‘But’ how do Children judge it on a Scale?

Leen Janssens (Leen.Janssens@ppw.kuleuven.be)
Laboratory of Experimental Psychology, Tiensestraat 102
3000 Leuven, BELGIUM

Annelies Van den Broeck (Annelies.Vandenbroeck@student.kuleuven.be)

Walter Schaeken (Walter.Schaeken@ppw.kuleuven.be)
Laboratory of Experimental Psychology, Tiensestraat 102
3000 Leuven, BELGIUM

Abstract

This study examined children’s comprehension of the conventional implicature induced by *but*, combined with *so* and *nevertheless*, in ‘p but q’ sentences constructed as distancing-contrastive connections. Based on the Pragmatic Tolerance Hypothesis of Katsos and Bishop (2011), a three-point scale was used as response format. Using a scale instead of a binary judgment task can reveal more insight in which factors are considered most important when processing ‘p but q’ sentences. The results indicated that the content of the p- and q-arguments plays a very important role when children process ‘p but q’ sentences. However, their use of the three-point scale also indicated that they are sensitive to the pragmatic meaning of *but*, *so* and *nevertheless*. These results must be interpreted cautiously since the children seemed to use the middle value on the scale around 30% of the time in each sentence category, which was not in line with our predictions. This might indicate that children experience a general incomprehension with this type of sentences and answer with the middle value on the scale because they simply don’t know the answer.

Keywords: conventional implicature; *but*; scale; content

Introduction

Over the past few decades, considerable experimental research has been devoted to scalar implicatures. Grice introduced the term *implicature* in the 1967 William James lectures to offer an explanation for how it is possible that an utterance can mean more than what is literally said. Scalar implicatures are a subcategory of conversational implicatures and are based on a scale of informativity. For example, on the scale <all, most, many, some> the use of the more informative *all* logically entails that *some* is also true. However, in an utterance such as *‘Some Belgians like to drink beer’*, the pragmatic meaning of *some* causes the hearer to interpret this utterance as *‘Some but not all Belgians like to drink beer’* even though the logical meaning of *some* is *‘some and perhaps all’.* According to Grice (1989), people follow a set of maxims in communication in order to understand each other correctly. That’s why the consensus applies that whenever a speaker uses a weak term such as *some*, a stronger term such as *all* does not hold. The speaker would not have been optimally informative if a stronger term applied.

Developmental conversational implicature research has shown that children are less pragmatic than adults. For example, Noveck (2001) found that 89% of the 7-to-8-year-olds in his study agreed with statements such as *‘Some giraffes have long necks’*, compared to only 41% of the adults. Similarly, with respect to propositional connectives, Braine and Rumain (1981) presented evidence showing that deductively competent 7- and 9-year-old children favor a logical interpretation of *or* (*‘p or q and perhaps both’*) over an implicit one (*‘p or q but not both’*). Adults on the same task were equivocal, though they tended to favor exclusive interpretations (Braine & Rumain, 1981). However, these (and other) studies claiming that children lack pragmatic competence have been criticized by Katsos and Bishop (2011). In their implicature studies, Katsos and Bishop (2011) argued that earlier studies mostly employed tasks that cannot differentiate between actual implicature derivation and mere sensitivity to violations of informativeness. The majority of studies concluding that children are more logical than adults used binary judgment tasks in which participants were instructed to judge an utterance as ‘true’ or ‘false’. Katsos and Bishop (2011) argued that children might not reject underinformative sentences because they are tolerant to violations of informativeness. However, this doesn’t mean that they are not sensitive to these violations. In order to test this Pragmatic Tolerance Hypothesis, Katsos and Bishop (2011, Experiment 2) instructed their participants to judge on a ternary scale how well a fictional character described certain situations. They found that children’s performance did not differ from adults’. Underinformative utterances were judged by both groups with the middle value on the scale. This shows that children understand that using for example *some*, when *all* would have been a more informative description, is not optimal. However, in a binary judgment task they would not penalize such a description as false whereas adults would. In previous research (e.g. Noveck, 2001) this falsely led to the conclusion that children lack pragmatic competence.

