CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Structural and interactional aspects of adverbial sentences in English mother-child interactions: an analysis of two dense corpora

Laura E. DE RUITER[1,2]* iD, Heather C. P. LEMEN[2], Elena V. M. LIEVEN[2], Silke BRANDT[3], and Anna L. THEAKSTON[2]

[1]Tufts University, USA,, [2]University of Manchester, UK,, and [3]Lancaster University, UK
*Corresponding author: Laura E. de Ruiter, Tufts University, Eliot-Pearson Department of Child Study and Human Development, 105 College Ave., 02155 Medford, MA (USA) E-mail: laura.deruiter@tufts.edu

Abstract

We analysed both structural and functional aspects of sentences containing the four adverbials "after", "before", "because", and "if" in two dense corpora of parent-child interactions from two British English-acquiring children (2;00–4;07). In comparing mothers' and children's usage we separate out the effects of frequency, cognitive complexity and pragmatics in explaining the course of acquisition of adverbial sentences. We also compare these usage patterns to stimuli used in a range of experimental studies and show how differences may account for some of the difficulties that children have shown in experiments. In addition, we report descriptive data on various aspects of adverbial sentences that have not yet been studied as a resource for future investigations.

1. Introduction

Once children move beyond the earlier stages of language learning and start producing multi-clause sentences, this allows them to make the relationship between what they are referring to in the world more linguistically explicit (Braunwald, 1985). Using relative clauses, they learn to add more information about the referent. Using adverbial clauses, they learn to talk about the relationship between events in the world. For instance, temporal adverbials like *before* and *after* sequence these events, and causal adverbials such as *because* and *if* refer to the causal relationships between events as well as to their temporal relationship. Thus, the study of how children learn to use these structures provides an important insight into the developmental interaction between their grammatical knowledge and how this relates to real world semantics and discourse organisation.

Although children's acquisition of sentences with adverbial clauses (henceforth 'adverbial sentences') has been an active research field since the early 1970s

CrossMark

(e.g., Amidon, 1976; Carni & French, 1984; Clark, 1971; Emerson, 1979), studies have yielded conflicting results on children's ability to comprehend these complex sentences and the age at which they are able to do so. While children start producing some adverbial sentences around the age of 3;0 (Diessel, 2004), in some experimental studies they show difficulties in comprehension until much later ages (Emerson & Gekoski, 1980; Johnson & Chapman, 1980; Pyykkönen, Niemi & Järvikivi, 2003). They misinterpret the temporal order, or reverse cause and effect in causal sentences.

However, as has been argued elsewhere for the comprehension of relative clauses (Brandt, Kidd, Lieven & Tomasello, 2009), the sentences typically used in comprehension experiments can be very different from the sentences that preschool children actually hear and/or use in everyday interaction with their caregivers. Thus, the conflicting findings on children's spontaneous production of adverbials and their comprehension of adverbial sentences in different experimental settings may reflect differences in the extent to which test sentences mirror those used in spontaneous speech. A usage-based approach would start from the adverbial sentences that children actually hear and produce and attempt to relate these to the patterns of children's comprehension in experimental settings.

In this study, we analyse both structural and functional aspects of sentences containing the four adverbials *after*, *before*, *because*, and *if* in a dense corpus of parent–child interactions from two British English-acquiring children. The aim of the paper is to analyse the relationship between the input and children's own production of adverbials in spontaneous speech to (i) examine whether children's patterns of learning are in line with a usage-based approach to acquisition and (ii) provide detailed information about the nature of spontaneously produced adverbial sentences and how they compare to the test sentences used in a range of experimental studies.

The structure of this article is as follows. We first outline the syntactic, semantic, and pragmatic factors that are thought to underpin children's performance and summarise the experimental evidence. We then consider the general role of input frequency and prototypicality (i.e., usage patterns) in language acquisition, and provide an overview of how experimental studies looking at the acquisition of adverbial sentences have differed in the stimuli they used. In the next section we review existing corpus studies of adverbial sentences produced by children and their caregivers. Then, after providing information about the current data set and the coding procedure, we report on the characteristics of adverbial sentences in our spontaneous speech data, relate them to some of the experimental findings from previous research, and discuss from a broad perspective how discrepancies between natural data and experimental stimuli may account for diverging findings. In appendices C and D, we provide descriptives on additional structural and interactional aspects of adverbial sentences in the data set which we intend as a resource for the research community. Appendix E gives an overview of the various tasks used in experimental studies as well as more information on structural aspects of their stimuli (specifically, types of subject noun phrases and verb types used). The article concludes with a discussion and potential avenues for further research.

1.1. Factors affecting children's comprehension of adverbial sentences

Adverbial sentences vary along a number of syntactic, semantic, and pragmatic dimensions, and studies have produced conflicting evidence on how these affect children's comprehension. Here we summarise evidence for the effects of iconicity,

clause order, semantic complexity (i.e., the meaning of different adverbials) and pragmatic function.

Iconicity

Other things being equal, the suggestion is that adverbial sentences are easier to produce and understand when they are iconic: that is, when the order of the clauses reflects the order of events being referred to (Clark, 1971). Thus, *before* sentences should be easier to process if the main clause comes first *(He pats the dog before he jumps the gate)* whereas the reverse would be the case for *after (After he pats the dog, he jumps the gate)*. This also applies to *because* and *if* adverbial clauses where the cause should precede the effect *(If you step in the puddle your shoes will get wet; Because you stepped in the puddle, your shoes are wet)* and therefore the subordinate clauses should precede the main clause.

Many experimental comprehension studies have examined the effects of iconicity. Although some have concluded that children have an easier time understanding sentences containing temporal connectives when the clauses are presented in iconic order (e.g., Blything, Davies & Cain, 2015; Clark, 1971; De Ruiter, Theakston, Brandt & Lieven, 2018; Hatch, 1971), other studies did not find an advantage for iconic order effects (Gorrell, Crain & Fodor, 1989). Conclusions about the impact of iconic ordering on comprehension of causal and conditional sentences (e.g., *If he falls, he cries really hard* vs. *He cries really hard if he falls*) are even more varied. While some studies found an advantage for iconic ordering with these kinds of sentences (De Ruiter *et al.*, 2018), some did not (Corrigan, 1975) and still others found that the extent to which children use ordering information to process these sentences may vary by task (Emerson, 1979).

Syntactic clause order

Diessel (2004, 2005) suggested that, from a processing perspective, listeners should find isolated complex sentences easier to process if they occur in main-subordinate order (e.g., *He eats a green pear after he drinks some water*). The underlying assumption is that main-subordinate orders are less taxing for working memory (see also Hawkins, 1994). Listeners can first process the main clause fully, and later attach the subordinate clause to their mental representation. In subordinate-main sentences, in contrast, the initial adverbial (e.g., *After he drinks some water he eats a green pear*) signals immediately that the sentence is complex (the *after* clause 'needs' a main clause to form a complete sentence); the listener needs to keep the subordinate clause in working memory, and can process the sentence fully only after the entire sentence has been heard. Two studies found that children showed better understanding with main-subordinate orders (Amidon & Carey, 1972; Johnson, 1975), but others found no difference (Amidon, 1976; De Ruiter *et al.*, 2018).

Relatedly, the majority of studies has used adverbial sentences in statement form (e.g., *The girl rang the bell before the boy hit the bunny*, Feagans, 1980), but Carni and French (1984) used them as questions (e.g., *What happened before Jane sat in her little seat?*). Questions could be considered more difficult, but Carni and French do not report lower levels of comprehension compared to other studies.

Semantic factors

One semantic factor that affects comprehension of adverbials is the specific adverbial type. For example, Clark (1971) argued that certain adverbials such as *before* are

semantically simpler than others (e.g., *after*) which makes them easier for children to learn.[1] However, regarding the difference between the two connectives *before* and *after*, previous research has, again, produced divergent results. Beginning with Clark (1971), and in line with her findings, several studies have found moderate to strong advantages for *before* (Blything *et al.*, 2015; Blything & Cain, 2016; De Ruiter *et al.*, 2018; Feagans, 1980; Johnson, 1975), including faster response times in a picture-selection task to sentences containing *before* (Blything & Cain, 2016), while others did not observe a significant difference between the two (Amidon, 1976; Amidon & Carey, 1972; French & Brown, 1977; Gorrell, Crain & Fodor, 1989), and one study found the opposite: that is, *after* being acquired earlier/being easier than *before* (Carni & French, 1984).

Another semantic factor that may impact on comprehension is the number of dimensions of meaning encoded by different adverbial types. Temporal connectives such as *before* and *after* solely express a temporal relationship between the clauses. But in order to interpret causal and conditional connectives such as *because* and *if*, listeners must interpret both temporality and causality or conditionality (De Ruiter *et al.*, 2018; Emerson & Gekoski, 1980). In addition, conditionality can be of different types (simple, hypothetical, counterfactual). Indeed, De Ruiter *et al.* (2018) reported that children took longer to interpret sentences when the connectives expressed an additional meaning – in this case, causal or conditional – over and above the temporal ordering of the events.

Pragmatic function

Another aspect of variability of adverbial sentences is the pragmatic function they fulfil. According to a model proposed by Sweetser (1990), causal and conditional clauses like those headed by *because* and *if* can serve various functions in discourse (see also e.g., Haegeman, 1984; Pander Maat & Degand, 2001; Redeker, 1990; Van Dijk, 1979; Zufferey, Mak & Sanders, 2015). In Sweetser's (1990) model (see also Kyratzis, Guo & Ervin-Tripp, 1990) *because*- or *if*-clauses can perform a Content, Speech-Act or Epistemic function. A Content sentence expresses a "real-world" cause or sufficient condition (e.g., *Your shoes are wet because you stepped in a puddle/Your shoes will get wet if you step in a puddle*). In a Speech-Act sentence, the speech act is performed in the main clause, while the subordinate clause either explains the speech act (causal sentences; e.g., *Don't step in puddles, because you are getting your shoes wet*) or provides the conditions for it (conditional sentences; e.g., *Don't get your shoes wet, if you insist on stepping in puddles*). In an Epistemic sentence, the main clause constitutes a conclusion, which is supported by evidence in the subordinate clause (e.g., *You were stepping in puddles, because your shoes are all wet/You were stepping in puddles, if your shoes are wet*). Diessel and Hetterle (2011) showed that, cross-linguistically, the Speech-Act function is frequently used for different adverbials by adults. While Sweetser does not extend her model to temporal terms like *before* and *after*, Diessel (2008) shows that *before*-sentences can, and do, perform the Speech-Act function (e.g., *Uhm well before we get into the detailed discussion of all of this have you got something else, Mary?*) (ibid.: 473). However, he found that most temporal adverbials are Content sentences (based on the numbers reported,

---

[1]As well as discussing main-subordinate clause order and iconicity, Clark (1971) argued for a hierarchy of semantic features in which the feature 'time' dominates but *before* has a '+prior' temporal feature while *after* has a '–prior' feature, and thus negative polarity, which should make it later for children to acquire.

approximately 94% in his corpus), suggesting that this pragmatic variation is rare within the temporal connectives.

