Assessing Truth and Speaker Knowledge
When Utterances Are Not Maximally True

Laura Simon-Pearson and Kristen Syrett

1. Introduction and Background

Speakers frequently deliver utterances expressing propositions that do not adequately capture the state of the actual world. Such utterances need not be blatantly false statements; for example, they may be exaggerations (hyperboles) or metaphors, or may be partially true in that they express some information that is true but some information that is false. This last type of utterance is especially interesting, because when an utterance contains both true and false information, or expresses a proposition that holds true only in a restricted part of the world, it may be unclear how to assign a truth value to such an utterance, or how to evaluate what such an utterance indicates about the speaker’s knowledge of the world. There are two ways in which an utterance could be both true and false: conjunction of a true and false utterance, and plural descriptions violating homogeneity.¹

Speaking strictly at a level of propositional logic, if a true proposition is conjoined with one that is false, then the output yields a false valuation. For example, to say of (1a) and (1b) below that the sentence in (2a) *The circles are black and the squares are white* is clearly false. Correspondingly, such a truth value should be uncontroversially assessed as false by adults.

(1) Situations to consider for the sentences in (2)

(a) [Image of circles]
(b) [Image of squares]

¹ There are other situations that give rise to truth value gaps and gluts, which we do not deal with in this paper. One such example involves vagueness and fuzzy boundaries inherent to context-dependence (e.g., with gradable adjectives, or judgments related to the Sorites paradox).
The circles are black and the squares are white.

b. The circles are black and the squares are black, too.

c. The circles are white.

d. The circle is white.

The corresponding propositional logic truth table for conjunction is shown in (3). Only when both utterances are true (e.g., if (2b) were applied to (1a-b)) can the entire utterance (or the proposition expressed by the utterance) be true. As long as either or both of the propositions are false, the entire conjunction is false.

![Truth table for conjunction](image)

While a wide body of research to date has investigated young children's ability to assign true and false values to individual propositions, little to no work has been devoted to investigating how children evaluate the truth values of [TÅF] conjunctions such as the one considered above. If we seek to determine how close children's semantic representations are to those of adults, then assessing their truth value judgments for conjunctions should be among our priorities, since valuations of such conjoined propositions is a cornerstone of propositional logic.

It is also possible for an individual proposition to be considered both true and false, in that it is only partially true of the world. An example of such a proposition is one in which the subject expresses a plural definite description (e.g., *the circles*) that makes reference to a plurality as a whole, but the verb phrase predicates a property that holds of only a proper subset of the plurality. For example, if one says of (1c) that *The circles are white* (2b), this statement is only true of some, but not all, of the circles in (1c). The sentence is therefore only true of a proper subset of the set of objects. The predicate ‘be white’ is a distributive predicate that obligatorily applies to atomic individuals. Plural definite descriptions are theoretically claimed to pick out the maximal element of a plurality: the sum of all atomic individuals (Link 1983). Thus, the predicate should apply to all or none of the individual members of the plurality. If it does not, the members are not homogenous with respect to this property (or, there is a violation of homogeneity), and a truth value cannot be assigned. Such an
utterance is certainly not uncontroversially true, but neither is it clearly judged to be false.

Propositions like this are thus known to give rise to truth value gaps (Breheny 2005; Büring and Križ 2013; Križ 2015; Löbner 1987; Magri 2014; Schwarzschild 1994; Spector 2013), in a pattern that may be reminiscent of cases where the presupposition of existence is not met with a singular definite description (e.g., (2c) The circle is white uttered in (1b)). The utterance in (2c) is also pragmatically underinformative of (1c), because it is only partially true of the scene. Consequently, it fails to fulfill the Gricean Maxim of Quantity, which directs speakers to make their contribution as informative as required for the current purposes of the conversational exchange (Grice 1989).

There is experimental evidence that adults judge such utterances to be ‘not completely false’ (Križ & Chemla 2015). Preliminary experimental evidence also suggests that four-year-old native French speakers judge similar sentences in French as ‘not true’ (Tieu, Križ & Chemla, under review). However, in French, determiners are obligatory, and so it is not clear how children acquiring English – a language that allows both bare plurals (circles) and noun phrases accompanied by overt determiners – assess such utterances, as compared to those with bare plurals in similar contexts.

