DUTCH CHILDREN’S INTERPRETATION OF DEFINITES AND INDEFINITES IN DISCOURSE

ANGELIEK VAN HOUT, LIANNE DE REE, MARLIES DE REE

1. Introduction

It is well-known that children make mistakes with article choice, specifically, they incorrectly produce definite NPs instead of indefinite NPs when referring to new referents in a discourse (Karmiloff-Smith, 1979; Maratsos, 1974). The explanations for the overuse of the have postulated cognitive, semantic and pragmatic problems. None of the explanations, however, address the optionality of the and a in children’s language—children alternately use the and a in certain discourse contexts. Nor do they carry over to indefinites. This paper contributes new data on the comprehension of definite and indefinite noun phrases in Dutch. We explain preschoolers’ non-target-like interpretation of indefinites as a failure to simultaneously take speaker and hearer perspectives into account, leading to a failure to draw scalar implicatures. We model our explanation in bidirectional Optimality Theory.

2. Semantics-Pragmatics of Definite-Indefinite NPs

Heim (1982) and Kamp (1981) posit a semantic difference between definite and indefinite NPs: indefinites involve existential quantification while definites are referential. Definite NPs carry the presuppositions of existence and uniqueness and typically refer to referents previously introduced in the domain of discourse which are part of the common ground, (1a). If no existing referent is available, a new referent is introduced that must be “bridged” to an existing one, (1b) (Clark, 1977). Indefinites lack both presuppositions; they typically introduce new referents in the domain of discourse. If only the speaker, but not the hearer knows the referent, the uniqueness presupposition is not warranted, and a
(specific) indefinite must be used, (1c). If neither speaker nor hearer knows the referent, the existence presupposition is not warranted and a (non-specific) indefinite must be used, (1d).

(1) a. A bird and a cat were sitting in a tree. The bird flew away.
   b. A tree is standing by the river. The leaves are orange.
   c. [Do you have any pets?] Yes, a cat.
   d. Do you know a gas station around here?

Horn (2006) proposes that definites carry uniqueness as a conventional implicature and argues that “while definites conventionally signal uniqueness (…), indefinites only conversationally implicate non-uniqueness” (p.18). For Horn articles form a scalar pair <the, a>, with the as the strong element and a as the weak one, which leads to a scalar implicature on which “the use of a weaker value (...) a suggests that, for all the speaker knows, no stronger value on the same scale (...) the) could have been substituted” (p. 18).

3. Acquisition of Articles

Preschoolers make mistakes with article choice. Maratsos (1974) targets elicitation of indefinite NPs: A lady had lots of boys and girls. She told them to be quiet. After she went to bed, one of them started to laugh and make noise. Who was making noise? (Target: a boy/girl). Children overused the by 58%. When the target was the (A lady had a boy and a girl, etc.), the children correctly provided it.

Maratsos (1974) attributes the overuse to egocentrism; children do not care about the referential knowledge of their listener. Karmiloff-Smith (1979) argues that children use definites deictically, instead of referring to “current context” and the listener’s knowledge of it. Wexler (2003) argues that children initially misconstrue the semantics of definite articles; they lack the uniqueness presupposition. In the “making noise” story, children infer there exists a child who makes noise, and so they answer the boy (girl). Schaeffer and Mathewson (2005) attribute children’s overuse of the to lack of a pragmatic principle which states that speaker and hearer assumptions are always independent, and thus in principle different. Children who do not make a distinction between speaker and hearer assumptions, cannot properly establish speaker and hearer belief states and make mistakes with article choice.

The explanations miss one important generalization. Any explanation that posits a cognitive deficit, lack of a pragmatic principle or semantic
property, predicts not just overuse of *the*, but vast overuse to the exclusion of *a* all together, contrary to fact. Children overproduce *the* to some extent (between 17-58% across studies), next to many target-like instances of *a*. Studies on children’s article use so far have focused on either production or comprehension, or they have looked at either definite or indefinite NPs. We now present two studies that systematically compare the interpretation of definite and indefinite NPs in a discourse.