There is some evidence that pragmatic differences may result in differences in comprehension for children. Corrigan (1975) found that children (aged 3–7) were more accurate with sentences that expressed physical or affective causality (which most closely align with Sweetser's Content) compared to sentences that express concrete logical causality (which most closely aligns with Sweetser's Epistemic). No study we are aware of provides any comparison of either of these forms against the frequently produced Speech-Act form; however, nor are we are aware of any that compared the pragmatic forms for *if*.

Frequency and prototypicality

While iconicity, clause order, semantic and pragmatic factors may each have an influence on comprehension, it is also possible that there are interactions between these factors that could reflect differences in usage patterns. A usage-based account would predict that, other things being equal, children would follow the usage-patterns that they hear. Factors that would affect this are the relative frequencies of the different constructions, relationships between form and meaning (whether these are one-to-one or more complex) and the semantic complexity of the construction. For example, Brandt *et al.* (2009) studied English and German children's comprehension of object relative clauses (e.g., *the dog that the cat chased*), another type of complex sentence that has been assumed to be difficult for structural reasons (specifically, in these languages object relatives are assumed to be more difficult than subject relative clauses, e.g., *the dog that chased the cat*, due to their non-canonical patient-agent word order). They found that children understood these types of sentences as well or even better than subject relative clauses when the sentences had the prototypical properties found in spoken discourse – in this case, inanimate head nouns and pronominal subjects (e.g., *the car that we bought*). In other words, not all object relatives are equally difficult. If there are similar, prototypical features of the adverbial sentences that children hear – for instance, with respect to the clause order in which particular connectives occur – this could have an effect on children's ease of processing in experimental studies. More specifically, test sentences may be more or less difficult to understand as a function of the extent to which they mirror patterns in children's input.

There have – to our knowledge – not been any investigations of the links between input frequencies of adverbial sentences and children's comprehension of these sentences in experiments.

To begin to allow us to understand why there are somewhat mixed findings from previous studies, it is important to know the details of the adverbial sentences that children actually hear in their input as well as the context in which they produce these sentences. This will inform our understanding of the kinds of experimental contexts in which we might expect children to perform relatively well, and those that are likely to pose greater challenges. In the next section, we consider the main characteristics of adverbial sentences outlined above in child speech and their relationship with input.

1.2. Corpus studies to date

Early corpus studies of adverbial sentences were concerned with the (order of) emergence of various connectives (Bloom, Lahey, Hood, Lifter & Fiess, 1980;

Braunwald, 1985) in early child language, with the aim of explaining this emergence in terms of mainly semantic factors. Others were focussed on the development of children's ability to express a particular semantic relationship such as causality or conditionality (Bowerman, 1986; Hood, Bloom & Brainerd, 1979; Kyratzis *et al.*, 1990). However, these early studies did not look systematically at the input children received. This is problematic as it is possible that the order of emergence of these forms in the children's speech could be predicted by the frequency with which they appear in the speech that they hear. Thus, learning may reflect simply amount of exposure rather than anything deeper about the semantic or pragmatic properties of the sentences themselves. Another possibility is that the children's seeming lack of use of certain forms simply reflects sampling biases. Even 'naturalistic' data is biased as to the contexts in which it is collected and, the smaller the sampling time frame, the more this is a problem. Recordings are unlikely to take place during mealtimes, bath times or outside the house for obvious reasons of reducing the amount of ambient noise. Less frequently used sentence types overall or those that occur more often in unsampled contexts are less likely to occur in a given sample of speech. However, only by ruling out frequency-driven/sampling explanations is it possible to determine the role of other factors such as semantic complexity and, the denser the corpus, the greater the possibility of doing this (Tomasello & Stahl, 2004).

The most comprehensive corpus study was conducted by Diessel (2004), who analysed all types of complex sentences (i.e., complement clauses, relative clauses, adverbial clauses, and coordinate clauses) in data from eight American English-speaking children between 1;8 and 5;1 and their caregivers to determine the developmental pathways from simple to complex constructions. In the course of his analysis, Diessel also looked at the frequencies of adverbial sentences in the mothers' speech and correlated these with the mean age of appearance of these adverbial sentences in the children's speech. He found that many of the earliest produced connectives were the ones that appeared most frequently in the mothers' speech. For example, *when, because* and *if* were among the most frequently produced connectives by the mothers (13.7%, 13.1% and 10.8%, respectively) and appeared in the children's speech at 2;10, 2;5 and 3;0, respectively. Comparatively, *before* and *after,* which each only accounted for about 2% of connectives in the input, were not produced by children until 3;2 and 3;4, respectively. While these findings provide evidence in favour of a relationship between input and production, the pattern was not entirely consistent across all connectives. For example, children produced the connectives *so* and *but* earlier than they produced *when* and *if*, despite the latter two occurring more frequently in the input.

Thus, the raw frequency with which children heard particular adverbials in their input could not fully explain their order of acquisition. Diessel considered that syntactic factors such as clause order, and semantic/pragmatic factors may also have an impact. This is illustrated by the following example. The children's later-produced *if-* and *when*-sentences appeared in both main-subordinate and subordinate-main order in the input, while the earlier-produced *if-* and *when*-sentences appeared only in main-subordinate order. Diessel argued that subordinate-main ordering is more difficult for children for both processing and discourse-related reasons. First, as discussed above, it has been argued that a subordinate-main structure is more difficult to process due to the demands on working memory. Second, following others before him (Chafe, 1984; Ford & Thompson, 1986) Diessel argued that the basic function of initial subordinate clauses is "to present information that is pragmatically presupposed providing a

thematic ground for new information asserted in subsequent clauses" (Diessel, 2013: 343). He argues that, as promoting this type of discourse-level coherence is not yet likely to be of concern for young children, subordinate-main sentences do not appear in speech until later.

On the other hand, although only 3% of the sentences in Diessel's (2004) child data were in subordinate-main order, they appeared primarily with the two conditional adverbials *if* and *when*, and this pattern of usage aligns with broad trends observed in adult speech: conditional sentences mainly appear in subordinate-main order, causal sentences mainly appear in main-subordinate order, and temporal sentences vary, appearing in the order that reflects the chronology of the events (iconic order) (Diessel, 2005, 2013). As Diessel's (2004) study did not look at the frequency of use of different clause orders for different adverbial sentences in the mothers' data, it is difficult to establish the extent to which the effects of syntax, semantics and pragmatics can be separated out from the frequency of use of specific adverbial sentences with specific clause orders and to fulfil specific pragmatic requirements in the input children hear. Moreover, it is still an open question whether such differences between adverbial sentences might be able to explain some of the experimental findings.

Another aspect of adverbial sentences not studied by Diessel (2004) is that of how the clauses relate to each other pragmatically (i.e., whether they function as Content, Epistemic or Speech-Acts). However, a study by Kyratzis *et al.* (1990), who used Sweetser's (1990) framework to analyse the causal speech (sentences containing the connectives *because* or *so*) of 21 children ranging in age from 2;7–11;1 in conversation with their friends and family, found that young children's causal sentences do vary pragmatically. Specifically, they found that children 3;6 and under only produced Speech-Act sentences. Furthermore, while children 3;7 and above did produce sentences in all three categories, more than half were Speech-Act sentences and only between about 15–24% expressed Content causality, which is the type that is typically tested in comprehension studies (Emerson, 1979; Lucia A French, 1988; Homzie & Gravitt, 1977; Johnston & Welsh, 2000; Kuhn & Phelps, 1976). Although Kyratzis *et al.* (1990) did not report detailed patterns in the mothers' speech, they did comment that "a preliminary analysis of the *adults'* uses of causals in this corpus revealed that a vast majority were also Speech Act-Level causals" (p. 210). It is, therefore, possible that many experimental stimuli are not representative of the kinds of sentences children typically hear (ibid.).

To summarise, experimental studies have produced differing results with respect to the age at which children understand different types of adverbial sentences and with respect to the factors that influence comprehension. Corpus studies have provided some information about children's acquisition of these sentences and their patterns of usage in the input that children hear which sometimes align with the results of experimental studies (e.g., *if*-sentences in iconic subordinate-main order are both more frequent and comprehended better). However, experimental findings are often contradictory, because stimuli are not comparable across studies, or are not well-controlled within studies. Moreover, corpus studies to date have not provided sufficient information regarding both the structural and pragmatic properties of adverbial sentences in child-directed speech to allow a more detailed evaluation of conflicting experimental findings. More detailed information about the patterning of adverbial sentences in children's early speech and their input is needed to shed light on the contradictory findings, and to inform the design of future studies.

1.3. The present study

Using data from a dense corpus of parent-child interactions from two British English-acquiring children, we analyse both structural features (e.g., clause order,) and functional (pragmatic) aspects of sentences containing the four adverbials *after*, *before*, *because*, and *if*.[2] The denser sampling of these corpora allows us to check the relative frequencies of the various measures with more confidence than is allowed by the relatively thin sampling of previous studies. Tomasello and Stahl (2004) calculated that 'traditional' child language corpora (which collect data for 1–2 hours every 2–3 weeks) probably only capture 1–2% of a child's input on a rough estimate. As these authors point out, this means that relatively rare phenomena may not be captured for many weeks, or even months, after they actually occur in either the adult or the child's speech. This makes the calculation of relative frequencies and orders of emergence very difficult. The dense corpora analysed in our study captured between 5 to 10 hours of data in any one week yielding between 10–20% of the child's input – again on a rough estimate (Lieven & Behrens, 2012). This allows us to conduct more detailed analyses of both form and function than is possible when the number of utterances available in less dense corpora is very low. We first present new and more detailed data on adverbial sentences in child-directed speech and their relation to children's own productions of these sentences, and discuss the extent to which these data may be able to explain some of the, sometimes conflicting, experimental findings outlined above. We focus on those factors that have received attention in experimental research. Descriptive data on additional aspects of adverbial sentences that have so far not been studied, such as the form of subjects and the argument structure of the clauses that may be useful for future investigations, are presented in Appendices C and D.

2. Data and coding

The data come from two high-density developmental corpora (Lieven, Salomo & Tomasello, 2009), the Thomas and the Gina corpus, both of which are available on the CHILDES website (MacWhinney, 2000). The Thomas data spans the years from 2;6 to 4;11[3], totalling 254 one-hour-long recordings (for more details, see https://childes.talkbank.org/access/Eng-UK/Thomas.html). The Gina corpus is smaller. It spans the years from 3;01 to 4;7, with 118 one-hour-long recordings in total (for more details, see https://childes.talkbank.org/access/Eng-UK/MPI-EVA-Manchester.html). Figure 1 shows the mean length of utterance (MLU) for both children for each recorded month.

