Children's comprehension of contrastive connectives*

Jennifer SPENADER

Artificial Intelligence Department, University of Groningen

ABSTRACT

Production studies show connective acquisition by age 3;0, but comprehension studies show errors until 9;0 or older. To further investigate this gap, two comprehension tasks were carried out with 78 Dutch children between the ages of 7;0 and 10;1, testing contrastive *maar* 'but' and causal *want* 'because' connectives for comparison. An existing context choice task and a task that tested children's ability to interpret pronouns dependent on the connective were used. Children did well on the context choice task for *want* 'because', but performed far below chance with *maar* 'but'. In the pronoun interpretation task the youngest children performed at or near chance with both connectives, but show gradual improvement with age. Task complexity may partially explain the results but in general, they are consistent with previous experiments showing contrastive connective comprehension is acquired very late compared to their correct production. Further, the pronoun interpretation task may give a more accurate picture of understanding.

Contrastive Causal Connectives

Production studies have shown that by the age of three most children are able to express relations of additivity, causality, or adversity correctly – important milestones in communicative development. Understanding these complex relations is equally important, forming the basis for learning new information about the world and correctly interpreting other linguistic phenomena, such as pronouns.

Contrastive connectives like *but,* however, show a surprising asymmetry in existing acquisition research. While investigations of natural production show early acquisition, controlled comprehension studies show children struggle until past age 9;0 with contrastive connectives. The differences between production and comprehension results are puzzling, and until we have an account of the extent and causes, we cannot understand how children eventually learn to participate fully in the communication of complex thoughts.

How robust is the apparent comprehension–production asymmetry for contrastive connectives? It's possible that existing production studies may be too generous in their estimation of children's abilities. Another possibility is that existing comprehension tasks may have been confusing, or too complex by requiring children to reason about connective meaning. Because most comprehension experiments have confined their investigation to (sets of) only contrastive connectives, it is difficult to distinguish task complexity from connective complexity.

The current study uses two tasks to directly investigate the effect of task and connective, comparing the Dutch contrastive connective *maar* 'but' to the positive causal connective *want* 'because'. The first task, the context choice task, asks children to choose appropriate contexts to precede a sentence with one of the target connectives. The task is a simplified version of an earlier experiment with French, so new Dutch data can be directly compared with existing results. The second task, the pronoun interpretation task, is new. It uses stimuli where the interpretation of a pronoun systematically varies depending on the connective used. Its design has advantages over previous tests in that it does not ask children to judge infelicitous sentences and it avoids asking children to reason about connective meaning directly. By using two tasks with two different connectives with the same children, we can see how task effects differ from connective effects and determine if contrastive connectives are in fact much more difficult than other connective types.

Seventy-eight Dutch children between the ages of seven and ten were tested with both tasks. For the context selection, all children performed below chance with *maar* 'but'. However, children performed near ceiling with *want* 'because', suggesting that previous poor results with the task are

in fact at least partly due to difficulties with the contrastive connective. For the pronoun interpretation task children also performed much better on *want* 'because' items than *maar* 'but' items, with accuracy increasing with age. However, with this task children did do better than chance with *maar*, though their performance is still far from adult-like, even with *want*. Together, both tasks are consistent with previous findings that comprehension of contrastive connectives is acquired very late.

The following sections give an overview of the theoretical view of contrastive connectives and the existing experimental work before explaining the motivation behind the tasks used.

Theories of contrastive connectives. One of the main uses of a contrastive connective like English *but* is to imply something in the first conjunct that is denied in the second conjunct. But even within this usage *but* can have multiple communicative functions:[1]

(1) The pen is new, but it writes poorly.
(2) John's a wug but he can play the ukulele.
(3) Mike offered Ed his bike but he preferred to walk.

In (1), the connective *but* marks the fact that an expectation inferred from the first conjunct ('that the pen is new'), based on world knowledge about pens ('new pens write well'), is contradicted in the second conjunct. Example (2) shows that even without any expectations about John or wugs, knowing what *but* means leads to the inference that wugs usually can't play the ukulele. In (3), we also see that knowing that the second event is unexpected is essential for correctly interpreting the pronoun *he* in the second clause as referring to Ed.

Frege (1948) proposed an early approach to *but,* arguing that it is logically the same as conjunctive *and*, but with some extra contrastive content. Grice (1975) makes a similar conclusion but within a pragmatic account, arguing that *but* is similar to *and*, but contributes an additional contrastive conventional implicature.[2] Early work within the argumentative theory of Anscombre and Ducrot (1977) argues that the second conjunct of a *but*-connected sentence suggests a conclusion (or states one outright) that

---

[1] There are several additional usages of *but* and its Dutch equivalent *maar* that do not have a causal element. There is a semantic opposition, or additive negative usage, e.g., *John is tall but Bill is short* (Lakoff, 1971). There is also a corrective usage that denies and revises a previous utterance, e.g., *He is not happy, but sad*. Finally, in conversation but and maar can also mark turn-taking. These usages are not discussed further in this paper. For an overview see Spooren (1989) or Foolen (1993).

[2] The category conventional implicature was always problematic as it was never sufficiently defined in Grice (1975). More recent work by Potts (2005) extends and clarifies this category, but in this newer work *but* is no longer considered a conventional implicature trigger.

is counter-oriented to a conclusion that is likely to be made from the first conjunct. More recent logical approaches describe contrastive connectives as appropriate when the first conjunct has raised a defeasible implication that is denied by the second conjunct (see Winter & Rimon, 1994; Spenader & Maier, 2009).

The major feature that distinguishes *but* from simple additive connectives like *and* is that the information in the second conjunct is perceived as the main message. The content of the second conjunct is what is either attributed to the speaker, or added to the common ground. In Spooren's words, it PERCOLATES to the speaker's perspective. Spooren (1989) also found empirical evidence for this in a series of experiments in Dutch with *maar* 'but'. This makes *but* asymmetrical. This contrastive content in the first conjuct must be either an implication of the first conjunct (e.g., (4)) or it has to be attributed to another individual (5). It cannot be simply asserted because this would attribute the speaker with conflicting beliefs, e.g., (6).

(4)   Bill owns five ukuleles but he can't play.
(5)   John said that Bill can play the ukulele but he can't.
(6)   ??Bill can play the ukulele but he can't.

If the first conjunct contributes an implication it can be a known one or it can be new. Unknown implications need to be accommodated by the hearer, such as in (2). Often, further distinctions between contrastive relations are made. One common distinction has to do with whether or not the second conjunct contradicts the implication of the first conjunct directly, or if the second conjunct contributes an implication and the contradiction is between the two implications. Logical approaches simply call the first type 'direct contrast' and the second type 'indirect contrast' (see Winter & Rimon, 1994; Spenader & Maier, 2009). Spooren (1989), on the other hand, calls direct cases 'Denial of expectation' and indirect cases 'Concessive Opposition'. Whether this distinction is categorical or one of degree is a matter of discussion. One major conclusion of Spooren's work is that these two cases are better treated as two points on a continuum. Another conclusion is that indirect contrast is more complex than direct contrast.

How to compare the complexity of connectives and their uses is not clear, but one answer is given by the cumulative cognitive complexity approach developed by Sanders, Spooren, and Noordman (1992). Their approach deconstructs coherence relations into features of additivity vs. causality, and negative polarity vs. positive polarity. Negative polarity is explained as the presence of some type of contrast. Denial of expectation uses of *but* signal coherence relations that are both causal and negative. The semantic

features are also ordered according to their cognitive complexity. Additive relations are considered cognitively less complex than causal relations, and positive relations are considered less complex than negative relations. Thus, negative causal relations like those marked with *but* will be cognitively the most complex type of relation, while positive additive relations such as those marked by conjunctive *and* are the least complex, with a positive causal marker like *because* somewhere in between. Note, however, that this approach does not deal with more fine-grained differences, and both direct and indirect contrast are in the same category of negative causal relations.

Much theoretical work on, e.g., *but* has been done on English, while the experimental work presented here was done in Dutch with *maar*. However, the functions used in the current study are also found in both languages, so the results should be comparable. Spooren (1989) is an extensive overview of the different functions and analyses of *but*, which he finds to be almost uniformly the same as Dutch *maar*. The work of Sanders *et al.* (1992) was also based on work on Dutch connectives, including *maar*, but they consider their approach and conclusions to be valid for English *but* as well.

The next sections summarize the existing empirical research, which together suggest children master contrastive connective production long before comprehension.