Besides conversational implicatures, Grice (1989) also distinguished the category of conventional implicatures. This paper will deal with this far less investigated category of conventional implicatures. Unlike conversational implicatures conventional implicatures (a) are related to the conventional meaning of words, (b) are immediate conclusions from utterances, (c) cannot be cancelled and (d) are related to the form of an utterance, not the content. In an utterance such as *'He's old but he's smart'*, *but* conventionally implies a contrast. The use of *but* elicits the inference that 'old' and 'smart' contrast each other even though this is not explicitly expressed.

The experiments in this paper focus on the conventional implicature induced by the conjunction *but*. Our experiments build further on Janssens and Schaeken (2013) and Janssens, Drooghmans and Schaeken (in press). However, because of the important findings of Katsos and Bishop (2011) concerning conversational implicatures we will apply a ternary scale instead of a binary judgment task. This allows us to test certain predictions about children's understanding of this conventional implicature that cannot be discovered by using a binary judgment task.

In Janssens and Schaeken (2013) 'p but q' utterances, constructed as distancing-contrastive connections, were examined. In a distancing contrast, *but* connects two parts of a complex speech act (Van Belle & Devroy, 1992) and the second part is disassociated from the first part, without denying what is being expressed in the first part (Haeseryn et al., 1997). For example:

(1) Hannah: "I really like these beautiful earrings, but they are very expensive."

In a 'p but q' construction, the speaker endorses that *p* is true (Van Belle, 2003). However, because *but* is used, the inference from the p-argument is cancelled in favor of the inference from the q-argument. In (1), the p-argument elicits the conclusion that Hannah will buy the earrings whereas the q-argument elicits the opposite conclusion that she will not buy the earrings. The conventional meaning of *but* causes the q-argument to outweigh the p-argument so the appropriate conclusion from a 'p but q' sentence is inferred from the q-argument. Consequently, from (1) the conclusion follows that Hannah will not buy the earrings. If the two arguments trade places ("*they are very expensive, but I really like these beautiful earrings"*) the opposite conclusion will be drawn that Hannah will buy the earrings. This shows that the conventional meaning of *but* provides more weight to the q-argument irrespective of the content of the arguments. According to Anscombre and Ducrot (1977), every argument is determined by a certain positive or negative value ascribed to its content, which they labelled the 'axiological value'. The axiological value we ascribe to the arguments of an utterance is dependent on cultural-specific common sense views (Van Belle & Devroy, 1992). In (1), the p-argument is oriented towards a positive conclusion (Hannah will buy the earrings) and the q-argument is oriented towards a negative conclusion (Hannah will not buy the earrings). That's why we label the p-argument in (1) as the positive argument and the q-argument as the negative argument.

The conclusion from a 'p but q' construction can be introduced by words such as *so* or *nevertheless*. The pragmatic meaning of these two words leads to opposite conclusions. *So* elicits the conclusion from *q* and therefore confirms the expected conclusion inferred from the pragmatic meaning of *but* (*I really like these beautiful earrings, but they are very expensive. So I will not buy them.*). In contrast, according to Van Belle (2003), whenever *nevertheless* -used as a conjunctive adverb- follows a 'p but q' sentence, it reverses the argumentative orientation again. The expected conclusion from *q* is overruled and the reader is redirected towards the conclusion inferred from *p* (*I really like these beautiful earrings, but they are very expensive. Nevertheless I will buy them).* Note that *nevertheless* is used here as a translation of Dutch *toch*.

The adult participants in Janssens and Schaeken (2013) were presented with short stories that ended with a 'p but q' sentence. Both sensible (Se) and irrelevant (Ir) arguments were administered. In (1), both arguments are sensible in a context in which a woman is standing in a jewelry store. In this same context, uttering *"I really like these beautiful earrings, but I have a brother"* clearly contains an irrelevant q-argument. The irrelevant arguments were unrelated to the context of the stories and their purpose was to examine whether the pragmatic meaning of *but* is understood irrespective of the content of the arguments.

