CONSERVATIVITY AND QUANTIFICATION

DO DUTCH CHILDREN OBEY CONSERVATIVITY IN A PICTURE-VERIFICATION TASK?

Bachelor-Thesis

Anna de Koster, s1975315, a.m.b.de.koster@student.rug.nl,
Advisors: Jennifer Spenader, Bart Hollebrandse

Abstract: Conservativity is a universal property of natural language determiners. Because of this universality it is possible that young children can apply it to quantifier interpretations. But do we have to learn conservativity or is it part of the innate structure of language? Using a picture verification task with sentences using the conservative determiner *al* (all) and the non-conservative quantificational adverb *alleen* (only), we tested 16 children between the ages of 4-6. Previous research has shown that *al* is more difficult for children in contrary to *alleen* with example sentences like *'Al de konijnen dansen'* and *'Alleen de konijnen dansen'*. However, in this previous experiment we found some peculiarities with the determiner *alleen.* To examine these peculiarities we did a variation on the previous experiment. As a novel element, rather than using existing animals we used non-existing monsters for the pictures, to rule out the relation with the real world. We also left out the determiner in the target-sentences, for example: *'Alleen de konijnen dansen'* became *'Alleen konijnen dansen'*. We found in this follow-up experiment that sentences with *alle* (all) are difficult for children to verify, but sentences with *alleen* (only) also are difficult. The removal of the definite article didn't influence the results, and children instead seem to be fairly consistent. The results neither show that children follow conservativity, nor that they do not follow conservativity and demand a more complex explanation.

1. Introduction

1.1 Conservativity

In Generalized Quantifier Theory (Barwise and Cooper, 1981; Mostowksi, 1957) quantificational determiners describe a relationship between two different sets. *All* and *Some* are two examples of those determiners.

(1) All bunnies are dancing.
(2) Some bunnies are dancing.

Sentences (1) and (2) both refer to two sets, the set of bunnies and the set of dancers, which we will call Set A and Set B. The relationship that holds between these two sets is given by the quantifier.

Set A is always evoked by the N in the subject NP, e.g. *bunnies*. This means that this set consists of bunnies. Set B arises from the VP and contains 'dancing things'. Sentence (1) is true iff all members of Set A also are members of Set B. Sentence (2) is true iff at least one member of Set A is also a member of Set B.

Conservativity means that for Sentence (1) you only need to determine if all the members of Set A aren't within the intersection of Set A and B, to verify if the sentence is true. Members of Set B that are not in the intersection of Set A and Set B are irrelevant to verifying the truth of the sentence. For Sentence (2), you only have to make sure the intersection between Set A and Set B is not empty (has at least one member). In other words, to verify Sentences (1) and (2), we only have to look at bunnies and dancing bunnies. Things that are dancing and aren't bunnies can be ignored. Members of Set B that are not in the intersection of Set A and B are irrelevant. As you can see in Figure 1, the dancing panda does not affect the truth-value of Sentence (1).

Figure 1. For Sentence (1), Set A is the group of three bunnies and Set B is the group of dancing things (the three bunnies and the panda). The intersection of Set A and Set B only includes the dancing bunnies.

Conservativity is a property that is shared by all natural language determiners. Several investigations of determiners in numerous languages have not yet found a convincing example of a non-conservative determiner (Barwise and Cooper, 1981). So we can state that conservativity seems to be a universal property of all natural language determiners.

1.2 Spreading

When looking at previous work on children's interpretation of quantificational determiners, particularly relevant is previous research on the well-known phenomenon called *spreading*. Spreading refers to the fact that children, up to the age of 7;0, frequently consider members of Set B that are irrelevant for determining the truth-value of a sentence. Roeper and De Villiers (1993), state in their paper that children show evidence of bound variable readings. They tend to overgeneralize the bound variable structure to contexts where it is unnecessary. Roeper and De Villiers used Sentence (3) as an example.

(3) Every child sat on a horse.

Because of the phenomenon spreading, children will reject Sentence (3), when they are shown a picture with three children on a horse and one extra horse without a child on it. To justify the response, the children would refer to the fact that there is an extra horse without a child. Children include the extra horse in their verification when it is actually irrelevant. This is called spreading, because the determiners 'spreads' its influence over both the subject and the object in the sentence.

