Contents lists available at ScienceDirect

# Journal of Pragmatics

journal homepage: www.elsevier.com/locate/pragma

# A comparison of the pragmatic patterns in the spontaneous *because-* and *if-*sentences produced by children and their caregivers

# H.C.P. Lemen, E.V.M. Lieven, A.L. Theakston\*

ESRC International Centre for Language and Communicative Development, School of Health Sciences, University of Manchester, Manchester, M13 9PL, UK

# ARTICLE INFO

Article history: Received 15 May 2020 Received in revised form 7 July 2021 Accepted 15 July 2021 Available online 20 September 2021

Keywords: Language acquisition Complex adverbial sentences Illocutionary acts Corpus analysis

# ABSTRACT

Findings from corpus (e.g. Diessel, 2004) and comprehension (e.g. De Ruiter et al., 2018) studies show that children produce the adverbial connectives because and if long before they seem able to understand them. However, although children's comprehension is typically tested on sentences expressing the pragmatic relationship which Sweetser (1990) calls "Content", children also hear and produce sentences expressing "Speech-Act" relationships (e.g. De Ruiter et al., 2021; Kyratzis et al., 1990). To better understand the possible influence of pragmatic variation on 2- to 4- year-old children's acquisition of these connectives, we coded the because and if Speech-Act sentences of 14 British Englishspeaking mother-child dyads for the type of illocutionary act they contained, as well as the phrasing following the connective. Analyses revealed that children's because Speech -Act sentences were primarily explanations of Statements/Claims, while their if Speech -Act sentences typically related to permission and politeness. While children's becausesentences showed a great deal of individuality, their if-sentences closely resembled their mothers', containing a high proportion of recurring phrases which appear to be abstracted from input. We discuss how these patterns might help shape children's understanding of each connective and contribute to the children's overall difficulty with because and if. © 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC

BY license (http://creativecommons.org/licenses/by/4.0/).

# 1. Introduction

The age at which children comprehend sentences containing the adverbial connectives *because* and *if* appears to be surprisingly late relative to the age at which they first produce them. More specifically, *because* generally appears in the speech of young children around the age of two-and-a-half and *if* first appears around the age of three (Bloom et al., 1980; Diessel, 2004), but some studies have concluded that children do not fully understand the relationships these connectives express until they are nine years old (e.g. Emerson and Gekoski, 1980). Various methods, such as matching sentences with pictures (e.g. De Ruiter et al., 2018; 2020; Emerson, 1979; Emerson and Gekoski, 1980; Kuhn and Phelps, 1976), elicited production/sentence completion (e.g. French, 1988; Johnston and Welsh, 2000), act-out tasks (e.g. Amidon, 1976; French, 1988), judgments of acceptability/truth (e.g. Emerson, 1980; Johnson and Chapman, 1980; Peterson and McCabe, 1985) and retelling stories (e.g. Amidon, 1976; De Ruiter et al., 2018, 2020), iconicity (e.g. De Ruiter et al., 2018; Emerson, 1979), explicit connective use (e.g. Homzie and Gravitt, 1977) and familiarity with the relationships expressed (e.g. French, 1988;

\* Corresponding author.

https://doi.org/10.1016/j.pragma.2021.07.016







E-mail address: anna.theakston@manchester.ac.uk (A.L. Theakston).

<sup>0378-2166/© 2021</sup> The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4. 0/).

Johnston and Welsh, 2000) impact children's understanding of these connectives and the relationships they signal. Despite these differing perspectives and methodologies, based on the types of test items used, it appears that almost all of the studies from this existing body of research share one core assumption: that children's understanding of these connectives is based on only one particular type of pragmatic meaning (see Donaldson, 1986 for related arguments).

According to a model by Sweetser (1990), however, these connectives can express three different pragmatic relationships (Content, Epistemic, Speech—Act) (see also e.g. Haegeman, 1984; Kyratzis et al., 1990; Pander Maat and Degand, 2001; Redeker, 1990; Van Dijk, 1979; Warchał, 2010; Zufferey et al., 2015 for related theories). Furthermore, while the sentences used to test children's comprehension of these connectives primarily express the Content relationship (e.g. Amidon, 1976; De Ruiter et al., 2018; 2020; Emerson, 1979; 1980; Emerson and Gekoski, 1980; French, 1988; Homzie and Gravitt, 1977; Johnson and Chapman, 1980; Johnston and Welsh, 2000; Kuhn and Phelps, 1976; Peterson and McCabe, 1985), corpus data show that the Speech—Act type is also frequent in input and production, occurring in higher or similar proportions to the Content type in children's naturalistic speech (De Ruiter et al., 2021; Kyratzis et al., 1990). This provides support for the idea (e.g. De Ruiter et al., 2021; Donaldson, 1986; Kyratzis et al., 1990) that the types of causal and conditional relationships young children hear and produce spontaneously do not necessarily align with the kinds on which they are tested.

In terms of how these input/production patterns may be expected to influence comprehension, it has been argued that children first acquire the forms they find most meaningful (e.g. Slobin, 1985). Additionally, a usage-based approach would predict a relationship between the distributional properties of the input and a child's own understanding and use of a form (e.g. Kirjavainen et al., 2009; Lieven et al., 2009; Tomasello, 2001). As such, we would expect that investigating patterns in how children hear and produce these connectives would be critical to understanding acquisition. However, the existing literature provides little information on these pragmatic patterns, instead often focusing on semantic understanding of the connectives (e.g. whether children make "mistakes", such as confusing cause with effect, McCabe and Peterson, 1985; or tracking developmental patterns in children's ability to use *if* to express different semantic concepts, such as predictive and hypothetical relationships, Reilly, 1986). Furthermore, when studies have investigated these pragmatic patterns, this tends to be in broader categories (e.g. including both *because* and *so* in the same study, Kyratzis et al., 1990). However, given that *because*- and *if* Speech–Act sentences are produced so frequently in the speech by, and to, young children (De Ruiter et al., 2021; Kyratzis et al., 1990), failing to understand how this pragmatic type, in particular, is used risks overlooking salient patterns in Speech–Act sentences, this study aims to provide a fuller picture of how 2- to 4- year-old children hear and use *because* and *if* and the extent to which meaningful patterns may be associated with this pragmatic type.

In this paper, we will first present theoretical accounts of pragmatic language use and findings from corpus data through which we establish typical pragmatic usage of the connectives *because* and *if*, focusing on Speech–Act usages. Then, through analysis of *because*- and *if* Speech–Act sentences produced by 14 English-speaking mothers and their 2- to 4-year-old children, we investigate the extent to which these patterns exist in the speech of young children relative to the input they hear. We will then offer a discussion on these overall patterns in children's speech and implications for future research.

#### 2. A pragmatic approach to because and if

#### 2.1. General pragmatic patterns

Although there have been many studies investigating children's comprehension of the connectives *because* and *if*, the stimuli for these studies are typically designed to test understanding of semantic cause-effect or conditional relationships (e.g. "she hears the doorbell, if/because she presses the button", De Ruiter et al., 2018, 2020) (see also e.g. Amidon, 1976; Emerson, 1979, 1980; Emerson and Gekoski, 1980; French, 1988; Peterson and McCabe, 1985), rather than the functional meaning of the connectives. As such, the causal/conditional relationships in these sentences are usually between states/events, and thus align most closely with what Sweetser (1990) calls Content relationships. However, according to Sweetser's (1990) model (see also Kyratzis et al., 1990), *because* and *if* can express three different types of pragmatic relationships between clauses:

1. **Content**: presents a "real-world" explanation (*because*) or sufficient condition (*if*) for an event/state. **Causal example**: John came back because he loved her (Sweetser, 1990, p. 77).

**Conditional example**: If Mary goes, John will go (Sweetser, 1990, p. 114).

2. **Epistemic**: the main clause expresses a conclusion that the speaker draws based on evidence expressed in the subordinate clause.

Causal example: John loved her, because he came back (Sweetser, 1990, p. 77)

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

3. **Speech-Act**: the main clause is a speech act and the subordinate clause provides the speaker's reason for the speech act (*because*) or conditions associated with its performance (*if*).

Causal example: What are you doing tonight, because there's a good movie on (Sweetser, 1990, p. 77).

Conditional example: If you went to the party, did you see John? (Sweetser, 1990, p. 120).

At this point we would like to offer clarification on the terminology that we will use in the paper. Although most authors (including Sweetser, 1990) use the term "speech act" for the main clause of Speech–Act sentences (e.g. "What are you doing

tonight" in the above causal example from Sweetser (1990, p. 77)), to avoid confusion between Speech–Act (the pragmatic category) and speech act (the main clause of Speech–Act sentences), we will hereafter use the term "illocutionary act" for the latter. Briefly, the idea of illocutionary acts is tied to speech act theory, wherein Austin (1962) explained that utterances contain both locutionary and illocutionary acts, and the latter is the "performance of an act in saying something as opposed to performance of an act of saying something" (pp. 99–100). Searle (1969) later used the term "speech act" to align with Austin's (1962) definition of illocutionary acts before adopting the latter label, himself. As such, while a theoretical discussion on the similarities/differences between speech acts and illocutionary acts is not the focus of this paper, we feel that this terminology is appropriate (for the purposes of this paper) to refer to the pragmatic functions performed by these main clauses.

While other approaches also acknowledge that these connectives express different pragmatic meanings (e.g. Haegeman, 1984; Pander Maat and Degand, 2001; Redeker, 1990; Van Dijk, 1979; Warchał, 2010; Zufferey et al., 2015), the differences are often framed in alternative ways. Addressing the pragmatic differences via a scalar approach, Pander Maat and Degand (2001) show that there are differences in the degree to which the information contained in causal sentences is related to the reality of events in the world around the speaker. In their model, they provide six categories of causal relationships, which differ in terms of speaker involvement (defined by Pander Maat and Degand, 2001 as the amount a speaker is involved in constructing the causal relationship that is expressed). These categories, ordered in ascending order of speaker involvement, are: non-volitional, volitional, causality-based epistemic, non-causal epistemic, speech act type 1 and speech act type 2 (Pander Maat and Degand, 2001). While this perspective differs from the three-category model Sweetser (1990) proposed, there is overlap between the frameworks: the former two categories in Pander Maat and Degand (2001) generally align with Sweetser's (1990) Content, the middle two with Sweetser's (1990) Epistemic, and the final two with Sweetser's (1990) Speech–Act. Within Pander Maat and Degand's (2001) framework, as speaker involvement increases, the amount that the causal relationship expressed is tied to "real-world" states/ events decreases. That is, they argue that at one end of the continuum there are non-volitional sentences, where the speaker merely "reports" (p. 217) a real-world cause–effect relationship, and at the other end there are speech act relationships, which "appear in discourse in response to the interactional needs of a specific/potential interlocutor, not to present facts or draw conclusions concerning the real world" (p. 225). Thus, Pander Matt and Degand (2001) show that information contained in Content sentences, and to a lesser extent Epistemic sentences,<sup>1</sup> is more restricted by events which are separate from the speaker; Speech-Act relationships, by contrast, are entirely of the speaker's own construction, semantically and pragmatically.