Each 'p but q' sentence was followed by two possible *so*-conclusions ('*so* conclusion from *p*' and '*so* conclusion from *q*') or by two *nevertheless*-conclusions ('*nevertheless* conclusion from *p*' and '*nevertheless* conclusion from *q*'). The participants were instructed to indicate the appropriate conclusion. Janssens and Schaeken (2013) expected the appropriate pragmatic conclusion following *so* to be the conclusion inferred from *q* and the appropriate conclusion following *nevertheless* to be the inferred conclusion from *p*. The general outline of the results showed that adults understand the pragmatic meaning of *but*. However, the content of the arguments plays a non-negligible role. Whenever an irrelevant argument was combined with a sensible argument, the participants practically always inferred the conclusion from the sensible argument, irrespective of the pragmatic inference from *but*, *so* and *nevertheless*. The importance of the content was confirmed in a second experiment in which participants were asked to justify their answer. As expected, participants mostly referred to the content of the arguments whenever they did not provide the appropriate conclusion. More evidence showing the importance of the content was found in the fact that 82% appropriate *so*-conclusions were given when two sensible arguments were presented. This means that 18% of the answers was based on the inappropriate p-argument which the participants probably judged as a better argument than *q*. Another finding was that inferring the appropriate

nevertheless-conclusion is a lot more difficult than inferring the appropriate *so*-conclusion. Only 48% appropriate *nevertheless*-conclusions were given when two sensible arguments were presented. This could indicate that the expected pragmatic meaning of *nevertheless* might not be the right one. On the other hand, this finding can be explained by the fact that the inference stemming from *nevertheless* is opposite to the inference stemming from *but*. In order to make the appropriate inference from *nevertheless*, the inferred conclusion from *but* has to be cancelled. It seems plausible that this would require effort and therefore leads to a higher percentage of inappropriate answers.

Janssens et al. (in press) performed the same experiment as Janssens and Schaeken (2013) but with children aged 8 to 12. Additionally, they measured working memory (WM) in order to see whether WM is involved in processing the conventional implicature stemming from *but*. The children's results showed the same pattern as the adult-data but the percentages of appropriate answers were lower. Moreover, no significant effect of WM was found.

In this paper we apply the methodology of Katsos and Bishop (2011) on children between the ages of 8 and 12. If children understand the pragmatic meaning of *but*, *so* and *nevertheless* but are also sensitive to the content of the arguments, we expect them to choose the middle value on the scale when they have to judge the appropriateness of a conclusion from a 'p but q' construction in which there is a conflict between the pragmatic answer and the answer based on the content. Since both the content and the conventional meaning of *but* can play a role in judging conclusions from 'p but q' sentences, different predictions can be made for each of the categories. A schematic view is presented in Table 1. This table depicts which of the two arguments (*p* or *q*) gets most weight based on (1) the content, (2) *but* and (3) the conclusion-word (*so* or *nevertheless*). We can see in Table 1 that the content as well as *but* and the conclusion-word guide the reader towards the conclusion from *q* in the 'IrSe *so*' sentences. That's why we predict a lot of optimal answers on the scale and no neutral (middle) answers. If the content is very important for children and they are rather tolerant with respect to *but*, then we also expect very few neutral answers for the 'SeIr *nevertheless*' sentences. When both arguments are sensible, the content should not play a role. When these sentences are combined with *nevertheless*, then *but* and *nevertheless* lead to opposite conclusions. This might lead to doubt, but also to inappropriate answers, depending on which of the two factors is more important. If children are not at all sensitive to the pragmatic meaning of *but*, *so* and *nevertheless*, we would expect many neutral answers for both the *so*- and the *nevertheless*-conclusions. Also, if children truly lack this sensitivity, we would expect no neutral answers for the 'IrSe *nevertheless*'- and the 'SeIr *so*' sentences. In both cases, the content guides them towards the inappropriate conclusion and this would not be corrected by *but*, *so* or *nevertheless*.

Table 1: Indication of which argument has more weight for every sentence category.

Sentence-category	Content	But	Conclusion-word
SeSe_So	=	q	q
SeSe_Nevertheless	=	q	p
IrSe_So	q	q	q
IrSe_Nevertheless	q	q	p
SeIr_So	p	q	q
SeIr_Nevertheless	p	q	p

Se=sensible; Ir=irrelevant

Experiment 1

Method

Participants Sixty-six Dutch speaking children (31 boys and 35 girls) between the ages of 8 and 10 years with a mean age of 9.1 participated in this study. They were recruited from five classes of two different schools.