Roeper and De Villiers (1993) argue that these (false) interpretations seem to depend on a cognitive achievement; the ability to make pairings or construct isomorphic correspondences. They also state that the bound variable reading consists in more than just correspondences provided by the non-verbal context, because for the adult grammar, syntax places significant constraints on the bound variable reading.

They gave a linguistic explanation about spreading using Sentence (3). They proposed that the quantifier had scope over the entire clause instead of only the noun immediately following the quantifier. They assumed that the quantifier wasn't only attached to the subject noun (Sentence (4)), but that it was attachted to the sentence as a whole (Sentence (5)). By scoping over the entire clause, both subject and object nouns could be quantified over. This could then lead to false interpretations, because the children use the object of Sentence (3), a horse, also in their interpretation, when it actually is irrelevant.

(4) [CP[Every child] sat on [a horse].]

(5) [CP Every [child] sat on [a horse].]

Adults can eliminate these bound variable readings and thereby choosing the right interpretation of Sentence (4). In contrary to the interpretations of young children. Young children can't see the difference between Sentences (4) and (5), which will lead to bound variable readings and thus to the false interpretation of Sentence (5)

1.3 Non-Conservative quantifiers

Another possible explanation for spreading is that children (up to the age of 7 years old) treat the conservative determiner 'all' as non-conservative. They refer to both set A and set B, when verifying the truth of the sentence. When we look at Sentence (1) again in combination with Figure 1, spreading would mean that children would also consider the dancing panda in their verification. This is also exactly what you would predict if children are non-conservative.

This paper considers both the interpretations of *al* (all) and *alleen* (only). *Al* is a conservative natural language determiner, but not all expressions in natural language describing two sets are conservative. The quantificational adverb only is such an example. To verify a sentence with the adverb only, one must consider the members of Set B which are not in the intersection of Set A en Set B. When such members exist, the truth-value of the sentence becomes false.

(6) Only bunnies are dancing.

For example, for Sentence (6) to be true in the situation depicted in Figure 1, there should be no members in Set B (the dancers) which are not in the intersection of Set A and Set B (set of dancing bunnies). The presence of the dancing panda

makes the sentence false, because it is another individual that is dancing (Set B).
We expect that children are able to correctly verify sentences and pictures with *only*, if they do not proceed from the assumption that determiners are conservative and instead assume that both sets must be considered. In fact, we predict that children will be better at *only* than at *all*. On the other hand if conservativity is innate, we predict that children will correctly verify sentences with *all* and that they will incorrectly verify sentences with *only.*

In section 1.1 we stated that all natural language determiners are conservative. So why isn't *only* conservative? When we compare the adverb *only* to true determiners, there is a difference. *Only* can appear in a number of different syntactic positions. It also associates with focus, which leads to focus-dependent interpretations. Because of this, the adverb *only* can't be considered as a counterexample to the claim that conservativity is a universal property of natural language determiners. But if *only* is in a sentence-initial position, it acts very similarly to a true determiner and functions semantically like a generalized quantifier (de Mey, 1991). Another similarity between a true determiner and *only* is the fact that the verification of sentences with *only* requires considering both Set A and Set B. Unlike natural language determiners, like *all* and *some*, *only* is non-conservative. To determine the truth-value of a sentence starting with *only* one most examine members of Set B that are not in the intersection of Set A en Set B.

Previous research about non-conservative determiners brings us to Hunter and Litz (2012). They tested the learnability of non-conservative determiners in children. In their experiment, 20 children (aged 4;5-5;6) were taught the meaning of a new determiner, *gleeb*. Half the children were taught a conservative interpretation of *gleeb* (similar to *Not every X is Y)* and half the children were taught a non-conservative interpretation (similar to *More than just X are Y)*. The results showed that children trained on the conservative variant performed significantly better than chance (χ^2=74.160, d*f*=5, p<0.0001), compared to the children trained on the non-conservative variant that weren't significantly better than chance (χ^2=6.640, d*f*=5, p>0.2488). However, 80% of the children that learned the non-conservative determiner performed better than chance, compared to 90% of the children that learned the conservative determiner. Looking at it in this way, the results seem more similar. So it is still quite comparable.
These results seem to show that conservative determiners are easier to learn than non-conservative determiners, but non-conservative determiners still are learnable in a very short training period. Why are conservative determiners easier to learn? Is it because they are simpler to verify, because the children only have to look at Set A? Or is it because of the children's experience with conservativity? Did this experience made it easier for them to assign a conservative meaning to the determiner? These questions remain unanswered in Hunter and Litz's paper. Our study seeks to extend our knowledge of children's understanding of conservativity.