# 2.2. Functional trends in adults' Speech-Act because and if

Given, then, that Speech–Act sentences are entirely subject to a speaker's discretion in both form and function, it is interesting to note that adults seem to use them in particular ways. Diessel and Hetterle (2011) argue that, for adults cross-linguistically, *because*-clauses are often independent from the main clause. They also argue that these clauses function to offer "justifications or explanations of the controversial statement" (p. 46). Ford (1993) presents a similar account, arguing that *because*-clauses are used when speakers want to expand upon utterances which are perceived as posing a difficulty to an interaction (see also Ford and Mori, 1994). Furthermore, she explains that these *because*-clauses can express different types of information, including background information and more detail (Ford, 1993). Therefore, *because* Speech–Act clauses can vary in content, but have a particular function: to provide the listener with more information about any illocutionary act which may be viewed as contentious within a discourse.

In contrast to *because*-clauses, Speech–Act *if* has an association with politeness (e.g. Brown and Levinson, 1987; Sweetser, 1990; Van der Auwera, 1986; Warchał, 2010). For example, Brown and Levinson (1987) show that *if*-clauses often function as hedges, which are associated with politeness. Furthermore, Sweetser (1990) discusses an entire subset of *if* Speech–Act sentences, which she calls "politeness conditionals" (p. 118), but also explains that even "given conditionals" (e.g. "If (as we both know) you were at the party, how's Harry these days?, Sweetser, 1990, p. 129), seem more polite than illocutionary acts without the conditional subordinate clause. The association with politeness also appears to be entrenched in the form, as well as function: both Sweetser (1990) and Van der Auwera (1986) suggest that Speech–Act *if*-clauses like "*if I may say so*" (Sweetser, 1990, p. 118) and "*if I may say you to*" (Van der Auwera, 1986, p. 199) are idiomatic expressions of politeness (see also Brown and Levinson, 1987). Specifically with regard to *if I may say so*, Sweetser (1990) claimed that it has "become so idiomatic that it no longer has any genuine conditional value; for most speakers it simply marks politeness rather than carrying its literal meaning" (pp. 118–119). Therefore, although any illocutionary act can be justified (*because*-clauses) or have any conditions attached to its performance (*if*-clauses), there is evidence to suggest that adults use Speech–Act *because* and *if* to serve specific, and relatively consistent, functions in discourse, which may not be entirely reliant on semantic notions of cause or condition.

# 3. Pragmatic patterns in children's speech: how are because and if used?

# 3.1. General pragmatic patterns in children's because and if

There is evidence that children are sensitive to the fact that *because* and *if* can express all three of Sweetser's (1990) pragmatic functions (Content, Speech–Act and Epistemic). For example, Kyratzis et al. (1990) analysed both *because*- and *so*-sentences from

<sup>&</sup>lt;sup>1</sup> Pander Maat and Degand (2001) argue that, while causality-based epistemic sentences are still tied to events in the real-world (i.e. the speaker draws a conclusion about something based on evidence they observe in the world), noncausal epistemic are only based on regular patterns in the world, on which speakers can draw conclusions.

21 children, aged 2;7–11;1 in naturalistic speech data. They found that, while children aged 3;7–6;6 produced causal sentences in all three pragmatic categories, Speech–Act was the most common and was the only type produced by children 3;6 and younger. This pattern also seems to hold in some other languages; of the codable causal utterances produced in naturalistic speech (via connectives *want* and *omdat: "because"; dus: "so"; daarom: "that's why"*) by 12 Dutch children, aged 1;6–5;6, Speech–Act sentences were produced most frequently (Evers-Vermeul and Sanders, 2011), and data from a single French-speaking child between the ages of 0;10-4;01 showed that she favoured the Speech–Act function with *parce que* ("*because*"; Sekali, 2012). These studies suggest that, at least for English-, Dutch- and French-speaking children, usage of adverbial sentences coordinated with a causal connective is not strongly tied to the cause-and-effect Content relationships that are normally tested in studies such as De Ruiter et al. (2018, 2020) or Emerson (1979, 1980; Emerson and Gekoski, 1980).

As well as corroborating Kyratzis et al.'s (1990) finding that children produce all three pragmatic types with *because* but favour the Speech–Act function, De Ruiter et al. (2021) found evidence that children also express all three pragmatic functions with *if*. Using dense naturalistic corpus data from two English-speaking mother-child dyads, we investigated a wide range of syntactic and pragmatic elements for *because*, *if*, *before* and *after*. In contrast to the patterns we found in children's *because*-sentences, we found that Content-sentences were the most frequently (about 75%) produced pragmatic type for children's *if*-sentences. De Ruiter et al. (2021) also provided some context for these patterns in the children's data by comparing them to the pragmatic proportions produced by the children's mothers (primary caregivers). For both connectives, we found that the mothers' patterns were generally similar to the proportions reported for children, with both mothers producing more Speech–Act with *because* and Content with *if*. To confirm these findings, we later coded the pragmatic function of *because*- and *if*-sentences produced by a further 12 mother-child dyads (using corpus data from Rowland and Theakston, 2009; Theakston and Rowland, 2009) and found that the patterns held for both groups and connectives, except for children's *if*. Specifically, when the data from the additional children were considered, children produced almost as many *if* Speech–Act sentences as *if* Content (52.6% Content vs. 45% Speech–Act), despite the mothers still showing a preference for *if*Content (69.2%). Therefore, a preference for *because* Speech–Act was relatively consistent across mothers and children, but there was more individual variation in children's pragmatic usage of *if*, where children did not consistently show the same preference for Content as their mothers.

#### 3.2. Functional trends in children's Speech-Act because and if

In terms of evidence of how Speech–Act sentences are used (as discussed in terms of patterns in adults' speech in Section 2.2), we are not aware of any studies directly comparing functional patterns in young English-speaking children's *if*-sentences, specifically, and what evidence does exist is inconsistent. For example, while McCabe et al. (1983) found that young children primarily use *if* to bribe or threaten, Bowerman (1986) found no evidence of this. Additionally, without more information about the bribes or threats in McCabe et al. (1983), it is not clear how frequently these were Speech–Act sentences rather than Content sentences (e.g. compare "I will let you have those books if you let me play with that new toy" (Content) versus "I've got two books you can have, if you are going to let me play with that new toy" (Speech–Act)). As such, it is difficult to tell whether this function is closely associated with one pragmatic type over another, if indeed it can be clearly associated with one at all.

There is, however, some evidence that, like adults (e.g. Diessel and Hetterle, 2011; Ford, 1993; Ford and Mori, 1994) children also use *because* to explain utterances which may be deemed contentious in a discourse. Although we have found no study in English which has examined functional patterns in young children's *because*-sentences, exclusively, Hood and Bloom (1979) found that children's causals, in general, related to the issuing of directives (such as "could you read this to me cause I don't *know how*"; p. 12) or the stating of intentions (such as "*I want some milk cause I have a cold*"; p. 14), both of which often expressed a negative meaning or relationship. Diessel (2004) also shows that young children use both *because* and so to express information about their own interests/goals. Furthermore, Kyratzis et al. (1990), found that most of children's Speech–Act causals "justified control acts" (p. 209). However, the findings of these studies were based on production of both *because* and *so*, meaning it is not entirely clear how much these patterns can be ascribed to Speech–Act *because*, alone. Additionally, Kyratzis et al.'s (1990) coding scheme stated that, while the main clauses for Epistemic and Content sentences were assertions, the main clause of the Speech–Act sentences were responses, interrogatives and direct and indirect imperatives (p. 208), suggesting the possibility that their coding of Speech–Act sentences might have been biased towards those types of utterances (a criticism also raised by Evers-Vermeul and Sanders, 2011). Interestingly, there does appear to be some cross-linguistic evidence that children's *because* sentences function like those of adults. In discussion of the results of her study investigating how Italian-speaking children use perché (*because*) with both their teachers and their peers, Orsolini (1993) argued that

*because* has an indexical nature. It displays the link between "dispreferred" or "unexpected" actions and claims, on the one hand, and communicative acts that inform the addressee of the speaker's intentions and knowledge on the other (Orsolini, 1993, p. 116).

While these studies provide some indication that children's *because* Speech—Act sentences serve a specific function, differences in methodology (e.g. connectives included in the study, coding scheme differences, first language of the participants) means this data paints only a vague picture about how English-speaking children use Speech—Act *because* and whether there are consistent and/or salient functional patterns associated with its usage by, and to, young children.

Overall, we appear to have conflicting patterns for the two connectives with regards to how their usage aligns with functional patterns in input. In adult speech, Speech–Act *if*-clauses are strongly associated with politeness, sometimes taking an idiomatic form (e.g. Brown and Levinson, 1987; Sweetser, 1990; Van der Auwera, 1986). As such, they are arguably well suited to be

abstracted verbatim. Despite this, children's proportions of *if* Speech—Act are less consistent (De Ruiter et al., 2021), suggesting that, if children do use *if* Speech—Act to express politeness in this way, they do not necessarily produce it with the same frequency as their mothers and that there is more individual variation with its usage, in general. By contrast, although the overall proportion with which young children produce *because* Speech—Act sentences is similar to what they hear in the input (De Ruiter et al., 2021; Kyratzis et al., 1990), *because* Speech—Act sentences are very speaker- and discourse-specific and can vary greatly in the information they express (e.g. Ford, 1993; Pander Maat and Degand, 2001). This means that it seems unlikely that all of children's *because* Speech—Act sentences (which account for the majority of their total *because*-sentences; De Ruiter et al., 2021; Kyratzis et al., 1990) are sentences they have directly copied from their input. However, without knowing more about how children produce *because* and *if* Speech—Act sentences, and how this relates to input, we are not able to draw conclusions about whether these connectives express any consistent or salient functional meaning for children.