Implicature Task The implicature task was based on Janssens and Schaeken (2013) but the design was adapted. The children were presented with 24 context stories. Each of the stories described a person in doubt about something. For example:

> *Peter's best friend is flying to Egypt to go on a diving holiday. He asks if Peter wants to come along. Peter is in doubt whether he will join his best friend or not.*

Each short story was followed by a 'p but q' sentence with two contrastive arguments expressing doubt. For example:

> *Peter thinks: "I'm afraid of flying, but I would like to learn how to dive."*

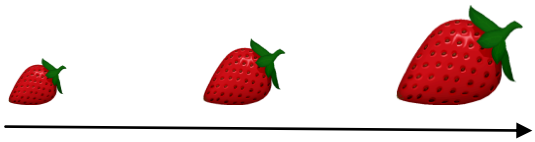
In the example above, both the p– and the q-argument are sensible arguments. The p-argument in this example is the negative argument (leading towards the negative conclusion '*I will not join my best friend on his trip')* and the q-argument is the positive argument ('*I will join my best friend on his trip'*). However, as in Janssens and Schaeken (2013), we also included irrelevant arguments in this experiment. The irrelevant arguments are not oriented towards a certain positive or negative conclusion. If the children understand the pragmatic meaning of *but* then these irrelevant arguments acquire a certain axiological value simply because they are contrasted with another (positive or negative) argument. An example of a combination of an irrelevant and a positive argument is:

> *Peter thinks: "I like to eat chicken, but I would like to learn how to dive."*

After each 'p but q' sentence, a certain conclusion was expressed by a fictional character, 'Mr. Coleman'. This conclusion was introduced by either *so* or *nevertheless*.

Whenever *so* follows a 'p but q' sentence we expect the conclusion from the q-argument but when *nevertheless* follows, we expect the conclusion from *p*. The conclusions that were presented could be of four different kinds: '*so* conclusion from *p*', '*so* conclusion from *q*', '*nevertheless* conclusion from *p*' or '*nevertheless* conclusion from *q*'. After a certain conclusion was expressed, the children had to indicate on a scale how appropriate they judged the conclusion. Based on Katsos and Bishop (2011), we used a three-point scale with different sized strawberries. The children were instructed to reward a good conclusion with the biggest strawberry, a bad conclusion with the smallest strawberry and a conclusion that was not completely bad nor good with the medium-sized strawberry. E.g.:

Mr. Coleman says: "So Peter will join his best friend on his trip to Egypt."



The 24 stories represented an item from every combination of our 3x2x4 design. There were three possible argument-combinations (SeSe, SeIr, IrSe), two different axiological value combinations (negative-positive, positive-negative) and four conclusion types ('*so q*', '*so p*', '*nevertheless q*', '*nevertheless p*').

Procedure The task was administered to the children as a pen-and-paper task which they performed individually in their classroom at school. The task was introduced by a preliminary story about mister Coleman who comes from America and wants to learn Dutch. The children were told that Mr. Coleman would utter several conclusions based on each story and that they had to reward Mr. Coleman with different sized strawberries, dependent on how appropriate the uttered conclusion was.

Results

In the analyses, we did not make a distinction between positive and negative arguments. When analyzing them separately, we did not find significant differences. That's why, in our analyses, we collapsed them. We recoded the children's answers into appropriate (three points), neutral (two points) and inappropriate (one point) answers. First we looked at the percentages of neutral answers for each of the argument-conclusion combinations. These percentages are displayed in Table 2, together with the percentages of appropriate answers. There were no significant differences in the number of neutral answers between the different categories (X^2=1.21; *df*=2, *p*=.55). This is not in line with our expectations. We expected almost no neutral answers for the 'IrSe *so*' sentences and the 'SeIr *nevertheless*' sentences. However, since the number of neutral answers was evenly distributed over the different categories, this allowed us to sum up the scores in every category.