In this research, we therefore contrast two cases of conjunction: one in which an unambiguously true conjunct in the first position is conjoined with an unambiguously false conjunct in the second position, and one in which the true conjunct is conjoined with a second conjunct that violates homogeneity. We predicted that participants who abide by the rules of propositional logic will evaluate the first case as false tout court. We predicted that the responses to the second case should be varied, but will not elicit as many rejections. However, we also predict that these responses will be modulated by the type of expression in subject position, and that the amount of rejections of conjunctions violating homogeneity will increase if a plural definite description appears in subject position, as opposed to a bare plural. It is an open question whether children acquiring English (a language that allows both bare plurals and plural definite descriptions) recognize that the plural definite description differs from the bare plural in picking out the maximal element, and if so, that a statement with a plural definite description in subject position that violates homogeneity may lead to a truth value gap – or at a minimum is not clearly true?

We further asked what the cost of such an utterance is on the assessment of the speaker’s knowledge. We predicted that participants might take a false or ‘not true’ valuation as evidence of degraded knowledge on the speaker’s part, and therefore, when asked directly about the speaker’s knowledge in an embedded question (e.g., Does X know which circles are white?), indicate as much. Previous research on children’s and adults responses to embedded questions reveals that delivery of assertions containing false information lead to a high probability that the speaker will be said to not know the embedded proposition (Cremers & Chemla 2014; Cremers, Tieu, & Chemla 2016), although children are more tolerant of false information than adults are (Cremers, Tieu, & Chemla 2016). The connection between violations of
homogeneity and assessments of speaker knowledge has not, to our knowledge, been investigated. Thus, we included this as a second goal in this research.

2. Experiment
2.1. Participants

48 children (3;04-6;03, M: 4;06; 22 females and 26 males) and 74 adult native English speakers participated. Children were recruited from nearby preschools, and adults received course credit in an Linguistics or Cognitive Science course for their participation. All participants were monolingual speakers of English, and either the parents (for children) or the participants themselves (adults) provided informed consent for participation. There were 4 conditions, with 12 children and approximately 18 adults in each condition. Participant ages and genders were balanced across conditions. The target age range of the children was identified based on children’s ability after age four to possess Theory of Mind, and thus to attribute thoughts and propositions to another speaker that may not align with their own (i.e., represent the mental state of another speaker), and to pass false-belief tasks (Leslie, Friedman, & German, 2004; Wimmer, & Perner 1983).

2.2. Materials and Procedure

Participants were shown a series of stories following a Truth Value Judgment Task design (Crain & Thornton 1998). For children, two experimenters administered the experiment. The first experimenter read the script in a consistent but lively and animated way, and the second played the role of the puppet, who delivered a target sentence at the end of the story. The narrating experimenter than asked the child what the puppet knew based on the statement it had delivered.

For adults, one experimenter administered the experiment, and a character appeared on screen after the story to administer the target sentence. Adults then circled their responses on their response sheet. The experimenter recorded the responses from the children. Sessions with children were also video recorded (with parental consent) for later verification and transcription of responses. Participants were run individually in a quiet room in the preschool (children) or (for adults) in sets of 1-3 participants in the laboratory, equipped with clipboards and personal response sheets.

There were 12 trials: eight test trials and four control trials, all with similar structure. The full list of target sentences appears in the Appendix. Two training items preceded the experimental session proper, so that participants were comfortable with the procedure and what was being asked of them. The entire session took approximately 12-15 minutes.

