Journal of Semantics Advance Access published August 10, 2012

Journal of Semantics, 0, 2012: 1–20
doi:10.1093/jos/ffs014

Conservativity and Learnability of Determiners

TIM HUNTER
Yale University

JEFFREY LIDZ
University of Maryland

Abstract

A striking cross-linguistic generalisation about the semantics of determiners is that they never express non-conservative relations. To account for this one might hypothesise that the mechanisms underlying human language acquisition are unsuited to non-conservative determiner meanings. We present experimental evidence that 4- and 5-year-olds fail to learn a novel non-conservative determiner but succeed in learning a comparable conservative determiner, consistent with the learnability hypothesis.

1 INTRODUCTION

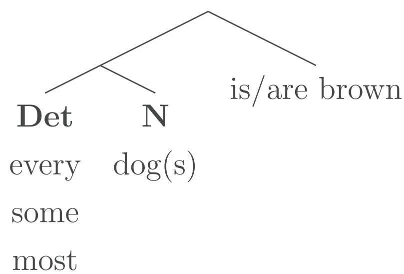
Testing children's abilities to acquire novel words tells us about the word meanings that children are likely to entertain as hypotheses, and therefore to some extent about the range and limits of the word meanings permitted by the language faculty. We examine children's learning of novel determiner meanings, in order to investigate whether a well-established typological generalisation might derive from a constraint on language learning. Specifically, all attested natural language determiners are conservative (defined below), and we compare children's abilities to learn a conservative determiner with their abilities to learn a non-conservative one. Striking as the typological generalisation may be, it does not logically entail any asymmetry in the status of conservative and non-conservative determiners in the learner's hypothesis space; in principle one can imagine alternative explanations based on some pragmatic or functional reason. We find, however, that children succeed in learning a novel conservative determiner but fail to learn a novel non-conservative determiner, which is consistent with the hypothesis that the typological generalisation results from constraints on children's hypothesis space of determiner meanings.

The rest of the article proceeds as follows. In Section 2, we review the relevant background concerning determiners and conservativity. In Section 3, we discuss some related findings concerning non-adult-like interpretations of quantificational expressions, which serve to emphasise that the nature of the conservativity generalisation remains unclear. In Section 4, we define two novel determiners, only one of which is conservative, and then in Section 5, present an experiment comparing children's abilities to learn these two determiners; the results show that children succeed only in the case of the conservative determiner. We conclude briefly in Section 6.

2 DETERMINERS AND CONSERVATIVITY

The class of determiners includes words such as 'every', 'some' and 'most'. These words can occur in the syntactic frame illustrated in (1).[1]

(1)



In the framework of generalised quantifier theory (Mostowski 1957), sentences with this form express a relation between two sets: the set of dogs, and the set of brown things. If we represent these sets by DOG and BROWN, respectively, the truth conditions of the three sentences abbreviated in (1) can be expressed as in (2).

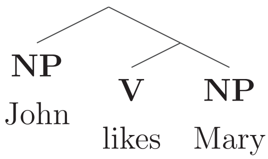
(2) 'every dog is brown' is true iff $\text{DOG} \subseteq \text{BROWN}$
'some dog is brown' is true iff $\text{DOG} \cap \text{BROWN} \neq \emptyset$
'most dogs are brown' is true iff $|\text{DOG} \cap \text{BROWN}| > |\text{DOG} - \text{BROWN}|$

An analogy can be made between the syntactic role of determiners and that of a transitive verb such as 'like'. A determiner expresses a relation between two *sets*, much as a transitive verb expresses a relation between two *individuals*: (3) indicates that a particular relation holds

[1] We remain agnostic about many of the details of the syntax of these sentences, and for this reason limit our attention to quantifiers in subject positions. What is important is just that 'determiner' is defined distributionally as something that combines with a noun to form a noun phrase.

between John and Mary.