Table 2: Percentages of neutral and appropriate (between brackets) answers for each argument-conclusion combination (Experiment 1)

Sentence	So	Nevertheless
Sensible-Sensible	31(41)	25(41)
Sensible-Irrelevant	29(25)	30(47)
Irrelevant-Sensible	28(55)	25(18)

When we look at the results of the *so*-conclusions, the children scored highest on the appropriateness scale for the 'IrSe' sentences (79%). This differed significantly from 'SeSe' (71%; Wilcoxon Signed Ranks test, *n*=55; *Z*=-3.66; *p*<.001) and from 'SeIr' (59%; Wilcoxon Signed Ranks test, *n*=66, *Z*=-5.28; *p*<.001). These last two categories also differed significantly from each other (Wilcoxon Signed Ranks test, *n*=56, *Z*=-4.18; *p*<.001). When we look at the results of the *nevertheless*-conclusions, the children scored highest on the appropriateness scale for the 'SeIr' sentences (75%). This differed significantly from the 'IrSe' sentences (54%; Wilcoxon Signed Ranks test, *n*=58; *Z*=-5.32; *p*<.001) and marginally significantly from the 'SeSe' sentences (69%; Wilcoxon Signed Ranks test, *n*=54; *Z*=-1.89; *p*=.059). These last two categories also differed significantly from each other (Wilcoxon Signed Ranks test, *n*=53; *Z*=-5.18; *p*<.001).

In order to compare *so*-conclusions with *nevertheless*-conclusions, we have to look at the 'SeSe' sentences. We found no significant difference in performance between these two categories (Wilcoxon Signed Ranks test, *n*=49; *Z*=-.93; *p*=.35).

Discussion

The general outline of the results of the 8-to-10-year-olds seems to be in line with previous findings in Janssens et al. (in press). However, the introduction of a three-point scale enabled us to examine children's sensitivity to *but* in another way. The fact that children provide an inappropriate answer about half the time for the 'SeIr *so*'- and the 'IrSe *nevertheless*' sentences means that they provide an appropriate or neutral answer the other half of the time. As a consequence, this implies that, despite the importance of the content of the arguments, children are clearly sensitive to the pragmatic meaning of *but* and the conclusion-words. However, Table 2 shows that the percentage of neutral answers is around 30% in each category. This is contrary to our expectations since we expected practically no neutral answers for the 'IrSe *so*' sentences. Because in these sentences not only the content, but also *but* and *so* guide the reader towards the conclusion from *q*, it is surprising that so many neutral answers were provided. This might point out that children could experience a general feeling of incomprehension and therefore prefer the middle value on the scale. Therefore we investigated slightly older children, aged 10 to 12, in Experiment 2. Perhaps a more clear answer pattern might emerge in older children. After all, childhood

can be seen as a time where major changes are present in the development of different areas such as language, pragmatic- and logical understanding (Berk, 2010). We wondered whether there would be an age effect: will the older children in this experiment be more pragmatic than the younger children in Experiment 1 and will their use of the scale provide a clearer image of their understanding of *but*, *so* and *nevertheless*?

Experiment 2

Method

Participants, Materials and Procedure The 61 Dutch-speaking children who participated in this experiment were aged 10 to 12 with a mean age of 11.3. Two participants were excluded from the analyses due to missing data. The remaining children were 36 boys and 23 girls. They were students from the same schools as the children in Experiment 1 and were recruited from four different classes. All materials and the procedure were exactly the same as in Experiment 1.