4. How do Children Interpret Definite and Indefinite NPs?

Do children associate definites with mentioned referents? Do they introduce new referents for indefinites? One experiment tested anaphoric associations of definite and indefinite NPs with an antecedent in previous discourse (section 4.1); the other experiment tested bridging associations with an antecedent in previous discourse (section 4.2).

4.1 Mentioned referent experiment

Nineteen Dutch learners participated (mean age 4;9, range 4;1-5;4). We used a truth-value judgment task giving a two-sentence description of a picture with two similar characters. The first sentence established one of the two characters as the discourse topic; the second sentence used the same noun, varying definiteness. We varied who carried the property expressed by the second sentence, either the referent mentioned in the first sentence or the non-mentioned referent. There were two variables: NP (definite/indefinite) and Referent (mentioned/ non-mentioned), yielding four conditions, with three items for each condition. The different conditions and target answers are illustrated in figures 1-4. The pictures were colored; here L (=left) and R (=right) indicate the targeted character.

Fig. 1: Mentioned—Definite

![A mouse is kicking a ball and the mouse has a long tail](image1)

YES

Fig. 2: Mentioned—Indefinite

![A girl is wearing a blue dress (L) and a girl is eating an apple](image2)

YES
Fig. 3: Non-mentioned—Definite  Fig. 4: Non-mentioned—Indefinite

A boy is holding a flag and
the boy is carrying a ball  NO  A man is holding an umbrella and
a man (R) is wearing a red coat  YES

The results of the Mentioned Referent experiment are plotted in figure 5 as percent yes-answers (i.e., acceptance).

Fig. 5: Results Mentioned Referent experiment

A multiple ANOVA reveals a main effect of referent ($p<.0001$) and an interaction effect of article by referent ($p=.001$), but not of article ($p=.12$). Paired-samples t-tests show an effect of article in the non-mentioned referent conditions ($t=4.400$, $p<.0001$), but not in the mentioned referent conditions. They also showed effects for referent (for indefinite: $t=-4.164$, $p=.01$ and for definite: $t=8.270$, $p<.0001$). These results will be discussed in combination with the results of the Bridging experiment in section 4.3.

4.2 Bridging experiment

The same children were also tested on conditions that involved bridging. In this experiment the second sentence of the discourse used a new noun which had to be associated to one of the characters. A definite
NP had to be bridged to the mentioned referent and an indefinite NP could be associated with either referent. The design was the same as that of experiment 1. The different conditions are illustrated in figures 6-9.

Fig. 6: Mentioned–Definite

A girl is running down the street and the hair is long

YES

Fig. 7: Mentioned—Indefinite

A boy is picking an apple and a cap (L) is yellow

YES

Fig. 8: Non-mentioned–Definite

A rabbit is eating a carrot and the tail is long

NO

Fig. 9: Non-mentioned—Indefinite

A woman is walking down the street and a hat is blown away

YES

Fig. 10: Results Bridging experiment

A girl is running down the street and the hair is long

YES

A boy is picking an apple and a cap (L) is yellow

YES

A rabbit is eating a carrot and the tail is long

NO

A woman is walking down the street and a hat is blown away

YES
The results are presented in figure 10. A multiple ANOVA gives only a main effect of referent \( (p<.0001) \), and no other main or interaction effects. Paired samples t-tests show that this effect holds for both articles (indefinite: \( t=4.025, p=.001 \), definite: \( t=6.877, p<.0001 \)).

4.3 Discussion of the two experiments

Do children associate definites with mentioned referents? The children indeed associated definites with a mentioned referent. Moreover they did the same for indefinites. Even though an indefinite is not the best article to refer to a mentioned referent, it is not false to do so, at worst, it is infelicitous. Children accepted an indefinite as correct in our truth-value judgment task. One would need a felicity-judgment task to see if children know what counts as the best form to refer to mentioned referents.