In each story, there were two sets of entities who were deciding which of two locations to go to (e.g., a bookstore and a library for books, a rock or a tree for a game of hide and seek, etc.). In all trials, each of the event participants in
each set ended up going to one of the two locations. In the test trials, all of the members of one set went to one location, and following this, the members of the second set either all went to the second location (the ‘homogenous’ condition), or were split between the two locations (the ‘non-homogenous’ condition), as shown in Figure 1. Four of the test items were ‘homogenous’ and four were ‘non-homogenous’. In the ‘homogenous’ trials, the second conjunct was always false of the second set of entities, while in the ‘non-homogenous’ trials, it was true of some but false of others (as shown in Figure 1). The homogeneity factor was within subjects, so that each participant saw trials of both types. In this way, we manipulated whether the second conjunct of the target sentence was entirely false of the entire second set of entities (yielding a [T∧F] conjunction) or only a proper subset of the second set (violating homogeneity), thereby allowing us to assess how both children and adults evaluate the target sentences in both of these ‘gappy’ contexts, and as a follow-up, how they take such a statement to be a reflection of the speaker’s (puppet’s) knowledge.

<table>
<thead>
<tr>
<th>Homogenous condition</th>
<th>Non-homogenous condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(The) boys went to the art room, and (the) girls went to the art room, too.</td>
<td>(The) red dinosaurs went to the bookstore, and (the) green dinosaurs went to the bookstore, too.</td>
</tr>
</tbody>
</table>

Figure 1: Homogenous and non-homogenous conditions (within subject), with corresponding target conjunction sentences

Paired with these within-subject context manipulations were two between-subject manipulations on subject type in order to probe the influence of set (non)homogeneity and linguistic interaction with it. Half of the participants heard a bare plural (boys, red dinosaurs) in subject position, while the other half heard a plural definite description (the boys, the red dinosaurs) in subject position. This condition was between subjects, because we wanted to avoid any spreading of the influence of definite determiner presence/absence across items. (Recall that such a manipulation would not have been possible for Tieu, Križ & Chemla (ms.), since their experiment was in French, which does not permit bare plurals here.) No matter what the subject was, all of the members referred to in the first conjunct (e.g., the boys or the red dinosaurs above) went to their location before the members of the second set made their decisions and went to their locations. The order of events in the story was therefore appropriately captured by the order of events in the target sentences, thereby adhering to the Gricean Maxim of Manner (Be orderly.).
After each story, the puppet delivered the target sentence. Participants were asked whether the puppet was right or wrong, based on the story they had been told. Children pointed to one of two faces on a sheet: a happy face or a frowny/sad face. Adults circled either YES or NO on their response sheet. Because we thought it more likely that participants would select the frowny/sad face, based on the propositional logic of conjunction, our dependent measure was stated in terms of rejection of the target sentence, rather than acceptance.

Control items were structured similarly to the test items, and so that statements were uncontroversially true or false. There were two sentences that involves a bare plural or plural definite description in subject position, and two that involved an existential quantifier in the determiner phrase, which either composed directly with the NP in the bare plural condition, or else was part of a partitive ($X$ of the $Y$) in the plural definite description condition. Control sentences are also included in the appendix.

Once participants heard the story and assessed the puppet’s target sentence recapping the events, the narrator experimenter posed a question to the participant regarding the puppet’s knowledge. Participants were asked whether the puppet (or the character on the screen, for adults) knew who went to the first location (e.g., Does $X$ know (who/which $N$) went to the [art room/bookstore]?). (The type of $wh$- expression was manipulated between subjects, but this factor did not end up being significant.) All participants were asked to assess the speaker’s knowledge on a ternary scale, as in Figure 2, allowing for us to capture this judgment in a more fine-grained way. A similar ternary rating strategy was successfully pioneered by Katsos & Bishop (2011). Participants were told to award the puppet/character a big yellow star if he definitely knew, a medium blue star if he ‘kinda’ knew, and a small red star if he definitely didn’t know at all. The experimenter delivering these options emphasized the choices with emphatic prosody and hand gestures to highlight the differences in ratings. Neither children nor adults had difficulty with this part of the task.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>definitely doesn’t know</td>
<td>kind of knows</td>
</tr>
</tbody>
</table>

Figure 2: Ternary reward system used for speaker knowledge assessment

Since the entire conjunction always featured a first conjunct that was true of a plurality with respect to that location, but a second conjunct that was either false of the second plurality with respect to that same location or partially true/false, participants were in effect assessing the cost of the speaker’s having uttered this false or homogeneity-violating conjunct in deciding whether or not the speaker’s knowledge was degraded as a result of this report (which was a reflection of, and therefore correlated with, the speaker’s false belief).
2.3. Results

We present the results for each of the two dependent measures obtained in the task separately. We first present the results from the truth value judgment task in terms of percentage of rejection of the puppet’s initial assertion. We then turn to the knowledge assessment task and present the distribution of participants’ choice of one of the three stars in response to the question about what the puppet knows.