Spontaneous production studies. Studies on the naturalistic acquisition of connectives have generally looked at multiple connectives. Braunwald's (1985) diary study of an English-speaking child found that most connectives were correctly produced by 3;0, though *but* was one of the last acquired. Bloom, Lahey, Hood, Lifter, and Fiess (1980) examined connective production in four English-speaking children between ages 2;0 and 3;2 in the Childes corpus (MacWhinney, 2000). They defined acquisition as five correct uses of a connective in at least two consecutive recordings. Causal connectives like *because* and *so* appeared at around 2;0, and the causal contrastive use of *but* at around 2;8. All connectives studied were produced by 3;0. Their results suggested a one-dimensional prototypical acquisition order: ADDITIVE < TEMPORAL < CAUSAL < ADVERSATIVE, and formed part of the basis for the multidimensional cumulative cognitive complexity account developed in Spooren and Sanders (2008) and Sanders *et al.* (1992).

Evers-Vermeul and Sanders (2009) directly tested prediction of the cumulative cognitive complexity account (Spooren & Sanders, 2008) using Dutch Childes data from twelve children between the ages of 1;5 and 5;6. Acquisition was liberally defined as the first correct usage of the

connective that was not a repetition of an adult utterance. Additive connectives were acquired first, followed by negative additives and positive causals, with children varying as to which were acquired first. Finally, the negative causal usage of *maar* 'but' was acquired last, consistent with the theoretical predictions.

One problem with these naturalistic results is that, because children only produce connectives infrequently in the transcript data, acquisition criteria are necessarily defined rather weakly. This means that results only identify the onset of connective production. Identifying connective mastery is almost impossible outside of controlled experiments, because they can often be felicitously omitted (see, e.g., Behrens, 2006).

We can also learn more by looking at production errors. French (1988) studied the spontaneous production of 96 children between the ages of 3;9 and 9;6. French exhaustively categorized all uses of *but,* identifying even pragmatic usages such as topic marking and turn-taking functions. She concluded that children rarely made errors, but when they did, in half of the cases simple conjunction would have been more appropriate. Peterson (1986) also studied the spontaneous use of *but* in English children between the ages of 3;6 and 9;6, and classified the errors. She also found that half the errors were due to misusing *but* as a simple conjunction, and in the other half a positive causal connective like *so* or *because* would have been more appropriate, suggesting the children understand that *but* has a causal function.

In summary, naturalistic production studies suggest that children may be able to correctly produce contrastive connectives as early as 3;0, but the evidence often comes from a small number of observations of correct usages for each child. Elicitation studies also found that children correctly use contrastive connectives with very few errors.

Controlled comprehension studies. Controlled comprehension studies of contrastive connectives are few, and all use different methods. They do share two consistent findings. First, even older children struggle with contrastive connectives. Second, children's errors suggest they treat contrastive connectives as simple conjunctions,[3] judging the plausibility of the co-occurrence of the two conjuncts.

An early study by Kail and Weissenborn (1984) tested French and German children's comprehension of contrastive connectives. Children were first given a context story about cowboys and Native Americans, where it is strongly suggested that Native Americans are better at rodeo

---

[3] Existing research alternatively claims that children treat *but* as *and*, ignore it, or only judge the compatibility of the conjuncts. All three of these explanations are simply alternative ways to describe interpretations as simple conjunction.

events. Children were then asked to judge sentences for acceptability. A sentence like *Joe is a Native American but he didn't win the rodeo* should be judged as acceptable given the story, while *Joe is a Native American but he won the rodeo* should be judged as unacceptable because there should be no contrast given the story. The seven-year-olds performed far below chance on the unacceptable sentences (French 22·2% correct and German 16·6% correct), but above chance for acceptable sentences (French 58·3% and German 61·1%.) The nine-year-olds performed above 90% correct for the acceptable sentences for both languages, but they only achieved an accuracy of 69·4% (for French) and 58·3% (for German) for the unacceptable sentences.

Kail and Weissenborn (1984) concluded that errors were cases of children simply judging plausibility, using a conjunction strategy. For sentences like *Joe is not a Native American but he won the rodeo*, children incorrectly rejected it, explaining it with comments like "If Joe is not a Native American then he cannot win" (see Kail & Weissenborn, 1984, p. 154).

A similar conclusion is made by Dragon, Berendes, Weinert, Heppt, and Stanat (2015) in their study of German schoolchildren's ability to identify appropriate and inappropriate usages of negative causal connectives. They found that 75% of erroneous responses could be explained as cases of children simply judging the appropriateness based on the semantic plausibility of the two clauses.

Making conclusions based on contrastive statements seems to be particularly difficult. Champaud and Bassano (1994) used a context choice task and a conclusion choice task to study children's comprehension of five French negative causal connectives (*mais*, *bien que*, *pourtant*, *quand même*, *même si*) and the positive additive connective *et* 'and'. Children aged 8;1–11;3 were presented with a target sentence like (7), and asked to choose the best context, C1 or C2, with C1 as the correct response. They were also asked to choose an appropriate conclusion from four possibilities, where (a) and (c) are correct, and (b) and (d) are incorrect.

(7)   He's a Native American but he lost at the rodeo
       C1 Native Americans are usually good at rodeo.
       C2 Native Americans are usually hopeless at rodeo.
       a He is disappointed.
       b He is pleased.
       c Everybody was astonished that a Native American lost.
       d Nobody was astonished that a Native American lost.

In general, children performed significantly better on the context choice task (82% correct) compared to the conclusion task (76% correct), but there were significant differences in the results by connective. On the

context choice, task the older children (mean age 10;10) scored above 90% for all connectives except for *mais*, where they only achieved 83% correct. For *mais* the younger children only had 71% accuracy.

For the conclusion choice task, the children performed worse overall, but again *mais* was an exception. The older children were near adult-like (96% correct) while the younger children were not far above chance (63% correct). Champaud and Bassano (1994) suggest that these exceptional results with *mais* showed different results because it was the only ambiguous connective, having both a semantic opposition and denial of expectation usage. Again, children's explanations for incorrect answers suggested a conjunction strategy.

Recent work on Dutch (Janssens & Schaeken, 2013; Janssens, Drooghmans, & Schaeken, 2015) also asked participants to choose an appropriate conclusion beginning with either *toch* (a negative causal connective, roughly translatable as English *nevertheless*, but more casual than *nevertheless*), or *dus* (a positive causal connective marking results, equivalent to English *so*). But because both adults and children (age range 8;0–12;0) performed poorly, it further suggests that conclusion tasks are quite difficult. The authors do note that children were very sensitive to whether or not the two arguments were plausibly related.

However, a number of studies suggest that errors cannot entirely be explained by a conjunction strategy. Kail (1980) presented French children with two types of inappropriate sentences. In the first type, (8a), the connective is inappropriate because the situation in the first conjunct would be expected to cause the situation in the second conjunct, so there is no contrast or expectation that is denied. In the second type, (8b), the two conjuncts are semantically compatible, but causally unrelated, so it should be rejected.

(8a)   The pen is new, but it writes well.
(8b)   The pen is red, but it writes well.

If children apply a conjunction strategy, then they should be equally likely to incorrectly accept both types. Instead, Kail (1980) found that six-year-olds (mean age 6;8) incorrectly accepted cases like (8a) 80% of the time, but incorrectly accepted unrelated sentences like (8b) significantly less often, 62% of the time. This difference in rate of acceptance suggests that children recognize that *but* requires some sort of causal relationship.

A recent study in German also supports the conclusion that children are doing more than simply treating contrastive connectives as conjunctions. Knoepke, Richter, Isberner, Naumann, Neeb, and Weinert (2017) tested appropriate and inappropriate usages of causal connectives (*deshalb, daher, darum,* and *denn* 'therefore') and the negative causal connective *trotzdem*

617

'nevertheless'. The events described in the conjuncts were always related. In inappropriate cases, e.g., (10) the situation in the first conjunct would be expect to lead to the opposite of the situation presented in the second conjunct, similar to (8a):

(9)   Sandra war niet müde. Trotzdem ging sie ins Bett.
       'Sandra was not tired. Nevertheless, she went to bed.'
(10)   *Das Wetter war gut. Trotzdem setzte Laura eine Sonnenbrille auf.
       *'The weather was good. Nevertheless, Laura put on her sunglasses.'

The children (age range 6;9–11;4) performed poorly with contrastive items, but were almost adult-like with the positive causal connectives. They systematically rejected coherent items like (9), while accepting incoherent ones, like (10). This result is consistent with a conjunction strategy. However, Knoepke *et al*. (2017) also included eight fillers with completely unrelated conjuncts, half with a positive causal connective and half with *trotzdem*. A conjunction strategy should lead to rejection of all these items. However, children incorrectly accepted these fillers, while correctly rejecting the positive causal connective fillers, indicating that at least some of their difficulty is related to the contrastive connectives themselves.