1.4 Research Question

The fact that conservativity is seen as a universal property of natural language determiners evokes a major question: How does this property relate to the learnability of those determiners? Is conservativity part of the innate structure of language or should it be learned? In other words, will children interpret determiners conservatively from a young age and thereby have trouble with non-conservative interpretations? Or aren't children conservative and will they have trouble with understanding conservative determiners?
Summarized, in the rest of this article we will investigate the assumption that children initially interpret determiners conservatively and have trouble with non-conservativity. We will compare children's interpretations of the conservative quantificational determiner *al* ('all') with their interpretations of the non-conservative quantificational adverb *alleen* ('only').

If we see conservativity as a universal property of quantificational determiners, than the expectation will be that conservativity leads interpretations from an early age on. Thus, the interpretation of the conservative determiner *all* is expected to be mastered from an early age on, whereas the interpretation of *only* is expected to be mastered difficultly for children. But, if conservativity isn't innate, the expectation will be that sentences with *all* are difficult for children, contrary to *only* which then will be easier to understand.

2. Previous Experiment

The current experiment is an extension of a previous experiment testing the same research questions (Hollebrandse et al, 2013, ms).

2.1 Participants

Forty-five Dutch children between the ages of 4 and 6 (mean age 5;4; range: 4;4-6;3) participated in the experiment. They were tested individually at their schools by one experimenter and an assistant.

2.2 Materials and design

In this previous research a sentence-picture verification task was used in a 2x2 design. This 2x2 design consisted of the factors PICTURE (Set A- and Set A+) and QUANTIFIER (*alle* 'all' versus *alleen* 'only').

2.2.1 Picture

There were two types of pictures in this experiment: Set A+ and Set A- pictures.

Figure 2a. Set A+ picture with dancing bunnies and a dancing panda.

Figure 2b. Set A- picture with two dancing bunnies, one not-dancing bunny and one not-dancing panda.

Figure 2a is an example of a set A+ picture. Set A+ pictures show all Set A characters (bunnies) performing the action denoted by the intransitive verb (dancing), plus an additional character (panda) performing the same action. Figure 2b is an example of a Set A- picture. These pictures show only some of the members of Set A (bunnies) performing the action denoted by the intransitive verb (dancing). Crucially, one member of Set A is not performing this action. In Figure 2b there is one bunny that isn't dancing. Note that this picture also contains another character that is not performing the action denoted by the intransitive verb, to maximize the similarity between the two picture types. Every picture contains two types of monsters.

2.2.2 Quantifier

The used sentences were of the form *all A B* (universally quantified Sentence 7a), or *only A B* (Sentence 7b with the quantificational adverb *only).* All test sentences used intransitive verbs.

(7) a. Al de konijnen dansen.
"All the bunnies are dancing."

b. Alleen de konijnen dansen.
"Only the bunnies are dancing."

2.3 Results

The results in this previous experiment confirmed our hypothesis about the sentences with *all*, but there were some peculiarities in the results of the sentences with *only.*

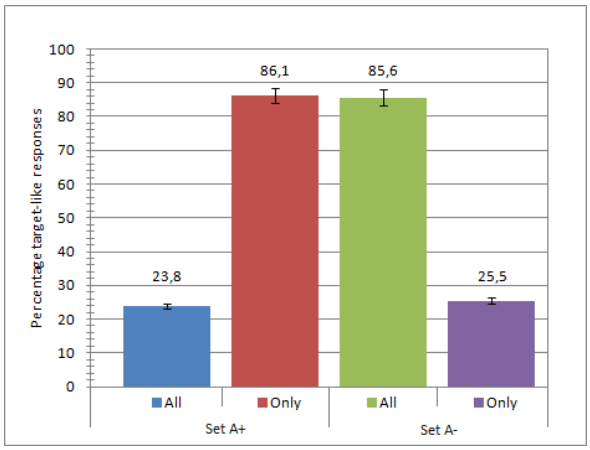

Figure 3. Results of the previous experiment: Accuracy as proportion of target responses (error bars indicate standard error).