# 4. The present study

#### 4.1. Framework

Given the high proportions with which Speech–Act *because-* and *if-*sentences appear in the speech of young children (De Ruiter et al., 2021; Kyratzis et al., 1990), as well as their associations with particular functions in adult speech (e.g. Diessel and Hetterle, 2011; Ford, 1993; Sweetser, 1990), there is reason to believe that these connectives may have a functional meaning for children over and above the cause-effect/sufficient condition meanings they express in Content form. As such, we expect that investigating more specific patterns in how children hear and use this pragmatic type will have two key benefits. First, in identifying patterns in how children produce Speech–Act sentences, we will gain a better understanding of the patterns associated with children's usage of this pragmatic type. Second, by comparing whether children's patterns resemble their mothers' in terms of a.) function and b.) form (phrasing), we will develop a better idea of whether children are abstracting broader functional patterns or simply copying utterances directly from input. This information will give us a better idea of whether there are salient patterns associated with the Speech–Act usage which may influence children's understanding of what these connectives mean. In line with this, the research questions for this study are:

- 1. What types of illocutionary acts do children and caregivers produce in their because- and if Speech–Act sentences; and
- 2. Is there evidence of recurring phrases following the connective in children's Speech–Act sentences which may be indicative either of copying directly from the input or their learning of idiomatic phrases?

To address these research questions, we analysed the types of illocutionary acts co-occurring with *because* and *if*, as well as the phrasing following the connectives in both child and caregiver speech.

# 4.2. Methods

# 4.2.1. Corpus

Data from two mother-child dyads, Thomas and Gina, from the Max Planck corpus (Lieven et al., 2009) were analysed. These data can be found on the CHILDES website (MacWhinney, 2000). The total data available contains 379 h of recording for Thomas, recorded during the time he was aged 2;00:12–4;11:20, and 118 h for Gina, recorded while she was aged 3;00:01–4;07:29. To avoid including data from a developmental period when children typically do not produce complex sentences (Bloom et al., 1980; Diessel, 2004), and to provide an approximately equal number of utterances for the two children, Thomas's data before the age of 2;06:12 was not included in the present study. A summary of the data included is in Table 1.

#### Table 1

Recorded hours for Thomas and Gina corpora.

Thomas				Gina			
Age	No. hours	Frequency of recordings	Mean MLU	Age	No. hours	Frequency of recordings	Mean MLU
2;06:12-3;02:12	154	$5 \times 1$ h recordings each week, every week	2.59	3;00:01-3;01:11	30	$5\times1$ h recordings per week, every week	2.89
3;03:02-3;11:06	43	$5 \times 1$ h recordings, 1 week per month	3.64	3;02:00-3;11:06	40	$5 \times 1$ h recordings, 1 week per month	3.39
4;00:02-4;11:20	57	$5 \times 1$ h recordings, 1 week per month	3.60	4;00:00-4;01:11	29	$5 \times 1$ h recordings per week, every week	3.93
				4;02:29-4;07:29	19	Multiple (between 1 and 4) recordings one or two weeks a month	3.59
Total	254	Mean	3.28	Total	118	Mean	3.45

The mothers' data was taken from the six-week period following their child's third birthday. This resulted in 26 h of recording for Thomas's mother and 30 for Gina's mother. The recordings for Thomas and Gina were done at their home or at the child study centre at the University of Manchester over 1-h sessions. The mothers and children engaged in a number of activities, including

mealtimes, play-time and general conversation. Both children came from two-parent families living in a major urban area of the UK and from middle-class backgrounds. In addition, data from twelve mother-child dyads (seven female children) in an additional corpus of data (Rowland and Theakston, 2009; Theakston and Rowland, 2009) were coded. For each mother-child dyad in this corpus, there are 22 x 1 h-long recordings, two in every three-week period, plus three further hours of recordings at each of the start, middle and end of the study period. The children ranged in age at the first recording from 2;08–2;11 (mean 2;10) and were between 3;04–4;01 (mean 3;06) on the final recording. MLUs on the first recording ranged from 2.41 to 3.79 (mean 3.22) and from 2.92 to 4.15 (mean 3.43) on the final recording. Recordings took place in the family home during a play context.

#### 4.2.2. Procedure

Only complex adverbial *because*- and *if*-sentences (i.e. those with a clearly identifiable main and subordinate clause and where *because* or *if* functioned as a connective between the clauses) were coded. Isolates (where there was no main clause, such as *because I like to*, Thomas, 3;03:07) and where *because* or *if* appeared in non-complex adverbial forms (such as "*Is that because of Fireman Sam's helmet*" [Thomas' mother, 3;00:03] or *I'll see if I can find the milk* [Gina's mother; 3;00:08]) were removed. Additionally, only sentences that were functionally and structurally interpretable were included, so some utterances that were incomplete or that contained a significant number of words that could not be transcribed were removed. That is, if the speaker finished enough of the utterance that both the function and key structural/semantic elements could be coded, the item was kept; if these things could not be identified or key elements were not able to be transcribed, the utterance was removed from the dataset. The remaining data were then coded for pragmatic type. Using the model proposed by Sweetser (1990), each item was coded as Content, Speech–Act or Epistemic, as detailed in De Ruiter et al. (2021).<sup>2</sup> The pragmatic type coding scheme is provided in Appendix 1.

#### 4.2.3. Coding

All sentences which were coded as Speech-Act were then coded for the following:

4.2.3.1. Illocutionary act type. To determine functional patterns in these sentences, the specific illocutionary act performed in the main clause was identified. The labels used were modified and reduced from Snow et al. (1996) whose coding scheme captured some extremely fine distinctions between categories (e.g. differentiating between "agree to carry out act requested or proposed by other" and "agree to do for last time"). While this level of detail is useful in some analyses, it was overly specific for the broader patterns of illocutionary act use we were investigating. Therefore, the categories were collapsed to reflect 13 broad illocutionary acts: Ask, Agree, Approve/Praise, Command, Disagree, Disapprove, Permit, Promise/Offer, Request/Suggest, State intent, State/ Claim, Threaten, Warn/Advise. A coding scheme, including examples from the current dataset, is found in Appendix 2.

4.2.3.2. Post-connective phrasing recurrence. To determine whether the connectives were used to introduce any consistent/ recurring phrasing, we next investigated the post-connective verb phrase (VP) recurrence frequency. More specifically, to determine whether any children were using the connectives to introduce phrasing which was either copied directly from their mothers (which may vary by child) and/or may be considered idiomatic (which we would expect to then appear frequently in more than one corpus), the first full VP following the connective was coded.

We decided to use VPs (rather than simply coding the first five or six, for example, words following the connective) because, while idiomatic phrases, by nature, typically have a set form, they may be subject to some variation. For example, as argued by Reagan (1987, p. 418), both "*Pull Barbara's leg*" and "*Pull Kathy's leg*" are variations of the idiomatic form "*Pull X's leg*". The same seems to be possible with idiomatic phrases in *if*-clauses. For example, Van der Auwera (1986, p. 199) states that the subordinate clause in "*If I can speak frankly, he doesn't have a chance*" is idiomatic. Arguably, however, it would have the same meaning if the speaker had produced the subordinate clause *If we can speak frankly* instead. Thus, the first occurring VP was selected to provide a point for comparison of specific phrasing, without being impacted by variation in subject or other discourse information preceding the verb. The verb form was recorded ignoring variation in person, number, tense and polarity (e.g. "*yeah*. *[*+ *I*] (*be*)*cause a sun's nice*"; Gina, 3;10:02 was coded as "*be nice*"), but VPs were recorded in full to capture specific phrasing for later comparison (e.g. "*they don't*! (*be*)*cause these two are allowed on the bus*", Bob; 3;06:00, was coded as "be allowed on the bus", rather than just "be allowed"). This allowed us to order the data alphabetically around a lemma form of a given verb to evaluate specific phrasing patterns and consider frequencies.

As the purpose of this coding was to establish patterns in wording associated with the connective and not analyse semantic or syntactic features of the subordinate clauses, specifically, these VPs were not necessarily the main VP of the subordinate clause headed by the connective. For example, where a further complex adverbial sentence appeared after the connective, only the VP in the first appearing clause after the connective was coded (e.g., "because normally when we see them round the corner they're here in a few minutes"; Thomas's mother, 3;00:07, becomes "see them round the corner"). Additional discourse markers/particles following, but not included in, the VP were excluded (e.g. the "then" at the end of "<I'll put it away> [<] if Gina's not gonna [: going to] do it then", Gina's mother; 3;00:25 was excluded and the VP was coded as "be going to do it"), as were tag questions following the VP (e.g. "okay. (be)cause you like pink" don't you ?" was coded as "like pink").

To ensure that the wording for all VPs could be evaluated, we removed additional sentences where the first postconnective VP was uncodable due to unclear or incomplete information, for example, *er no*. (*be*)*cause I need* +//. [+ IN]

 $<sup>^{2}</sup>$  The results of this pragmatic type coding were reported in De Ruiter et al. (2021). Minor discrepancies between the data reported in De Ruiter et al. (2021) and here result from further checking and correcting of the pragmatic coding during the coding phase for this study.

(Billy, 3;01:02). This resulted in the removal of an additional 135 sentences across the entire dataset (14 (1.4%) from the mothers' *because*; 115 (6.4%) from the children's *because*; 2 (0.7%) from mother's *if*; 4 (1.9%) from children's *if*)).

#### 4.2.4. Reliability

Approximately 15% of the illocutionary acts were coded by an independent researcher. The average free marginal kappa was .71 for *because*-sentences and .76 for the *if*-sentences, which is substantial agreement (Landis and Koch, 1977).

# 5. Results

After those with unclear illocutionary acts were removed, a total of 3267 Speech–Act sentences were coded. Within the children's data, there were 1785 *because*-sentences and 214 *if*-sentences. Within the mothers' data, there were 976 *because*-sentences and 292 *if*-sentences (see Table 2  $^{3}$ ).