Both children come from middle-class backgrounds in the North of England, and their primary caregivers were their mothers. For the analysis of the children's speech, we analysed the complete data set. For a representative analysis of the input, we selected a slice of six weeks, starting with the children's third birthdays. We chose this period because it is around this age that children typically start producing complex sentences involving adverbials other than *because* (e.g., Diessel, 2004). The

---

[2]For consistency, the four adverbials will always be presented in this order (temporal adverbials first, then causal, then conditional) in all tables and figures. The only exception is Table 1, in which the adverbials are presented in the order of acquisition.

[3]The Thomas corpus actually starts at 2;00, but because children do not usually produce these connectives before two-and-a-half years (Bloom *et al.*, 1980; Diessel, 2004), and in order to keep the number of recordings of the two corpora comparable, we only used data starting at 2;06.

Table 1. Ages and first occurrences of complete complex sentences for each adverbial in Thomas' and Gina's speech, in order of acquisition.

Type	Thomas			Gina	
	Age	Sentence		Age	Sentence
if	2;09:18	play something else if you like.		3;00:04	if you look I have too.
because	2;10;21	because you 0are [*] crying you're baking.		3;00:04	yeah. [+ SR] (be)cause I'm sharing.
before	3;00:16	<before you get in> [?] I dropped a leaf.		3;05:03	sit down now on the train (be)fore it goes.
after	3;00:26	Mummy , after [?] I driving a@p (de)livering my [//] you're all my fishes I have another sweet.		3;06:02	after you've finished yeah +… [+ IN] I need that mirror to put in my bag.
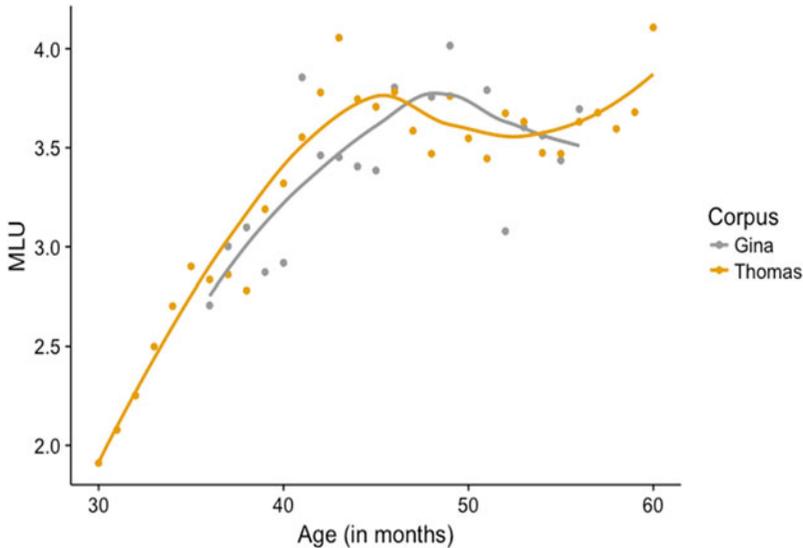
Figure 1. Scatter plot showing the mean length of utterance (MLU) for Gina and Thomas at a given age (in months).

period contains 26 recordings in the Thomas corpus, and 30 recordings in the Gina corpus. Because we were interested in the range of meanings mapped to *after, before, because*, and *if*, we extracted utterances from the database with all occurrences of these words whether or not they occurred in adverbial clauses. We included an analysis of other uses of the four words such as in phrasal verbs (e.g., *to go after someone*), because the frequency of specific form-meaning mappings can influence acquisition, with clear 1:1 mappings between form and function typically being easier than forms serving multiple functions (e.g., Bates & MacWhinney, 1987). If, for example, the word *after* is used often, but only rarely used as a conjunction, we may expect its conjunctive use to be difficult for children.

We included 20 lines preceding and five lines following each occurrence to provide context. Overall, we analysed 5631 utterances (3247 from the children, 2384 from the mothers). Each utterance was then coded for 26 semantic, morphosyntactic and pragmatic variables (see Appendix A for coding scheme).[4] We first coded whether an utterance was a complex sentence, an isolated clause, an incomplete utterance, whether it was sung[5], or whether the word was used in a different construction (variable COMPLEX). Note that a considerable proportion of the sentences – in particular, those with *because* – contained elliptical main clauses in response to requests or questions (e.g., "No, (be)cause you can't put that on him"; Gina at 3;00:12). These sentences were coded as complex, if they contained at least some elements of the main clause (usually "no" or "yes"), but not if they consisted only of

---

[4]The coding scheme contains additional variables such as a running utterance number, the corpus name, or the child's age.

[5]All instances of "sung" come from children or adults singing the song "If you're happy and you know it". These were coded as a separate function, because they are scripted and therefore not spontaneous utterances as part of a dialogue.

the subordinate (adverbial) clause. In that case, they were coded as isolated. If sentences with elliptical main clauses were counted as isolated, that would reduce the proportion of complex sentences in the data, especially for the children (10.1% of their *because*-sentences were elliptical). We are including these elliptical sentences to capture specific pragmatic meanings (see variable PRAGMATICTYPE below), in line with other studies investigating children's production of the different pragmatic types (Evers-Vermeul & Sanders, 2011; Kyratzis *et al.*, 1990).

Next, we coded for type of adverbial (i.e., *after, before, because, if*; variable TYPE), and whether the utterance contained combinations of the four subordinators (e.g., "and you've got to have tea <u>before</u> you go out <u>because</u> you're a tired boy"; Thomas' mother at 3;01:06; variable MULTISUB). This last variable was coded to gauge to what extent naturally occurring sentences may differ from experimental stimuli, which typically only use one adverbial per sentence. For each (complex) sentence, clause order (needed to determine iconicity) was coded (variable CLAUSEORDER). All complex sentences were also coded for whether they were a question or not (variable QUESTIONYN) and whether they were a reply to a question (variable REPLYQUESTIONYN). Finally, to study the distribution of the different discourse functions, we coded for PRAGMATICTYPE (Content, Speech-Act, or Epistemic) (see Appendix A for the coding scheme).

In addition to the variables above that are directly relevant for the evaluation of the experimental findings, we coded the utterances for a number of additional variables, which may be of interest to other researchers. Full details can be found in Appendices C and D.

The same coding scheme was used for both the child and the adult data. Two trained researchers (the first and the second author) coded the data. We tested the reliability of the coding scheme by having trained research assistants code a random sample of the data ($\sim$ 15%) and measuring agreement across all raters using free-marginal multirater kappa (κ) (Randolph, 2005). Unlike other agreement measures like Cohen's kappa (Cohen, 1960) or Fleiss' kappa (Fleiss, 1971), this measure does not assume that raters know a priori how many cases they should assign to each category of a variable, which is appropriate for our data. The overall agreement for all variables was .84, and .81 for all utterances that were actual adverbial sentences, as opposed to other uses of the words (see below). The mean interrater agreement for pragmatic coding (variable PRAGMATICTYPE) was $\kappa = .83$.

3. Results

The results are mostly descriptive, providing absolute frequencies and proportions. We use chi-square tests (using the Holm adjustment for multiple comparisons, where necessary) in some cases in order to test for significant distributional differences (e.g., between adults and children, or between the two mothers). Our main analyses collapse across the two mothers and the two children, but we also report any major differences between them. We first present results describing the overall pattern of use of the four adverbials in the input and the children's speech before turning to their structural (clause order, questionhood) and pragmatic (discourse function) properties, and discussing to what extent the distributions may shed light on the experimental findings.

Two factors thought to affect children's comprehension of adverbials are: (i) their consistency of form–meaning mapping (see discussion in De Ruiter *et al.*, 2018); and (ii) the frequency with which specific structures occur in their input. Thus, although

the focus of this article is on complex sentences with adverbials, it is informative to put this into the context of the overall usage of the four connectives. Figure 2 shows the absolute frequencies of the four connectives in both the children's and the mothers' speech, and how often they occurred in different uses (repetitions and recasts are included in this count).

The two temporal connectives, *after* and *before*, are relatively rare and, especially for *after*, occur more often in other constructions such as phrasal verbs ("it says 'please look after this bear'", Thomas' mother at 3;0:18) or adverbial phrases ("do you want a hot chocolate before bed?", Gina's mother at 3;1:06), both in the mothers' and in the children's speech. This is relevant, given the relative prominence that these two adverbials have received in the experimental literature and because a few experimental studies, including more recent ones with larger samples, have found that children understand *before* better/earlier than *after* (Blything *et al.*, 2015; Clark, 1971; De Ruiter *et al.*, 2018; Feagans, 1980; Johnson, 1975), despite its apparently low overall frequency of use. *Because* and *if* are much more frequent, with *because* being the most frequently used conjunction. Interestingly, in the children's data, *because* often occurs as an isolated clause providing a reply to a question ("'cause I don't want to"; Gina at 3;0:26), and is relatively more frequent as an isolated clause than the other adverbials in the input too. We return to the use of adverbial sentences as replies to questions in section 3.1.2. Figure 3 shows the relative proportion of the various uses for both children over time. Both children show a marked increase in the use of complex sentences and reduction in isolated clauses between 36 and 40 months.

Looking at the emergence of complex sentences, Thomas produced his first *if*-sentence at 2;09:18, although the next one did not appear in the sampled data until almost 1.5 years later. His first *because*-sentence occurred at age 2;10:21. The two temporal conjunctions emerged only later: *before* at age 3;00:16 and *after* at age 3;00:26. In Gina's data, we found the first complete sentences with *because* and *if* both at age 3;00:04. Note, however, that Gina's data collection starts only at age 3;00:01, so it is quite possible that she produced *because-* and *if*-sentences before that age. Her first *before*-sentence appeared at age 3;05:03, and her first *after*-sentence at age 3;06:02.

Both for Thomas and for Gina the earliest produced conjunctions (in complex sentences) were thus those that were most frequent in their mothers' speech, echoing Diessel's (2004) findings. Although apparently late acquisition of the temporal forms could reflect the likelihood of sampling these lower frequency forms, a comparison of the children's and mothers' data suggests that usage changes over development. Figure 4 shows the relative proportion of the four adverbials in the complex sentences of both the mothers' and the children's speech. Children produced a significantly higher proportion of *because*-sentences than mothers did (0.73 vs. 0.59, p < .0001), and a significantly lower proportion of *if*-sentences (0.24 vs. 0.35, p < .0001), and *before*-sentences (0.02 vs. 0.05, p < .001). There was no significant difference between adults in children for *after*.

