, and instead turned such questions into fronted questions. We further compared our results for English with what is known so far about the acquisition of French *wh*-questions because both languages allow alternative *wh*-strategies and display approximately the same percentage of *wh*-in-situ in child-directed speech. Overall, our results suggest that a structural economy-based approach alone is not sufficient to account for children's behavior through the course of acquisition. We suggest, following Gagliardi & Lidz (2014), that children may not learn everything available to them in the input, pointing to a need to differentiate between input and intake. Lower-frequency, lower-consistency variants—per the input filter hypothesis—are more likely to be viewed as noise, hence taking longer for children to learn them.

Note

Twenty-seven children (including seven pilot participants) and 14 adult controls had participated in the within-subject study when the lab was forced to close due to COVID-19. We had originally aimed for a higher number of participants. An online version subsequently piloted with adult participants failed to yield good results even after multiple modifications, possibly because the experiment requires in-person interaction to be pragmatically plausible (see description). Given the current worrisome evolution of the COVID-19 pandemic in the United States and the fact that all human subject research is banned for the foreseeable future at our institution, we anticipate that further data collection may not be achievable in the coming months or year. We have therefore opted to use nonparametric statistical tests, which are more conservative than parametric ones. Note also that the effect sizes found in this study are also relatively large, justifying a smaller sample size. Thus, despite a sample size smaller than we wished, we believe that the study still has sufficient statistical power to make meaningful contributions to the field of language acquisition.

ORCID

An D. Nguyen http://orcid.org/0000-0003-0271-8457

References

Adli, Aria. 2004. Y a-t-il des morphèmes intonatifs impliqués dans la syntaxe interrogative du Français? Le cas du qu-in-situ. In T. Meisenburg & M. Selig (eds.), *Nouveaux départs en phonologie: les conceptions sub- et suprasegmentales*, 199–215. Tübingen: Narr.

Adli, Aria. 2006. French *wh*-in-situ questions and syntactic optionality: Evidence from three data types. *Zeitschrift für Spachwissenschaft* 25. 163–203.

Authier, Marc. 1993. Nonquantificational *wh* and weakest crossover. *Linguistic Inquiry* 24. 161–168.

Banfield, Ann. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan Paul.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Becker, Misha & Megan Gotowski. 2015. Explaining children's *wh*-in situ questions: Against economy. In Elizabeth Grillo & Kyle Jepson (eds.), *Proceedings of the 39th annual Boston University Conference on Language Development [BUCLD 39]*, 127–137. Somerville, MA: Cascadilla Press.

Bellugi, Ursula. 1965. The development of interrogative structures in children's speech. In K. Riegel (ed.), *The development of language function*, 103–137. Ann Arbor, MI: University of Michigan Press.

Blakemore, Diane. 1994. Echo questions: A pragmatic account. *Lingua* 94. 197–211.

Boersma, Paul & David Weenink. 2019. *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.52. http://www.praat.org/ .

Boucher, Paul. 2010. *Wh*-questions in French and English: Mapping syntax to information structure. In C. Breul & E. Göbbel (eds.), *Comparative and contrastive studies of information structure*, 101–137. Amsterdam: John Benjamins.

Braunwald, Susan. 1985. The development of connectives. *The Journal of Pragmatics 9*(4). 513–525.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Chang, Lisa. 1997. *Wh-in-situ phenomena in French*. Vancouver, BC: University of British Columbia master's thesis.

Cheng, Lisa Lai-Shen & Johan Rooryck. 2000. Licensing *wh*-in-situ. *Syntax* 3. 1–19.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth L. Hale & S. Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: The MIT Press.

Clark, Eve. 1978. Awareness of language: Some evidence from what children say and do. In R. J. A. Sinclair & W. Levelt (eds.), *The child's conception of language*, 17–43. Berlin: Springer Verlag.

Comyn, Marie. 2013. *Wh-in situ is accompanied by which formal features*? Ghent, Belgium: Ghent University master's thesis.

Coveney, Aidan. 2002. *Variability in spoken French: A sociolinguistic study of interrogation and negation*. Bristol/Portland: Elm Bank. Reprint from 1996 original.

Crain, Stephen & Mineharu Nakayama. 1987. Structure dependence in grammar formation. *Language* 63(3). 522–543.

Crisma, Paola. 1992. On the acquisition of *wh*-questions in French. *Geneva Generative Papers* 1. 115–122.