#### Table 2

Summary of total becau	se- and if-utterances	utterances by connec	tive and Speech-Act	utterances by speaker.
	· · · · · · · · · · · · · · · · · · ·	,	· · · · · · · · · · · · · · · · · · ·	

Speaker	Total no. because- and if-utterances	No. because	% Because Speech-Act	No. if	% if Speech–Act
INPUT					
Alice	124	83	76%	41	37%
Billy	112	70	63%	42	33%
Bob	11	8	63%	3	33%
Gina	637	400	77%	237	27%
Helen	142	73	68%	69	26%
Ivy	50	28	43%	22	14%
Jack	75	32	62%	43	26%
Lucy	73	40	70%	33	30%
Mary	179	96	67%	83	39%
Olga	141	81	68%	60	33%
Rebecca	183	88	60%	95	37%
Sid	59	29	62%	30	20%
Steve	93	64	59%	29	24%
Thomas	761	465	47%	296	19%
Mean input	188.6	111.2	63.2%	77.4	28.4%
CHILDREN					
Alice	4	3	100%	1	100%
Billy	128	103	69%	25	4%
Bob	121	114	78%	7	57%
Gina	554	427	78%	127	24%
Helen	150	126	79%	24	38%
Ivy	116	90	83%	26	58%
Jack	42	30	53%	12	67%
Lucy	140	98	77%	42	36%
Mary	268	195	62%	73	30%
Olga	260	207	71%	53	60%
Rebecca	59	41	88%	18	72%
Sid	347	272	70%	75	17%
Steve	125	110	78%	15	47%
Thomas	806	593	76%	213	21%
Mean children	222.9	172.1	<b>75.</b> 7%	50.8	45.0%

Descriptive data were used to explore the patterns in the speech of the different groups, with one-sided Wilcoxon signed rank comparisons used to assess the significance of any observed differences. The analysis was done in R (R Core Team, 2018) version 3.5.1 ("Feather Spray") using the coin package (Hothorn et al., 2019), with the default Pratt method (Pratt, 1959) for zeros and ties.

#### 5.1. Illocutionary acts

The illocutionary acts produced by the mothers for each connective were compared to those produced by the children. *Because* illocutionary acts are reported first, followed by *if.* For both connectives, results are presented first for children and then for mothers.

# 5.1.1. Because

The most common illocutionary acts children produced with their *because*-sentences were State/Claims (M = .382; SD = .116), followed by Commands (M = .271 SD = .089) and Request/Suggests (M = .083, SD = .05). Disapprove, Promise/ Offer and Threaten occurred rarely with *because* in the children's data (none accounted for more than 4% of any child's productions). Although there was some individual variation within the children, the trends were largely consistent (see Fig. 1).

<sup>&</sup>lt;sup>3</sup> De Ruiter et al. (2021) reports pragmatic coding for the Thomas and Gina data, repeated here, but provides only average values for the additional 12 dyads, full details reported here.



Fig. 1. Children's because illocutionary acts.

The primary illocutionary act performed with the mothers' *because*-sentences was Command (M = .368; SD = .102), followed by State/Claim (M = .145, SD = .086) and Request/Suggest (M = .124; SD = .078). Like with the children, there was a relatively high degree of consistency between the mothers. Disapprove, Disagree and Threaten were rarely produced by any of the mothers with *because* (none accounted for more than 5% of any mother's *because*-illocutionary acts) (see Fig. 2).



Fig. 2. Mothers' because illocutionary acts.

In comparison to their mothers, the children produced proportionately more State/Claims (.382 vs. .145, Z = 3.296, p = .001) and Disagrees (.044 vs. .006, Z = 2.826, p = .017). The output of the Wilcoxon tests for *because* are summarised in Table 3, with Fig. 3 providing a visual comparison. Significance (p) values have been adjusted using the Bonferonni correction for multiple comparisons.

# Table 3

Summary of Wilcoxon signed rank comparison of child and mother illocutionary acts with because.

	Children		Mothers		Significanc	e	
Illocutionary act	Mean	SD	Mean	SD	Y/N	р	Z
Agree	.021	.019	.021	.026	N	1	0.126
Approve/Praise	.012	.016	.058	.037	Ν	1	-3.078
Ask	.028	.023	.080	.056	Ν	1	-2.449
Command	.271	.089	.368	.102	Ν	1	-2.668
Disagree	.044	.036	.006	.011	Y	0.017	2.826
Disapprove	.007	.008	.012	.019	Ν	1	-0.159
Permit	.036	.029	.040	.056	Ν	1	0.282
Promise/Offer	.013	.013	.033	.029	Ν	1	-2.328

#### Table 3 (continued)

	Children		Mothers		Significanc	e	
Illocutionary act	Mean	SD	Mean	SD	Y/N	р	Z
Request/Suggest	.083	.050	.124	.078	N	1	-2.166
State intent	.061	.043	.053	.049	Ν	1	0.628
State/Claim	.382	.116	.145	.086	Y	0.001	3.296
Threaten	0	0	.003	.008	N/A	N/A	N/A
Warn/Advise	.043	.085	.058	.049	Ν	1	-1.507



Fig. 3. Because Illocutionary act types – children versus mothers.

# 5.1.2. If

We next examined the children's and mothers' *if*-sentences to establish whether they showed similar patterns of usage. Different patterns emerged from what was found with the *because*-sentences. Permit was the most frequent *if*-illocutionary act for the children (M = .398, SD = .239), followed by State/Claim (M = .188, SD = .258) and Request/suggest (M = .109, SD = .108). No child produced any Disagrees with *if* and only one child produced any Threats (accounting for only 3% of her *if* illocutionary acts). Fig. 4 shows the patterns of children's *if* illocutionary acts.





Like the children, Permit was the most frequently produced *if*-illocutionary act for the mothers (M = .199, SD = .096), followed by Ask (M = .166, SD = .255) and State/Claim (M = .155, SD = .124, see Fig. 5).



Fig. 5. Mother *if* illocutionary acts.

After applying the Bonferonni correction for multiple comparisons, there were no differences between the proportions in which children and mothers produced the different illocutionary acts with *if* (see Table 4 and Fig. 6).

#### Table 4

Summary of wheekon signed rank companson of chine and mounci moculionary acts with
--

	Children	Mothers		Significance		
Illocutionary act	Mean SD	Mean	SD	Y/N	р	Z
Agree	.024 .063	.016	.025	N	1	-0.169
Approve/Praise	.016 .038	.026	.055	N	1	-0.808
Ask	.039 .070	.166	.255	N	1	-2.199
Command	.063 .076	.068	.097	N	1	0.031
Disagree	0 0	.006	.024	N/A	N/A	N/A
Disapprove	.006 .015	.008	.022	N	1	-0.576
Permit	.398 .239	.199	.096	N	0.229	2.103
Promise/Offer	.105 .104	.074	.062	N	1	1.211
Request/Suggest	.109 .108	.128	.113	N	1	-0.565
State intent	.020 .033	.022	.033	N	1	-0.035
State/Claim	.188 .258	.155	.124	N	1	0.063
Threaten	.002 .009	.038	.069	N	1	-1.992
Warn/Advise	.031 .069	.093	.153	Ν	1	-2.066
Approve/Praise Ask Command Disagree Disapprove Permit Promise/Offer Request/Suggest State intent State/Claim Threaten Warn/Advise	.016         .038           .039         .070           .063         .076           0         0           .006         .015           .398         .239           .105         .104           .109         .108           .020         .033           .188         .258           .002         .009           .031         .069	.026 .166 .068 .006 .008 .199 .074 .128 .022 .155 .038 .093	.055 .255 .097 .024 .022 .096 .062 .113 .033 .124 .069 .153	N N N/A N N N N N N N	1 1 N/A 1 0.229 1 1 1 1 1 1 1 1	-0. -2. 0.03 N/A -0. 2.10 1.2 -0. -0. 0.00 -1. -2.



Fig. 6. If Illocutionary act types comparison - children versus mothers.

#### 5.2. Phrasing recurrence

We next looked at post-connective phrasing patterns to see if children were using the connectives to introduce the same phrasing as their mothers or whether there were any patterns which may be indicative of more generally accepted idiomatic usage. To explore whether these patterns exist, we compared the most frequently produced VPs appearing after the connective in these Speech–Act sentences for each connective across the dyads.

# 5.2.1. Because VPs

Overall, the children produced 1446 different post-connective VPs in their 1670 *because* Speech–Act sentences and mothers produced 908 different post-connective VPs in their 962 *because* Speech–Act sentences. Table 5 shows the most frequently repeated *because* post-connective VPs for both speakers within each dyad, including the number of times each speaker produced that VP. Where a speaker did not have any repeated forms (and therefore no form(s) was/were more frequent than any others), an N/A is recorded. No *because* post-connective VP was produced more than six times by any speaker, nor accounted for more than 5% of any child's total *because* Speech–Act sentences. Only one dyad (Gina and her mother) shared their most frequently produced form, *look* (e.g. *no*. *(be)cause look*, Gina; 4;00:27). However, for Gina and her mother, this VP only accounted for about 1% of each speaker's *because* Speech–Act sentences, and for Gina this was tied with three other VPs. *Look* was also the most frequent *because* post-connective VP for five other speakers (but not for their interlocutors), although it was never produced more than six times by any speaker. This means that several speakers produced it, but not repeatedly. Therefore, for *because*, there was no evidence that any of the children were consistently copying the post-connective phrasing their mothers used. Additionally, although *look* appears to be a phrase that speakers sometimes use following *because* in their Speech–Act sentences, there is not any strong evidence this or any other post-connective phrasing has any consistent or idiomatic usage in *because* Speech–Act sentences.