In experimental studies children tend to do better on *before* than *after* (Blything *et al.*, 2015; Clark, 1971; De Ruiter *et al.*, 2018; Feagans, 1980; Johnson, 1975), and not better on *because* and *if* compared to temporal adverbials (De Ruiter *et al.*, 2018). We have already suggested that this may be because although *after* is more frequent than *before* in the input, it has a wider range of (non-temporal) meanings than *before*. Conversely, Figure 2 shows that *before* occurs more frequently in

Overall frequencies of types and use as a conjunction
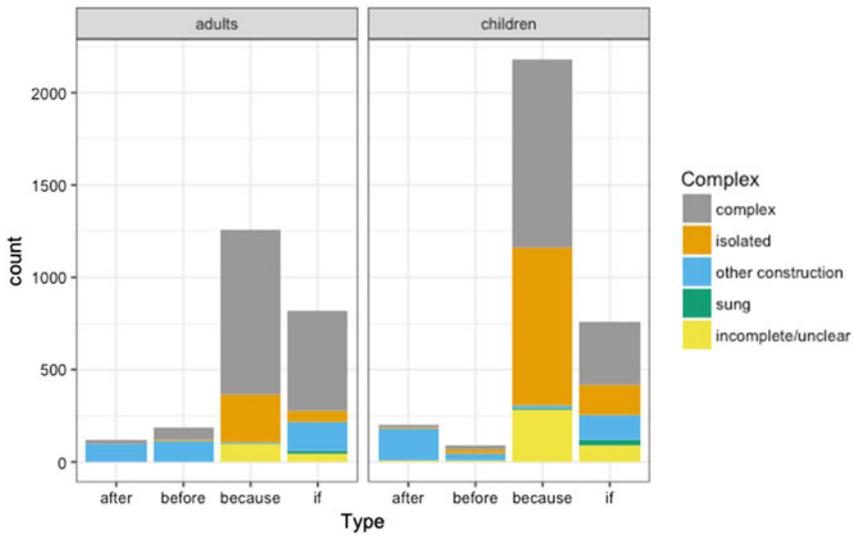
Figure 2. Bar chart showing absolute numbers of occurrences of each type of connectives, indicating the various uses for both adults and children. N = 5631.
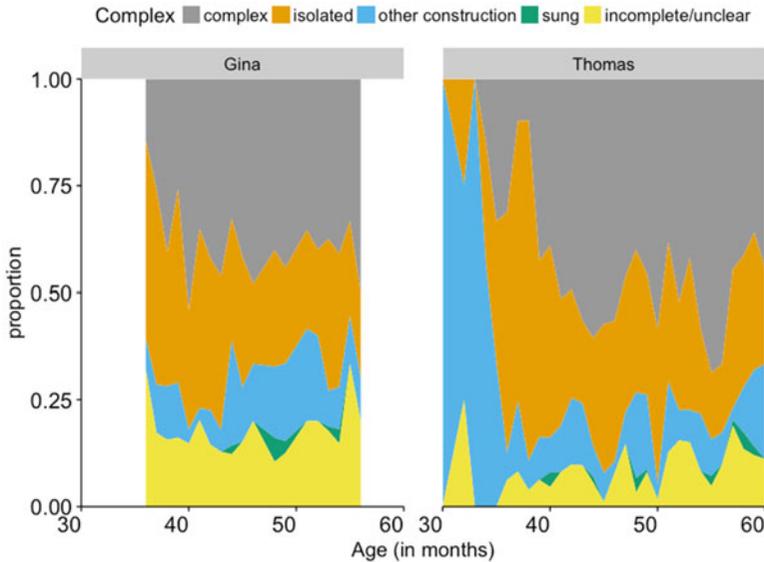
Figure 3. Stacked area chart showing the proportion of different uses of the four adverbials over time for both children. Proportions for each use were averaged over months (e.g., all instances from 3;00:1 up to 3;00:30 provide the data for the data point at 36 months). N = 3247.
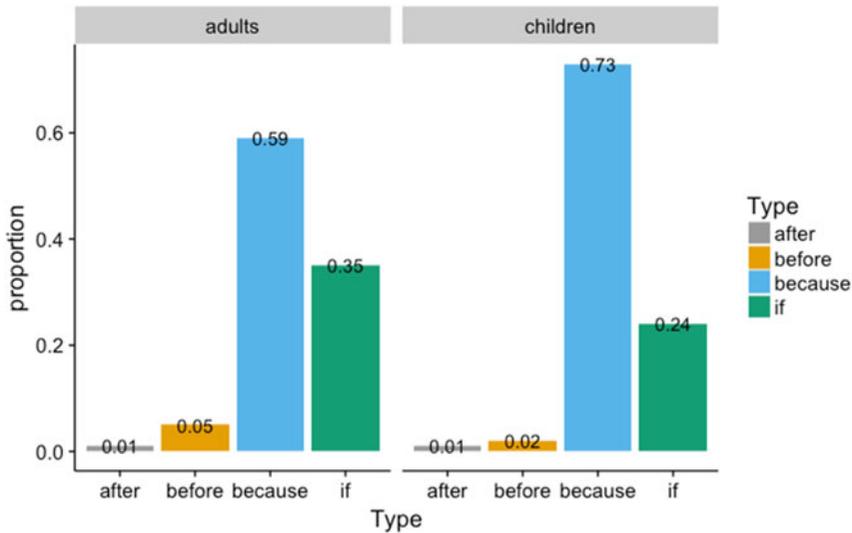
Figure 4. Bar chart showing the proportion of the four different adverbials for both adults (left panel) and children (right panel) in complex sentences only. N = 2924.

complex adverbial sentences than does *after,* which is likely also to be a factor in its better comprehension. However, frequency alone cannot account for the fact that children do not perform better on *because* and *if* in experiments despite the much higher frequency of usage both by adults and children. We return to the semantic and pragmatic factors that may account for this in section 3.2.1 below.

It is worth noting that, in the mothers' speech, approximately 11% of all complex sentences were combinations of two or more of the four conjunctions (MULTISUB variable), such as: "because the hippopotamus knows that if the crocodile goes to see the elephant who's going to squirt some water there'll be water everywhere" (Thomas' mother at 3;01:01).[6] This kind of syntactic complexity is bound to present an additional challenge for children. Thus, the raw frequency of use of the various conjunctions, even if only complex sentence types are considered, is unlikely to directly map onto their ease of acquisition.

3.1. Structural aspects

3.1.1. Clause order and iconicity

As outlined in the Introduction, the effects of clause order and iconicity have been the topic of many experiments with conflicting results. Results have varied as to whether iconicity determines ease of comprehension (Blything *et al.*, 2015; e.g., Clark, 1971; De Ruiter *et al.*, 2018) and/or whether the order of main and subordinate clauses is also involved. Figure 5 shows the proportion of main-subordinate and subordinate-main orders for the four adverbials. Both children and their mothers show the same type-specific clause order preferences: for *after* and *if*, the preferred clause order is subordinate-main, while for *before* and *because*, it is main-subordinate.

---

[6]The majority of complex sentences containing multiple subordinators were sentences that combined because and if; only very few combined temporal connectives with others.
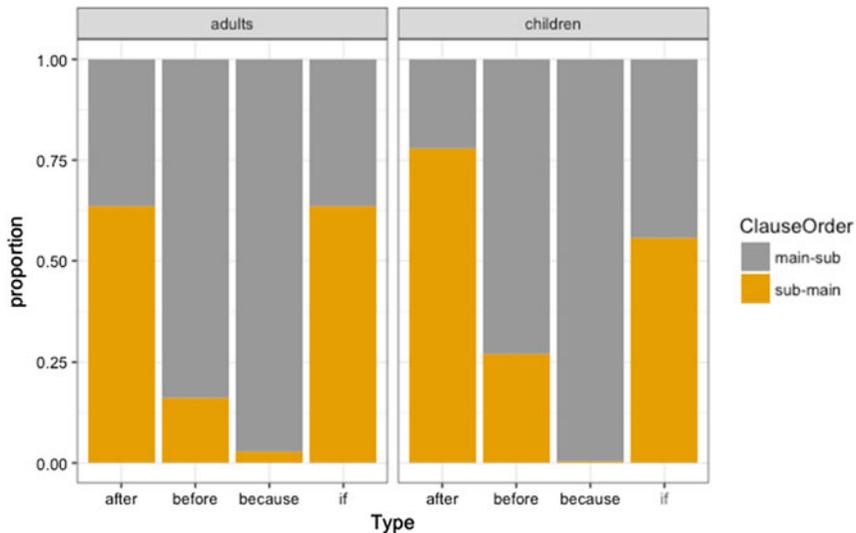
Figure 5. Bar chart showing the proportion of main-subordinate (main-sub) and subordinate-main (sub-main) orders for the four connectives for both adults and children. N = 2924.

Note that the clause order preference for the temporal adverbials, *after* and *before*, is iconic (i.e., the order of the clauses reflects the order of events being referred to). This supports Clark's (1971) original findings as well as those of a number of other experiments (Blything *et al.*, 2015; De Ruiter *et al.*, 2018) and suggests that the mutual influence of input frequency and iconic semantic mappings renders the understanding of temporal adverbial sentences in non-iconic order more difficult. Determining iconicity is less straightforward for *if-* and *because*-sentences.[7] Although, in purely temporal terms, subordinate-main order is iconic for *if* and is the preferred order in the corpus, this is not the case for *because* where subordinate-main is the iconic order but is very infrequent for both mothers and children. We will take this issue up again when we discuss the functional uses of those sentences (see section 3.2.1). Children and mothers differed only in that mothers used *because*-sentences significantly more often in subordinate-main orders (p < .0001), albeit still at a very low rate.

3.1.2. Questions and replies

With respect to questions, we coded all complex sentences for whether they were a syntactic question (e.g., "Did you want that orange juice before we start?", Gina's mother at 3;00;19), a pragmatic question – i.e., a non-interrogative sentence that is an indirect speech act of questioning (e.g., "You're tidying up before Dimitra comes?", Thomas' mother at 3;00;15) – or not a question (see Appendix A for the coding scheme). As can be seen in Table 2, for both groups, the majority of utterances were

---

[7]The pragmatic variation makes it additionally complicated as the more simultaneous nature of Speech-Act and Epistemic relationships means that ordering does not apply in the same way to these pragmatic types (e.g., Degand & Pander Maat, 2003)

Table 2. Absolute and relative frequencies of adverbial sentences formulated not as a question, formulated as a pragmatic question, and formulated as a syntactic question. N = 2924.

	Adults		Children	
	N	Proportion	N	Proportion
Not a question	1395	0.915	1342	0.959
Pragmatic question	22	0.014	4	0.003
Syntactic question	108	0.071	52	0.037

Table 3. Absolute numbers and relative frequencies of clause orders for each pragmatic type in because-sentences for both children and mothers. N = 1920.

	CLAUSEORDER			
	main-subordinate		subordinate-main	
PRAGMATICTYPE	N	proportion	N	proportion
Content	403	0.96	18	0.04
Epistemic	186	0.99	1	0.01
Speech Act	1304	0.99	7	0.01

not questions (0.96 children, 0.91 adults), but the small difference between adults and children was still significant (p < .0001).