2.3.1. Assessments of Truth Values

We begin with the truth value judgment aspect of the task. The results are presented in Figure 3.

![Statement Rejection Rates by Condition](image)

**Figure 3. Percentage of rejection for the Truth Value Judgment task part of the test trials**

Both children and adults behaved as propositional logic would predict with the homogenous \([T \land F]\) conjunctions: they consistently rejected these assertions, regardless of the subject type. This pattern adds to the evidence that children at the age tested here have a capacity to calculate truth values that parallels that of adults.

The results become more interesting, however, in the non-homogenous trials: while children hover around chance level in responding to the puppet’s statement, adults tend to accept the utterance and assess it as true when the subject is a bare plural, but are more resistant to accepting the utterance and saying that it is true in the case of the definite description in subject position. A logistic regression revealed there was a significant main effect of homogeneity, such that participants for both age groups were more likely to reject the statement when it obeyed homogeneity \((\beta = 5.5446, \alpha = -4.1728, SE = 0.4697, z = 11.805, p < .001)\). There was also a significant interaction of homogeneity and
subject type ($\beta = 3.6212, \alpha = -2.7324, \text{SE} = 0.3155, z = 11.477, p< .001$). As predicted, adults were least likely to reject the statement when it did not obey homogeneity, and featured a bare plural subject.

This pattern is perhaps what one would predict, given two things. First, the semantics of definite descriptions pick out the maximal element of a plurality while bare plurals do not and may instead allow for an existential interpretation. Second, previous results investigating children’s and adults’ responses to plural definite descriptions has shown that children do not seem to appreciate the ‘maximality’ of plural definite descriptions until after 4-5 years of age, and in the face of homogeneity violations, adults exhibit variable responses (see e.g., Caponigro et al. 2012 and related references therein).

### 2.3.2. Assessments of Speaker Knowledge

We now turn to the knowledge assessment task. The results showing the assignment of stars in each of the conditions and by both groups of participants is presented in Figure 4.

![Evaluation of Speaker Knowledge](image)

**Figure 4. Percentage of star awarded for the Speaker Knowledge Assessment task of the test trials, where yellow star (top): definitely knew, medium blue star (middle): ‘kinda’ knew, small red star (bottom): definitely didn’t know at all**

As a backdrop, we note that participants willingly awarded the largest yellow star for control trials where the puppet delivered an unambiguously true assertion, and a red star for those where the assertion was false (a conjunction where both conjuncts were false). As the figure above indicates, both children and adults took the statements containing false information as a sign of the speaker’s degraded knowledge. There was a penalty for delivering a false assertion – either one that was completely false (as in the homogenous trials) or one that was partially true/false (as in the non-homogenous trials).
In the homogenous trials, both sets of participants were reluctant to award the largest yellow star for the test trials. Interestingly, children were more likely to award the smallest red star than either of the other stars, and more often than adults did; adults were drawn to awarding the medium blue star. Thus, adults seemed to be more attuned to the partial knowledge, whereas many children took the false assertion as a fatal blow. In the non-homogenous trials, children in both subject conditions and adults in the bare plural condition (but not the definite description condition) were more willing to award the largest yellow star, but not so much so that this became the main option. These patterns linked to the effect of homogeneity were supported by simple $\chi^2$ tests. Only adults distinguished between the two subject conditions – the plural definite description on the one hand and the bare plural on the other – and just in the non-homogenous condition, just as we predicted, based on the theoretical work on definite descriptions and previous research comparing children’s and adults’ knowledge of maximality tied to plural definite descriptions.