In four experiments, Cain and Nash (2011) asked eight- and ten-year-olds to rate appropriate and inappropriate usages of connectives. They found that their ten-year-olds performed similar to adults, but eight-year-olds were significantly different. The eight-year-olds still clearly differentiated between appropriate and inappropriate usages for all categories. In a reading time experiment, Cain and Nash also found that both eight- and ten-year-olds showed slower reading times for inappropriate *and* items, suggesting the incorrect use of the connective lead to greater processing difficulty. Inappropriate usage of adversative connectives resulted in the slowest reading times for both groups. These results seem to suggest that eight-year-olds are sensitive to correct and incorrect usages of connectives, but are still not adult-like.

In summary, results from previous studies suggest that children only begin to correctly understand contrastive connectives quite late, though studies differ as to when. Some find very poor results up to the age of 11;0, while others find good (though not adult-like) results already at age 8;0. However all comprehension results are unexpectedly late given the early accuracy (around 3;0) found in spontaneous production studies. There are a number of well-researched asymmetries in child language where production has been shown to precede comprehension, including the Delay of Principle B Effect (Chien & Wexler, 1990), and understanding SVO word order (Chapman & Miller, 1975) (see Hendriks

618

& Koster, 2010, and Hendriks, 2014, for more examples), but a gap of five or more years is remarkable.

Even if the robustness of some of the claims of correct production can be questioned, it seems unlikely that stricter criteria would add five or more years to the estimated age of acquisition. Instead, it seems more likely either that children in fact cannot understand contrastive coherence relations or that some particular features of testing comprehension inadvertently mask their abilities. Experimental artifact is one of the most common explanations to account for possible production–comprehension asymmetries (Hendriks & Koster, 2010). Careful selection of experimental methods might help us determine if the apparent gap between the production and comprehension data is real.

Choice of tasks. If poor comprehension results stem from overly complex tasks, then using a task with a simpler or more natural design should improve results. To better distinguish task difficulty from connective difficulty we should also test a cognitively simpler connective alongside the contrastive connective. The causal connective *because* is a good choice. It shares the fundamental causal feature with *but*, but is considered simpler because it lacks the negative/adversative meaning (Sanders *et al*., 1992; Spooren & Sanders, 2008). Because *but* and *because* are both causal, it is easy to create items with the same two events where either connective can be used. *Because* appears earlier than negative causal connectives (Evers-Vermeul & Sanders, 2009; Van Veen, Evers-Vermeul, Sanders, & Van den Bergh, 2009), and in comprehension studies with children (Evers-Vermeul & Sanders 2009; Van Veen *et al*., 2009; Knoepke *et al*., 2017) and adults (Knoepke *et al*., 2017), even young children show adult-like comprehension.[4] Finally, *because* is not ambiguous in Dutch (*want*).

A simple comprehension task would seem more natural than tasks that ask for truth-value judgments or ask children to compare alternatives. For this reason, a task where the interpretation of a pronoun is dependent on understanding the connective was developed. For example:

(11a)   Diego asked Sponge Bob to tidy up, because he didn't feel like doing it.
(11b)   Diego asked Sponge Bob to tidy up, but he didn't feel like doing it.
(11c)   Who didn't feel like tidying up?

---

[4] Note, however, that one aspect of *because* has been shown to cause difficulty up to the age of 11;0 (Johnson & Chapman, 1980): the order required by the connective is different than the order of the events, with the result-event preceding the cause. Note that neither task used relies on children correctly understanding this temporal ordering requirement.

In (11a), correctly interpreting the connective should lead to the answer *Diego* because an aversion to tidying up is a good reason to ask someone else to do it.[5] For (11b), the correct answer is *Sponge Bob*, because the expectation that Sponge Bob might help tidy up is violated (otherwise why would Diego ask him?). Children's knowledge of connectives can then be tested by simply (indirectly) asking how they interpreted the pronoun in the sentences. If children's poor performance on existing comprehension tasks compared to their production abilities is due in part to the unnaturalness of previous tasks, then this task should give us results similar to production results.

Finally, errors in this task are not ambiguous between a misunderstanding of the connective versus a conjunction strategy. The information in the two conjuncts is equally compatible and causally possible. A conjunction strategy should result in responses similar to chance.

One potential drawback is that pronoun interpretation strategies may also affect responses. A well-documented adult strategy is to interpret an ambiguous subject pronoun as the referent of the subject of the previous sentence (Crawley, Stevenson, & Kleinman, 1990; Smyth, 1994; Grosz, Weinstein, & Joshi, 1995; Chambers & Smyth, 1998). Children also show a similar subject antecedent preference (Wykes, 1981, for age 5;0, and Song & Fisher, 2005, for age 3;0). For this reason, items need to be balanced for the preferred grammatical role of the antecedent.

By using a new task and testing Dutch children, the current research is less linked to previous controlled comprehension experiments. In order to increase comparability to earlier work, an existing task was replicated, for Dutch with the same children. The truth-value judgment tasks used with Kail (1980), Kail and Weissenborn (1984), Dragon *et al*. (2015), and Knoepke *et al*. (2017) all require children to evaluate inappropriate sentences, something which seems to be particularly difficult. Further, they also rely in part on children's world knowledge based expectations.

The context choice task by Champaud and Bassano (1994) avoids these problems. Like a picture selection task, it presents two choices. This may be cognitively less demanding than asking for an absolute judgment because the task makes clear that only one interpretation is possible (Hendriks, 2010). Simplifying the context choice task, items with *want* 'because' as well as *maar* 'but' were used. Items were constructed to be unrelated to world knowledge, so responses cannot be based simply on co-occurrence plausibility. This task does have two drawbacks. First, incorrect responses to contrastive items could be interpreted as participants simply treating the connective as a conjunction. Second, choosing a

---

[5] Note that these types of questions would then fall into Spooren's (1989) 'Denial of Expectation' category.

620

sentence to occur before the target sentence is an artificial and rather awkward task. But because the aim is to link the new results to earlier work, these problems are acceptable. Champaud and Bassano (1994) had poor results for children at age eight, but found adult-like performance at age ten. For these reasons we decided to test ages seven to nine.

The Current Study

Task 1: the context choice task. The context choice task was based on the first task used in Champaud and Bassano (1994), but rather than using multiple concessive connectives as in the original study, only *maar* 'but' and *want* 'because' were used. However, first note that Dutch has two connectives that can be translated as 'because': *want* and *omdat*. In terms of meaning, *want* seems more subjective, while *omdat* is considered more objective. However, in many contexts both are possible, but there are a number of reasons to prefer *want* for the current study. *Want* seems to be acquired earlier (Evers-Vermeul & Sanders, 2009, 2011), and is intonationally integrated earlier. This difference is attributed to a syntactic difference: *want* is a coordinating conjunction, and has verb-second word order in its second conjuct, whereas *omdat* is a subordinating conjunction and requires verb-final word order. Because of this, *want* is more comparable to *maar*, also a coordinating conjunction.

Twelve different story items were created. Each item identified an individual, and then made a general statement either about what they can or cannot do, or about what they always or never do. Each item was thus a direct contrast (Winter & Rimon, 1994; Spenader & Maier, 2009), or Denial of Expectation contrast, using the definition of Spooren (1989). Four different versions of each item were made, crossing two factors: causal relationships with *want* 'because' or contrastive relationships with *maar* 'but', with or without negation.

(12a)   Sven is good at hopscotch because he comes from Sweden.
(12b)   Sven isn't good at hopscotch because he comes from Sweden.
(12c)   Sven is good at hopscotch but he comes from Sweden.
(12d)   Sven isn't good at hopscotch but he comes from Sweden.

Participants were presented with one of the story items and then asked to choose the best context sentence from two choices. One context sentence stated a positively oriented generalization, C1, and one context sentence stated an opposite, negative generalization, C2. Consider the following example. Subjects would be presented with (13), and then given the choice of two context sentences, C1 or C2. Whether or not the choice of context sentence was consistent with the target sentence was the dependent variable. Additionally, for three of the story items, two possible conclusion items were created, e.g., a and b.

621

(13)   Sven comes from Sweden, but he's not good at hopscotch.
       C1 Swedes are usually good at hopscotch.
       C2 Swedes are usually not good at hopscotch.
       a. He also thinks it's really fun.
       b. He also doesn't really like it.

The conclusion items served as control items to ensure that the children understood the task, and choosing the correct conclusion did not require knowledge of *want* 'because' or *maar* 'but'. The same two events were used in the *maar* and the *want* versions, but the order was different. For *want* sentences, the effect precedes the cause, while for *maar* sentences the cause precedes the effect. Four balanced lists were created, and two versions of each list were made by reversing the order of the lists. Each child only saw one version of each item, with six *want* questions and six *maar* questions.