Of all questions posed, the majority were syntactic questions using *because* (e.g., "Can I have some more [/] more of this [*] chocolate things in (be)cause I've ate all of them"; Gina at 4;06:00), for both adults (0.44) and children (0.64). Second most frequent were syntactic questions using *if* (e.g., "If the bin truck was dead or the trucks were smashed um [/] (.) um how [/] how would your rubbish get away?", Thomas at 4;03:06). These accounted for 0.27 of questions in adults, and 0.22 in the children. All other categories occurred rarely (see Table 8 and Table 9 in Appendix B for the detailed results). Overall then, mothers asked more questions than their children did. This is not surprising, as mothers asking many questions is an attested pattern in child-directed speech (e.g., Hoff-Ginsberg, 1991; Newport, 1977). But when children asked questions, they resembled the mothers in the use of syntactic vs. pragmatic means. The overall rarity of adverbial sentences as questions suggests that they may not be an ideal way of probing children's understanding, although the one study that did use this method for temporal adverbials (Carni & French, 1984) does not stand out as reporting a lower level of comprehension than others. A systematic comparison of methods could shed more light on this issue. As we suggested above, a larger difference between children and mothers emerged for the use of adverbial sentences as replies to questions.[8] While for both groups the

---

[8]Although replies to questions are an interactional aspect rather than a structural one, we decided to present the data on posing questions and replying to questions together in one section.

Table 4. Absolute numbers and relative frequencies of clause orders for each pragmatic type in if-sentences for both children and mothers. N = 878.

	ClauseOrder			
	main-subordinate		subordinate-main	
PragmaticType	N	proportion	N	proportion
Content	226	0.33	454	0.67
Epistemic	6	0.21	22	0.79
Speech Act	114	0.69	52	0.31

majority of utterances were not replies to questions (0.91 adults, 0.79 children), children used adverbial sentences as replies significantly more often (in 0.2 of the cases, in contrast with only 0.08 for the mothers, p < .0001). Making children respond with adverbial sentences to questions thus may be a more natural way of gauging their comprehension. Only a few studies have used *when-* or *why-*questions to do this (Amidon, 1976; Kun, 1978; Peterson & McCabe, 1985). Their observations suggest that when visual aids are provided, even pre-schoolers can demonstrate comprehension.

3.1.3. Other structural aspects
Analyses of the types of subjects in main and subordinate clauses and the verb types used by the mothers and their children can be found in Appendices C and D.

3.2. Interactional and pragmatic aspects
3.2.1. Pragmatic type
The last analysis of the data concerned the pragmatic types of *because-* and *if-*sentences (i.e., Content, Speech-Act, Epistemic). This is important because experiments with these forms almost exclusively use Content sentences (e.g., De Ruiter *et al.*, 2018; Emerson, 1979; Emerson & Gekoski, 1980; French, 1988; Johnson & Chapman, 1980; Kuhn & Phelps, 1976), but it is unclear whether this reflects what children hear and produce in naturalistic interactions. Temporal sentences with *before* and *after* can also sometimes be Speech-Act sentences, but this was not expected to account for a large portion of the data (Diessel, 2008). In our corpus data, all but four sentences (two *before*-Speech-Act sentences in the children's data and two in the mothers') were Content sentences, confirming this prediction.

We coded each sentence for whether it expressed a Content relationship (e.g., "Clock hand came off. (be)cause it was so windy.", Gina at 4;00:10), a Speech-Act relationship (e.g., "You can put your police helmet on, if you like.", Thomas' mother at 3;00:10), or an Epistemic relationship (e.g., "He won't reach my other strawberries because it's at the top.", Thomas at 3;10:03). Figure 6 shows the proportions of the three different pragmatic types for both mothers and their children. The patterns are very similar across speakers. For *because*-sentences, Speech-Acts were the most frequent type (between 0.46 and 0.78), while most *if*-sentences expressed Content causality (between 0.73 and 0.8). There were only two significant differences between mothers or between a mother and her child: in *because*-sentences, Thomas' mother used
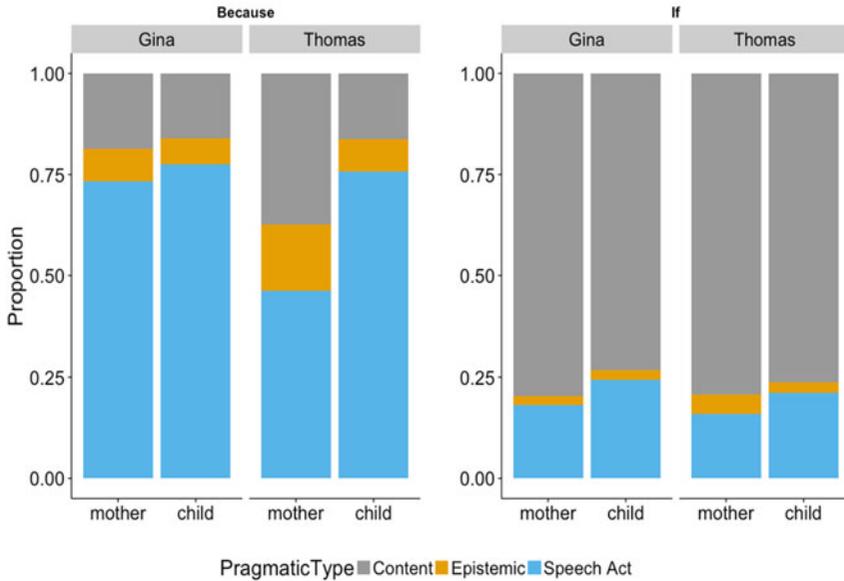
Figure 6. Proportion of Content, Epistemic, and Speech Act causality for both mothers and their children in *because*-sentences (left panel) and *if*-sentences (right panel). N = 2798.

significantly fewer Speech-Acts than both Gina's mother (0.46 vs. 0.73, p < .0001), and Thomas (0.46 vs. 0.76, p < .0001).

    Therefore, aligning with Diessel and Hetterle's (2011) findings in adult speech as well as Kyratzis *et al.*'s (1990) findings in child speech, Speech-Act is the most frequent function for *because* clauses in both child speech and child-directed speech. As noted earlier, the only study which provides any comparison of children's comprehension of *because* based on these kind of functional differences (Corrigan, 1975) overlooks Speech-Act causality, providing only a comparison of causal sentences which most closely align with the Content and Epistemic functions. [9] For *if,* we know of no studies which compare children's understanding on the basis of the different pragmatic forms.

    Looking at clause order, we found that for *because*-sentences, which were overwhelmingly in main-subordinate order, the little variation that was there was due to Content causality: out of only 26 sentences in subordinate-main order, 18 were Content uses (see Table 3). The same pattern has been reported by Kyratzis *et al.* (1990).

---

[9]Although Corrigan (1975) used different categories than those in Sweetser's (1990) model, the logical relationship between the main and subordinate clauses in concrete logical sentences can be seen to align with Epistemic causality. Similarly, the function of explaining the relationship between states/events described in both Corrigan's affective and physical causality align with Sweetser's Content causality. Corrigan (1975: 196) provides the following examples from each of categories used in the study: affective – *Peter cried because Jane hurt him*; physical – *She stayed home because she was sick*; concrete logical – *John had a white block because there were only white ones*.

Table 5. Absolute and relative frequencies of the different conditionals (simple, hypothetical, counterfactual) in children's and adults' *if*-sentences. N = 787.

	Adults		Children	
CONDTYPE	N	PROPORTION	N	PROPORTION
simple	429	0.799	307	0.9
hypothetical	94	0.175	29	0.085
counterfactual	14	0.026	1	0.003
unclear	0	0	4	0.012

For *if*-sentences, we found that Speech Act were more often in main-subordinate order, while Content and Epistemic uses were more often in subordinate-main order (see Table 4).

Note, however, that there were only 28 cases of Epistemic uses overall. Speech-Act uses occurred significantly more often in main-subordinate order than both Content uses (p < .0001) and Epistemic uses (p < .0001).

To summarise, these pragmatic patterns show that, unlike with *before* and *after,* for both *because* and *if* children hear and produce significant functional variation in how the clauses relate to one another. However, by far the greatest use of *because* by the children is in Speech-Act sentences (see also Kyratzis *et al.*, 1990), while experimental studies use Content sentences almost exclusively (e.g., De Ruiter *et al.*, 2018; Emerson, 1979; Emerson & Gekoski, 1980; French, 1988; Johnson & Chapman, 1980; Kuhn & Phelps, 1976). For *if* the use of Content sentences in experiments is matched by the high frequency of Content sentences with *if* in the corpus data. Among the *if*-sentences, Speech-Act uses stand out in that they occur more often in main-subordinate order. It is also worth noting that for *because*, in particular, both children were almost identical in their pragmatic proportions, despite differences in input patterns. We return to these points in the Discussion.

3.2.2. Types of conditionals in if-sentences

For all *if*-sentences (N = 878), we coded whether the sentence was a simple (or indicative) conditional (e.g., "I'm gonna get it on now if you don't let me", Gina at 4;01:04), a hypothetical (or subjunctive) conditional (e.g., "I'm sure if it was very dark that the dust(b)in wagon man would put his flashing lights on", Thomas' mother at 3;00:10), or a counterfactual conditional (e.g., "If Purdie had done that at your party she would have won a prize", Thomas' mother at 3;00:15).

Table 5 shows the absolute and relative frequencies of the types of conditionals for both adults and children. For both groups, simple conditionals were most frequent, but the percentage was significantly higher in children (0.9 vs. 0.799, p = .0006). Conversely, the mothers produced more hypothetical *if*-sentences (0.175 vs. 0.085, p = .0011). Test items in studies investigating comprehension of *if* tend to use simple conditionals (e.g., Amidon, 1976; De Ruiter *et al.*, 2018), aligning with the types that children hear and produce most frequently. This arguably overlooks the fact that almost a fifth of children's *if* input is in hypothetical form – which may complicate meaning for children. This will be further considered in the Discussion.

Table 6. Absolute and relative frequencies of the three types of *if*-conditionals across the three pragmatic types for both mothers and children. N = 878.

	CONDTYPE							
	counter-factual		hypothetical		simple		unclear	
PRAGMATICTYPE	N	prop.	N	prop.	N	prop.	N	prop
Content	12	0.02	114	0.17	551	0.81	3	0.00
Epistemic	1	0.04	4	0.15	23	0.82	0	0.00
Speech-Act	2	0.01	5	0.03	158	0.95	1	0.01

We also looked at the distribution of different *if*-conditionals across the three pragmatic types (Content, Epistemic, Speech Act). All pragmatic types occur most often with simple conditionals (see Table 6). Content conditionals show more variation with respect to *if*-conditionals than Speech-Act Conditionals, but there is no discernible difference between Content and Epistemic uses.

4. Discussion

We analysed adverbial sentences containing the conjunctions *after*, *before*, *because*, and *if* in two dense corpora of mother-child interaction. We used the data to find answers to two questions. First, what is the relationship between the input children receive, and their own production? Second, to what extent can the data help explain results from comprehension studies?