Cronel-Ohayon, Stephany. 2004. *Etude longitudinal d'une population d'enfants dysphasiques*. Geneva, Switzerland: University of Geneva dissertation.
Culbertson, Jennifer. 2010. Convergent evidence for categorial change in French: From subject clitic to agreement marker. *Language* 86(1). 85–132.
Cuza, Alejandro, Lauren Miller, Rocio Tattam & Mariluz Vergara. 2019. Structure complexity effects in child heritage Spanish: The case of the Spanish personal *a*. *International Journal of Bilingualism* 23(6). 1333–1357.
de Villiers, Jill (see Villiers)
Demetras, Martha. 1989. *Working parents' conversational responses to their two-year-old sons*. Tucson, AZ: University of Arizona dissertation.
Demuth, Katherine, Jennifer Culbertson & Jennifer Alter. 2006. Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech* 49. 137–174.
Déprez, Viviane, Kristen Syrett & Shigeto Kawahara. 2013. The interaction of syntax, prosody, and discourse in licensing French *wh*-in-situ questions. *Lingua* 124. 4–19.
Dickinson, David & Patton Tabors. 2001. *Beginning literacy with language: Young children learning at home and school*. Baltimore: Paul Brookes Publishing.
Durrleman, Stephanie, Theodoros Marinis & Julie Franck. 2016. Syntactic complexity in the comprehension of *wh*-questions and relative clauses in typical language development and autism. *Applied Psycholinguistics* 37(6). 1501–1527.
Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1). 67–88.
Gagliardi, Annie & Jeffrey Lidz. 2014. Statistical insensitivity in the acquisition of Tsez noun classes. *Language* 90(1). 58–89.
Gass, Susan. 1997. *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.
Goodwin, Anthony, Deborah Fein & Letitia Naigles. 2015. The role of maternal input in the development of *wh*-question comprehension in autism and typical development. *Journal of Child Language* 42(1). 32–63.
Gotowski, Megan. 2017. *Wh*-in situ production in child French. *Linguistica Atlantica* 36(2). 99–109.
Grice, Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and semantics*, vol. 3, *Speech acts*, 41–58. New York: Academic Press.
Gualmini, Andrea, Sarah Hulsey, Valentine Hacquard & Danny Fox. 2008. The question-answer requirement for scope assignment. *Natural Language Semantics* 16(3). 205–237.
Hamann, Cornelia. 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(1). 143–189.
Hendriks, Petra & Charlotte Koster. 2010. Production/comprehension asymmetries in language acquisition. *Lingua* 120 (8). 1887–1897.
Hopp, Holger, Michael T. Putnam & Nora Vosburg. 2019. Derivational complexity vs. transfer effects: Long-distance *wh*-movement in heritage and L2 grammars. *Linguistic Approaches to Bilingualism* 9(3). 341–375.
Hout, Angeliek van, Alma Veenstra & Sanne Berends. 2011. All pronouns are not acquired equally in Dutch: Elicitation of object and quantitative pronouns. In *Proceedings of the 4th Conference on Generative Approaches to Language Acquisition North America [GALANA 2010]*, 106–121. Somerville, MA: Cascadilla Proceedings Project.
Hudson Kam, Carla & Elissa Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1. 151–195.
Hudson Kam, Carla & Elissa Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1). 30–66.
Humphreys, Gina. 2012. *Linking sentence production and comprehension: The neural mechanisms underlying production and comprehension control processes*. York, UK: University of York dissertation.
Hupp, Julie & Melissa Jungers. 2013. Beyond words: Comprehension and production of pragmatic prosody in adults and children. *Journal of Experimental Child Psychology* 115(3). 536–551.
Ito, Kiwako, Nobuyuki Jincho, Utako Minai, Naoto Yamane & Reiko Mazuka. 2012. Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year-olds. *Journal of Memory and Language* 66. 265–284.
Jakubowicz, Celia. 2005. The language faculty: (Ab)normal development and interface constraints. Paper presented at GALA 2005, September 8–10, Siena, Italy.
Jakubowicz, Celia. 2011. Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua* 121(3). 339–351.
Jakubowicz, Celia & Nelleke Strik. 2008. Scope-marking strategies in the acquisition of long distance *wh*-questions in French and Dutch. *Language and Speech* 51(1–2). 101–132.
Kuczaj, Stan. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16. 589–600.
Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.
Lambrecht, Knud. 2000. When subjects behave like objects: A markedness analysis of sentence focus constructions across languages. *Studies in Language* 24. 611–82.

Larrivée, Pierre. 2019. Historical pragmatics, explicit activation, and *wh*-in-situ in French. In I. Feldhausen, M. Elsig, I. Kuchenbrandt & M. Neuhaus (eds.), *Romance languages and linguistic theory 15: Selected papers from "Going Romance" 30, Frankfurt*, 113–133. Amsterdam/Philadelphia: John Benjamins.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum.

Mathieu, Eric. 2004. The mapping of form and interpretation: The case of optional *wh*-movement in French. *Lingua* 114. 1090–1132.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, and Morgan Sonderegger. 2017. *Montreal Forced Aligner* [Computer program]. Version 0.9.0. http://montrealcorpustools.github.io/Montreal-Forced-Aligner/

Myers, Lindsy. 2007. *Wh-interrogatives in spoken French: A corpus-based analysis of their form and function*. Austin, TX: University of Texas at Austin dissertation.