Participants. Seventy-eight native Dutch-speaking children (M = 8·34 years, age range: 7;0–10;1) from four schools were tested individually in a quiet room at their school. Eight Dutch adults were also tested (M = 27·88, age range: 21–36).

Procedure. The target sentences were displayed in a presentation book, glued in the middle of each page, with one item per two pages. A Velcro strip was placed above the target sentence. Context sentences were printed on laminated paper strips with Velcro on the back. Children were asked to complete the stories by choosing the best context sentence and attaching it to the book with Velcro. Control conclusion statements also had two choices and were similarly fixable with Velcro.

The children were told that they were being asked to help create very simple stories for younger children. The experimenter first read the target sentence, and then read the two possible context sentences, allowing the children to hold them. Then the children were asked to choose a good beginning for the story. If there was a conclusion, the experimenter then read those sentences and gave the children the sentences to choose. Finally, the experimenter read the entire story and asked if the children were satisfied. The experimenter emphasized that, because of the Velcro, children could easily revise their responses. If this happened, the experimenter read the new sequence again. All responses were recorded on paper by a second experimenter, and the audio for all sessions was recorded for reference.[6] The entire test took about seven minutes and the children were quite enthusiastic.

---

[6] In accordance with agreements with the parents and schools, these recordings were erased after responses were added to the database.

Results Task 1: context choice task. Seven children made one or more errors on the control conclusion questions and were removed from further analysis. seventy-one children remained (M = 8·4 years).

For all children combined, the mean correct response for *want* 'because' was 91% (SD = 0·28), while the mean correct response for *maar* 'but' was 26% (SD = 0·44). From this data alone it is clear that children are much worse with the negative causal connective compared to the positive causal *want*.

To determine significant effects, responses were analyzed using logistic mixed-effect models (Baayen, Davidson, & Bates, 2008). The mean accuracy by age group data are presented in Figure 1.[7] The Akaike information criterion (AIC) values were compared to determine which model best fits the data, with a complex model being preferred over a simpler model only if the AIC value was lower with a difference of at least two. The fixed effects were CONNECTIVE, AGE in months (using normalized z-scores), TEST VERSION,[8] NEGATION, e.g., the presence of negation in the context sentence, and the random effects of PARTICIPANT, ITEM, and SCHOOL. Model testing began with a maximal model with all fixed factors and random factors. A comparison between models including the random factor Item and models without it showed no significant difference. This indicates that there were no significant differences between different test items, so this factor was removed from the model. Models with School were also not significantly different from models without, so this factor was also excluded from the final model. The model that explained significantly more variance retained the fixed factors Connective and Age only, with no significant interaction and a random slope of Connective for Participant. Children were significantly better at *want* 'because' than *maar* 'but' ($p < ·001$). Further, children get progressively better with age with both connectives ($p < ·001$). The final model is presented in 'Appendix A', Table A1.

Adult results were compared to the nine to ten age group using Fischer's exact test.[9] Adults were significantly better at *maar* 'but' than children ($p < ·001$), and the effect was large (Cramer's V = 0·655). For *want* 'because' there was no significant difference between adult participants and the nine to ten age group ($p = ·276$) or the eight-year-olds ($p = ·085$). However,

---

[7] For model comparison we treated age (in months) as a numerical value, but the ages in Figure 1 are presented by age group.

[8] Four schools in total were tested, three in a rural area and one in a small city. The experiment done at the first three schools wasn't balanced for negation across lists. In the task run at the final school, some list items were reordered. Tests indicated these differences had no significant effects.

[9] We used Fischer's exact test because adults made no errors with *want* 'because'. In such cases, mixed-effect linear models do not converge because of the separation problem.
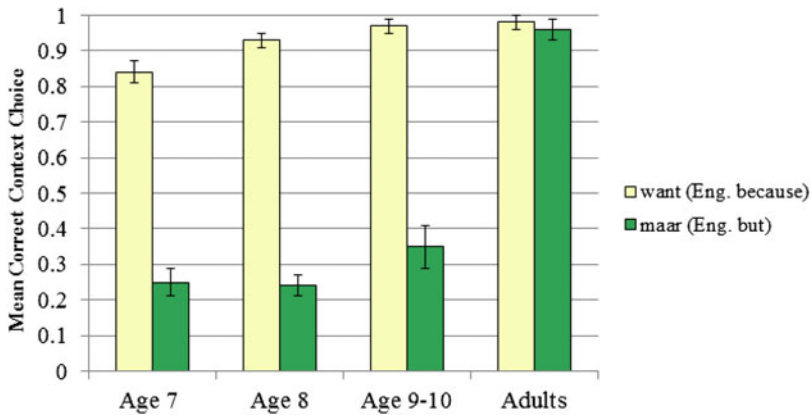
623

Fig. 1. Results: context choice task, with children divided into three age groups. Bars show mean accuracy and standard error. Adults included for comparison.

adults were significantly better than seven-year-olds ($p < \cdot 05$), but the effect was small (Cramer's V $= 0 \cdot 229$).

Were children who revised their initial response more or less likely to be correct? Recall that the potential to revise a response was emphasized in the instructions, and physically in the task by using Velcro to attach the sentences in the book. Unfortunately, revised responses were only annotated for the last school tested.[10] For the 26 children for which this data was retained, 14 changed their answer for one question, and the total number of revised responses was 24. Six children changed their answer for two or more questions, and one child changed his answer four times, the maximum. The 12 times a 'because' (Dutch *want*) answer was changed, the resulting response was correct. For the 12 times a 'but' (*maar*) answer was changed, eight times the resulting response was correct. In sum, children did make use of this option and, in general, a changed response was much more likely to be correct, especially for 'because' questions.

Discussion: context choice task. The children showed high accuracy with the *want* 'because' items, and this also suggests that the problems with the *maar* 'but' items are not simply inherent to the task itself. The Dutch children's results are much worse than their French counterparts who did a similar task. Recall that Champaud and Bassano (1994) showed that with *mais* 'but', the closest French equivalent to *maar* 'but', French eight-year-olds had an 83% accuracy rate.

---

[10] The other schools were tested more than a year earlier and this feature was not annotated at that time, and the recordings have since been discarded.

624

The accuracy of the two youngest groups is very far below chance. This suggests that the children are not simply guessing, but actually have an incorrect hypothesis about what *maar* 'but' actually means.

Task 2: pronoun interpretation task. The pronoun interpretation task uses sentences where the interpretation of pronouns differs depending on the connective, to test knowledge of the meaning of discourse connectives.

Sixteen test items were created with four versions, crossing the two factors Connective Type (*maar* 'but' or *want* 'because') and the Grammatical Role of the preferred antecedent type, e.g., subject or object.

(14a)   Ernie asked Big Bird for money because he didn't have any money. (*he* = Ernie)
(14b)   Ernie asked Big Bird for money because he had enough money (*he* = Big Bird)
(14c)   Ernie asked Big Bird for money but he didn't have any money (*he* = Big Bird)
(14d)   Ernie asked Big Bird for money but he had enough money. (*he* = Ernie)

Each item had the same structure. The first conjunct introduced two characters with the same gender using well-known children's characters (e.g., *Big Bird, Dora*). The second conjunct began with a pronoun, followed by a new event that, together with the connective information, should disambiguate the pronoun to a preferred interpretation based on the relationship between the events (e.g., in (14), asking someone for money, and having vs. not having money). After each target item participants were presented with an appropriate question, such as "Who didn't have any money / Who didn't have enough money" for (14). The response to the question was the dependent variable.

Nine different verbs were used (*ask, beg, call, comfort, encourage, give, help, lend,* and *offer*). Note that some verbs have strong continuation preferences, leading to expectations about whether the subject or object is the more likely antecedent to a continuation beginning with a pronoun. However, the verbs used here are not typically considered 'biased' because either they have no consistent bias or their biases are too weak.[11]

Four balanced lists were created and randomized. A version of each list in reverse random order was also created. Lists were also roughly balanced across ages and schools.

---

[11]   Verbal biases are often researched under the term 'implicit causality'. Six of the nine verbs we used have slight object biases according to Hartshorne and Snedeker (2013), ranging from 55% for *beg* to 75% for *ask*. But implicit causality experiments tend to use verbs with 90% or higher continuation biases, with verbs that are in the Stimulus-Experiencer or Experiencer-Stimulus class preferred. None of our verbs belong to either of these classes so verb bias is unlikely to have an effect on responses. However, to be sure we included BIAS as a potential factor in our analysis.