Our findings show that children's production of constructions containing *after, before, because*, and *if* closely reflects that of their mothers. The children's earliest and most frequently produced conjunctions are those that their mothers use most frequently (*because* and *if*), while those that are relatively rare in their mothers' speech, both overall and as conjunctions (*after* and *before*), emerge later, and are produced only infrequently. The majority of experimental studies has been conducted with *after* and *before*. Our finding that these conjunctions were quite rare is in line with what Diessel (2004) found in his corpus analysis. Given how little exposure children have to adverbial sentences with *after* and *before*, it is quite surprising that children perform as well in comprehension studies as they do. While the picture for younger children is mixed, four- to five-year-olds typically show accuracy rates around 60 to 80%, depending on the task (Amidon, 1976; Blything *et al.*, 2015; De Ruiter *et al.*, 2018). Still, the patterns we found may explain why more studies suggest that children have more difficulty with *after* than they do with *before*: *after* is not only overall very rare, it is also more often used in contexts other than adverbial sentences, such as in phrasal verbs (e.g., '"please look after this bear"). As has been argued for other linguistic forms and functions, clear 1-to-1-mappings between form and function are typically easier to acquire than forms that serve multiple functions (e.g., Bates & MacWhinney, 1987). *Before* would therefore be expected to be easier to learn than *after*.

Turning to *because* and *if*, our data and that of others show that children hear a lot of *because*- and *if*-sentences, yet have been found to perform no better with them than with temporal sentences (De Ruiter *et al.*, 2018), or show similar accuracy rates only at a later age (Emerson, 1979, 1980; Emerson & Gekoski, 1980). Mere input frequency of the adverbial forms themselves does not seem to account for the experimental findings. However, we suggest that more fine-grained usage patterns may explain the findings to some extent, if pragmatic function is considered. We found that *because*-sentences are primarily used for Speech-Acts (e.g., "Don't go on it yet (be)cause I need your help here", Gina's mother at 3;0.22). In contrast, experiments typically ask children to interpret *because*-sentences with Content causality (e.g., De Ruiter *et al.*, 2018; Emerson, 1979). If experiments test only one type of relationship, they may underestimate children's ability to comprehend these forms in other pragmatic contexts.

As an aside, we note that the pragmatic type of *because*-sentences was the only aspect for which we found differences between the two mothers: Thomas' mother used more Content causality compared to Gina's mother. This confirms an

impression that we gained already during coding. Thomas' mother often explained things to her son, while Gina's mother did this less often, and used *because*-sentences more often with speech acts (e.g., "Now be careful with these scissors, madam, because they're very sharp"; Gina's mother at 3;0:18). It is interesting that his mother's pattern is not reflected in Thomas' speech. His patterns are more similar to Gina and Gina's mother in that respect[10]. The tendency for children to use more Speech-Act causality is probably because young children are less able or less inclined to explain things to their parents than to give reasons for their actions, as is done with Speech-Act *because*-sentences (e.g., "I don't wanna open the book (be)cause you're doing my hair", Gina at 3;01:11), and they may learn to use *because*-sentences for this purpose first. This aligns with Kyratzis *et al.*'s (1990) suggestion that "the Speech Act-Level function of causals emerges earlier ontogenetically, since it is a practical one in terms of getting things accomplished in the child's world" (p. 210).

An additional difference between corpus data and experimental findings emerged for *because*-sentences with respect to clause order preferences. Overall, the children showed the same clause-order preferences for the four conjunctions as their mothers, with *after*- and *if*-sentences occurring predominantly in subordinate-main order, and *before*- and *because*-sentences in main-subordinate order. For *after*, *before*, and *if* the preferred orders are iconic, but for *because* the preferred order does not reflect iconicity (recall that *because*-sentences are iconic in subordinate-main order e.g., "Because it was cold, I put on a hat", but they are overwhelmingly produced in main-subordinate order, e.g., "I put on a hat because it was cold"). In comprehension studies, children find iconic orders with *because* easier than non-iconic orders in general, despite the fact that iconic *because*-sentences are rare in natural discourse. It appears that when children are confronted with Content uses of *because* (which are less frequent and thus less familiar than Speech Act uses), they find these easier to understand when the cause precedes the effect.

While pragmatic function differences can to some extent explain why children do not find *because*-sentences easier in experiments than temporal sentences despite *because*-sentences occurring so frequently in natural discourse, it is less clear why children do not perform better with *if*-sentences, where the most frequent pragmatic type (Content) is also that used in experiments. Again, the distributional properties of the input provide some possible explanations. *If*-sentences are less frequent in child-directed speech than *because*-sentences (537 vs. 897 occurrences in our data), and have the added complexity of occurring in different types of conditionals (simple, hypothetical, and counterfactual). In our data about 30% of Content *if*-sentences in the mothers' speech were hypotheticals or counterfactuals. We also note that the children in our data produced significantly fewer *if*-sentences than their

---

[10]Given that the two children differ in their similarity to input, to ensure that the findings about pragmatic type from the two mother-child dyads could be considered generalisable to a wider population, we coded an additional 12 mother-child dyads (Rowland & Theakston, 2009; Theakston & Rowland, 2009) using the same coding scheme for pragmatic types. The combined analysis of all 14 dyads revealed the following patterns: **Children** – *because*: .152 Content (SD = .068), .093 Epistemic (SD = .065), .755 Speech-Act (SD = .112); *if*: .526 Content (.249), .024 Epistemic (SD = .026); .45 Speech-Act (SD = .258); **Mothers** - *because*: .22 Content (SD = .069), .151 Epistemic (SD = .067), .629 Speech-Act (SD = .092); *if*: .692 Content (SD = .058), .024 Epistemic (SD= .037); .284 Speech-Act (SD = .075). Thus, the patterns for *because* are very similar, but the larger data set suggests that in children's speech, *if* Speech-Acts may be more frequent than the Thomas and Gina corpus indicates.

mothers, and that Speech-Acts dominate in children's speech overall. Furthermore, an analysis of a larger sample (see footnote 9) found that children use *if*-sentences for Speech-Act conditionality more often than their mothers do. Still, these explanations are tentative, and more research on children's understanding of different sentence types with different pragmatic types is needed. We also did not look at the different types of speech acts that the mothers and the children used (e.g., commissive, directive, assertive or as questions) (Searle, 1975). Future investigations could analyse in more detail what mothers and children do with these frequent pragmatic uses of *because* and *if*.

Our results also raise important issues for clinicians. Language impaired individuals often struggle to produce complex sentences (Marinellie, 2004; Nippold, Mansfield, Billow & Tomblin, 2008). Paying attention to the ways in which children hear these types of sentences in their everyday life could be used to inform the intervention programs used to help their development of more complex language.[11]

A final aspect in which adverbial sentences in child-directed speech differ from those used in experimental settings is context. In conversation, all adverbial sentences are embedded in the surrounding linguistic and non-linguistic context. In experimental settings, sentences are usually presented without context. This means that children have to construct a mental representation of the sentence without any scaffolding, which is something that they almost never have to do when interacting with their caregivers. It is likely that testing children on isolated sentences presents a greater challenge, and thus is more difficult than interpreting sentences in context. Indeed, recent research suggests that even minimal context improves children's comprehension of adverbial sentences significantly (De Ruiter, Brandt, Lieven & Theakston, 2020). Thus, even when adverbial sentences are constructed structurally and pragmatically in such a way as to reflect patterns in child-directed speech, children are likely to find them harder when given the task of interpreting them in isolation.

Our results suggest that while formal analyses of the syntactic, semantic and pragmatic features of these constructions are useful in setting up the framework for investigating how children learn them, a usage-based approach is crucial in identifying the actual learning path. Without an analysis of what children are hearing and how it relates to their production, we are likely to be misled, as well as to design experiments that do not match what they're used to. Thus, the two fundamentals of a usage-based theory: the importance of distributional patterns and the nature of form-to-function mappings are strongly supported by the analyses presented in this paper.

5. Conclusion

In the main, children's usage of adverbial sentences follows that of their parents in terms of frequency, structure and pragmatics. Deviations from this pattern are interesting for what they tell us about development. Initially the children use more isolated adverbial clauses and elliptical structures, often to answer their mothers' questions, which may be one means of learning how to produce full complex sentences (Bloom *et al.*, 1980; Diessel, 2004). Despite the fact that utterances with *after* and *before* are better comprehended in experimental studies than those with

---

[11]We are grateful to a reviewer for drawing our attention to this point.

because and *if*, they are much less frequent overall. Most *because* utterances in the corpora were Speech-Acts while in experiments they are exclusively about Content causality. While this can explain the difficulties that children have with *because* in experiments, explanations for difficulty with *if*, which appears in Content sentences in both corpora and experiments, probably relate to the multiple available form-meaning mappings for conditional forms. Thus, in comparing mothers' and children's usage we have been able to separate out the effects of frequency, cognitive complexity and pragmatics in explaining the course of acquisition of complex adverbials as well as explaining some of the differences between naturalistic usage and experimental results.

List of Appendices (B-E are in Supplementary Materials)
Appendix A: Coding Scheme for main paper
Appendix B: Supplementary tables for main paper
Appendix C: Coding scheme for additional analyses
Appendix D: Tables of additional analyses
Appendix E: Overview of tasks, noun phrase and verb types used in experimental studies.

References

Amidon, A. (1976). Children's understanding of sentences with contingent relations: Why are temporal and conditional connectives so difficult? *Journal of Experimental Child Psychology*, *22*(3), 423–437. https://doi.org/10.1016/0022-0965(76)90106-5

Amidon, A., & Carey, P. (1972). Why five-year-olds cannot understand before and after. *Journal of Verbal Learning and Verbal Behavior*, *11*(4), 417–423. https://doi.org/10.1016/S0022-5371(72)80022-7

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. *Mechanisms of Language Acquisition*, 157–193.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, *7*(2), 235–261. https://doi.org/10.1017/S0305000900002610

Blything, L. P., & Cain, K. (2016). Children's processing and comprehension of complex sentences containing temporal connectives: The influence of memory on the time course of accurate responses. *Developmental Psychology*, *52*(10), 1517–1529. https://doi.org/10.1037/dev0000201

Blything, L. P., Davies, R., & Cain, K. (2015). Young Children's Comprehension of Temporal Relations in Complex Sentences: The Influence of Memory on Performance. *Child Development*, *86*(6), 1922–1934.

Bowerman, M. (1986). First steps in acquiring conditionals. In E. C. Traugott, A. T. Meulen, J. S. Reilly, & C. A. Ferguson (Eds.), *On Conditionals* (pp. 285–308). Cambridge University Press.

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: An investigation of young German and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics*, *20*(3), 539–570. https://doi.org/10.1515/COGL.2009.024

Braunwald, S. R. (1985). The development of connectives. *Journal of Pragmatics*, *9*(4), 513–525. https://doi.org/10.1016/0378-2166(85)90019-0

Carni, E., & French, L. (1984). The Acquisition of Before and After Reconsidered: What Develops? *Journal of Experimental Child Psychology*, *37*, 394–403. https://doi.org/10.1016/0022-0965(84)90011-0

Chafe, W. (1984). How People Use Adverbial Clauses. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 437–449. https://doi.org/10.3765/bls.v10i0.1936

Clark, E. V. (1971). On the acquisition of the meaning of "before" and "after." *Journal of Verbal Learning and Verbal Behavior*, *10*, 266–275.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Corrigan, R. (1975). A Scalogram Analysis of the Development of the Use and Comprehension of "Because" in Children. *Child Development*, *46*, 195–201.