Figure 3 gives the results. For Set A+ pictures with the quantifier *all,* was the percentage target-like responses significantly lower (according to a mixed effect linear model) than the Set A+ pictures with the adverb *only*, following our hypothesis. But, when looking at Set A-, it becomes clear that the percentage target-like responses for sentences with *only* are significantly lower than the sentences with *all*, unlike our expectation.

We expected the *only* sentences to be easier than the *all* sentences and this should result in a higher percentage for the *only* sentences in comparison to the *all* sentences, but this is not the case.
We think this difference is due to some difficulties we experienced. These difficulties and the changes in the method and materials we made due to these problems will be deepened and explained in the following section.

3. Follow-up Experiment

3.1 Participants

Sixteen Dutch children between the ages of 4 and 6 (mean age 5;5; range: 4;5-6;11) participated in the experiment. They were tested individually by the author and an assistant in their homes.

3.2 Materials and design

We again used a sentence-picture verification task in a 2x2 design with each of the participants. The 2x2 design consisted of the factors PICTURE (set A+ versus set A-) and QUANTIFIER (*alle* 'all' versus *alleen* 'only').

3.2.1 Picture

In this follow-up experiment we chose to use 21 non-existing monsters as objects in the pictures instead of existing animals as was done in the previous experiment, because of the fact that children have a wild imagination and make connections with the real world. In the previous experiment we used existing animals like the bunnies and the panda in Figures 2a and 2b, but we noticed that young children invent stories about the animals and think about the animals existing in our world.
In the previous experiment we used the picture in Figure 2a, belonging to Sentences (7a and b).
Some children would give the answer '*no*' to these sentences without even looking at the pictures. After we asked why they said no, they explained to us that they also had a bunny at home and their bunny couldn't dance. So the sentence must be wrong. This problem led to the usage of non-existing monsters. As previously mentioned we used 21 different monsters that we created ourselves to rule out any relation with any world. Not only the real world, but we also chose not to use existing characters for example Pokémons, because children would know too much about them and their characteristics. We think it is necessary that the children know little to nothing about the characters to avoid the influence on the interpretation of the sentences.
We also chose 21 non-existing names (like Kroepie, Noppie, Hopje etc.) for the same reasons as mentioned above.

Figure 4a and 4b show the new A+ and A- pictures we created with the monsters.

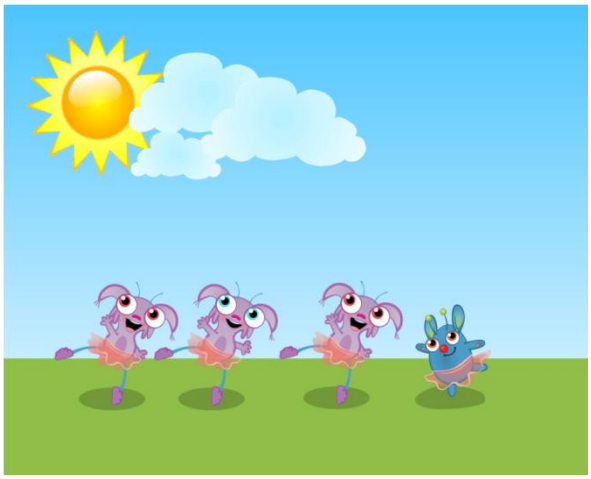

Figure 4a. Set A+ picture, with dancing Kroepies and a dancing Loempa.