Results

The results of the older children are similar to those of the younger children in Experiment 1. We inspected the distribution of the neutral answers to see if it was permitted to sum up the scores. The percentages of the number of neutral answers are displayed in Table 3, together with the percentages appropriate answers. As in Experiment 1 there was no significant difference in the number of neutral answers between the different categories (X^2=.66; *df*=2, *p*=.72). This allowed us to sum up the scores in Experiment 2 as well and perform the same analyses as in Experiment 1.
When we look at the results of the *so*-conclusions, the 10-to-12-year olds scored highest on the appropriateness scale of the 'IrSe' sentences (87%). This differed significantly from the 'SeSe' sentences (79%; Wilcoxon Signed Ranks test, *n*=45; *Z*=-3.83; *p*<.001) and from the 'SeIr' sentences (57%; Wilcoxon Signed Ranks test, *n*=56: *Z*=-6.13; *p*<.001). These last two categories also differed significantly from each other (Wilcoxon Signed Ranks test, *n*=52; *Z*=-5.65; *p*<.001).
When we look at the results of the *nevertheless*-conclusions, the same pattern emerges as in Experiment 1. The children scored highest on the appropriateness scale of the 'SeIr' sentences (78%). This differed significantly from the 'SeSe' sentences (67%; Wilcoxon Signed Ranks test, *n*=59; *Z*=-3.51; *p*<.001) and from the 'IrSe' sentences (49%; Wilcoxon Signed Ranks test, *n*=54; *Z* =-5.99; *p* <.001). These last two categories also differed significantly from each other (Wilcoxon Signed Ranks test, *n* =51; *Z* =-5.06; *p* <.001).
In contrast to Experiment 1, we did find a significant difference when we compared *so* with *nevertheless* for the 'SeSe' sentences (79% vs. 67% respectively; Wilcoxon Signed Ranks test, *n* =55; *Z* =-3.78; *p* <.001).
We performed Mann-Whitney U tests in order to explore the difference in performance between the two age groups. Generally, the older children provided more appropriate answers than the younger children but this difference was only significant for the 'SeSe *so*'– and the 'IrSe *so*' sentences. The results of these Mann-Whitney U tests are displayed in Table 4.

Table 3: Percentages of neutral and appropriate (between brackets) answers for each argument-conclusion combination (Experiment 2)

Sentence	So	Nevertheless
Sensible-Sensible	31(53)	32(35)
Sensible-Irrelevant	23(24)	21(57)
Irrelevant-Sensible	22(70)	19(14)

Table 4: Results of Mann-Whitney U tests and percentages for the comparison between the two age groups.

Sentence	*U*	*p*	8-10 years	10-12 years
SeSe_So	1365.5	.003	71	79
IrSe_So	1329.5	.002	79	87
SeIr_So	1778.5	.400	59	57
SeSe_Nevertheless	1783.5	.410	69	67
IrSe_Nevertheless	1616	.096	54	49
SeIr_Nevertheless	1676.5	.180	75	78

Se=sensible; Ir=irrelevant

General Discussion

This paper aimed to examine children's understanding of the conventional implicature stemming from *but*, *so* and *nevertheless* in 'p but q' sentences constructed as distancing-contrasts. Instead of using a binary judgment task as in Janssens et al. (in press), a three-point scale was used. The use of a ternary response format was inspired by Katsos and Bishop (2011) who provided evidence that binary judgment tasks can conceal children's pragmatic competence. In line with their Pragmatic Tolerance Hypothesis they showed that children are equally aware of pragmatic violations as adults but are more tolerant for these violations. We expected a three-point scale to shed light on children's understanding of conventional implicatures as well. The results of Janssens et al. (in press) had shown that children seem to have a general understanding of the pragmatic meaning of *but*, *so* and *nevertheless* but are very sensitive to the content of the arguments. The use of a three-point scale enables the children to answer with the middle value on the scale whenever they experience a conflict between the conclusion based on the content of the arguments and the conclusion based on the pragmatic meaning of the instruction-words. Two different age groups were examined: a group of 8-to-10-year-olds (Experiment 1) and a group of 10-to-12-year-olds (Experiment 2). The results of both age groups were similar. The percentages

appropriate answers seemed to be higher for the older children in most cases but this difference was only significant for the 'SeSe *so*'– and the 'IrSe *so*' sentences. More importantly, we made predictions based on Table 1 which allow us to gain insight into children's sensitivity to the conventional implicature from *but* on the one hand and the content of the arguments on the other hand.