(3)



The transitive verb 'like' combines first with 'Mary' and then with 'John', resulting in a sentence that expresses a relation between the two corresponding individuals. If we ignore the linear order of the trees and consider only the hierarchical relations, we see that the determiners in (1) likewise combine first with 'dog(s)' and then with 'is/are brown', resulting in a sentence that expresses a relation between the two corresponding sets. We call 'Mary' and 'dog(s)' the *internal* arguments, and call 'John' and 'is/are brown' the *external* arguments.

Standard approaches to natural language semantics (e.g. Heim & Kratzer (1998); Larson & Segal (1995) among many others) postulate that knowing the meaning of a determiner consists in knowing which of all the conceivable two-place relations on sets the determiner expresses, just as knowing the meaning of the transitive verb 'like' consists in knowing that it expresses "the liking relation" on individuals. Thus the three determiners in (1) are associated with the following three relations on sets:

(4) $\mathcal{R}_{\text{every}}(X)(Y) \equiv X \subseteq Y$
$\mathcal{R}_{\text{some}}(X)(Y) \equiv X \cap Y \neq \emptyset$
$\mathcal{R}_{\text{most}}(X)(Y) \equiv |X \cap Y| > |X - Y|$

and so the sentence 'every dog is brown', for example, in which the internal argument of 'every' denotes the set DOG and the external argument of 'every' denotes the set BROWN, is true if and only if $\mathcal{R}_{\text{every}}$(DOG)(BROWN) is true.

When the determiners of the world's languages are analysed in this way, a surprising generalisation emerges (Barwise & Cooper 1981; Higginbotham & May 1981; Keenan & Stavi 1986): every attested determiner expresses a relation that is conservative, as defined in (5).[2]

[2] Two apparent counterexamples are 'only' and 'many'. Closer examination quickly shows that 'only' is not a determiner, as defined distributionally. While at first 'only dogs are brown' looks superfically like 'some dogs are brown', 'only' can appear in many other positions where 'some' and 'every' cannot, e.g. 'dogs only/*some/*every are brown', and 'dogs are only/*some/*every brown'. The case of 'many' is less clear, complicated by context-dependence, but can also plausibly be made to fit with the conservativity generalisation; see for example Keenan & Stavi (1986) and Herburger (1997).

A two-place relation on sets $\mathcal{R}$ is *conservative* if and only if the following biconditional is true:

(5) $\mathcal{R}(X)(Y) \Longleftrightarrow \mathcal{R}(X)(X \cap Y)$.

For example, consider the English determiner ‘every’. This determiner is conservative[3] because the relevant biconditional holds.

$$\mathcal{R}_{\text{every}}(X)(Y) \Longleftrightarrow X \subseteq Y \Longleftrightarrow X \subseteq (X \cap Y) \Longleftrightarrow \mathcal{R}_{\text{every}}(X)(X \cap Y)$$

To think about this more intuitively we can express the crucial biconditional in natural language. Since the requirement entails that $\mathcal{R}_{\text{every}}$(DOG)(BROWN) holds if and only if $\mathcal{R}_{\text{every}}$(DOG)(BROWN ∩ DOG) holds, and since (BROWN ∩ DOG) is the set of brown dogs, the crucial biconditional is ‘every dog is brown if and only if every dog is a brown dog’. This is trivially true, and so ‘every’ is conservative.

Another intuitive view of what it means for ‘every’ to be conservative is that in order to determine whether a sentence like ‘every dog is brown’ is true, it suffices to consider only dogs. The brownness or otherwise *of dogs* is relevant, but the brownness of anything else is not. Barwise & Cooper (1981) call this ‘living on the internal argument’, since DOG is the set denoted by the internal argument of ‘every’ in this sentence. Other members of the set denoted by the external argument, BROWN, can be ignored.