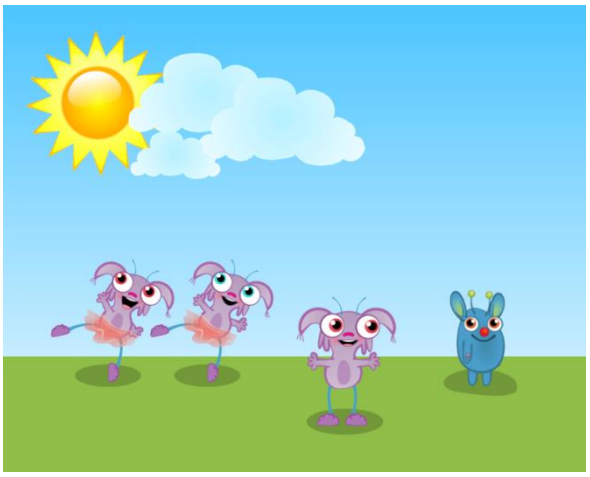

Figure 4b. Set A- picture, with two dancing Kroepies, one not-dancing Kroepie and one not-dancing Loempa.

To avoid any other type of influence, we also chose to use night and day pictures. Every monster is once in Set A (Figure 5a) and once the additional character (Figure 5b). So every monster occurs twice with the exception of two monsters, they only occur once as additional character.

Figure 5a. Day A+ picture: The yellow monsters are Set A and performing the action calling.

Figure 5b. Night A+ picture: The yellow monster is the additional character.

The main point of this distinction is that the monsters only perform one action in daytime and one other action at nighttime. There is a very obvious difference, so the children can't confuse the two actions. It is a precaution for the fact that the children could say: No the Kroepies are not dancing, because they are also drawing. With the clear extinction between night and day, we hope to avoid this kind of mix-ups.

We also have ensured that the two monsters in one picture are very different. They are easily distinguishable by color, size and other characteristics. The names of the monsters in one picture are also very different.

3.2.2 Quantifier

The sentences we used were of the form *all A B* (universally quantified Sentence 8a), or *only A B* (Sentence 8b with the quantificational adverb *only*). All test sentences used intransitive verbs and monster-names instead of animals.

(8) a. Alle Kroepies dansen.
"All Kroepies are dancing."

b. Alleen Kroepies dansen.
"Only Kroepies are dancing."

In the previous experiment the sentences were a bit different (returning to Sentence 7a and 7b). They were of the form *al de* 'all the' and *alleen de* 'only the', with the quantificational determiner or adverb being followed by a determiner.

(7) a. Al **de** konijnen dansen.
"All the bunnies are dancing."
b. Alleen **de** konijnen dansen.
"Only the bunnies are dancing."

This was to restrict interpretation to the situation shown in the picture and to avoid any interference with the real world. However, adding the determiner wasn't very helpful and led to other problems.
We noticed that putting the determiner in place like Sentence (7a and 7b), caused problems with the Set A- picture (Figure 2b) in combination with Sentence (7b). When verifying Sentence (7b) in combination with the Set A- picture, looking to just Set A (the two bunnies that are dancing) and to other things that are dancing is sufficient. So the bunny in Figure 2b that isn't dancing is irrelevant for validation, but because of the extra determiner *alleen* ***de***, it seems to be that all bunnies have to dance to make the sentence true. This problem will be avoided, when the determiner is left out and you just ask to verify Sentence (8b). In this way the emphasis on the Set A character that isn't performing the action will be less and people won't get distracted by it.

The restriction to the situation shown in the picture and the avoidance of interference with the real world are also gone, when the determiner is left out. We solved this consequence by using monsters and night and day pictures, as explained in section 3.2.1.

3.3 Design of the experiment

Using a laptop, participants were asked to verify the two types of test-sentences like Sentences (8a and 8b) with the two different types of pictures A- and A+ (Figure 4a and 4b). This resulted in four possible combinations (Table 1).

Table 1. All four possible conditions, with their truth-values

Quantifier/Picture	A+	A-
All	True	False
Only	False	True

We included 4 test-items of each condition, so 16 in total, extended with 8 filler items that were used to see if the children were following the experiment. The test session started with 3 practice-items. After each item, we asked the children to justify their responses, for the same reason as the usage of the fillers. The participants were presented with one picture at a time while a recorded sentence was played. They were instructed to verify whether the sentence matched the picture.

If children proceed from the assumption that all determiners are conservative, we expect them to give target responses for *all*-sentences with Set A+ pictures. If they apply the same assumption to sentences with *only*, we expect these children to also accept these sentences with Set A+ pictures; this should result in non-target responses.