The scale-data included evidence that children are aware of the pragmatic meaning of *but*, *so* and *nevertheless*. We found that the children in both experiments provided the neutral answer about one third of the time for the 'IrSe *nevertheless*'- and the 'SeIr *so*' sentences and even a considerable amount of appropriate answers. This is evidence that children are sensitive to the pragmatic meaning of *but*, *so* and *nevertheless*. If they would have been exclusively sensitive to the content then we would have expected almost 100% inappropriate answers. The answers on the 'SeIr *nevertheless*' sentences also indicate that children are sensitive to the pragmatic meaning of *but*. Both the content of the arguments and the use of *nevertheless* elicit the conclusion from *p*. However, *but* elicits the conclusion from *q* and this conclusion has to be cancelled in order to reach the appropriate conclusion. The fact that 30% (Experiment 1) and 21% (Experiment 2) neutral answers were given, suggests that sensitivity to the implicature from *but* causes doubt.

Apart from evidence showing that children are sensitive to the pragmatic meaning of the instruction-words, the scale-data also indicated that the content of the arguments has a lot of influence on children's answers. When both arguments are sensible, none of the two arguments outweighs the other. When these sentences are combined with *nevertheless*, then *but* and *nevertheless* lead to opposite conclusions. In both experiments, one third of the answers were inappropriate which indicates that the pragmatic meaning of *nevertheless* is not that easy to grasp. The neutral answers (25% in Experiment 1 and 32% in Experiment 2) are interpreted as evidence that the children notice the conflict between the conclusion based on *but* and the conclusion based on *nevertheless*. As a consequence, the results from the 'SeSe *nevertheless*' sentences seem to suggest that children generally understand the pragmatic meaning of *but* and *nevertheless* but this understanding is far from perfect. When two sensible arguments are combined with *so*, both *but* and *so* elicit the conclusion from *q*. This means that the neutral (31% in both experiments) as well as the inappropriate answers (28% in Experiment 1 and 16% in Experiment 2) are evidence that the content of the arguments is very important for children and sometimes outweighs the answer based on the pragmatic meaning of *but* and *so*.

In contrast to the results showing that children are sensitive to the pragmatic meaning of *but*, *so* and *nevertheless* as well as to the content of the arguments, the results of the 'IrSe *so*' sentences pose a bigger problem to interpret. Both the content and the instruction-words elicit the conclusion from *q* so we would have expected almost exclusively appropriate answers. The fact that 'only' 55% (Experiment 1) and 70% (Experiment 2) appropriate answers were provided suggest that children's performance on these 'p but q' sentences is far from optimal and the use of the middle answer on the scale might rather express a certain general incomprehension. So, this latter finding causes us to interpret our scale-data cautiously. Further research with the use of scales on conventional implicatures seems necessary. It might be useful to apply this scale-format on adults. This would allow us to compare their responses with children's responses. Consequently we could get a clearer view on how to interpret these results.

Acknowledgments

We would like to thank all schools who participated for allowing us to perform our experiments. Leen Janssens is a research assistant from the Fund for Scientific Research Flanders (FWO). This research was carried out with the financial support of the National Council for Scientific Research – Flanders, Belgium (FWO grant G.0634.09)

References

Anscombre, J., & Ducrot, O. (1977). Deux 'mais' en français? *Lingua, 43*, 23–40.

Berk, L.E. (2010). *Development through the lifespan* (5th ed.). Boston: Allyn and Bacon.

Braine, M., & Rumain, B. (1981). Development of comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology, 31*(1), 46-70.

Grice, P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & van den Toorn, M. C. (1997). *Algemene Nederlandse Spraakkunst*. Groningen/Leuven: Wolters Noordhoff.

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.

Janssens, L., Drooghmans, S., & Schaeken W. (in press). But: do age and working memory influence conventional implicature processing? *Journal of Child Language.*

Katsos, N. & Bishop, D.V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, *120*, 67-81.

Noveck, I.A. (2001). When children are more logical than adults: experimental investigations of scalar implicature. *Cognition, 78*(2), 165-188.

Van Belle, W. (2003). *Zwijgen is niet altijd toestemmen. De rol van inferenties bij het interpreteren en argumenteren.* Leuven: Uitgeverij Acco.

Van Belle, W., & Devroy, G. (1992). *Tegenstellende en toegevende connectoren. Een argumentatieve beschrijving* (Preprint 143, voorlopige publicatie). Katholieke Universiteit Leuven, Faculteit Letteren, Departement Linguïstiek.