We can now observe that both ‘some’ and ‘most’ are also conservative: to determine whether ‘some/most dogs are brown’ it is safe to ignore any brown things are not dogs. Alternatively, we can note that both of the following biconditionals are true: (i) ‘some dogs are brown if and only if some dogs are brown dogs’, and (ii) ‘most dogs are brown if and only if most dogs are brown dogs’.

For comparison, consider a fictional determiner ‘equi’. The relation that this determiner expresses is illustrated in (6) (sometimes known as the ‘Härtig Quantifier’; see also Crain et al. 2005: 182).

(6) a. $\mathcal{R}_{\text{equi}}(X)(Y) \equiv |X| = |Y|$
b. ‘equi dogs are brown’ is true iff |DOG| = |BROWN|

So ‘equi dogs are brown’ is true if and only if the number of dogs (in the relevant domain) is equal to the number of brown things. Note that brown things that are not dogs *are* relevant to the truth of this sentence. To verify this claim it does not suffice to consider only dogs, so ‘equi’

[3] We systematically overload the term ‘conservative’, using it to apply both to relations as defined in (5) and to determiners that express such relations.

does not 'live on' its internal argument. We can also observe the falsity of the crucial biconditional: $|DOG| = |BROWN| \iff |DOG| = |DOG \cap BROWN|$, or 'the number of dogs is equal to the number of brown things if and only if the number of dogs is equal to the number of brown dogs'. Thus 'equi' is not conservative.

The absence of non-conservative determiners is problematic for standard theories of semantics, on at least one view of what these theories aim to account for: ideally, it would be desirable for the mechanics of a semantic theory to allow determiners with all and only the meanings that the human language faculty allows. Following familiar reasoning about the relationship between innate properties of the language faculty and linguistic typology, a reasonable hypothesis to consider is that the lack of non-conservative determiners in the world's languages derives from the (in)ability of the human language faculty to associate the structure in (1) with the claim that a non-conservative relation holds between the set of dogs and the set of brown things. The overwhelming majority of current theories, however, are equally compatible with conservative and non-conservative determiners, essentially predicting that the language faculty should be able to associate the structure in (1) with either kind of relation (but see Pietroski 2005; Bhatt & Pancheva 2007; Fox 2002 for some exceptions).[4]

In this article, we investigate whether children allow the structure in (1) to express non-conservative relations. If children permit (1) to be associated with a non-conservative meaning, then a semantic theory which permits non-conservative determiners would appear to be an accurate reflection of the workings of the human language faculty, and the lack of non-conservative determiners in natural languages would need to be explained by something else. However, in the experiment we report below, we find no evidence that children consider non-conservative meanings for novel determiners, supporting the

[4] To elaborate, a theory of semantics might in principle allow the words of a certain syntactic category too small a range of possible meanings, or too large. A theory might allow determiners *too small* a range of meanings by, for example, requiring that the structure [[Det X] Y] is associated with a claim that a relation expressible in first-order predicate logic holds between X and Y. This would incorrectly exclude the meaning we need to associate with 'most', which requires a more powerful logic. Alternatively, a theory might allow determiners *too large* a range of meanings by permitting the structure [[Det X] Y] to either express a two-place relation between the set X and the set Y, or a three-place relation between the set X, the set Y and some fixed other set. We never see the human language faculty making use of this latter three-place option, so we suppose that the option is not there and prefer theories that do not allow it. The case of conservativity is analogous: if we never see the human language faculty making use of the ability to learn non-conservative determiners, we would prefer theories that do not allow it.

Keenan & Stavi (1986) quantify the degree to which allowing non-conservative meanings would allow 'too large' a hypothesis space (if this flexibility is indeed unnecessary): given a domain of n entities, there are $2^{(4^n)}$ possible determiner meanings, only $2^{(3^n)}$ of which are conservative.

hypothesis that the language faculty is ill-equiped to associate the structure in (1) with a non-conservative relation and strengthening the case that semantic theories should be revised to reflect this. Of course, the children's failure to learn a non-conservative determiner meaning in our experiment is not equivalent to observing that no child in any experiment could ever learn a non-conservative determiner meaning. But the results are consistent with the hypothesis that the underlying cause of this failure is the non-conservativity of the putative determiner's meaning, whether this meaning is completely unlearnable in some sense or just difficult to activate in these tasks.