Omaki, Akira & Jeffrey Lidz. 2015. Linking parser development to acquisition of syntactic knowledge. *Language Acquisition* 22(2). 158–192.

Palasis, Katerina. 2009. *Syntaxe générative et acquisition: le sujet dans le développement du système linguistique du jeune enfant*. Nice, France: University of Nice Sophia Antipolis dissertation.

Palasis, Katerina, Richard Faure & Frédéric Lavigne. 2019. Explaining variation in *wh*-position in child French: A statistical analysis of new semi-naturalistic data. *Language Acquisition* 26(2). 210–234.

Papafragou, Anna, Kimberly Cassidy & Lila Gleitman. 2007. When we think about thinking: The acquisition of belief verbs. *Cognition* 105(1). 125–165.

Park-Johnson, Sunny. 2017. Cross-linguistic influence of *wh*-in-situ questions by Korean-English bilingual children. *International Journal of Bilingualism* 21(4). 419–432.

Perfors, Amy. 2012. When do memory limitations lead to regularization? An experimental and computational investigation. *Journal of Memory and Language* 67(4). 486–506.

Perfors, Amy, Joshua Tenenbaum & Terry Regier. 2011. The learnability of abstract syntactic principles. *Cognition* 118 (3). 306–338.

Perkins, Lauren, Naomi Feldman & Jeffrey Lidz. 2017. Learning an input filter for argument structure acquisition. In *Proceedings of the 7th Workshop on Cognitive Modeling and Computational Linguistics*, 11–19. Valencia, Spain: Association for Computational Linguistics.

Plunkett, Bernadette. 2002. Null subjects in child French interrogatives: A view from the York corpus. *Romance corpus linguistics: Corpora and spoken language*, 441–452. Tübingen, Germany: Narr.

Pozzan, Lucia & Virginia Valian. 2016. Asking questions in child English: Evidence for early abstract representations. *Language Acquisition* 24(3). 209–233.

Prévost, Philippe, Laurice Tuller, Marie Barthez, Joelle Malvy & Frédérique Bonnet-Brilhault. 2017. Production and comprehension of French *wh*-questions by children with ASD. *Applied Psycholinguistics* 38(5). 1095–1131.

Prévost, Philippe, Nelleke Strik & Laurie Tuller. 2014. *Wh*-questions in child L2 French: Derivational complexity and its interactions with L1 properties, length of exposure, age of exposure, and the input. *Second Language Research* 30(2). 225–250.

Reis, Marga. 2012. On the analysis of echo questions. *Tampa Papers in Linguistics* 3.1–24.

Roeper, Thomas & Jill de Villiers. 2011. The acquisition path for *wh*-questions. *Handbook of Generative Approaches to Language Acquisition. Studies in Theoretical Psycholinguistics*, vol. 41, 189–246. Dordrecht: Springer.

Sachs, Jacqueline. 1983. Talking about the there and then: The emergence of displaced reference in parent–child discourse. In K. E. Nelson (ed.), *Children's language*, vol. 4, 1–28. Hillsdale, NJ: Lawrence Erlbaum.

Schneider, Jordan, Lauren Perkins & Naomi Feldman. 2019. A noisy channel model for systematizing unpredictable input variation. In Megan M. Brown & Alexandra Kohut (eds.), *Proceedings of the 44th Annual Boston University Conference on Language Development [BUCLD 44]*, 533–547. Somerville, MA: Cascadilla Press.

Schwab, Jessica, Casey Lew-Williams & Adele Goldberg. 2018. When regularization gets it wrong: Children over-simplify language input only in production. *Journal of Child Language* 45(5). 1054–1072.

Seidl, Amanda, George Hollich & Peter Jusczyk. 2003. Early understanding of subject and object *wh*-questions. *Infancy* 4 (3). 423–436.

Sekerina, Irina & John Trueswell. 2012. Interactive processing of contrastive expressions by Russian children. *First Language* 32(1–2). 63–87.

Sharwood Smith, Michael. 1986. Comprehension versus acquisition: Two ways of processing input. *Applied Linguistics* 7. 239–256.

Shlonsky, Ur. 2012. Notes on *wh* in situ in French. In L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro & C. Poletto (eds.), *Functional heads*, 242–252. Oxford: Oxford University Press.

Singleton, Jenny & Elissa Newport. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49(4). 370–407.

Snedeker, Jesse. 2008. Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language* 58(2). 574–608.

Strik, Nelleke & Ana Pérez-Leroux. 2011. Jij doe wat girafe?: *Wh*-movement and inversion in Dutch-French bilingual children. *Linguistic Approaches to Bilingualism* 1(2). 175–205.