But on the other hand, if conservativity isn't part of the innate structure of language and children will have to learn that determiners are conservative, we expect them to make errors with *all*-sentences. In particular, the *all*-sentences in combination with the Set A+ pictures. Because of the presence of the additional character (see spreading, section 1.2). We also expect them to correctly reject *only*-sentences in this case.

In the Set A- pictures there is one monster too few performing the relevant action. So these pictures test the exhaustivity requirement of *all.* If children master this requirement, they are expected to reject *all*-sentences with Set A- pictures. For example, if one Kroepie isn't dancing, children should answer that the statement that all Kroepies are dancing is false.

Finally, if the adjustments we made after the previous experiment have not solved the problem with *only* in combination with A-, we expect that the children would still incorrectly reject *only*-sentences for A- pictures. And thereby we expect that they don't make a distinction between the determiner *all* and the adverb *only*. They would treat *all* like *only* or vice versa.

4. Results

Each of the participants was asked to verify two types of sentences with two types of pictures. Figure 6 gives the results.

In combination with Set A+ pictures (Figure 4a), children incorrectly rejected *all*-sentences, as expected, but they also correctly rejected the *only*-sentences as expected if children are not treating *all* and *only* as conservative. For the Set A- pictures (Figure 4b) they correctly rejected *all*-sentences, but incorrectly rejected *only*-sentences. These results are just like the results of the previous experiment.

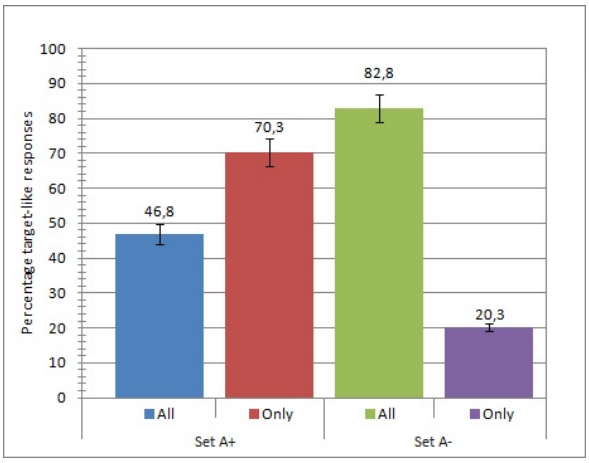

Figure 6. Accuracy as proportion of target responses (error bars indicate standard error).

We analysed the results using mixed effect linear models, using PICTURE and QUANTIFIER as predictors, and target interpretation as the response variable. We found a significant effect for both QUANTIFIER and PICTURE and also a significant interaction (Table 2). All noticeable differences in Figure 6 turned out to be significant.

Whereas adults would accept *all* and reject *only* with Set A+ pictures, the children significantly rejected *all* incorrectly and rejected *only* correctly with the Set A+ pictures. For Set A- pictures the children significantly rejected *all* correctly, but incorrectly rejected *only*. Whereas adults would accept *only* in combination with Set A- pictures. This shows us that the children overwhelmingly rejected all four conditions.

Table 2. Fixed effects of the maximally best fitting logistic mixed-effects model
Formula
Response ~quantifier * picture + (1 | participant)

Predictor		Estimate	SE	z-Value	p-Value
Intercept	**	0.8622	0.2736	3.151	0.00162
Quantifier	**	-0.9874	0.3709	-2.662	0.00777
Picture	***	-2.2291	0.4140	-5.384	7.27e-08
Quantifier:Picture	***	3.9265	0.5864	6.696	2.15e-11

5. Discussion

5.1 Set A+

The results of the children's responses with the Set A+ pictures in combination with both *all* and *only* suggests that children interpret the determiner *all* non-conservatively. Their responses on the sentences with the adverb *only* were target-like, which requires a non-conservative interpretation. So for the results of the sentences in combination with the Set A+ pictures, we see both non-conservative interpretations.

In section 1.3 we gave a possible explanation for the phenomenon spreading: Another possible explanation for spreading is that children (up to the age of 7 years old) treat the conservative determiner 'All' as non-conservative. They refer to both set A and set B, when verifying the truth of the sentence.