The results from the children revealing that they were overwhelmingly selected the medium and small stars may be somewhat surprising in light of the results from Cremers, Tieu, and Chemla (2016), where children were more willing than adults to overlook a false assertion and attribute knowledge to a speaker who delivered the false report. (See also Moyer (2016).) However, in those paradigms, the speaker did not deliver a conjunction with both a true and false assertion, and the child was not asked to evaluate the truth value of the utterance prior to being asked about the speaker’s knowledge. In addition, the children were simply asked to deliver a ‘yes’ or ‘no’ response without an intermediate option. Likewise, adults in that same task were inclined towards outright rejection, but were also not given a ‘partial’ option. The current manipulations seem to have uncovered children’s sensitivity to false assertions, and may also indicate that without a more fine-grained scale, children may be inclined towards acceptance, and adults inclined towards rejection, but with an intermediate option present, both participants are drawn to this partial option, which indicates degraded knowledge on the part of the speaker.

In support of their decision to award the puppet the medium blue or small red star in the homogenous condition, children provided justifications that highlighted the falsity of the second conjunct, or corrected the puppet’s assertion, as shown in (4) and (5). The child’s age, the condition, and the target test sentence from the Appendix appear after the child’s utterance.

(4) Justifications in the homogenous condition
   a. “Because the squirrels choosed the cave and they [the bunny rabbits] choosed the log” (age 5;09, DD condition, target A.1.6)
   b. “Rabbits goed in the log and squirrels goed in the case.” (age 4;05, BP condition, target A.1.6)
   c. “You were right when you said the penguins going into igloo but you weren’t right because the seals are going sliding down the hill.” (age 4;07, DD condition, target A.1.7)
d. “Um, only boys went to there and girls went to there. Why did you thought that?” (age 3;06, BP condition, target A.1.8)
e. “Cause he thought that the dogs went to the tree, too, with the cats” (age 4;03, DD condition, target A.1.9)

(5) Justifications in the non-homogenous condition
a. “No, I think half of them, the green guys, went to the bookstore” (age 4;10, DD condition, target A.2.10)
b. “Girls went, two girls went in the red tent, and three girls goed into the lellow tent, three boys goed in the red tent.” (age 4;05, BP condition, target A.2.11)
c. “Well, actually, Mr. Rabbit, some grasshoppers went into the yellow car, too, and they [pointing] only went into the purple car” (age 4;07, DD condition, target A.2.12)
d. “Two grasshoppers in the purple car, and three grasshoppers in the yellow car, and three ladybugs in the yellow car” (age 4;06, BP condition, target A.2.12)
e. “Because three of the girls went on the red rug and two of the girls went on that rug.” (age 5;01, DD condition, target A.2.13)

Adults’ willingness to assign a medium blue star the vast majority of the time may be seen as in line with claims by Beck & Rullman (1999), George (2013), Guerzoni & Sharvit (2007), Heim 1994), and Sharvit (2002) regarding the availability of weak exhaustive answers to embedded questions with know, or reflect that they are resistant to allowing knowledge to be attributed to the speaker in the case of a false report, which may be taken to reflect a false belief (Spector & Egre 2015). (See also Klinedinst and Rothschild (2011).) If adults required that the speaker deliver an assertion that was entirely true, then they should have awarded the character with a small red star for every utterance that contained false information. They were also somewhat forgiving of the false information, as long as it was paired with a true assertion. Whether the pattern of results would still be the same had the false assertion appeared sequentially as the first conjunct is a separate question that we cannot answer with these data.

3. Discussion

Three major findings arise from the current research. First, children of the age tested compute the truth conditions of conjunction in a manner consistent with propositional logic. When a true conjunct is conjoined with a false conjunct, the entire assertion is deemed false. Thus, children’s semantic treatment of these propositions is highly similar to that of adults. Second, when a conjunction includes a conjunct that violates homogeneity, both children and adults are significantly less likely to reject the entire utterance. However, adults are more likely to accept such an utterance, and more likely to do so when the subject of the conjuncts is a bare plural, as opposed to a plural definite description.
Children of the age tested do not appear to distinguish between the two subject forms in this task, or appreciate that a plural definite description selects the maximal element of the plurality – a finding consistent with previous literature. These findings are in keeping with previous research on children’s developing comprehension of plural definite descriptions. However, the data also indicate that children do not simply or uniformly interpret either subject type existentially, since such an interpretation should lead to these utterances being evaluated as true (particularly the conjuncts with bare plural subjects), and an assessment that the speaker has the relevant knowledge. An open question is whether or not the (or a) key difference between children and adults is in children’s willingness to restrict the domain of quantification, a possibility we are pursuing in follow-up work.