#### Table 5

Most frequently repeated because post-connective verb phrase(s) for each speaker

Corpus	Most frequently repeated post-connective VP(s) in because Speech-Act sentences (N	N = number of times produced)
	Child	Mother
Alice	N/A	N/A
Billy	N/A	N/A
Bob	<b>Be omega+ranger</b> (e.g. yeah ! [+ SR] (be)cause I'm omega+ranger; 3;05:09)	N/A
	<b>Be sick</b> (e.g. yes . (be)cause I'm sick; 3;02:08)	
	<b>Be two</b> (e.g. yeah ! (be)cause you're two,aren't you?; 3;05:09)	
	<b>Can find me</b> (e.g. in the grey grass . [+ SR] (be)cause you can't find me; 3;06:00)	
	Like them (e.g. didn't you forgot [*] I want sweeties ? (be)cause I like them; 3;06:16)	
	(N = 2; 2.5%  each)	
Gina	<b>Be a big girl (</b> e.g. I won't burn my fingers . (be)cause I'm a big girl; 3;01:06);	Look (e.g. &wha [//] what about the crocodile ?
	<b>Be nice</b> (e.g. yeah . [+ 1] (be)cause a sun's nice; 3;10:02);	(be)cause look . the crocodile's gonna [: going to] go
	<b>Look</b> (e.g. no (be)cause look I don't even come up to there; 4;01:11);	snap@o; 3;00:14) (N = 4; 1.3%)
	<b>Need it</b> (e.g. where's my phone (be)cause I need it ?; 4;00:26)	
	(N = 4; 1.3% each)	
Helen	<b>Be the teacher</b> (e.g. I am . (be)cause I am the teacher; $3;05:29$ ) (N = 3; $3.5\%$ )	N/A
Ivy	<b>Do</b> (why's this one not stand because that one does ?; 3;08:03)	N/A
	(N = 3; 4.5%)	
Jack	N/A	<b>Look</b> (e.g. shall we push ? because, look . $3;01:29$ ) (N = 2; 10%)
Lucy	Be going to sleep (e.g. we have to be quiet because this baby's going to sleep; 2;11:05);	<b>Be wet</b> (e.g. no . (be)cause it's wet; 3;01:23) (N = 2;
	Look (e.g. no . don't do it on that page (be)cause look; 3;01:18)	7.4%)
	(N = 2; 3%  each)	
Mary	<b>Be very good</b> (e.g. I'll do it for you # (be)cause I 0am [*] very good; 3;01:22);	Look (e.g. watch this . now now . (be)cause, look . look
	<b>Be a big girl</b> (e.g. yeah . (be)cause I'm a big girl; 3;05:06);	look look; 3;01:07) (N = 2; 3.1%)
	<b>Be quite difficult</b> (e.g. you open my box (be)cause it's quite difficult; 3;01:22);	
	<i>Like elephants</i> (e.g. <you could=""> [/] you could be an elephant (be)cause you like elephants; 3;04:08);</you>	
	<b>Do work</b> (e.g. it's not working . (be)cause some don't work; 3;00:03)	
	(N = 2; 1.8% each)	
Olga	Be the mummy (e.g. okay . (be)cause I'm the mummy; 2;09:03);	N/A
	<b>Be her mummy</b> (e.g. yeah . [+ SR] (be)cause I'm her mummy; 2;10:21) (N = 3: 2.2% each)	
Rebecca	N/A	N/A
Sid	<b>Look</b> (e.g. I'll stop you from getting down . (be)cause, look !: 3:02:25)	N/A
biu	(N = 3; 1.8%)	
Steve	<b>Be broken</b> (e.g. no . [+ 1] (be)cause [/] (be)cause it's broken; 3;03:00) (N = 4; 4.9%)	N/A
Thomas	<b>Look</b> (e.g. you have not . (be)cause look; $4;04:03$ ) (N = 6; 1.4%)	<b>Can see</b> (e.g. I like glass because without glass <i can't="" see=""> [&gt;]; 3;01:12);</i>

(continued on next page)

#### Table 5 (continued)

Corpus	rpus Most frequently repeated post-connective VP(s) in <i>because</i> Speech–Act sentences (N = number of times produced)		
	Child	Mother	
		<b>Be upstairs</b> (e.g. you can't play with your shop now because it's upstairs and we're downstairs; 3;00:18); <b>Will break</b> (e.g. don't [/] don't put anymore in, Thomas, because it'll break; 3;01:03); <b>Get dirty</b> (e.g. you shouldn't start puting food in toys, Thomas . because then they get dirty; 3;00:24) (N = 2; 0.9% each)	

# 5.2.2. If VPs

Table 6 shows the most frequently produced post-connective VPs in *if* Speech–Act sentences for each speaker. For *if*, the children produced 91 different post-connective VPs in their 210 *if* Speech–Act sentences and the mothers produced 196 different post-connective VPs in their 290 *if* Speech–Act sentences. In three dyads the same post-connective VP (*want*) was the most frequently produced by both speakers, accounting for between 23.5% and 65.6% of these speakers' *if* Speech–Act sentences. There were also more general patterns across the data. Specifically, the VPs *want, like* and *want to* were all produced by several speakers (*want*: 11 children, 9 mothers; *like*: 8 children, 7 mothers; *want to*: 7 children, 5 mothers), and for 20 of the 28 speakers, one of these was the most frequently produced post-connective VP in their *if* Speech–Act sentences (see Table 6). No other post-connective VP was produced more than 4 times in the overall *if* datasets for either children or mothers (as such, no others accounted for more than 2% of the children's dataset or 1% of the mothers').

#### Table 6

Most frequently repeated if	post-connective	verb phrase for	each speaker.
-----------------------------	-----------------	-----------------	---------------

Corpus	Most frequently repeated post-connective VPs in <i>if</i> Speech–Act sentences				
	Children	Mothers			
Alice	N/A	<b>Want</b> (e.g. I'll tell you if you want; 3;01:25) (N = 4; 26.7%)			
Billy	N/A	<b>Want</b> (e.g. yeah . if you want; 3;01:07) (N = 5; 35.7%)			
Bob	<i>Want</i> (e.g. come on ! you [/] you can be one if you want; 3;06:00) (N = 2; 50%)	N/A			
Gina	<i>Want</i> (e.g. you can talk to me if you want; 4;05:28) (N = 13; 43.3%)	<b>Want</b> (e.g. you can put some on your feet as well if you want; $3;00:04$ ) (N = 16; 25%)			
Helen	<b>Want to</b> (e.g. so let's put everybody in . if they want to they can all be in; $3;05:15$ ) (N = 2; 25%)	<b>Like</b> (e.g. yeah . put it on there if you like; 3;01:28) (N = 5; 27.8%)			
Ivy	<b>Want</b> (e.g. you can swim through if you want; $3;02:03$ ) (N = 4; $30.8\%$ )	N/A			
Jack	<i>Like</i> (e.g. you [/] you [/] you can buy them if you like; 3;01:08) (N = 3: 37.5%)	N/A			
Lucy	Want (e.g. Mum . [+ SR] have this fork if you want; 2;11:15) (N = 7; 46.7%)	N/A			
Mary	Want (e.g. okay . if they want; 2;11:19) (N = 11; 50%)	<i>Like</i> (e.g. you could swap if you like; $3;01:22$ ) (N = 14; 45.2%)			
Olga	<i>Want</i> (e.g. or anyway if you want . tip it in this way on my spoon; 3;00:02) (N = 21; 65.6%)	<b>Want</b> (e.g. you can have that pizza for free if you want; $3;01:28$ ) (N = 5; 25%)			
Rebecca	<b>Want</b> (e.g. 10 [*] put my dinosaur there if you want; 3;01:00) (N = 5; 38.5%)	<b>Want</b> (e.g. there's a blue one as well if you want ?; 3;01:07) (N = 8; 23.5%)			
Sid	Like (e.g. you can break it in . <if like="" you=""> [&gt;] ?; 3;02:25) (N = 2; 15.4%)</if>	N/A			
Steve	<b>Want to</b> (e.g. he [/] he can have [/] have this if he want [*] to; 3:01:15) (N = 3; 42.9%)	N/A			
Thomas	<b>Want</b> (e.g. you're going to be the delivery lady . if you want; 4;10:08) (N = 8; 18.6%)	Like (e.g. but you have it if you like; 3;01:11) (N = 6: 10.7%)			

Overall, these recurring forms (*want, like, want to*) account for a large proportion of the data for both groups. Their usage is more consistent in the children's data accounting for just over half of all *if* Speech–Act sentences; for mothers, they account for almost a quarter (see Table 7).

#### Table 7

Proportions of frequently recurring if post-connective VPs in the data.

	Children	Mothers
Want	.345 (SD = .282)	.128 (SD = .124)
Like	.104 (SD = .119)	.079 (SD = .132)
Want to	.080 (SD = .126)	.025 (SD = .049)
Total	.529	.232

Given the extensive overlap in post-connective verb phrases used in *if*-sentences, we then examined the subjects of these VPs and the illocutionary acts with which they co-occurred to determine any more specific patterns of usage. Of the 194 post-connective VPs that took one of these forms across the two speaker groups, 175 (90.2%) had the subject *you* (mothers': M = 1.00, SD = 0; children: M = .80, SD = 0.229). This means that these recurring phrases in the mothers' speech were either the specific phrases "*if you want*", "*if you like*", or "*if you want to*" and the children also preferred this phrasing, although they did occasionally change the subject (e.g., *okay*. *if they want*; Mary, 2;11:19). With regard to illocutionary acts, all three of these VPs were primarily produced with Permits, Request/Suggests and Promise/Offers (see Fig. 7), suggesting these recurring forms also have a very limited function.



Fig. 7. Comparison of illocutionary acts produced with the recurring if post-connective VPs (want, like, want to).

To summarise, while mothers and children both use *because* to explain Commands, children produce it to explain State/ Claims and Disagreements proportionately more frequently than their mothers. With *if*, there were no significant differences between children and their mothers in the proportional use of any illocutionary act. While there were no consistent patterns in the post-connective phrasing of *because*-sentences, there was evidence of repetitive phrasing in the *if* post-connective VPs of both mothers and children, with these forms accounting for more than half of the children's *if* Speech–Act sentences. As such, it seems that while children's *because*-sentences broadly reflect their mothers' pragmatically, they differ slightly in primary function and have little resemblance in form. By contrast, their *if* Speech–Act sentences are similar to their mothers' in terms of both function and repetitive phrasing.

#### 6. Discussion

In this study, we aimed to provide more insight into functional patterns in the Speech–Act sentences (Sweetser, 1990) that children hear and produce. Specifically, we coded *because* and *if* Speech–Act sentences for the illocutionary act performed in

the main clause, as well as the post-connective phrasing. The analyses revealed some clear, although opposing, patterns for the two connectives, which we will discuss below.

With *because*, both mothers and children produced a large number of Commands (children 27.1%, mothers 36.8%). However, while this was the most frequently produced illocutionary act for the mothers, it was not for the children. Instead, State/Claim was the most frequently occurring illocutionary act in the children's speech (38.2%) and they produced these proportionately more frequently than their mothers (14.5%). Additionally, although it was low frequency overall, the children produced proportionately more Disagrees with *because* than their mothers (4.4% vs. 0.6%). Thus, although there are related patterns in mothers' and children's productions of *because* Speech—Act sentences, children do show some patterns of usage that are different from their mothers.