3 CHILDREN'S SYMMETRIC INTERPRETATIONS OF QUANTIFICATIONAL SENTENCES

The question of whether children can associate determiners with non-conservative meanings remains open, despite various much-discussed findings of non-adult-like 'symmetric' interpretations of quantificational sentences. Inhelder & Piaget (1964: 60–74) found that some children will answer 'no' to a question like (7) if there are blue non-circles present. When prompted, these children will explain this answer by pointing to, for example, some blue squares.

(7) Are all the circles blue?

Taken at face value it appears that these children are understanding (7) to mean that (all) the circles *are* (all) the blue things, as if $\mathcal{R}_{\text{all}}(X)(Y) \equiv X = Y$. This is a clearly non-conservative relation, since answering (7) on this interpretation requires paying attention to non-circles. But there is little or no reason to think that these children have associated this non-conservative relation with the determiner 'all'.

Importantly, similar 'symmetric responses' have been observed with questions like (8) involving transitive predicates. Some children will answer 'no' to (8) if there are elephants not being ridden by a girl (Philip 1991, 1995); see Geurts (2003); Drozd (2005/2006) for review.

(8) Is every girl riding an elephant?

Consider the interpretation of 'every' that these children are using. If 'every' is analysed as a determiner with a conservative meaning, then answering (8) should require only paying attention to the set of girls (and which *of the girls* are riding an elephant), since this is the denotation of the internal argument 'girl'. Clearly 'every' is not being analysed in this way by the children for whom the presence of unridden elephants is relevant. However, these children are not analysing ‘every’ as a *non-conservative determiner* either. Such a determiner would permit meanings that required looking beyond the set of girls denoted by the internal argument and take into consideration the entire set denoted by the external argument, as the fictional ‘equi’ does; but crucially, the external argument would be ‘is riding an elephant’ and would therefore denote the set of *elephant-riders*, not the set of elephants.[5] So allowing the non-conservative relation $\mathcal{R}_{\mathrm{all}}$ above into the child’s hypothesis space would leave room for an interpretation of (8) on which the presence of non-girl elephant-riders triggers a ‘no’ response, but would do nothing to explain the relevance of unridden elephants. On the assumption then that these symmetric responses to (7) and (8) are to be taken as two distinct instances of a single phenomenon, this phenomenon is more general than (and independent of) any specific details of determiners and conservativity.

4 TWO NOVEL DETERMINERS: ‘GLEEB’ AND ‘GLEEB′’

The question we aim to address is whether children permit structures like (1) to have non-conservative meanings. To investigate this question, we attempted to teach children novel determiners. If children have no inherent restrictions on determiner meanings, then we would predict that they will be able to learn both novel conservative determiners and novel non-conservative determiners. However, if the typological generalisation that we observe reflects a restriction imposed by the language faculty, then we predict that children will succeed in learning novel conservative determiners, and will not succeed in learning novel non-conservative determiners.

In order to test these predictions we created two novel determiners, one conservative and one non-conservative. The conservative one, ‘gleeb’, expresses the relation $\mathcal{R}_{\mathrm{gleeb}}$ as illustrated in (9).

(9) a. $\mathcal{R}_{\mathrm{gleeb}}(X)(Y) \equiv X \not\subseteq Y \equiv \neg(X \subseteq Y)$
    b. ‘gleeb girls are on the beach’ is true iff GIRL$\not\subseteq$BEACH

So ‘gleeb girls are on the beach’ is the negation of ‘every girl is on the beach’: it is true if and only if the set of girls (GIRL) is *not* a subset of the set of beach-goers (BEACH), so we might paraphrase it as ‘not all girls are on the beach’. For example, it is true in the scene shown in

[5] The relevant distinction is collapsed in (7) because the denotation of the determiner’s external argument, the verb phrase ‘are blue’, is (on standard assumptions) the same as that of this verb phrase’s own complement ‘blue’, namely the set of blue things.