Participants. The same 78 children who took part in the context choice experiment took part in the pronoun experiment (M = 8;4, age range: 7;0–10;1). For this task, all children were retained. They were tested individually in a quiet room at their school. Fourteen native-speaking adults were also tested (M = 24·57, Age range: 18;0–32;0).

Procedure. The sentences were presented on a laptop visually with an audio-recording of each sentence read by a female native speaker. Because intonation might represent an additional clue to the intended referent, the files were spliced so that the text following the connective was the same recording for each version of each item. All responses were saved by the computer at testing. Additionally, audio-recordings were made for reference. Each child saw eight observations for each connective and eight observations for each grammatical role of the preferred antecedent. No child saw more than one version of an item. The entire task took about five minutes.

Results Task 2: pronoun interpretation task. Responses were again analyzed using mixed-effect logistic regression modeling. The fixed factors included were the factors CONNECTIVE (*want* 'because' or *maar* 'but'), GRAMMATICAL ROLE (subject or object), VERB BIAS (subject, object, or neutral), and AGE. Random effects were PARTICIPANT, ITEM, and SCHOOL. Models including School and Item were not significantly different from models without them, so these two random factors were removed. The maximally best model retained the features Connective, Age, and Grammatical Role, with Grammatical Role as a random slope for Participant. (See 'Appendix A', Table A2). No interactions were significant. Figure 2 shows the results per age group with adult results for comparison. Children were significantly better at *want* than *maar* ($p < ·01$). The overall mean score for the *want* sentences was 66% (SD = 0·47), while the overall mean for *maar* was 59% (SD = 0·49). Children also get significantly better with age ($p < ·001$). Further, if the correct antecedent was the subject of the previous sentence, they were more likely to get the response correct than if it was an object ($p < ·05$).

A Fischer's exact test was used to compare the results of the oldest children to adults. Adults were significantly better at both connectives than the nine- to ten-year-old group, but the effect size was small (For *want* 'because', $p < ·001$, Cramer's V = 0·245; for *maar* 'but', $p < ·001$, Cramer's V = 0·239).

Discussion: pronoun interpretation task. The most striking result is that results for both connectives are similar within age groups and no results
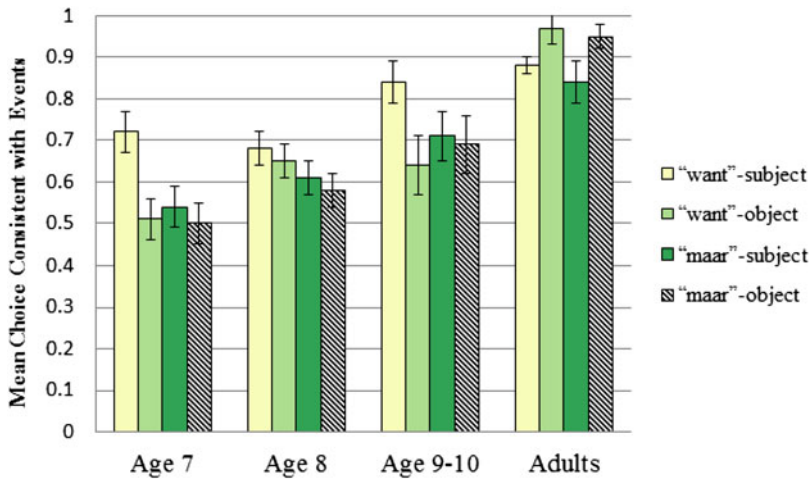
Fig. 2. Results: pronoun choice task, with children divided into three age groups Bars represent mean preferred responses consistent with the bias of the events and include standard error, distinguished by connective and whether the preferred interpretation had a subject or object antecedent. Adult responses included for comparison.

are below chance. Children are better at choosing the correct pronoun for sentences with *want* 'because' than with *maar* 'but', consistent with the context choice task. Only around age 8;0 do they start to make correct interpretations with *maar*. Development seems to be gradual, and *maar* is only slightly behind *want*. However, even the oldest children are still far from adult-like with *want*. They can apply connective information sometimes, but not consistently so.

Recall that the motivation for the pronoun interpretation task was to create a task that did not require judging the truth value of inappropriate sentences, and also did not require direct comparisons of answers. Indeed, children's overall performance with *maar* 'but' was better in the pronoun interpretation task than in the context choice task. But the pronoun interpretation task also involves other factors, such as pronoun interpretation strategies that children might rely on.

Children's better performance on items with a subject antecedent was expected, given earlier results that children prefer subjects as antecedents. Children unsure about the interpretation of the pronoun may fall back on a subject strategy, and by choosing more subjects than objects they get more items with subject antecedents correct than those with object antecedents.

Finally, recall that if children interpret 'but' as conjunctive 'and' in this task, the sentences become completely ambiguous, so the chance-level

627

accuracy of the seven-year-olds with 'but' are consistent with this analysis. But from age 8;0 there is improvement, suggesting that they are using the connective information sometimes.

Comparison of Results from the Two Tasks

The same participants took part in both experiments, which made it possible to compare results across tasks. Our original expectation was that performance on the pronoun interpretation task would be better than performance on the context choice task, because the context choice task is less natural and asks children to reason about connective meaning. The pronoun interpretation task tested connective knowledge indirectly, though it does also require the children to correctly interpret anaphors. It is certainly the case that for *maar* 'but' children do show much better performance with the pronoun interpretation task.

To see if there was a correlation between performance on the two tasks, the children's mean accuracy on each task was plotted and the correlation was measured using Spearman's rank correlation. For *want* 'because' there is a positive correlation between the two tasks, but it did not reach significance ($\rho = 0.21$, $p \leqslant .08$). For *maar* 'but' there was no correlation ($\rho = 0.08$, $p \leqslant .457$). Figure 3 shows the correlation between the two tasks for *want* and Figure 4 shows the correlation for *maar*. Ages 7;0–8;0 are represented with dots, while children from ages 9;0–10;0 are represented by triangles.

General Discussion

Results from the two experiments support the conclusions of earlier controlled comprehension experiments: comprehension of contrastive connectives is acquired very late (10+). In both experiments children did better with *want* 'because' than *maar* 'but', and in the context choice task, dramatically so. Because the results on the context choice task are positive for *want*, previous results for contrastive connectives that showed a great lag compared to production data seem to be genuine and not merely an artifact of the task.

However, children seem to have some knowledge of contrastive connectives. They improve with age in both experiments. During the experiment, many older children thought quite a long time about *maar* 'but' items. There were also a few older children with (near-)perfect performance, and many children were actively considering possible interpretations during the task.

The following sections address a number of remaining issues. First, why are our contrast choice task results for *maar* 'but' so much worse than results for the same task in French? Second, can we claim that the pronoun interpretation task results are a more accurate reflection of
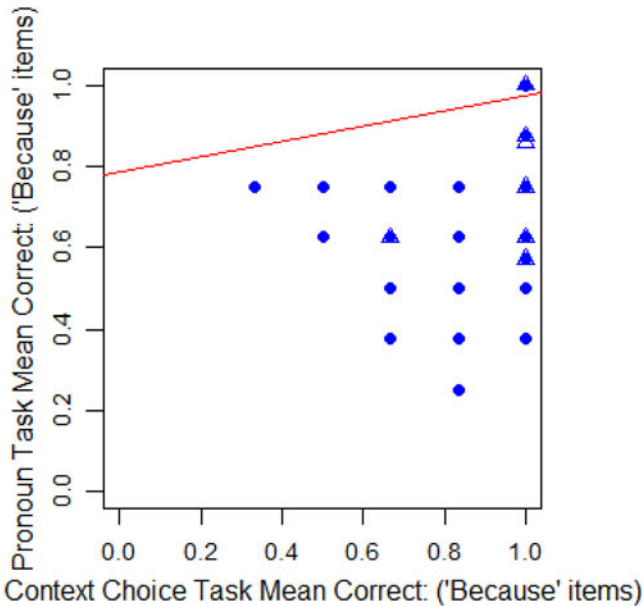
628

Fig. 3. Plot of correlation between the context choice task and the pronoun choice task for *want* 'because' items. Dots represent individual children aged 7;0–8;11, while triangles represent children aged 9;0–10;0.

children's abilities? Finally, what do our results suggest about how adult-like interpretations are acquired?

The difficulty of the context choice task. It was not unexpected that children would perform poorly with *maar* 'but', but it was surprising that the accuracy levels were actually below chance, given previous results with French. Children in Champaud and Bassano's (1994) experiment were older; the youngest group was already eight years old. Their task was more complex but otherwise identical. Yet still the average accuracy was above 80% for all groups, while the current study showed below chance responses at every age.