The specific illocutionary act patterns observed in the children's data seem to indicate that their because Speech-Act sentences serve a different function to that suggested by Kyratzis et al. (1990), who reported that children's causal Speech-Act clauses mainly accompanied "control acts". Although this seems to be at least partially corroborated in the present study by the high number of Commands the children produced, these were not the most frequently produced illocutionary acts in our children's because data. Rather, our data showed that children produced more explanations of Statements/Claims (e.g. yes . because it's cold, Helen, 3; 02:12) than any other illocutionary act with because. Thus, while the children's because Speech-Act sentences are regularly related to their own interests (thus aligning with patterns in children's causal speech, in general, as reported in Diessel, 2004; Hood and Bloom, 1979), their primary function may not be as "coercive" as Kyratzis et al. (1990, p. 210) suggest. Rather, because (along with the arguments/explanations it introduces) may primarily function to increase co-operative discourse. This idea is based on Ford and Mori (1994), who shows that in adult speech "causal connectors are used in the service of negotiating agreement (or managing disagreement) between interlocutors" (pp. 52–53) (see also Orsolini, 1993 for related arguments regarding Italian children's use of *perché* (because)). Additionally, Sweetser (1990) claimed that explaining one's utterance can help the speaker prevent being perceived as rude. With regard to child discourse, specifically, Kyratzis et al. (2010) give further support to the argument in their study of justifications in peer discourse. Although the boys in their study did not produce enough data for statistical comparisons, Kyratzis et al. (2010) argued that girls (aged 3; 7–5; 4) produced more causal connectives with justifications that "validated" (i.e. expanded upon or agreed with a peer's statement/idea) than "opposed" (i.e. rejected a peer's action or idea). Our maternal input data is also consistent with the idea that Speech–Act because has a broadly co-operative function. Although mothers primarily produced Commands, these did not appear to have the sole intent of controlling their children. Rather, the commands were often instructions aimed to help their children and/or prevent a generally negative consequence, while because was then used to introduce an explanation as to why this directive was applicable (e.g. put your cardigan on, babes . (be)cause I think you're getting a little bit of a cough < and cold> [>], Gina's mother, 3; 00:04). Although these kinds of explanations can be produced without a connective (e.g. as two independent sentences, such as put your cardigan on. I think you are getting a little bit of a cough), Kyratzis et al. (2010) argue (following Chafe, 1984) that the use of a connective "focuses attention on the reason and does not allow the main clause to be asserted strongly" (Kyratzis et al., 2010, p. 122) (although cf. Orsolini, 1993). One possibility is that the patterns observed in the input to children teach them that Speech-Act because enables them to draw attention to their explanation – and thus act more co-operatively in the discourse – even when producing illocutionary acts that are self-focused.

If this is the case, given the usefulness of this function, as well as the frequency with which because Speech-Act sentences are heard and produced by young children, it is possible that this pragmatic meaning is the one that is the most salient, and thus prioritised in acquisition (see Slobin, 1985). Although they argued that the primary function of Speech–Act causals in child speech is somewhat different than the function we have described here, Kyratzis et al. (1990) make a similar argument regarding the usefulness of this pragmatic type and its presence in the speech of young children. Evers-Vermeul and Sanders (2011) label this approach a "social-pragmatic complexity approach" (p. 1647), where there is a relationship between the usefulness of a connective's function and the ease with which a child acquires it, a theory that overlaps with Slobin's (1985) argument that children prioritise meaningful language. Ford (1993), drawing on Schiffrin, (1987, as reported in Ford, 1993), suggested that adults may sometimes use because to mean "what I have just said may be clarified through what I am about to say" (p. 135). In a related way, then, for young children, because may mean something like "the reason I just said that is ...". This, then, would give further support to Kyratzis et al.'s (1990) (see also De Ruiter et al., 2021) hypothesis that children's difficulty in understanding because reflects the fact that experimental studies primarily use Content sentences as their stimuli. This may be especially problematic when interpretation of *because* in these studies relies on an understanding of the cause-effect/temporal ordering in Content sentences (e.g. De Ruiter et al., 2018; Emerson, 1979; Emerson and Gekoski, 1980) (see related arguments in Donaldson, 1986). In these studies, understanding an ordering relationship is critical to being able to interpret the sentence correctly. In Speech-Act sentences, however, the sentence meaning is not bound by this same sort of ordering (i.e. both the illocutionary act and the explanation for it occur in the present discourse, Degand and Pander Maat, 2003; Pander Maat and Degand, 2001). As such, when the type of because-sentences that children hear/produce most frequently does not require understanding of an ordering relationship, it is perhaps not surprising that they struggle with experimental stimuli testing this understanding.

With regard to *if*, different patterns emerge. First, for both children and their mothers, the most frequently produced illocutionary act was Permit (children: 39.8%, mothers: 19.9%) and there were no significant differences in the proportional frequency with which any illocutionary act appeared in the children's data in comparison to their mothers'. Thus, we see a clear difference between *because* and *if*: while children's *because* Speech–Act sentences align with their mothers only in

terms of broad functional patterns, both the function and the form of their *if* Speech–Act sentences aligned with those produced by their mothers. Second, although we have argued above that because in because Speech-Act sentences has a cooperative function, the illocutionary acts it accompanied were primarily related to the child's own interest (State/Claims and Commands). By contrast, if illocutionary acts were often focused on their addressee. Permission accounted for approximately 40% of children's if productions, primarily occurring with post-connective phrasing in the form of "if you like/want (to)", which appear to relate to the listener's, rather than speaker's, interest. Thus, in alignment with Sweetser (1990), for children as well as adults, if Speech-Act sentences seem to be strongly associated with politeness. Finally, these recurring VPs constitute another clear difference between because and if. While there were no consistent patterns in the post-connective VPs produced with because, the forms if you want (to) and if you like appeared repeatedly in the speech of many speakers from both groups and were associated with specific illocutionary acts. Their consistency of usage suggests that these are "idiomatic" (e.g. Sweetser, 1990; Van der Auwera, 1986) forms, devoid of true conditional meaning and not specific to individual discourse. In contrast, for the other if post-connective VPs, there were no consistently recurring patterns; rather they were generally tailored to individual discourse in some way, such as by expressing given information (see Sweetser, 1990; Van der Auwera, 1986) (e.g. if you're looking for the trailer. I know where the car is; Thomas, 4; 04:03). Given the frequency with which these idiomatic forms were heard performing a specific function in input (accounting for 23% of the if Speech–Act sentences in input, mainly with Permits, Promises/Offers and Request/Suggests), it is possible to see how these could be acquired as entire phrases with a specific functional, rather than conditional, meaning. These may be idiomatic phrases which function solely to signal politeness, similar to Sweetser's (1990) arguments regarding "if I may say so"-clauses. More specifically, these appear to be set phrases which let the listener know that their preference is being prioritised in the present discourse.

This aligns with Tomasello's (2001) idea of "holophrases", which are either single words or set phrases that children abstract from their input and use to relate a particular communicative meaning. From this perspective, the inclusion of idiomatic if-sentences as Speech–Act sentences may overinflate the frequency with which children meaningfully hear and produce this pragmatic type (e.g. see Kirjavainen et al., 2009, p. 1097, for a related argument that excuse me is a "frozen phrase" which cannot be used to gauge children's understanding of the verb + me construction). That is, although these still constitute part of if Speech–Act sentence input and production, if they are produced/interpreted as chunks used for a specific function (i.e., rather than as wholly analysed clauses), children may not (fully) associate the Speech–Act function with these idiomatic utterances. If this is the case, although children produce if Speech–Act and Content sentences in almost equal proportions (De Ruiter et al., 2021), their production may not equate to the same level of comprehension of these two pragmatic types. Interestingly, in comparison to the figures reported in De Ruiter et al. (2021) showing children produce a higher proportion of if Speech–Act sentences than their mothers (but with a high degree of individual variation), when these idiomatic sentences are removed from the data, the proportions of children's if pragmatic types (i.e. Content, Epistemic, Speech–Act) match their mothers much more consistently (e.g. both children and their mothers produce .23–.24 Speech–Act and .72–.75 Content), and the individual variation (as indicated by the standard deviation) for the children's Speech-Act and Content types is reduced (Speech-Act: .258 vs .138; Content: .249 vs. .142). This suggests that the frequency with which individual children produced these idiomatic forms largely contributed to the high degree of variation in children's if Speech–Act sentences as reported in De Ruiter et al. (2021); when these forms are removed, children's pragmatic proportions favour the Content relationship with *if* far more consistently, showing a stronger relationship to input.

If this is the case, it means that the pragmatic type children are most commonly tested on in experiments with if is the kind they are most likely to use meaningfully/process entirely. Unlike with because, then, where we have suggested above that the salience and frequency of the Speech–Act type may contribute to children's difficulty with Content stimuli, pragmatic patterns seem less helpful in explaining the difficulty children have with *if* in comprehension studies such as De Ruiter et al. (2018). However, if is more complex semantically than because. While because contains the semantic aspects of causality and, at least for some pragmatic types (see Degand and Pander Maat, 2003; Pander Maat and Degand, 2001), ordering (e.g. Emerson and Gekoski, 1980), if can require understanding of additional concepts like hypotheticality, contingency and inference (Bowerman, 1986), and sentences expressing simple, hypothetical and counterfactual conditionality are all found (in varying levels of frequency) in the speech of, and to, young children (De Ruiter et al., 2021). Furthermore, unlike because sentences, which primarily occur in main-subordinate order, if-sentences are more varied, occurring in both mainsubordinate and subordinate-main (Diessel, 2004, 2005), and the proportions in which they occur in either order changes with pragmatic type, such that Content is more likely to occur in subordinate-main, while Speech-Act is more likely to occur in main-subordinate (De Ruiter et al., 2021). Thus, while the pragmatic variation and Speech-Act patterns discussed here do not seem to fully explain children's difficulty with demonstrating understanding of if, in and of themselves, they likely provide an additional level of complexity to a connective which is already very complicated to acquire (see De Ruiter et al., 2021 for related arguments). All this noise in the form-meaning and form-function mapping of if may simply cause children to have more difficulty with it than other connectives (see e.g. Slobin, 1982 for discussion on how acquisition is complicated by noise in form-function mapping).