Do children introduce new referents for indefinites? Strikingly about half the time the children rejected the new, non-mentioned referent (42% in the first experiment, 58% in the second one). Such overly restrictive answers suggest that the children expected the second sentence to refer to the discourse topic introduced in the first sentence, independent of the article. This behavior was not absolute. In both experiments children sometimes accepted new referents for indefinites and sometimes rejected them, suggesting guessing behavior. We explain this pattern in section 5.

The children showed another non-target-like pattern in allowing a definite to refer to a non-mentioned referent, at least to some extent (30%, 33%), even though most of their answers were target-like no’s. This shows that while they clearly associated definites with mentioned referents (93%, 96%), an old referent was not the only possible interpretation, as they sometimes accepted new referents as well. This over-acceptance of the for new referents mirrors—now in comprehension—the overproduction of definites for new referents which was found in previous studies.

How should we interpret this interpretation pattern for definites? Is the ambiguous for children, like we just argued for a? We doubt so. Whereas the acceptance rates for the combination of indefinites with new referents waver around 50%, most answers for definites with a new referent were in fact rejections. Instead we want to argue that for children the visual context sometimes suffices to establish an antecedent for a definite: even if the referent is not mentioned in the linguistic context (i.e., the previous discourse), it is clearly present for speaker and hearer in the visual context (i.e., the picture). When there is a “visually mentioned” referent, children occasionally accept the in interpretation and use the in production.
5. Guessing in children’s interpretation of indefinite NPs

How can we explain Dutch preschoolers’ guessing behavior in the interpretation of indefinites in contrast to their (mostly) target-like interpretation of definites? We do not want to claim that children lack target-like semantic rules (contra Wexler, 2003), pragmatic principles (contra Schaeffer & Matthewson, 2005), nor that cognitive biases overrule their grammar (contra Karmiloff-Smith, 1979; Maratsos, 1974). Rather, we propose that children’s grammar rules are target-like, but they do not simultaneously entertain speaker-hearer perspectives (de Hoop and Krämer, 2005/06; Hendriks and Spenader 2005/06). If a child cannot simultaneously entertain speaker and hearer perspectives, she cannot compute scalar implicatures.

Hearers not only need to determine the meaning of what is actually spoken, but also take into account what a speaker might have said, but did not. The intended meaning is derived by implicature, which Horn (2004) defines as “… an aspect of what is meant in a speaker’s utterance without being part of what is said. (…) Speaker S tacitly exploits pragmatic principles to bridge this gap and counts on hearer H to invoke the same principles for the purposes of utterance interpretation” (p.1). On our proposal a child follows only hearer considerations when she—as a hearer—selects a meaning for a certain form, and does not weigh these against possibly diverging speaker considerations. She does not bridge the gap between what is said and what is intended. Assuming with Horn (2006) that articles form a scalar pair <the, a>, use of a leads for adults to the exclusion of the uniqueness interpretation, because the is a better form to express uniqueness. We predict, and indeed find, that for children, use of a does not exclude the uniqueness interpretation and so they entertain two possible interpretations—unique and non-unique, mentioned and non-mentioned.

In OT (Prince and Smolensky, 2004), production and comprehension are modelled as different directions of use of the same grammar. In bi-directional OT (Blutner, 2000) the speaker’s perspective—the mapping from meaning-to-form as in OT syntax—is combined with the hearer’s perspective—the mapping from form-to-meaning as in OT semantics—in one joined model in which the same set of constraints apply. Certain form/meaning pairs win while others are blocked as a result of bi-directional reasoning, which is in effect how scalar implicatures are drawn.

De Hoop and Krämer (2005/06) argue that children initially do not reason bidirectionally but use unidirectional models—one for form-to-meaning mappings and another for meaning-to-form mappings (de
We assume that there are two violable constraints. The first is a faithfulness constraint which is based on the semantic/pragmatic rule for definites, i.e., definites must have unique reference, i.o.w. definites with a non-unique interpretation are banned. (2) The other constraint is a markedness constraint we call STRENGTH, which is inspired after Heim’s (1991) ideas that a speaker makes her contribution presuppose as much as possible. This constraint applies to production and penalizes the use of indefinite articles because they do not presuppose, (3).