Figure 1a, but false in the scene shown in Figure 1b. Since 'gleeb' is the 'negation' of the conservative determiner 'every', it is also conservative:[6] anything on the beach that is not a girl is irrelevant to the truth of the sentence in (9b), so 'gleeb' does live on its internal argument, and the biconditional 'not all girls are on the beach if and only if not all girls are girls on the beach' is true.

The novel non-conservative determiner, written 'gleeb′' but pronounced identically to the conservative determiner 'gleeb', expresses the relation $\mathcal{R}'_{\text{gleeb}}$ as illustrated in (10).

(10) a. $\mathcal{R}'_{\text{gleeb}}(X)(Y) \equiv Y \not\subseteq X \equiv \neg(Y \subseteq X) \equiv \mathcal{R}_{\text{gleeb}}(Y)(X)$
b. 'gleeb′ girls are on the beach' is true iff BEACH$\not\subseteq$GIRL

So 'gleeb′ girls are on the beach' is the 'mirror image' of 'not all girls are on the beach': it is true if and only if not all beach-goers are girls. For example, it is true in the scene shown in Figure 1b, but false in the scene shown in Figure 1a. Since the 'lived on' set (the beach-goers) is not expressed as the *internal* argument of 'gleeb′' in (10b), 'gleeb′' is *not* conservative.[7] To determine whether the sentence in (10b) is true, one cannot limit one's attention to the set of girls; beach-goers who are not girls are relevant. And the crucial biconditional, which we can paraphrase as 'not all beach-goers are girls if and only if not all beach-going girls are girls', is false since the first clause can be true while the second cannot.

Our experiment will compare children's ability to learn 'gleeb' with their ability to learn 'gleeb′', based on equivalent input. Note that the conditions expressed by these two determiners are just the 'mirror image' of each other, with the subset–superset relationship reversed. By any non-linguistic measure of learnability or complexity, the two determiners seem likely to be equivalent, since each expresses the negation of an inclusion relation. Thus there is no reason to expect a difference in how easily they can be learnt—*unless* there are constraints on the semantic significance of specifically being the internal or external argument of a determiner, since this is all that distinguishes 'gleeb' from 'gleeb′'. A finding that 'gleeb' and 'gleeb′' differ in learnability would therefore be difficult to explain by any means other than such a restriction on the way internal and external arguments of determiners are interpreted.

[6] Suppose that $\mathcal{R}$ is conservative, and that $\mathcal{R}^*(X)(Y) \equiv \neg\mathcal{R}(X)(Y)$. Then $\mathcal{R}^*(X)(Y)$ is equivalent to $\neg\mathcal{R}(X)(X \cap Y)$ by the conservativity of $\mathcal{R}$, and is therefore equivalent to $\mathcal{R}^*(X)(X \cap Y)$ by the definition of $\mathcal{R}^*$. Therefore $\mathcal{R}^*$ is also conservative.

[7] The fact that 'gleeb′' happens to live on its external argument makes it anticonservative—unlike 'equi', which is neither conservative nor anticonservative—but this is not relevant here.