Because the current task was actually a simplification of Champaud and Bassano (1994), a ready explanation is that our simplification made the task harder in some way. Recall that Champaud and Bassano tested a number of contrastive connectives at the same time, while the current experiment randomly presented children with causal items with *want* 'because' and contrastive items with one connective, *maar* 'but'. This might have been confusing, in particular because the correct choice with
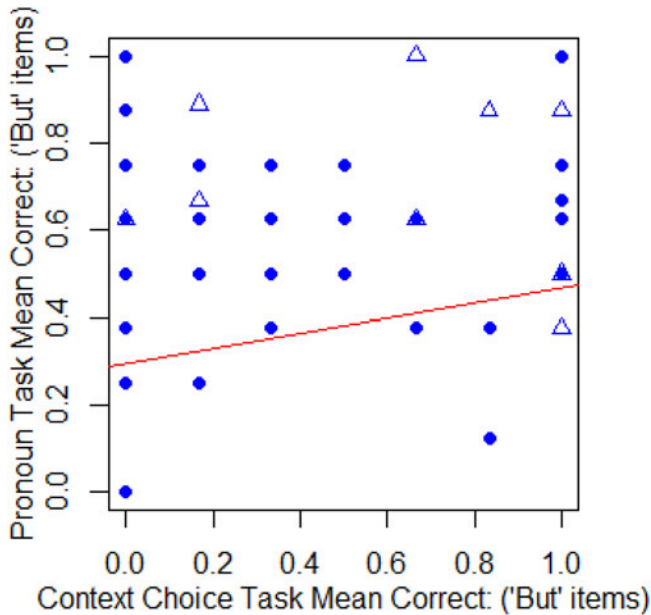
629

Fig. 4. Plot of correlation between the context choice task and the pronoun choice task for *maar* 'but' items. Dots represent individual children aged 7;0–8;11, while triangles represent children aged 9;0–10;0.

the causal items is the opposite of the correct choice with the contrastive items. The interpretation strategy for *want* items is simply to choose a context similar to the first conjunct of the target sentence. With *maar*, children must recognize that the predicated characteristic must be unexpected given the information presented, and then infer that if it is unexpected, then the opposite characteristic must be correct. Requiring children to randomly switch interpretation strategies where they must look for the same characteristic or look for the opposite characteristic might be more difficult than asking them to answer multiple questions that always require the same strategy. However, if this explanation is correct, it is still the case that the contrastive interpretation is harder because the need to switch strategies did not have an adverse effect on the *want* items.

The contrast in accuracy between *want* 'because' and *maar* 'but' is particularly striking. It may, however, be that some items with *want* were easier to get correct. In most items, children were told about a feature of an individual belonging to a group and they needed to match that with a general statement about people from that group. In a few items, children were told that an individual always or never did something: e.g., *Mr. Jansen always drinks coffee at breakfast*. Consider the following two items:

630

(15) TARGET 1 Piet houdt van schaatsen, want hij is matroos.
              'Piet likes ice skating, because he is a sailor.'
      TARGET 2 Piet houdt niet van schaatsen, want hij is matroos.
              'Piet does not like ice skating, because he is a sailor.'
      C1   Matrozen houden van schaatsen.
              'Sailors like to ice skate.'
      C2   Matrozen houden niet van schaatsen.
              'Sailors don't like to ice skate.'
(16) TARGET 1 Matt won the cooking competition, because he's a cowboy
      TARGET 2 Matt is a cowboy, but he didn't win the cooking competition.
      C1   Cowboys are good at cooking.
      C2   Cowboys are not good at cooking.

Half of the items were like example (15), where what the individual and the group like are presented with the same lexical items. The other half were instead like example (16). Examples like (16) encourage thinking more about the target statement's intention: if you win a cooking competition you are probably a good cook. For the first type of example, simply choosing the context with the same string (being mindful of the negation) will lead to the right answer for *want* 'because' but the wrong answer for *maar* 'but'. For examples like (16), this is not a possible strategy. However, if the number of matching words played a role, we should have seen an effect of negation because negated contexts share a longer string than non-negated contexts. This was not found in the statistical models.

Three changes could be made to avoid the two concerns above. First, each connective could be presented in a different block. This would no longer force children to juggle two strategies. Second, the experiment could be presented as a truth-value judgment task. The story could be presented as a two-sentence sequence, and children could be asked to judge its appropriateness. Because it requires judgments, this change unfortunately has some of the problems of being rather unnatural, as discussed earlier. A third improvement would be to make sure that context choices are all like example (16), so partial string matching would be ruled out as a possible strategy.

One final feature peculiar to spoken Dutch might have additionally contributed to the Dutch children's poor results compared to the French children in Champaud and Bassano (1994). During the creation of the stimuli we noticed a natural tendency to want to add the Dutch contrastive emphatic particle *wel* (roughly *actually* in English) to the positive target sentences. For example:

(17) DUTCH: Matt is een cowboy maar hij kan (wel) goed koken.
      GLOSS: Matt is a cowboy but he can (actually) good cook.
              'Matt is a cowboy but he's actually good at cooking.'

631

The addition of *wel* 'actually' is only possible for *maar* 'but' sentences with a positive argument, because *wel* cannot co-occur with negation. Recall also that the factor Negation was included in the model comparisons for the context choice task because negated sentences might be more difficult for children. However, Negation was not a predictive factor, so this indirectly indicates that positive sentences without *wel* did not differ significantly from sentences with negation. Still, *maar* often co-occurs with *wel* in spoken Dutch, and *maar . . . wel* 'but . . . actually' could be analyzed as a semantic unit. It would be useful to check the frequency of *maar . . . wel* in a corpus, and perhaps repeat the experiment with *maar . . . wel* to see if this change gives the children a boost in their interpretation of the positive sentences.

Evaluation of the pronoun interpretation task. This study was the first to use a pronoun interpretation task to test connective comprehension. Even though the scores for the pronoun task were much higher than for the context choice task, they were still rather low. This could in part be explained by the additional difficulty involved in anaphor interpretation. It also relies in part on children's knowledge of possible causal associations, e.g., that it makes sense to ask people who have money for money, or it doesn't make sense to ask people for money if you have enough money already. If children have difficulty with these types of generalizations, they may have difficulty with the task unrelated to their understanding of *want* 'because' and *maar* 'but'.

But it is promising that the pronoun interpretation task results are more consistent with production research. The pronoun interpretation task shows only a small gap between *want* 'because' and *maar* 'but', consistent with the naturalistic results from Evers-Vermeul and Sanders (2009) with Dutch. It is also consistent with the reading time results from Cain and Nash (2011) showing that children from the age of 8;0 react appropriately (showing slower reading times) for inappropriate usages of both causal and contrastive connectives, and that they show a greater slowdown for adversative connectives. However, it avoids presenting children with infelicitous items, and could be a useful complement to other testing methods.

How do children acquire contrastive connective interpretations? Previous production experiments showed high performance at a young age (around 3;0), while comprehension results showed poor performance until 9;0 or even older. Our results, even with the new pronoun interpretation task, still show poor comprehension with contrastive connectives, so a gap of five or more years remains. Recall that

632

one of the main explanations for production–comprehension asymmetries is task artifact (see Hendriks & de Koster, 2010). However, our results add to the previous results showing that adversative connectives are problematic for children across many different tasks and languages, so the asymmetry seems to be real.

An explanation for this gap might be found by examining some general differences between the nature of comprehension and production. In production, the children don't have to infer, because they know what they intend to say already. Johnson and Chapman (1980) argue that appropriate production relies most on context and knowledge about event probability. Children begin with context-dependent uses. Johnson and Chapman claim that children then improve at using connectives not because they better understand the lexical meaning of the connective itself, but because they become better at using it for context-independent statements. However, it is not clear if the development of these types of general communicative strategies really offer an explanation for such a specific lag in comprehension.

If children really do misunderstand contrastive connectives regularly, why don't we notice their confusion in everyday interactions? The explanation often given is that in natural communication the immediate context and knowledge about event probability are sufficient for children to make correct interpretations in most cases. Exposed to enough cases, the child can learn to associate sentences marked with *but* with unexpected events co-occurring. The unexpected association of the two events triggers attention to the connective. This explanation is then consistent with the results showing that children get better at recognizing appropriate usages faster than they get better at rejecting inappropriate usages, because those latter seem fine if the connective is ignored.