To summarise, our study has shown clear patterns in children's usage of *because*- and *if* Speech—Act sentences, supporting the idea that these sentences express particular functional meaning for children. For *because*, we argue that this is a way of achieving goals/promoting their ideas in a co-operative manner; for *if, it* is a way of expressing politeness. Furthermore, we also offer evidence that a large portion of children's *if* Speech—Act sentences contain idiomatic phrasing and, as such, may not express a meaningful conditional relationship. Therefore, of those *if*-utterances that do meaningfully express a conditional relationship, children's speech and input may favour Content relationships with *if*-sentences more than the patterns reported

in De Ruiter et al. (2021) initially suggest. However, while the patterns here tell us more about children's usage of this pragmatic type, without data on children's comprehension of these kinds of sentences, particularly in comparison to the more commonly assessed Content relationship (e.g. Amidon, 1976; De Ruiter et al., 2018; Emerson, 1979; 1980; Emerson and Gekoski, 1980; French, 1988; Peterson and McCabe, 1985), we cannot be certain whether the patterns presented here actually relate to children's comprehension, either of the Speech–Act function specifically, or of the connectives overall. As such, at present, we are exploring this via experiments designed to determine how comprehension of these connectives changes with pragmatic type. In doing this, we hope to provide more information, not only about children's understanding of what these connectives mean and how they typically function, but also about children's sensitivity to the pragmatic relationships these connectives express.

# **Declaration of competing interest**

None.

# Acknowledgements

Anna Theakston and Elena Lieven were supported by the ESRC International Centre for Language and Communicative Development (LuCiD), Economic and Social Research Council [ES/S007113/1 and ES/L008955/1], and Heather Lemen by a PhD studentship from the University of Manchester.

# Appendix 1. Pragmatic coding

# 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, p. 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, p. 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, p. 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)

\* \*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.

\*Sweetser, E. (1990). From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

# 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 der Auwera, 1986, p. 198).

e.g. If you get me some coffee, I'll give you a cookie (Sweetser, 1990, p.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 der Auwera, 1986, p. 198). e.g. If I may say so, that's a crazy idea (Sweetser, 1990, p.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)

\* Sweetser, E. (1990). From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

\*\*\* 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.

# Appendix 2. Illocutionary act coding scheme

**1. Agree** – the main clause offers agreement with an idea or statement that the other speaker has said. These do not offer any judgement on behaviour (such as "this is good/bad", etc), but simply state the speaker's agreement with an idea presented.

e.g.1 CHILD: I should get my pyjamas on.

MOTHER: Yeah. Because it is almost bedtime.

e.g. 2 That's right. Because we saw them playing together, didn't we?

e.g. 3 Yes, if that really is true

(Note: this may also include sentences where the main clause indicates a negative agreement, like: CHILD: I didn't like that book, did I?. MOTHER: No (you didn't) because you were overtired and not in the mood to give it a chance.)

Because example from dataset: yeah . because that's silly . (Helen; 3; 02:12)

If example from dataset: yes you do have rather a problem if your shoes don't fit . (Thomas' mother; 3; 00:10)

**2. Approve/Praise** — the main clause praises or approves of a state, event, behaviour, etc. described or performed. This may be of something the listener or someone else has said or done or may be a general value judgement of something (such as "I like this" or "this is nice").

e.g. 1 This is great *because* look at how much work they've put into this.

e.g. 2 Good boy! *Because* I didn't even have to ask you to pick up your toys.

e.g. 3 This is lovely, *if* you built that all by yourself.

Because example from dataset: *very nice*. (*be*)*cause Cinderella's got some birds, hasn't she*? (Mary; 3;02:04) If example from dataset: *that's good if you have little animals*. (Ivy; 3; 02:24)

**3. Ask** – the main clause asks a question.

e.g. 1 So, what should we eat for dessert, because you ate all your dinner?

e.g. 2 Do you think he will, *if* you are such an expert?

(Note: these are when the speaker actually asks a question of the listener. This does not include directives phrased as questions, such as "Can you pick your toys now, *because* it's time for bed?").

Because example from dataset: how ? (be)cause they don't have any eyes . (Mary; 3;01:14)

If example from dataset: *if they're not the arms what are they* ? (Steve's mother; 3;00:13)

**4. Command** – the main clause demands/orders a certain behaviour of the listener. Usually this relates to the present or immediate future. The listener, in these cases, is expected to comply. This may also include indirect commands, such as "can you get me that, *because* I need it".

e.g. 1 Don't do that *because* you'll get hurt.

e.g. 2 You need to hand that to me right now because I am tired of asking.

e.g. 3 Put away your blocks, *if* you are really finished with them.

This includes the forbidding of activities, such as:

e.g. 1 CHILD: Can I climb up there?

- MOTHER: No because you could hurt yourself.
- e.g. 2 You can't do that all by yourself because you are too young.

Because example from dataset: *cut all the stars ! (be)cause [/] (be)cause they are alien's stars !* (Sid; 3;04:20) If example from dataset: *and then wash it if it's tatty.* (Olga: 2:11:11)

**5. Disagree** – the main clause offers disagreement or refusal of an idea/fact stated by the other speaker. These do not relate to behaviour control; they solely express the speaker's disagreement with the truth of an idea.

e.g. 1 CHILD: This music is too loud.

MOTHER: No (it isn't) because I want to listen to it in the kitchen.

e.g. 2 MOTHER: We won't be late.

# CHILD: Yes we will, if I really have to finish all of my lunch first.

Because example from dataset: no. that's a little one (be)cause that's Daddy's . (Lucy; 2;11:05)

If example from dataset: *it isn't if the wheels won't turn* . (Jack's mother; 3;01:29)

**6. Disapprove** – the main clause primarily functions to express the speaker's disapproval at an event, state, behaviour, etc. This may be of something the listener or someone else has said or done or may be a general value judgement of something (such as "I don't like this" or "this is weird"). They are not related to behaviour control, they simply express a negative judgement of an event/situation/person, etc.

e.g. 1 This is not good because this is not what I asked for.

e.g. 2 It's ridiculous when they don't answer *because* they said they would be home.

e.g. 3 That's not okay, if you have hit your sister.

Because example from dataset: *she wasn't very nice, was she because she wouldn't share or anything*. (Billy's mother; 3; 01:09)

If example from dataset: that's [/] that's bad if you don't brush your teeth. (Gina; 4; 01:07)

**7. Permit** — main clause primarily functions to express permission for the listener to do something. This may be via either direct permission ("yes", "you can", etc) or by the lack of forbidding ("I don't mind", "if you want", etc). Generally, this is permission that has been specifically requested (see example 2) or related to an activity where the listener was not assumed to have permission before (see example 1). Typically, these are more related to a behaviour (e.g. the act of having a cookie) rather than a thing (e.g. being offered a cookie — which would be promise/offer).

e.g. 1 You can have them *because* I know you would like them.

e.g. 2 CHILD: Can I have these?

# MOTHER: I don't mind because they aren't mine.

e.g. 3 You can play with that toy now, *if* you like.

Because example from dataset: *and the tigers can go in there now because they want to go in there*. (Helen; 3; 02:26) If example from dataset: *Jess can play if she wants*. (Ivy; 3; 03:16)

**8. Promise/Offer** – main clause presents a promise or offer to the listener. This may be a promise or offer to do something in the future or the offer of something in the immediate discourse (such as "here is a balloon because I know you like them"). The primary function is to offer something (either a commitment/promise or an actual item).

e.g. 1 I will get them for you *because* you can't reach

e.g. 2 I will bring you home a present because they have some nice things there that you will like.

e.g. You can have **this** puzzle, *if* you want.

Because example from dataset: I'll help (be)cause I'm the biggest engines [\*]. (Steve; 3; 07:02)

If example from dataset: and here's some [//] another ticket for you. if you want to come back again (Thomas; 4; 02:02)

**9. Request/Suggest** - main clause requests or suggests behaviour in the present or future, but without the authority or urgency of a command. In these utterances, the listener could more likely refuse to comply. This also includes the requesting of permission to do something (such as "can I just see that for a minute, *because* I don't understand what you are saying") or requesting assistance.

e.g. 1 We should dance because the music is on.

e.g. 2 Could we go to the store because I want to see if the new book is in?

e.g. 3 We could call, *if* you think that's a good idea.

Because example from dataset: *let's play your game # because we 0 are [\*] going [?] to play your game.* (Helen; 3;02:12) If example from dataset: *if you want a thin one why don't you use this one ?* (Mary's mother; 3; 00:10)

**10. State intent** – main clause expresses the speaker's plans/commitment to perform an action, including behaviour that is ongoing or intended in the immediate future. Also includes an expression of negative commitment, such as in example 2.

e.g. 1 I am going to make dinner now *because* I think you are hungry.

e.g. 2 MOTHER: Are you going to clean up those toys?

# CHILD: NO, because I don't want to.

e.g. 3 I will do it later, *if* we are going out now.

Because example from dataset: I will . (be)cause if I don't it will burn me. (Lucy; 3;04:29)

If example from dataset: *I'm just gonna [: going to] take [/] take these if he wants them.* (Ivy; 3;02:03)

**11. State/Claim** — main clause is a fact or declaration about a state, event, behaviour, etc. This may include a speaker's judgement/opinion, so long as it does not primarily function to provide a positive or negative evaluation (as these would be approve/praise or disapprove, respectively). They can include stream of consciousness utterances about the current situation (e.g. I need to jump higher because I want to reach it) or a response to a question (e.g. MOTHER: which one do you want to wear? CHILD: I want to wear the blue one because I like the blue one).

e.g. 1 The doll's favourite colour is going to be yellow, *because* that is the colour of the sun.

e.g. 2 It doesn't matter *because* we can just get some more later.

e.g. 3 It is dark out, if you didn't notice.

Because example from dataset: *but I got new eyes (be)cause look at these*. (Thomas; 4;07:06) If example from dataset: *if you have a look at the bottom*. *there's some chalk*. (Billy's mother; 3;01:09) **12. Threaten** – main clause makes a threat.

e.g. 1 I won't give you any more help if you keep doing this because I am tired of you not listening.

e.g. 2 You won't be able to go to the party if you do that because only good boys can go.

e.g. 3 I will send you straight to your room, *if* you think you can talk to me like that.

Because example from dataset: well I'm not gonna [: going to] play with you until you do because (.) that is a buggy for dollies . (Gina's mother; 3;00:29)

If example from dataset: if you're not coming then I won't make you anything. (Gina; 4;00:26)

**13.** Warn/advise – main clause advises or warns the listener of any danger or negative consequences in the present or future, or provides advice, so as to avoid negative consequences for the listener in the future.

e.g. 1 Be careful *because* they could break.

e.g. 2 You should always be gentle with animals because they could hurt you if you hurt them.

e.g. 3 I hope you watch where you're going, if you are running everywhere.

Because example from dataset: *you shouldn't go underneath bikes (be)cause it's dangerous* (Mary; 3; 03:11) If example from dataset: *if you don't mind [?] there's a spider on you*. (Thomas; 4;02:08)

#### References

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

Austin, J.L., 1962. How to Do Things with Words: the William James Lectures Delivered at Harvard University in 1955. Clarendon Press.