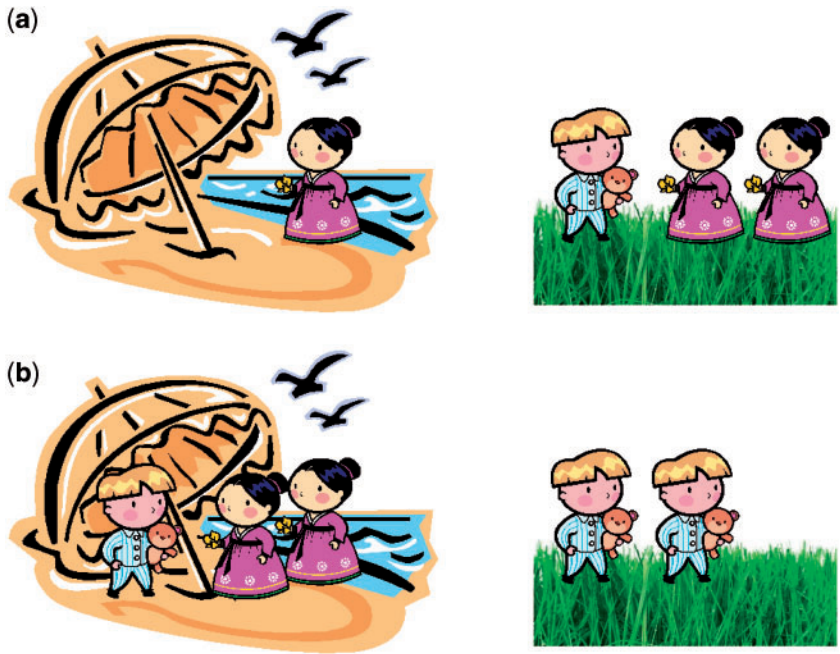


Figure 1 Two sample cards. In the conservative condition, the puppet would like only the card in (a): 'gleeb girls are on the beach' is true in (a), but false in (b). In the non-conservative condition, the puppet would like only the card in (b): 'gleeb′ girls are on the beach' is false in (a), but true in (b).

5 EXPERIMENT: CONSERVATIVITY AND LEARNABILITY

5.1 *Design and methodology*

Each participant was assigned randomly to one of two conditions: the conservative condition or the non-conservative condition. Participants in the conservative condition were trained on 'gleeb', and participants in the non-conservative condition were trained on 'gleeb′'; we then tested each participant's understanding of the determiner he/she was exposed to.

To assess the participants' understanding of these novel determiners, we used a variant of the 'picky puppet task' (Waxman & Gelman 1986). The task involves two experimenters. One experimenter controls a 'picky puppet', who likes some cards but not others. The second experimenter places the cards that the puppet likes in one pile, and the cards that the puppet does not like in a second pile. The child's task is to make a generalisation about what kinds of cards the puppet likes, and subsequently 'help' the second experimenter by placing cards into the appropriate piles.

The experimental session was divided into two phases: warm-up and target. During the warm-up phase, the experimenter ensured that the child could carry out the basic task of sorting cards into piles according to 'liking criteria': for example, in the first such warm-up item the child would be told 'The puppet only likes cards with yellow things on them', and then asked to sort a number of cards into 'like' and 'doesn't like' piles accordingly. The warm-up phase contained three items; the particular cards and the puppet's liking criterion differed from item to item.

The target phase used cards like those shown in Figure 1, and was divided into a training period and a test period. The child was told that the puppet had revealed to the experimenter whether he liked or disliked some of the cards, but not all of them. The child was told that the experimenter would sort what he/she could, but that the child would then have to help by sorting the remaining cards that the puppet was silent about. During the training period the experimenter sorted five cards, according to the criterion appropriate for the condition: in the conservative condition, the child was told that the puppet likes cards where 'gleeb girls are on the beach' (i.e. where not all girls are on the beach), and in the non-conservative condition, the child was told that the puppet likes cards where 'gleeb′ girls are on the beach' (i.e. where not all beach-goers are girls). The experimenter placed each card into the appropriate pile in front of the participant, providing either (11a) or (11b) as an explanation as appropriate.[8]

(11) a. The puppet told me that he likes this card because gleeb girls are on the beach

b. The puppet told me that he doesn't like this card because it not true that gleeb girls are on the beach.[9]

We avoided using the novel determiner in the partitive-like construction 'gleeb of the girls' because this seemed likely to bias in the direction of restricting the relevant domain to the set of girls, independently of conservativity (consider for example 'Of the girls, I have met Mary and Susan').