Usage-based theories of acquisition argue that children's acquisition can be explained to a great degree by the input that they receive. We could consider whether the contrastive use of *but* is perhaps so rare that this would interfere with children's acquisition process. However, available corpus data suggest that this is not the case. The Penn Discourse Treebank (PDTB) (Prasad *et al.*, 2008) is a one-million-word English corpus with texts from the *Wall Street Journal*, manually annotated with connective and discourse coherence relation types. For comparison, the connective *because* is used 858 times in the corpus, with 854 of those usages marking a causal relation (called 'reason' in the corpus). The connective *but* is used for a denial-of-expectation meaning 494 times (called 'contra-expectation')[12] (Prasad *et al.*, 2007). Even though the texts

---

[12] Note that *but* was also used with a meaning we would classify as semantic opposition meanings 824 times (aggregation of three categories, 'juxtaposition' (636), 'opposition' (174), and 'pragmatic contrast' (14)) (Prasad *et al.*, 2007).

633

these frequencies are taken from are quite different from spoken language and from the language children are exposed to, they do suggest that contrastive causal usages of *but* are not rare occurrences. It is therefore unlikely that we can ascribe children's late acquisition to a lack of examples in the environment.

It could, however, be the case that adults do not use the contrastive connective frequently in their interactions with children. We know that the frequency of linguistic constructions in child-directed input can correlate in part with children's output. For example, in an investigation of extensive recordings of a German child, Behrens (2006) showed that linguistic categories such as parts-of-speech were used in similar proportions in the child's language as found in the caretaker's speech. Van Veen *et al*. (2009) investigated the same child's connective usage frequencies.[13] They found that both short-term usage by adults in the same session, and long-term usage by adults across sessions, predicted the frequency of the child's individual connective use. Van Veen *et al*., however, did not find evidence that adults modified their usage for their child audiences (cf. Clark & Murphy, 1982). In a more recent paper, Van Veen, Evers-Vermeul, Sanders, and Van Den Bergh (2013) investigated the acquisition of the positive causal connective *because* in English and *weil* in German children and their caretakers in the Childes corpus (MacWhinney, 2000). They found evidence that it was not the use of connectives by adults that predicted acquisition, but the use of 'why'-questions. Adults asked 'why'-questions to 'scaffold' children's connective acquisition, and later children themselves produce 'why'-questions, eliciting causal connectives from the adults around them. Adults thus elicit responses that are naturally marked with causal connectives, and this correlates with children's acquisition rate. These results thus also suggest another potential explanation for why contrastive connectives are so difficult for children. There is no obvious or natural question that leads to a response with a contrastive connective.

Conclusions and Future Research

The results from the two tasks seem to confirm earlier controlled comprehension studies, that children have trouble interpreting contrastive connectives until age 9;0 and perhaps older.

By using the causal connective *want* 'because' with each task, task difficulty could be distinguished from connective difficulty. However, including *want* with *maar* 'but' in the same task may have made interpreting *maar* particularly difficult. The pronoun interpretation task is

---

[13] Note that they did a string search, so different usages of connectives were ignored.

a promising task that we can cautiously conclude seems to test connective interpretation with fewer confounds.

In future work, it would be useful to develop additional experiments that can distinguish the interpretation of *because* and *but* from interpretations based on co-occurrence plausibility and a conjunction strategy. Additionally, we should also test more contrastive connectives.

We also need to test even older children in order to see more evidence of acquisition, a conclusion shared by Janssens, Drooghmans, and Schaeken (2015). The production–comprehension gap still remains unexplained, and the results of the current study seem to even suggest that the lag is greater than first thought.

References

Anscombre, J.-C., & Ducrot, O. (1977). Deux mais en Francais. *Lingua*, *43*, 23–40.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Behrens, H. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes*, *21*(1–3), 2–24.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, *7*, 235–261.

Braunwald, S. R. (1985). The development of connectives. *Journal of Pragmatics*, *9*(4), 513–25.

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*(2), 429–41.

Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: a test of centering theory. *Journal of Memory and Language*, *39*(4), 593–608.

Champaud, C., & Bassano, D. (1994). French concessive connectives and argumentation: an experimental study in eight- to ten-year-old children. *Journal of Child Language*, *21*, 415–538.

Chapman, R. S., & Miller, J. (1975). Order in early two and three utterances: Does production precede comprehension? *Journal of Speech and Hearing Research*, *18*, 355–71.

Chien, Y.-C., & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, *1*(3), 225–95.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, *9*, 287–99.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, *19*(4), 245–64.

Dragon, N., Berendes, K., Weinert, S., Heppt, D.-P. B., & Stanat, P. (2015). Ignorieren grundschulkinder konnektoren? Untersuchung einer bildungssprachlichen komponente. *Zeitschrift für Erziehungswissenschaft*, *18*(4), 803–25.

Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives: how cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language*, *36*(4), 829–54.

Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains–on the acquisition of causal connectives. *Journal of Pragmatics*, *43*(6), 1645–62.

Foolen, A. (1993). *De betekenis van partikels: Een dokumentatie van de stand van het onderzoek met bijzondere aandacht voor maar [The meaning of particles: a research survey focusing on maar]*. Unpublished doctoral dissertation, Catholic University of Nijmegen.

Frege, G. (1948). Sense and reference. *Philosophical Review*, *57*(3), 209–30.

French, L. A. (1988). The development of children's understanding of 'because' and 'so'. *Journal of Experimental Child Psychology*, *45*, 262–79.

635

Grice, H. Paul (1975), Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and semantics, volume 3: speech acts* (pp. 41–58). New York: Academic Press.

Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, *21*(2), 203–25.

Hartshorne, J. K., & Snedeker, J. (2013). Verb argument structure predicts implicit causality: the advantages of finer-grained semantics. *Language and Cognitive Processes*, *28*(10), 1474–508.

Hendriks, P. (2010). What does eye-tracking reveal about children's knowledge of linguistic structure? In C. Jan-Wouter Zwart, & Mark de Vries (Eds.), *Structure preserved: studies in syntax for Jan Koster* (pp 149–56). John Benjamins.

Hendriks, P. (2014). *Asymmetries between language production and comprehension*. Dordrecht: Springer.

Hendriks, P., & Koster, C. (2010). Production/comprehension asymmetries in language acquisition. *Lingua*, *120*(8), 1887–97.

Janssens, L., Drooghmans, S., & Schaeken, W. (2015). 'But': Do age and working memory influence conventional implicature processing? *Journal of Child Language*, *42*(3), 695–708.

Janssens, L., & Schaeken, W. (2013). 'But': how do we reason with it: an experimental investigation of the implicature stemming from 'but'. *Journal of Pragmatics*, *57*. 194–209.

Johnson, H., & Chapman, R. (1980). Children's judgment and recall of causal connectives: a developmental study of 'because', 'so', and 'and'. *Journal of Psycholinguistic Research*, *9*(3), 243–60.

Kail, M. (1980). Etude génétique des présupposés de certains morphèmes grammaticaux. Un exemple: Mais. In *Approches de langage (pp. 53–62)*. Paris: Publications de la Sorbonne.

Kail, M., & Weissenborn, J. (1984). A developmental cross-linguistic study of adversative connectives: French 'mais' and German 'aber/sondern'. *Journal of Child Language*, *11*, 143–58.

Knoepke, J., Richter, T., Isberner, M.-B., Naumann, J., Neeb, Y., & Weinert, S. (2017). Processing of positive-causal and negative-causal coherence relations in primary school children and adults: a test of the cumulative cognitive complexity approach in German. *Journal of Child Language*, *44*(2), 297–328.

Lakoff, R. (1971). If's, and's and but's about conjunction. In C. J. Fillmore & D. T. Langendoen (Eds.), *Studies in linguistic semantics* (pp. 3–114). New York: Holt, Reinhart and Winston.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk*, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Peterson, C. (1986). Semantic and pragmatic uses of 'but'. *Journal of Child Language*, *13*, 583–90.

Potts, C. (2005). *The logic of conventional implicatures*. Oxford University Press.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. K., & Webber, B. L. (2008). The Penn discourse treebank 2·0. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*

Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., Joshi, A., Robaldo, L., & Webber, B. L. (2007). The Penn Discourse Treebank 2·0 annotation manual.

Sanders, T., Spooren, W., & Noordman, L. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1–35.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *23*(3), 197–229.

Song, H.-J., & Fisher, C. (2005). Who's 'she'? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, *52*(1), 29–57.

Spenader, J., & Maier, E. (2009). Contrast as denial. *Journal of Pragmatics*, *41*, 1707–26.

Spooren, W. P. M. S. (1989). *Some aspects of the form and interpretation of global contrastive coherence relations*. Unpublished doctoral dissertation, Nijmegen University.

Spooren, W., & Sanders, T. (2008). The acquisition order of coherence relations: on cognitive complexity in discourse. *Journal of Pragmatics*, *40*(12), 2003–26.