The interpretations of *all* in combination with the Set A+ pictures turned out to be non-conservative as expected. But did the children show spreading in this condition? During the experiment we noticed that almost all children referred to the member of Set B, while this member actually is irrelevant in this condition. They rejected the sentence, because they included the extra member in their verification. This indicates spreading, because the children interpret the determiner as if it spreads over both the subject (Set A) and the object (Set B).

5.2 Set A-

The responses with the Set A- pictures are not that explicit and show more complexity. When looking at the Set A- pictures in combination with *all*, we can conclude that the children correctly interpreted the determiner and the exhaustivity requirement.
However, this does not necessarily mean that they treated *all* as conservative. Actually, this particular condition cannot be used as evidence for or against the hypothesis that the children treat *all* as conservative. This is because the Set A- pictures in our study did not contain any members of Set B that are not in the intersection of Set A and Set B. So in Figure 4b, there is no other monster dancing than a Kroepie (Set A). The blue Loempa in this picture is just an additional character to make the two types of pictures more similar and does not belong to Set B. Hence, the Set A- pictures do not have any influence on the issue of conservativity in contrast to the Set A+ pictures.

So the condition of *all* in combination with Set A- pictures doesn't give us any information concerning conservative or non-conservative interpretations. On the other hand, the condition of *only* in combination with Set A- pictures does give us some information about conservativity.

When looking at this particular condition with *only* and Set A- pictures (Figure 6, rightmost bar), it is clear that the children incorrectly rejected this condition in contrast to our expectation. We expected the *only*-sentences to be easier to interpret than the *all*-sentences. The children should interpret these sentences non-conservatively, but the justifications the children gave us show us that they treat *only* like *all*. Almost all the children explained to us that this sentence was incorrect, because there was one monster that was not dancing. They were referring to the monster of Set A that is not performing the action, but this monster is

actually irrelevant for validation. The children were validating these *only*-sentences with the interpretation of the determiner *all.*

In their interpretations and explanations they did not seem interested in members of set B, they were just focusing on all the members of set A, including the one that isn't relevant. For the validation of this sentence in combination with Set A- pictures the children only have to look at the intersection of Set A and Set B and just Set B. In this picture it is clear that there are no members of Set B. In other words there are no other 'things' that are performing the action except for Set A members. Even though they didn't need to look at the monster of Set A that isn't performing the action, they still did. So they were not interested in Set B and the additional character which indicates conservative interpretations, but they still gave incorrect answers. Is there a possible explanation for these strange errors?

We noticed that *only* in combination with Set A- is the only condition where it is actually needed to know what Set A is to interpret the sentence, and this conditions turned out to be the condition where the children made the most errors.

Looking at the other three conditions; even if the names of the characters are not known, it is possible to say whether or not the sentence is true or false. For example: Set A+ picture (Figure 4a) with *only*. Consider this sentence: Only Kroepies are dancing; it can immediately be seen that there is more than one monster dancing. So it is possible to immediately reject the sentence, without even knowing the names. This is the case in all conditions; except for the Set A- condition in combination with *only.* In this condition it is necessary to know which monster is which. This fact could explain why this condition is so difficult for children to interpret.

However, in the previous experiment we used existing animals that were familiar to the children. This would mean that they knew the names of the animals. But, the results still showed that they incorrectly rejected this condition despite of the fact that they knew the names.

An informal survey of native speakers from the previous experiment confirms that even adults find this sentence in combination with the A- picture slightly odd. However, they still consistently judged the sentence as correctly matching the picture.

5.3 Possible negative effects of materials

In this follow-up experiment we chose to use non-existing monsters with non-existing names, to rule out the relationship with the real world (knowledge). We considered the fact that the names of all 21 monsters were too difficult to remember for the children. We tried to solve this by adding pictures of the monsters before each of the target-pictures and rehearsing the names with the children before each trial. During the experiment we didn't notice that the children had a problem with memorizing the monsters and recalling them.

However when looking at the Set A- pictures in combination with *only* it is not excluded that those errors are due to the memorization of the monsters. Especially for the fact that this condition is the only condition in which remembering the monster's names is crucial.