Having placed all the training cards (the cards that 'the puppet had told the experimenters about') in the appropriate piles, the experimenter turned the task over to the child for the test period. The experimenter handed five new cards to the child, one at a time, and asked the child to put the card in the appropriate pile, depending on whether or not the

[8] We do not distinguish between the conservative 'gleeb' and the non-conservative 'gleeb′' in writing (11), to illustrate that the explanations were homophonous across the two conditions.

[9] Negation was always expressed in a separate clause to avoid any undesired scopal interactions.

child thought the puppet liked the card. The experimenters recorded which cards the child sorted correctly and incorrectly according to the criterion used during training. The cards that the experimenter had sorted during the training period remained visible throughout the testing period.

The same training cards and the same testing cards were used in both conditions, though whether the puppet liked or disliked the card varied from one condition to the other. Table 1 shows, for each card, the number of girls and boys on the beach and on the grass, and whether each condition's relevant criterion is met or not. These were designed to be as varied as possible, while maintaining the pragmatic felicity of the two crucial target statements. The total number of characters on each card was also kept as close to constant as possible: either five or six for each card. The number of training cards that the puppet likes is the same in each condition (three), so the situation that the participant is presented with during the training phase is analogous across conditions.

The participants were 20 children, aged 4;5 to 5;6 (mean 5;0).[10] Each condition contained 10 children. Ages of those in the conservative condition ranged from 4;5 to 5;5 (mean 4;11), and ages of those in the non-conservative condition ranged from 4;11 to 5;3 (mean 5;1); the two groups did not differ significantly in age ($t = 1.4141$, $df = 18$, $p = 0.17$).

5.2 *Results*

The results indicate that children exposed to the novel conservative determiner showed significant understanding of it during the test phase, and that children exposed to the novel non-conservative determiner did not. The results are summarised in Table 2.

First we can consider how many cards children in the two conditions sorted correctly. If children never succeeded in learning the determiner's meaning, we would expect performance to be at chance. For each condition, participants were classified into six groups according to the number of test cards sorted correctly (zero to five), as shown in Figure 2,

[10] It is plausible that by this age, children are generally able to use 'real' English quantificational determiners in a manner that can be considered adult-like for the purposes of comparisons with this experimental setup. Detailed questions about their knowledge of quantificational determiners are difficult to answer, because many studies have found that they will behave in a non-adult-like manner in situations involving scalar implicatures or scopal ambiguities; see Gualmini (2003); Papafragou & Musolino (2003); Musolino & Lidz (2006), among many others. But the way in which 'gleeb' and 'gleeb′' are used in our experiment seems unlikely to involve any of these complications. Note also that whatever is responsible for the 'symmetric' interpretations discussed in Section 3 seems unlikely to interfere here since there is no plural in the determiner's external argument 'are on the beach'.

Card	beach		grass		'gleeb girls are on the beach'	'gleeb′ girls are on the beach'
	boys	girls	boys	girls		
Train 1	2	0	1	2	true	true
Train 2	0	2	3	0	false	false
Train 3	0	1	2	3	true	false
Train 4	2	3	0	0	false	true
Train 5	2	1	1	2	true	true
Test 1	3	0	0	2	true	true
Test 2	0	3	3	0	false	false
Test 3	2	3	0	2	true	true
Test 4	1	2	2	0	false	true
Test 5	1	2	0	2	true	true

Table 1 The distribution of girls and boys on each card in the experiment

Condition	Conservative	Non-conservative
Cards correctly sorted (out of 5)	mean 4.1 (above chance, $p<0.0001$)	mean 3.1 (not above chance, $p>0.2488$)
Subjects with "perfect" accuracy	50%	10%

Table 2 Summary of results

and the distribution compared with the grouping expected under the assumption of chance performance.[11] Children in the conservative condition performed significantly better than chance ($\chi^2 = 74.160$, $df = 5$, $p < 0.0001$), sorting an average of 4.1 cards correctly, whereas children in the non-conservative condition did not ($\chi^2 = 6.640$, $df = 5$, $p > 0.2488$), sorting an average of 3.1 cards correctly.