Finally, when asked to evaluate the knowledge of the speaker who delivers the conjunctive utterance as related to the events at hand, both children and adults perceive the speaker’s knowledge as degraded as a consequence of their having delivered a false report (which presumably reflects a false belief). Here, too, adults distinguish between subjects that are bare plurals or plural definite descriptions in the non-homogenous condition, and are more likely to award the middle option (‘kinda knows’) for conjuncts with bare plural subjects. Children are not willing to simply overlook false reports, and those utterances that obey homogeneity incur more of a penalty for the false report than those that violate it. The picture that emerges, then, is that children are not only sensitive to the presence of false reports and feed this information into truth value evaluations, but they are also sensitive to them as reflections of degraded speaker knowledge, and feed this information into assessments of what a speaker can be said to know. These features are presumably cross-linguistic universals. Where young children continue to diverge from adults is in their sensitivity to language-specific morphosyntactic cues to maximality.

Appendix: Target sentences and questions

Contrastive focus placement is indicated with capitalized syllables.

A. Test Items

A. 1. Homogenous

(6) Puppet: (The) RABbits went into the log, and (the) SQUIRrels went into the log, too. Am I right?
Exp: Does Mr. Rabbit know [who/which animals] went into the log?

(7) Puppet: (The) PENguins went to the igloo, and (the) SEALS went to the igloo, too. Am I right?
Exp: Does Mr. Rabbit know [who/which animals] went to the igloo?
(8) Puppet: (The) BOYS went to the art room, and (the) GIRLS went to the art room, too. Am I right?
Exp: Does Mr. Rabbit know [who/which children] went to the art room?

(9) Puppet: (The) CATS hid in the tree, and (the) DOGS hid in the tree, too. Am I right?
Exp: Does Mr. Rabbit know [who/which animals] hid in the tree?

A. 2. Non-Homogenous
(10) Puppet: The RED dinosaurs went to the bookstore, and the GREEN dinosaurs went to the bookstore, too. Am I right?
Exp: Does Mr. Rabbit know [who/which dinosaurs] went to the bookstore?

(11) Puppet: (The) BOYS went into the red tent, and (the) GIRLS went into the red tent, too. Am I right?
Exp: Does Mr. Rabbit know [who/which children] went into the red tent?

(12) Puppet: (The) LAdybugs got into the yellow car, and (the) GRASShoppers got into the yellow car, too. Am I right?
Exp: Does Mr. Rabbit know [who/which animals] got into the yellow car?

(13) Puppet: (The) BOYS sat on the blue rug, and (the) GIRLS sat on the blue rug, too. Am I right?
Exp: Does Mr. Rabbit know [who/which children] sat on the blue rug?

B. Control Items
Groups separate properties, True
(14) Puppet: (The) BLUE butterflies landed on the TUlips and (the) YELlow butterflies landed on the DAIsies. Am I right?
Exp: Does Mr. Rabbit know [who/which butterflies] landed on the daisies?

Groups separate properties, False
(15) Puppet: (The) PIGS dined at the sandwich shop and (the) COWS dined at the pizzeria. Am I right?
Exp: Does Mr. Rabbit know [who/which animals] dined at the pizzeria?
Groups mixed properties, True

(16)
Puppet: SOME (of the) birds chose to rest in the birdhouse, and SOME (of the) birds chose to rest in the hole in the tree. Am I right?
Exp: Does Mr. Rabbit know [who/which birds] chose to rest in the birdhouse?

Groups mixed properties, True

(17)
Puppet: SOME (of the) farm animals chose to rest under the tree, and SOME (of the) farm animals chose to go to the pond. Am I right?
Exp: Does Mr. Rabbit know [who/which farm animals] went to rest under the tree?

References


Spector, Benjamin. (2013). Homogeneity and plurals: From the strongest meaning hypothesis to supervaluations. Presentation at *Sinn und Bedeutung 18*.