(2) *DEF./NON-UNIQUE: Definites do not refer to non-unique referents.

(3) STRENGTH: Presuppose as much as possible.

In the interpretation tableaux 1 and 2 the input is a certain form, i.e., a definite or an indefinite NP, and the output is a meaning, i.e., a mentioned or non-mentioned referent. On unidirectional reasoning the constraints yield the following results, which represent, we claim, children’s interpretation process. For definites the mentioned referent is the winner, because the other candidate violates *DEF./NON-UNIQUE. For indefinites there are two winners. Neither constraint creates violations: *DEF./NON-UNIQUE does not apply to indefinites, and STRENGTH only applies to production. The latter result predicts guessing behavior, which is indeed what our results showed. Failure to reason bidirectionally thus only impacts the interpretation of indefinite NPs, not definite NPs.

Table 1: Children’s unidirectional comprehension of definites

<table>
<thead>
<tr>
<th>Input = the</th>
<th>*DEF./NON-UNIQUE</th>
<th>STRENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentioned ref.</td>
<td>*&lt;sup&gt;!&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Non-mentioned ref.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Children’s unidirectional comprehension of indefinites

<table>
<thead>
<tr>
<th>Input = a</th>
<th>*DEF./NON-UNIQUE</th>
<th>STRENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentioned ref.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-mentioned ref.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For adults, speaker and hearer considerations are simultaneously at work, which is modelled in bidirectional tableau 3. Optimization considers all possible form/meaning pairs $<f, m>$. In the first round $<\text{the}, \text{Mentioned ref.}>$ is a super-optimal pair, because it violates no constraints. The next round of optimization blocks pairs with the same form but a different meaning and pairs with the same meaning but another form (note strike-through). Another super-optimal pair remains: $<a, \text{Non-mentioned ref.}>$.

**Tableau 3 : Adults’ bidirectional comprehension-production grammar**

<table>
<thead>
<tr>
<th>Input = $&lt;f, m&gt;$</th>
<th>*DEF./NON-UNIQUE</th>
<th>STRENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt;\text{the}, \text{Mentioned ref.}&gt;$</td>
<td>$!$</td>
<td></td>
</tr>
<tr>
<td>$&lt;\text{the}, \text{Non-mentioned ref.}&gt;$</td>
<td>$!$</td>
<td></td>
</tr>
<tr>
<td>$&lt;a, \text{Mentioned ref.}&gt;$</td>
<td>$!$</td>
<td></td>
</tr>
<tr>
<td>$&lt;a, \text{Non-mentioned ref.}&gt;$</td>
<td>$!$</td>
<td></td>
</tr>
</tbody>
</table>

Bidirectional reasoning formally implements the pragmatic reasoning process of drawing scalar implicatures. The hearer reasons as follows: “The best form for referring to a mentioned referent is a definite. The speaker did not use a definite, but an indefinite. Apparently, the speaker did not have a mentioned referent in mind, hence she must be referring to a non-mentioned referent.” We claim that it is this pragmatic reasoning process that children do not do. Instead they rely on unidirectional optimization with the effects illustrated in tableaux 1 and 2.

5. Conclusions

Our comprehension study of definite and indefinite NPs finds that while Dutch preschoolers associated definites with old referents, they did not always accept indefinites with new referents. We explained this asymmetry as a failure to reason pragmatically offering an implementation in bidirectional OT. We also found that children over-accepted *the* for new referents to some extent, mirroring in comprehension *the* over-production from previous studies. We argued that the visual context causes this over-acceptance and over-use: for children the picture suffices to establish a referent as “known” and hence it can be referred to with a definite.

References

Hendriks, P. (this volume) “A unified explanation of production-comprehension asymmetries”.
de Hoop, H. and I. Krämer (2005/6) “Children’s Optimal Interpretations of Indefinite Subjects and Objects,” Language Acquisition 13, 103-123.