Strik, Nelleke. 2007. L'acquisition des phrases interrogatives chez les enfants francophones. *Psychologie française* 52(1). 27–39.
Strik, Nelleke. 2008. *Syntaxe et acquisition des phrases interrogatives en français et en néerlandais: une étude contrastive*. University Paris 8, doctoral thesis.
Takahashi, Mari. 1991. The acquisition of echo questions. In Thomas Maxfield & Bernadette Plunkett (eds.), *Papers on the acquisition of* wh*: Proceedings of the UMass Roundtable 1990*, 213–223. Amherst, MA: GLSA.
Valian, Virginia & Lyman Casey. 2003. Young children's acquisition of *wh*-questions: The role of structured input. *Journal of Child Language* 30. 117–143.
van Hout, Angeliek (see Hout)
Van Houten, Lori. 1986. Role of maternal input in the acquisition process: The communicative strategies of adolescent and older mothers with their language learning children. Paper presented at the Boston University Conference on Language Development, October 17–19, Boston.
Vieira, Clariana & Elaine Grolla. 2020. The pragmatics of *wh*-in-situ questions in Brazilian Portuguese: Data from child and adult language. In Megan M. Brown & Alexandra Kohut (eds.), *Proceedings of the 44th Boston University Conference on Language Development [BUCLD 44]*, 677–690. Somerville, MA: Cascadilla Press.
Villiers, Jill de. 1995. Questioning minds and answering machines. In D. MacLaughlin & S. McEwen (eds.), *Proceedings of the 19th annual Boston University Conference on Language Development [BUCLD] 19*, 20–36. Somerville, MA: Cascadilla Press.
Villiers, Jill de & Jennie Pyers. 2002. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development* 17(1). 1037–1060.
Weist, Richard & Andrea Zevenbergen. 2008. Autobiographical memory and past time reference. *Language Learning and Development* 4(4). 291–308.
Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.
Yang, Charles. 2016. *The price of productivity*. Cambridge, MA: The MIT Press.
Yip, Virginia & Stephen Matthews. 2000. Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism: Language and Cognition* 3(3). 193–208.
Yip, Virginia & Stephen Matthews. 2007. *The bilingual child: Early development and language contact*. Cambridge: Cambridge University Press.
Yuan, Boping. 2015. The effect of computational complexity on L1 transfer: Evidence from L2 Chinese attitude-bearing *wh*-questions. *Lingua* 167. 1–18.
Zuckerman, Shalom & Aafke Hulk. 2001. Acquiring optionality in French *wh*-questions: An experimental study. *Revue Québéquoise de Linguistique* 30. 71–97.

Appendix

Table A1. Overview of stimuli in the comprehension task.

Question	Answer for a PQ	Answer for an EQ
The map is where*	On the fridge in the kitchen	In the room with a plant
The boy bought what?*	A watermelon cake	A cake with fruit
Kate went on the ferris wheel with who?	With her mom	With her parent
The boy got what?	An ice-cream	A sweet treat
The two kids chose to play what?	A bounce house	A jumping game
The thief is caught by who?	A fireman	A man in uniform
The man got the apples from where?	7-11	A store
The girl got what?	A bunny	A white pet
The man went where?	To the hospital	To the white building
The squirrel is hiding where?	Behind a rock	Behind something big
The winner is who?	Batman	The man with a mask
The dog belongs to who?	The doctor	The girl in blue

*Practice trial.

Table B1. Overview of stimuli in the production task.

Prompt	Target
Let's ask Beeple about the food the boy is eating.*	The boy is eating what?/What is the boy eating?
Let's ask Beeple if he knows about the zoo.*	The two boys are going where?/Where are the two boys going?
Let's ask Beeple about the person the girl is hugging.	The girl is hugging who?/Who is the girl hugging?
Let's ask Beeple about the game the boy is playing.	The boy is playing what?/What is the boy playing?
Let's ask Beeple if he knows about the fruit on the table.	That is what on the table/What is that on the table?
Let's ask Beeple about the place the children are playing at.	The children are playing where?/Where are the children playing?
Let's ask Beeple about the person the fairy is talking to.	The fairy is talking to who?/Who is the fairy talking to?
Let's ask Beeple about the thing the boy has.	The boy has what?/What does the boy have?
Let's ask Beeple about the place the cat is at.	The cat is where?/Where is the cat?
Let's ask Beeple about the place the man gets his books from.	The man gets his book from where?/Where does the man get his book?
Let's ask Beeple about the person the girl gives the balloon to.	The girl gives the balloon to who?/Who does the girl give the balloon to?
Let's ask Beeple about the lady the kids are playing with.	The kids are playing with who?/Who are the kids playing with?

*Practice trial.