Bloom, L., Lahey, M., Hood, L., Lifter, K., Fiess, K., 1980. Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. J. Child Lang, 7 (2), 235–261. https://doi.org/10.1017/S030500090002610.

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

Brown, P., Levinson, S.C., 1987. Politeness: Some Universals in Language Usage. Cambridge University Press.

Chafe, W., 1984. How people use adverbial clauses. In: Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society, pp. 437–449. https:// doi.org/10.3765/bls.v10i0.1936.

- De Ruiter, L.E., Lemen, H.C., Lieven, E.V., Brandt, S., Theakston, A.L., 2021. Structural and interactional aspects of adverbial sentences in English mother-child interactions: an analysis of two dense corpora. J. Child Lang. 1–35. https://doi.org/10.1017/S0305000920000641.
- De Ruiter, L.E., Lieven, E.V., Brandt, S., Theakston, A.L., 2020. Interactions between givenness and clause order in children's processing of complex sentences. Cognition 198. https://doi.org/10.1016/j.cognition.2019.104130.
- De Ruiter, LE., Theakston, A.L., Brandt, S., Lieven, E.V., 2018. Iconicity affects children's comprehension of complex sentences: the role of semantics, clause order, input and individual differences. Cognition 171, 202–224. https://doi.org/10.1016/j.cognition.2017.10.015.
- Degand, L, Pander Maat, H., 2003. A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. In: Verhagen, A., van de Weijer, J. (Eds.), Usage Based Approaches to Dutch: Lexicon, Grammar, Discourse, vol. 1. LOT, pp. 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., Hetterle, K., 2011. Causal clauses: a cross-linguistic investigation of their structure, meaning, and use. In: Siemund, P. (Ed.), Linguistic Universals and Language Variation. Walter de Gruyter, pp. 23–54.

Donaldson, M., 1986. Children's Explanations: A Psycholinguistic Study. Cambridge University Press.

Emerson, H.F., 1980. Children's judgements of correct and reversed sentences with "if". J. Child Lang. 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. J. Exp. Child Psychol. 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. J. Pragmat. 43 (6), 1645–1662. https://doi.org/10.1016/j. pragma.2010.11.015.

Ford, C.E., 1993. Grammar in Interaction: Adverbial Clauses in American English Conversations. Cambridge University Press.

- Ford, C.E., Mori, J., 1994. Causal markers in Japanese and English conversations: a cross-linguistic study of interactional grammar. Pragmatics 4 (1), 31–61. https://doi.org/10.1075/prag.4.1.03for.
- French, L.A., 1988. The development of children's understanding of "because" and "so. J. Exp. Child Psychol. 45, 262-279. https://doi.org/10.1016/0022-0965(88)90032-X.

Haegeman, L., 1984. Pragmatic conditionals in English. Folia Ling. 18, 485–502. https://doi.org/10.1515/flin.1984.18.3-4.485.

Emerson, H.F., 1979. Children's comprehension of "because" in reversible and non-reversible sentences. J. Child Lang. 6 (May 2015), 279–300. https://doi.org/10.1017/S0305000900002300.

Homzie, M.J., Gravitt, C.B., 1977. Children's reproductions: effects of event order and implied vs. directly stated causation. J. Child Lang. 4, 237–246. https://doi.org/10.1017/S030500090000163X.

Hood, L., Bloom, L., 1979. What, when, and how about why: a longitudinal study of early expressions of causality. Monogr. Soc. Res. Child Dev. 44 (6), 1–47. https://doi.org/10.2307/1165989.

- Hothorn, T., Winell, H., Hornik, K., van de Wiel, M.A., Zeileis, A., 2019. Coin: Conditional Inference Procedures in a Permutation Test Framework (1.3.1). https://cran.r-project.org/package=coin.
- Johnson, H.L., Chapman, R.S., 1980. Children's judgment and recall of causal connectives: a developmental study of "because," "so," and "and. J. Psycholinguist. Res. 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 Lang. 20, 291–304. https://doi.org/10.1177/ 014272370002006004.
- Kirjavainen, M., Theakston, A., Lieven, E., 2009. Can input explain children's me-for-I errors? J. Child Lang. 36 (5), 1091-1114. https://doi.org/10.1017/s0305000909009350.
- Kuhn, D., Phelps, H., 1976. The development of children's comprehension of causal direction. Child Dev. 47 (1), 248–251. https://doi.org/10.2307/1128307. 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.
- Kyratzis, A., Ross, T.S., Köymen, S.B., 2010. Validating justifications in preschool girls' and boys' friendship group talk: implications for linguistic and sociocognitive development. J. Child Lang. 37 (1), 115–144. https://doi.org/10.1017/S0305000908009069.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33 (1), 159–174. https://www.jstor.org/stable/ 2529310.
- Lieven, E., Salomo, D., Tomasello, M., 2009. Two-year-old children's production of multiword utterances: a usage-based analysis. Cognit. Ling. 20 (3), 481–507. https://doi.org/10.1515/COGL.2009.022.
- MacWhinney, B., 2000. The CHILDES Project: Tools for Analyzing Talk, third ed. Lawrence Erlbaum Associates, Inc. Publishers, Mahwah, NJ.
- McCabe, A.E., Evely, S., Abramovitch, R., Corter, C.M., Pepler, D.J., 1983. Conditional statements in young children's spontaneous speech. J. Child Lang. 10 (1), 253–258. https://doi.org/10.1017/S0305000900005298.
- McCabe, A., Peterson, C., 1985. A naturalistic study of the production of causal connectives by children. J. Child Lang. 12 (1), 145–159. https://doi.org/10.1017/ S030500090006280.
- Orsolini, M., 1993. Because in children's discourse. Appl. Psycholinguist. 14 (1), 89–120. https://doi.org/10.1017/S014271640001016X.
- Pander Maat, H., Degand, L., 2001. Scaling causal relations and connectives in terms of speaker involvement. Cognit. Ling. 12 (3), 211–245. https://doi.org/10. 1515/cogl.2002.002.
- Peterson, C., McCabe, A., 1985. Understanding "because": how important is the task? J. Psycholinguist. Res. 14 (2), 199-218. https://doi.org/10.1007/ BF01067629.
- Pratt, J.W., 1959. Remarks on zeros and ties in the Wilcoxon signed rank procedures. J. Am. Stat. Assoc. 54 (287), 655–667. https://www.tandfonline.com/ doi/abs/10.1080/01621459.1959.10501526.
- R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.
- Reagan, R.T., 1987. The syntax of English idioms: can the dog be put on? J. Psycholinguist. Res. 16 (5), 417-441. https://doi.org/10.1007/BF01073271.

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

Reilly, J.S., 1986. The acquisition of temporals and conditionals. In: Traugott, E.C., Meulen, A.T., Reilly, J.S., Ferguson, C.A. (Eds.), On Conditionals. Cambridge University Press, pp. 309–332.

Rowland, C.F., Theakston, A.L., 2009. The acquisition of auxiliary syntax: a longitudinal elicitation study. Part 2: the modals and auxiliary DO. J. Speech Lang. Hear. Res. 52 (6), 1471–1492. https://doi.org/10.1044/1092-4388(2009/08-0037a.

- Schiffrin, D., 1987. Discourse Markers. Cambridge University Press.
- Searle, J.R., 1969. Speech Acts: an Essay in the Philosophy of Language. Cambridge University Press. https://doi.org/10.1017/CB09781139173438.
- Sekali, M., 2012. The emergence of complex sentences in a French child's language from 0;10 to 4;01: causal adverbial clauses and the concertina effect. J. Fr. Lang. Stud. 22 (1), 115–141. https://doi.org/10.1017/S0959269511000615.
- Slobin, D., 1982. Universal and particular in the acquisition of language. In: Wanner, E., Gleitman, L.R. (Eds.), Language Acquisition: the State of the Art. Cambridge University Press, pp. 128–170.
- Slobin, D., 1985. Crosslinguistic evidence for the language-making capacity. In: Slobin, D. (Ed.), The Crosslinguistic Study of Language Acquisition, Volume 2: Theoretical Issues. Lawrence Erlbaum Associates, Inc. Publishers, pp. 1157–1256.
- Snow, C.E., Pan, B.A., Imbens-Bailey, A., Herman, J., 1996. Learning how to say what one means: a longitudinal study of children's speech act use. Soc. Dev. 5 (1), 56–84. https://doi.org/10.1111/j.1467-9507.1996.tb00072.x.
- Sweetser, E., 1990. Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure. Cambridge University Press.
- Theakston, A.L., Rowland, C.F., 2009. The acquisition of auxiliary syntax: a longitudinal elicitation study. Part 1: auxiliary BE. J. Speech Lang. Hear. Res. 52 (December), 1449-1470. https://doi.org/10.1044/1092-4388(2009/08-0037.
- Tomasello, M., 2001. First steps toward a usage-based theory of language acquisition. Cognit. Ling. 11 (1–2), 61–82. https://doi.org/10.1515/cogl.2001.012.
  Van der Auwera, J., 1986. Conditionals and speech acts. In: Traugott, E.C., Meulen, A.T., Reilly, J.S., Ferguson, C.A. (Eds.), On Conditionals. Cambridge University Press, pp. 197–214.
- Van Dijk, T.A., 1979. Pragmatic connectives. J. Pragmat. 3 (5), 447-456. https://doi.org/10.1016/0378-2166(79)90019-5.
- Warchał, K., 2010. Moulding interpersonal relations through conditional clauses: consensus-building strategies in written academic discourse. J. Engl. Acad. Purp. 9 (2), 140–150. https://doi.org/10.1016/j.jeap.2010.02.002.
- Zufferey, S., Mak, W.M., Sanders, T., 2015. A cross-linguistic perspective on the acquisition of causal connectives and relations. Int. Rev. Pragmat. 7 (1), 22–39. https://doi.org/10.1163/18773109-00701002.

Heather Lemen completed her PhD at the ESRC International Centre for Language and Communicative Development at the University of Manchester. Her primary research area is children's pragmatic language development.

**Elena Lieven** is an Emeritus Professor at the University of Manchester and was Director of the ESRC International Centre for Language and Communicative Development. Her primary research areas include the relationship between children's language development and their linguistic input, as well as the differences in children's language environments.

Anna Theakston is a Professor at the University of Manchester and Co-Director of the ESRC International Centre for Language and Communicative Development. Her research focuses on various aspects of children's communicative development and its relation with their linguistic environment.