Van Veen, R., Evers-Vermeul, J., Sanders, T., & Van den Bergh, H. (2009). Parental input and connective acquisition: a growth curve analysis. *First Language*, *29*(3), 266–88.

Van Veen, R., Evers-Vermeul, J., Sanders, T., & Van Den Bergh, H. (2013). The influence of input on connective acquisition: a growth curve analysis of English 'because' and German 'weil'. *Journal of Child Language*, *40*(5), 1003–31.

Winter, Y., & Rimon, M. (1994). Contrast and implication in natural language. *Journal of Semantics*, *11*(4), 365–406.

Wykes, T. (1981). Inference and children's comprehension of pronouns. *Journal of Experimental Child Psychology*, *32*(2), 264–78.

Appendix A
Appendix B Experimental Items

Context choice task

Table A1

Context choice task: Model = Correct Choice ∼ Connective + Age (in months, normalized) + (1 + Connective | Participant)

Predictor	Estimate	Standard error	z-value	p-value
Intercept (*but*)	−1·9494	0·3911	−4·985	< ·001
Connective (*because*)	4·7619	0·5561	8·563	< ·001
Age (in months)	0·5653	0·1511	3·742	< ·001

Table A2

Pronoun task: Model = Correct ∼ Connective + Age (in months, normalized) + Grammatical Role (1 + Grammatical Role | Participant)

Predictor	Estimate	SE	z-value	p-value
Intercept (subject + *but*)	0·55329	0·11781	4·697	< ·001
Connective (*want*)	0·33469	0·12079	2·771	< ·01
Grammatical Role (*object*)	−0·35269	0·15321	−2·302	< ·05
Age (in months)	0·26732	0·06464	4·136	< ·001

The Context choice experiment had twelve items with four different versions crossing the factors of Connective (*because* or *but*) and Negation (whether or not negation was present). Three items also had a control question. All four conditions for the first item are presented in Dutch with English translations. For the following eleven items we present only the positive version with *because* in Dutch and its English translation.

1. Matt heeft de kookwedstrijd gewonnen, want hij is een cowboy.
   Eng: Matt won the cooking competition, because he is a cowboy.
   Matt heeft de kookwedstrijd verloren, want hij is een cowboy.
   Eng: Matt lost the cooking competition, because he is a cowboy.
   Matt is een cowboy, maar hij heeft de kookwedstrijd gewonnen.
   Eng: Matt is a cowboy, but he won the cooking competition.
   Matt is een cowboy, maar hij heeft de kookwedstrijd verloren.

637

Eng: Matt is a cowboy, but he lost the cooking competition.
Cowboys zijn erg goed in koken.
Eng: Cowboys are good at cooking.
Cowboys zijn erg slecht in koken.
Eng: Cowboys are very bad at cooking.

2. Mika is goed in afwassen, want hij is een Ninja.
   Eng: Mika is good at washing dishes, because he is a Ninja.
3. Lisa eet graag gevulde koeken, want ze komt uit Brabant.
   Eng: Lisa likes to eat filled cookies, because she is from Brabant.
4. Meneer Jansen drinkt koffie, want hij eet zijn ontbijt.
   Eng: Mr. Jansen drinks coffee, because he's eating his breakfast.
5. Zorg kan goed dansen, want hij is een marsmannetje.
   Eng: Zorg can dance well, because he is a Martian.
6. Anne heeft gevoel voor humor, want ze komt uit Nijmegen.
   Eng: Anne has a sense of humor, because she comes from Nijmegen.
7. Paul houdt van verstoppertje, want hij komt uit Canada.
   Eng: Paul likes to play hide-and-go-seek, because he comes from Canada.
8. Lee vindt de Wii cool, want hij komt uit China.
   Eng: Lee thinks the Wii is cool, because he comes from China.
9. Piet houdt van schaatsen, want hij is matroos.
   Eng: Piet likes to ice skate, because he is a sailor.
10. Sven kan goed hinkelen, want hij komt uit Zweden.
    Eng: Sven can play hopscotch well, because he comes from Sweden.
11. Tomoko kan tekenen, want ze komt uit Japan.
    Eng: Tomoko can draw well, because she comes from Japan.
12. Peter kijkt televisie, want het is maandagavond.
    Eng: Peter is watching TV, because it is Monday evening.

The pronoun interpretation task items

The pronoun interpretation task had sixteen items with four versions of each item, crossing the two factors Connective (*because* or *but*) and the Grammatical Role of the preferred antecedent (SUBJECT or OBJECT). The first item is given with all four forms and their English translations plus the two questions possible (depending on which ending the item had). The following fifteen items are given in shorthand form.

1. Ernie vroeg Pino om geld, want hij had genoeg geld.
   Eng: Ernie asked Pino for money, because he had a lot of money.
   Ernie vroeg Pino om geld, want hij had geen geld.
   Eng: Ernie asked Pino for money, because he didn't have any money.
   Ernie vroeg Pino om geld, maar hij had genoeg geld.
   Eng: Ernie asked Pino for money, but he had a lot of money.

638

Ernie vroeg Pino om geld, maar hij had geen geld.
Eng: Ernie asked Pino for money but he didn't have any money.
Wie had genoeg geld?
Eng: Who had a lot of money?
Wie had geen geld?
Eng: Who didn't have enough money?

2. Nemo smeekte Voldemort om een snoepje (want/maar) hij had veel snoep/hij had nooit snoep.
   Eng: Nemo begged Voldemort for a piece of candy, (because/but) he had lots of candy / he never had candy.

3. Shrek heeft Bert zijn fiets aangeboden, (want/maar) hij woonde ver weg/hij wilde liever met de bus.
   Eng: Shrek offered Bert his bike, (because/but) he lived far away / he preferred to take the bus.

4. Pikachu belde Mickey Mouse op over het huiswerk, (want/maar) hij was er heel goed in./hij snapte er niks van.
   Eng: Pikachu called Mickey Mouse up about the homework, (because/but) he was very good at it / he didn't understand any of it.

5. Kwik gaf Grover zijn knikkers, (want/maar) hij wilde met lego spelen/hij wilde er nu graag mee spelen.
   Eng: Huey gave Grover his marbles, (because/but) he wanted to play with Lego / he really wanted to play with them now.

6. Knorretje vroeg Ronald om advies, (want/maar) hij had al besloten wat er gedaan moest worden./hij wist de oplossing niet.
   Eng: Piglet asked Ronald for advice, (because/but) he had already decided what had to be done / he didn't know the solution.

7. Octo gaf Dagobert zijn jas, (want/maar) hij vond een t-shirt genoeg./hij had het erg koud.
   Eng: Octo gave Dagobert his coat, (because/but) he thought a T-shirt was enough / he was very cold.

8. Iejoor vroeg Tommie om een krijtje, (want/maar) hij had een hele doos vol./hij had er geen.
   Eng: Eeyore asked Tommy for a crayon (because/but) he had a whole box full / he didn't have any.

9. Patrick vroeg Koekie-monster om een vel papier, (want/maar) hij had een heel schrift./hij had geen papier.
   Eng: Patrick asked Cookie monster for a piece of paper, (because/but) he had a whole book / he didn't have any paper.

10. Barbar wilde Poeh troosten, (want/maar) hij was niet meer verdrietig./hij was te laat.
    Eng: Barbar wanted to comfort Pooh, (because/but) he was not sad anymore / he was too late.

639

11. Elmo heeft Pipo zijn auto geleend, (want/maar) hij wilde er ook graag mee spelen./hij wilde liever met playmobil spelen.
Eng: Elmo lent his car to Pipo, (because/but) he also really wanted to play with it / he wanted to play with playmobil instead.

12. Kermit wilde Plopp komen aanmoedigen, (want/maar) hij had een wedstrijd./hij had al gewonnen.
Eng: Kermit wanted to cheer Plopp on, (because/but) he had a race / he had already won.

13. Diego vroeg Sponge Bob om op te ruimen, (want/maar) hij had de meeste rommel gemaakt./want hij had geen zin.
Eng: Diego asked Sponge Bob to tidy up, (because/but) he had made most of the mess / he didn't feel like it.

14. Elmo vroeg Batman om eten te koken, (want/maar) hij kon veel beter koken./hij had geen zin.
Eng: Elmo asked Batman to cook something, (because/but) he was a better cook / he didn't feel like it.

15. Piet Piraat wilde Ash helpen, (want/maar) hij was al klaar./hij was een beetje moe.
Eng: Piet Pirate wanted to help Ash (because/but) he was already done / he was a bit tired.

16. Simba smeekte Elmo om een koekje, (want/maar) hij had veel koekjes/hij had er geen meer.
Eng: Simba begged Elmo for a cookie, (because/but) he had lots of cookies / he didn't have any more.