De Ruiter, L. E., Brandt, S., Lieven, E. V. M., & Theakston, A. L. (2020). Interactions between givenness and clause order in children's processing of complex sentences. *Cognition*. https://doi.org/10.1016/j.cognition.2019.104130

De Ruiter, L. E., Theakston, A. L., Brandt, S., & Lieven, E. V. M. (2018). Iconicity affects children's comprehension of complex sentences: The role of semantics, clause order, input and individual differences. *Cognition*, *171*, 202–224.

Degand, L., & Pander Maat, H. (2003). A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. *Usage Based Approaches to Dutch*, *1*, 175–199.

Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge University Press.

Diessel, H. (2005). Competing motivations for the ordering of main and adverbial clauses. *Linguistics*, *43*(3), 449–470. https://doi.org/10.1515/ling.2005.43.3.449

Diessel, H. (2008). Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics*, *19*(3), 465–490. https://doi.org/10.1515/COGL.2008.018

Diessel, H. (2013). Adverbial Subordination. In S. Luraghi & C. Parodi (Eds.), *Bloomsbury Companion to Syntax* (pp. 341–354). Continuum.

Diessel, H., & Hetterle, K. (2011). Causal clauses: A cross-linguistic investigation of their structure, meaning, and use. In P. Siemund (Ed.), *Linguistic Universals and Language Variation* (pp. 21–52). Mouton de Gruyter. https://doi.org/10.4324/9780203209493_Introduction

Emerson, H. F. (1979). Children's comprehension of "because" in reversible and non-reversible sentences. *Journal of Child Language*, *6*, 279–300. https://doi.org/10.1017/S0305000900002300

Emerson, H. F. (1980). Children's judgements of correct and reversed sentences with "if" Children's judgements of correct and reversed sentences with 'if'*. *Journal of Child Language*, *7*(7), 137–155. https://doi.org/10.1017/S0305000900007078

Emerson, H. F., & Gekoski, W. L. (1980). Development of comprehension of sentences with "because" or "if." *Journal of Experimental Child Psychology*, *29*(2), 202–224. https://doi.org/10.1016/0022-0965(80)90016-8

Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains—On the acquisition of causal connectives. *Journal of Pragmatics*, *43*(6), 1645–1662. https://doi.org/10.1016/j.pragma.2010.11.015

Feagans, L. (1980). Children's understanding of some temporal terms denoting order, duration, and simultaneity. *Journal of Psycholinguistic Research*, *9*(1), 41–57. https://doi.org/10.1007/BF01067301

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378.

Ford, C. E., & Thompson, S. A. (1986). Conditionals in discourse: A text-based study from English. *On Conditionals*, 353–372.

French, L. A. (1988). The Development of Children's Understanding of "Because" and "So." *Journal of Experimental Child Psychology*, *45*, 262–279.

French, L. A., & Brown, A. L. (1977). Comprehension of "before" and "after" in logical and arbitrary sequences. *Journal of Child Language*, *4*(2), 247–256. https://doi.org/10.1017/S0305000900001641

Gorrell, P., Crain, S., & Fodor, J. D. (1989). Contextual information and temporal terms. *Journal of Child Language*, *16*(3), 623–632. https://doi.org/10.1017/S0305000900010758

Haegeman, L. (1984). Pragmatic conditionals in english. *Folia Lingüística*, *18*, 485–502.

Hatch, E. (1971). The young child's comprehension of time connectives. *Child Development*, *42*(6), 2111–2113 ST-The young child's comprehension of. https://doi.org/10.2307/1127618

Hawkins, J. A. (1994). *A PerformanceTheory of Order and Constituency* (Cambridge, Vol. 73). Cambridge University Press.

Hoff-Ginsberg, E. (1991). Mother-Child Conversation in Different Social Classes and Communicative Settings. *Child Development*, *62*(4), 782–796.

Homzie, M., & Gravitt, C. B. (1977). Children's reproductions: Effects of event order and implied vs. Directly stated causation. *Journal of Child Language*, *4*(2), 237–246.

Hood, L., Bloom, L., & Brainerd, C. J. (1979). What, When, and How about Why: A Longitudinal Study of Early Expressions of Causality. *Source: Monographs of the Society for Research in Child Development*, *44*(6), 1–47. https://doi.org/10.2307/1165989

Johnson, H. L. (1975). The meaning of before and after for preschool children. *Journal of Experimental Child Psychology*, *19*(1), 88–99. https://doi.org/10.1016/0022-0965(75)90151-4

Johnson, H. L., & Chapman, R. S. (1980). Children's judgment and recall of causal connectives: A developmental study of "because," "so," and "and." *Journal of Psycholinguistic Research*, *9*(3), 243–260, https://doi.org/10.1007/BF01067240

Johnston, J. R., & Welsh, E. (2000). Comprehension of "because" and "so": The role of prior event representation. *First Language*, *20*, 291–304.

Kuhn, D., & Phelps, H. (1976). The Development of Children's Comprehension of Causal Direction. *Child Development*, *47*(1), 248–251. https://doi.org/10.2307/1128307

Kun, A. (1978). Evidence for preschoolers' understanding of causal direction in extended causal sequences. *Child Development*, 218–222.

Kyratzis, A., Guo, J., & Ervin-Tripp, S. (1990). Pragmatic conventions influencing children's use of causal constructions in natural discourse. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, *16*, 205–214.

Lieven, E., & Behrens, H. (2012). *Dense sampling*. In E. Hoff (ed.), *Research methods in child language: a practical guide* (pp. 226–239). Oxford: Blackwell.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, *20*(3), 481–507. https://doi.org/10.1515/COGL.2009.022

MacWhinney, B. (2000). *THE CHILDES project: The database* (Vol. 2). Psychology Press.

Marinellie, S. A. (2004). Complex syntax used by school-age children with specific language impairment (SLI) in child–adult conversation. *Journal of Communication Disorders*, *37*(6), 517–533.

Newport, E. (1977). Motherese: The speech of mothers to young children. In N. Castellan, D. Pisoni, & L. Gleitman (Eds.), *Cognitive theory* (Vol. 2). Erlbaum.

Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development. *American Journal of Speech-Language Pathology*.

Pander Maat, H., & Degand, L. (2001). Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics*, *12*(3), 211–245. https://doi.org/10.1515/cogl.2002.002

Peterson, C., & McCabe, A. (1985). Understanding "Because": How Important is the Task? *Journal of Psycholinguistic Research*, *14*(2), 199–218.

Pyykkönen, P., Niemi, J., & Järvikivi, J. (2003). Sentence structure, temporal order and linearity: Slow emergence of adult-like syntactic performance in Finnish *SKY Journal of Linguistics*, *16*, 113–138.

Randolph, J. J. (2005). Free-Marginal Multirater Kappa (multirater K [free]): An Alternative to Fleiss' Fixed-Marginal Multirater Kappa. *Online Submission*.

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, *14*(3), 367–381. https://doi.org/10.1016/0378-2166(90)90095-U

Rowland, C. F., & Theakston, A. L. (2009). The acquisition of auxiliary syntax: A longitudinal elicitation study. Part 2: The modals and auxiliary DO. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1471–1492.

Searle, J. R. (1975). Indirect Speech Acts. In *Speech Acts* (pp. 59–82). Brill. https://brill.com/view/book/edcoll/9789004368811/BP000004.xml

Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press. https://doi.org/10.1016/S0378-2166(96)90003-X

Theakston, A. L., & Rowland, C. F. (2009). The acquisition of auxiliary syntax: A longitudinal elicitation study. Part 1: Auxiliary BE. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1449–1470.

Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough?. *Journal of Child Language*, *31*(1), 101–121.

Van der Auwera, J. (1986). Conditionals and speech acts. In E. C. Traugott, A. T. Meulen, J. S. Reilly, & C.A. Ferguson (Eds.), *On Conditionals* (pp. 197–214). Cambridge: Cambridge University Press.

Van Dijk, T. A. (1979). Pragmatic connectives. *Journal of Pragmatics*, 3(5), 447–456. https://doi.org/10. 1016/0378-2166(79)90019-5

Zufferey, S., Mak, W. M., & Sanders, T. J. M. (2015). A Cross-Linguistic Perspective on the Acquisition of Causal Connectives and Relations. *International Review of Pragmatics*, 7(1), 22–39. https://doi.org/10. 1163/18773109-00701002

Appendix A: Coding scheme for analyses in the paper

CODING SCHEME

General procedure for extracting and coding complex sentences in the input and in the children's own productions

- Files:
  - mothers: first 6 weeks starting at age 3;00:00
  - children: all available data
- CLAN search strings: *after, before, because, if* in the *MOT tier and *CHI tier, respectively
- Include preceding 20 lines in the transcript, and the following two lines
- Copy all utterances (see below) into an excel spread sheet, number all utterances (variable 1, ITEMNUMBER)
- Exclude utterances <u>from analysis</u> (but still enter them into the excel sheet) if
  - they are sung nursery rhymes
  - they are incomplete (marked in the transcript by [+ IN]) and syntactically not analysable)– note that this does not exclude isolated subordinate clause utterances, which are to be labelled as such in the spread sheet (see below)
  - the two clauses are separated either by intervening, grammatically unrelated material from the speaker or a complete turn (not a backchannel response) from the listener.
- Complex sentences that are spread across two lines in CHILDES (e.g., by having the main clause on one line and the subordinate clause on the next line), or across three lines if there is a backchannel or overlapping speech, are considered one utterance. Original punctuation should be retained, however.
  - Code each utterance that is a coded as "complex" in the COMPLEX variable for all of the variables specified in the table below, using the labels specified in the "Label" column in the table below.
  - Code utterances that are coded as anything other than "complex" in the COMPLEX variable only up to and including the Type variable. All other fields are to be coded NA.
- Note that an utterance is repeated again as an additional item in the spreadsheet if it consists of multiple subordinate clauses.

Examples and further explanations are provided underneath the table.