Another issue that came up was the Set A+ picture in combination with *all*, in which four monsters are presented: three of Set A performing the action and one member of Set B also performing the action. Another possibility is an Set A+ picture with three monsters of Set A performing the action, but one additional monster that is not performing the action. Returning to Figure 4a, this would result in a picture with three dancing Kroepies and one NOT-dancing Loempa. This additional character that is not performing the action can also show if children use non-conservative interpretations, even though it isn't a member of Set B but just an additional character. This is because of the fact that children that use non-conservative interpretations are expected to include that character in their verification. We predicted that using a real member of Set B (so a different monster that is performing the action denoted by the VP) would rather result in non-conservative interpretations. S, that's why we chose to use members of Set B. However, we do expect that using an additional character instead of a real member of Set B would also result in non-conservative interpretations, because it still is an additional member that children could include in their verification.

5.4 Comparison with Hunter and Lidz

Looking back to and comparing Hunter and Lidz's (2012) results, we found that children had

trouble making conservative interpretations even for quantificational expressions with which they surely must have had extensive experience (which remains unclear in Hunter and Lidz's paper). This seems to contradict Hunter and Lidz's results that non-conservative determiners are harder to learn than conservative determiners. They also state that, if non-conservative natural language determiners are difficult to learn, then natural language semantic theories should rule out non-conservative relations.

Given our results we disagree with this argument. Even if all quantificational determiners are conservative, the most natural way to model the meaning of *only* is as a relationship between two sets, like non-conservative interpretations.

Also, if non-conservative quantifiers like *only* may not be lexicalized as true determiners, they are frequently used to express non-conservative meanings. These non-conservative meanings are easily interpretable by adults.

6. Conclusion and further research

Thus, the new results of the follow-up experiment show the same peculiarities as in the previous experiment (Hollebrandse et al.), which we tried to solve by leaving out the definite determiner (as explained in section 3.2.2), using non-existing monsters and night and day pictures.

Children still make the same errors with *only* in test sentences without the determiner. These errors plus the justifications of the children might suggest that they confuse *alleen* 'only' with *al* 'alleen'. According to these results, we can conclude that the adjustment of the determiner hasn't helped and that we have to dive deeper in the investigation of the adverb *only* in terms of non-conservativity. Children did show non-conservative interpretations and signs of spreading with the Set A+ pictures in combination with the conservative determiner *all.* They treat *all* non-conservative and this indicates that conservativity isn't part of the innate structure of language. However, our results are not sufficient to conclude that conservativity must be learned. Further research is necessary.

As mentioned in the previous section the memorization of the monsters could be a problem. So, in further research one can think about a new way to learn the children unknown monsters and names and thereby a new way to check if the children did remember the names. In that case it will become clear if the errors with the Set A- pictures in combination with *only* were due to the fact that the children had trouble with memorizing the names. We certainly want to recommend the usage of monsters instead of existing characters, because children have a very wild imagination and in this way their attention is kept on the situation shown in the picture.

Further research with other Set A+ and Set A- pictures could also give new insights about conservativity. In the previous section we mentioned that maybe an Set A+ picture with an additional character that is not performing the action could also lead to non-conservative interpretations. This would be additional evidence to the view that conservativity must be learned.

Our experiment showed that children definitely used non-conservative interpretations, instead of a required conservative interpretation. However, our results do not allow us to definitely answer the key question of whether or not conservativity is learned or innate. Clearly, further research is required.

References

Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy* (Vol. 4: 159-219).

Chomsky, N. (1986), Barriers, MIT Press, Cambridge.

Hollebrandse, B. et al. (2013). Dutch children's interpretation of quantificational determiners: Must the universal property of conservativity be learned?, ms.

Hunter, T. and Lidz, J. (2012). Conservativity and Learnability of Determiners. *Journal of Semantics,* Advanced Access.

de Mey, S. (1991). 'Only' as a Determiner and as a Generalized Quantifier. *Journal of Semantics* (8(1-2): 91-106).

Mostowksi, A. (1957). On a generalization of quantifiers. *Fundamenta mathematicae* (44:12-36).

Roeper, T. and de Villiers, J. (1993). The emergence of bound variable structures. In: T. Maxfield and B. Plunkett (Eds). *Papers in the acquisition of WH.* Amherst, University of Massachusetts, GLSA.