Variable number	Variable name	Description	Label	Examples
1	ITEMNUMBER	Running count of utterances as they are entered from the CLAN output		152
2	CORPUS	Indicates the corpus the utterance comes from		
3	AGE	Age of child at time of recording (file name) in the format YY;MM;DD		3;00:24
4	GROUP	Indicates whether the data comes from the children or the mothers	children	
			adults	
5	UTTERANCE	Utterance as copied from the transcript	*none (the utterance itself)*	it's a wonder she sent you a card after you said that
6	COMPLEX	Indicates whether utterance constitutes complex sentence or an isolated subordinate clause. N/A is used when utterance contains the search string, but the sentence is neither a complex sentence nor an isolated subordinate clause (e.g., when the preposition is part of an adverbial phrase)	complex	I have and that's where they're staying. because if you play with them all they'll start to melt „ won't they ?
			isolated	and because it was Mothering_Sunday +... [+ IN]
			incomplete	we've kept your cards and balloons up because last week Jake_Jimmy +//. [+ IN]
			sung	if you're happy and you know it clap your hands> [=! sings]. [+ R]

(*Continued*)

(*Continued.*)

Variable number	Variable name	Description	Label	Examples
			uncodable	if you press it siren away [//] again it's going away.
			N/A	Grandma and Grandad came over <u>after lunch</u> „ didn't they ?
7	Type	Type of subordinator	after	
			before	
			because	
			if	
8	MultiSub	Indicates for sentences whether they come from an utterance containing multiple subordinate clauses (of the four subordinators). N/A is used for all other sentences (i.e., non-complex ones).	Multisub	<u>I'm sure I've saw his taxi because</u> after I saw him walking this morning <u>I thought I saw him drive out in a taxi</u> well when we've put hot water in there to make hot tea we must keep it in the middle of the table , Thomas. because if it falls on the floor I'll be very [/] very upset
			singlesub	
			NA	
9	ClauseOrder	Indicating the order of main and subordinate clause. For isolated sentences or other uses of the adverbs/ conjunctions (N/A in variable Complex) use N/A.	main-sub	it's obviously raining because <Percy's er> [//] Percy's fire-'has gone out. no. (be)cause what happens when you start messing with water [>] and stuff. (*see below table for examples and further explanation*)

				sub-main	after we've wrapped them in pretty paper we'll wrap them again in brown paper.
				N/A	(be)cause you're gonna [: going to] lose it all like this.
26	PragmaticType	Indicates the pragmatic type based on Sweetser (1990) for *because*-and *if*-sentences. *Please see below table for more detailed coding instructions and more examples*		content	but he's had a very busy day (be)cause it's been his birthday oh I might break that chair if I sit on that.
				epistemic	I think it's the farmer (be)cause there's a tractor. they must be enjoying themselves if they're laughing
				speech act	you don't eat the orange bit [>]. (be)cause [<] that's wax. you can have one now if you want. stand [?] back (be)cause I'm doing a loud trick. you can go across here. because Bertie can fly +//. [+ IN]
25	CondType	For if-clauses, indicates the type of conditional. For laws of nature, logical deductions and predictions, clauses are marked "simple". For hypothetical events (events that might occur), "hypothetical" is used. For counterfactual events (events that are impossible or did not occur), "counterfactual" is used. For all other sentence types (i.e., before, after, because) and for non-complex sentences ("N/A" in variable Complex) use N/A.		simple	well Sonia and Isabel but they'll be very disappointed if they think it wasn't er +/. [+ IN]
				hypothetical	because if it was going past Mummy's house before Mummy had left her house to go to the platform (.) then either Mummy's running very [/] very

(*Continued.*)

Variable number	Variable name	Description	Label	Examples
				late and she would've missed the train anyway. or the train's going past too early.
			counterfactual	if I hadn't been a little bit more careful I think it could easily have bitten.
		For if-clauses when the type of conditionality cannot be determined	unclear	
		For complex sentences that are not if-sentences (i.e., *after, before, because*)	N/A	
27	QUESTIONYN	Indicates whether the utterance was a statement (no question), a question with question syntax (syntactic question) or a question with declarative syntax but transcribed with a "?" in the transcript (pragmatic question). Tag questions are not coded as questions	no question	
			syntactic question	oh Thomas_Henry <u>how are we going to get through</u> today because I'm very tired. <u>isn't that naughty</u> of somebody to put dust on Mummy's picture after she dusted it.?
			pragmatic question	you're tidying up before Dimitra comes ?
28	REPLYYN	Indicates whether the utterance is a reply to a question or a request, a reply to a self-posed question, or not a reply to a question/request.	no reply	

				reply	no. because the tape recorder is a bit noisy at the moment „ isn't it ?
				self-reply	why not ?- (be)cause he's big and he's got a beard.
29	REPEAT	Indicates whether the utterance is a verbatim repetition of the child's/mother's utterance, a recast or neither (i.e., a new independent contribution)		repetition	*See below table for examples*
				recast	
				independent	

Variable numbers, names, descriptions, labels and examples
Note that the variable numbers in the table below are not continuous. For ease of reading, the coding scheme has been divided: Below are the parts of the coding scheme that describe the variables analysed in this article. Appendix C provides the coding scheme for additional variables that are not discussed in the main paper. Analyses for some of these additional variables can be found in Appendix D.

Additional examples and explanations
COMPLEX

- Examples of "N/A" type utterances, i.e., utterances that are not complex sentences of the type that we are interested in nor isolated subordinate clauses:
  1. Adverbial phrases:
     - Grandma and Grandad came over <u>after lunch</u> „ didn't they ?; I think I need a drink <u>after all that talking</u>
     - and you had it <u>before your birthday</u>; Miss_Moppet ties up her head in a duster and sits <u>before the fire</u>
     - is that <u>because of</u> Fireman_Sam's helmet ?
  2. Infinite adverbial clauses:
     - after all that drivering [//] driving and delivering all those fishs you'll have another sweet „ will you ?
     - when you've touched Purdie you always wash your hands <u>before eating food</u>; they're bound to look again <u>before driving off</u> „ aren't they ?
  3. complement-taking verbs:
     - well I <u>don't know if</u> I've got any left.
     - <u>see if</u> you recognise anybody.
- I wonder if you should be wearing your Bob_the_Builder hat , Thomas to do this ?

MULTISUB

- Examples of utterances with multiple subordinate clauses (variable 6, MultiSub) and how they should be coded:
  1. "We went home because after we've had dinner we didn't feel very well."
     - Should be copied twice into the spread sheet
     - First utterance: MultiSub (because-after A), main clause: we went home, subordinate clause: because we didn't feel very well
     - Second utterance: MultiSub (because-after B) main clause: we didn't feel very well, subordinate clause: after we've had dinner
  2. "well you were sad because your cupboard went. because you used to play with that cupboard „ didn't you ?"
     - The first sentence should be treated as a complex sentence (MultiSub: complex), the second sentence as an isolated subordinate clause (MultiSub: isolated)
  3. "and if it's six o clock tonight that's an awful long time to wait because we [//] we're just having breakfast now „ aren't we ?"
     - Should be copied twice into the spreadsheet
     - First utterance: MultiSub: if-because A, main clause: that's an awful long time to wait, subordinate clause: if it's six o'clock
     - Second utterance: Multisub if-because B, main clause: that's an awful long time to wait, subordinate clause: because we're just having breakfast now, aren't we
  - Note that in cases in which two complex sentences follow each other or are conjoined by "and", but are not embedded in each other, these sentences are analysed separately and coded as MultiSub "N/A":
  1. "don't [/] don't draw near the pages. because Mummy-'has got some things written down and if you scribble on them (.) I'll think that I've used them."

- ○ Here the first complex sentence is "don't [/] don't draw near the pages. because Mummy-'has got some things written down"
- ○ The second complex sentence is "if you scribble on them (.) I'll think that I've used them."

CLAUSEORDER
- ○ As in the second example given for "main-sub" above, sentences that are of the structure "No/Yes/…, because" should be coded as main-sub. Typically there is an elided part:
    1. *MOT: right (.) would you like to eat that piece of bread ?
        *CHI: no (.) (be)cause it's yours.
        - ○ Here the ellipsis is "I wouldn't like to eat that piece of bread"
        - ○ In addition, all variables concerning the main clause should be coded as "N/A". Even though it is often possible to determine what has been elided, the principle is to avoid too strong interpretations, and not code something that has not been said.

REPEAT

- ○ An utterance should only be marked as "repeat" if the mother provides a verbatim repetition. If any changes occur, it is marked as "recast":
    1. *CHI: <u>you</u> better it before it melts.
        *MOT: <u>I</u> better eat it before it melts.

PRAGMATICTYPE
1. Because
The labels for these are CONTENT, EPISTEMIC and SPEECH-ACT (based on definitions given in Sweetser, 1990 and Kyratzis *et al.*, 1990).

- a. **CONTENT**: The subordinate clause provides a "real-world" cause for the event in the main clause. The function of these is to explain the specific cause of a state/event mentioned in the main clause.
    - e.g., He was barking. Because he wanted to get out. (Kyratzis et al, 1990: 206)
    - e.g., The chef set out the ingredients because he was about to start cooking.
    - e.g., but I'm just putting it on because I'm cold (Gina; 4;02:30) (from present dataset)
- b. **EPISTEMIC**: The subordinate clause provides an explanation of how a speaker arrived at the conclusion expressed in the main clause.
    - e.g., This is for gardening, because it's fat (Kyratzis *et al.*, 1990: 207)
    - e.g., The chef is about to start cooking, because he set out all the ingredients.
    - e.g., or perhaps it isn't Sue because she-'has got some new neighbours (Thomas' mother; 3;00:07) (from present dataset)
- c. **SPEECH-ACT**: The subordinate clause explains/justifies a speech act (illocutionary act) that is performed in the main clause (i.e., explains a speech/illocutionary act, instead of providing an explanation about how something occurred.)
    - e.g., Take the gloves off. Because they'll get dirty. (Kyratzis et al, 1990: 206)
    - e.g., Pass me the ingredients, because I am about to start cooking.
    - e.g., yeah. (be)cause I need to get them right. (Gina, 3;07:04) (from present dataset)

2. If
The labels for these are CONTENT, EPISTEMIC and SPEECH-ACT (based on Sweetser, 1990, with further support from Van der Auwera, 1986).

- a. **CONTENT**: The subordinate clause describes the sufficient conditions for a state or event. The main clause of these can be a speech/illocutionary act, provided the entire utterance is conditional (e.g., If you inherit, will you invest? (Van Auwera, 1986, p. 198).
    - e.g., If you get me some coffee, I'll give you a cookie (Sweetser, 1990: 114)
    - e.g., The chef sets out the ingredients if he is going to start cooking.
    - e.g., I'll turn you into a slug if you don't go now (Thomas; 4;10:05) (from present dataset)

b. **EPISTEMIC**: The subordinate clause provides the conditions (evidence) for drawing a conclusion that is expressed in the main clause. The function of these is to verbalise a deduction/inference.

e.g., If John went to that party, (then) he was trying to infuriate Miriam (Sweetser, 1990: 116)

e.g., The chef is going to start cooking, if he is setting out the ingredients.

e.g., there must be special crayons if they're fifty pound (Gina's mother; 3;00:12) (from present dataset)

c. **SPEECH-ACT**: The subordinate clause defines the conditions for a speech act (illocutionary act). Unlike Content sentences with speech/illocutionary acts in the main clause, in Speech-Act sentences it is the saying of the speech/illocutionary act, itself, that is conditional (e.g., If you saw John, did you talk to him? (Van Auwera, 1986: 198).

e.g., If I may say so, that's a crazy idea (Sweetser, 1990: 118)

e.g., I have set out the ingredients, if you are ready to start cooking.

e.g., I've got a sweet if he behaves (Thomas; 4;04:05) (from present dataset)