Alternatively, we can consider how many children in each condition performed 'perfectly', sorting all five test cards correctly. Of the children in the conservative condition, 5 out of 10 sorted all test cards correctly, whereas only 1 child out of 10 in the non-conservative condition sorted all test cards correctly, indicating a marginally significant dependency between conservativity of the determiner and success in learning ($p = 0.07$, Fisher's exact test).

[11] This is computed via the binomial distribution, i.e. the proportion of participants expected to give exactly k correct responses out of 5 is $\left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{5-k} \binom{5}{k} = \frac{1}{2^5}\binom{5}{k}$.

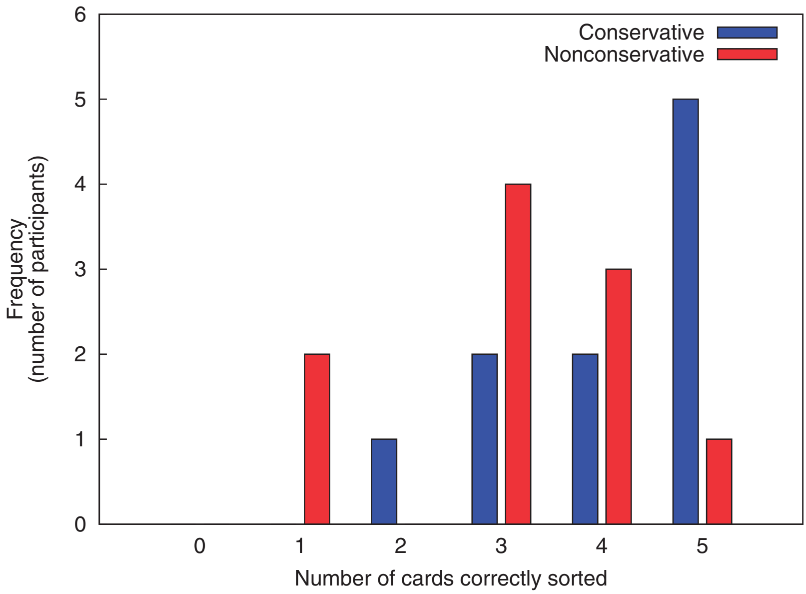


Figure 2 Distribution of participants in each condition according to how many cards were correctly sorted.

One might wonder whether participants may have ended up with above chance performance in the conservative condition 'by accident', by interpreting the novel word according to some other meaning that happens to line up reasonably well with the intended 'gleeb' on the particular test cards we constructed. The full details of the participants' responses, given in Table 3, do not support this scepticism. We have compared the responses of each participant with those that would be expected if various alternative potential interpretations are assigned to the novel word, and given the number of matching responses for each: 'all', 'none' and 'some' are the obvious determiners; 'some+' indicates the interpretation like that of 'some' but with the 'not all' pragmatic implicature enforced (i.e. 'some but not all'); and 'only' refers to the interpretation (not representable by a conservative determiner) that reverses the inclusion expressed by 'all' (i.e. 'only girls are on the beach'). The final column shows the number of responses that were consistent with the determiner the participants were trained on, either 'gleeb' or 'gleeb′' as appropriate.

The relevant question is whether there are participants in the conservative condition whose responses seem to be underlyingly driven by some incorrect (non-'gleeb') interpretation but look reasonably

Subject	Test 1	Test 2	Test 3	Test 4	Test 5	all	none	some	some+	only	gleeb
C-01	Yes	No	Yes	No	Yes	0	3	2	4	1	5
C-02	Yes	No	Yes	No	Yes	0	3	2	4	1	5
C-03	Yes	No	No	No	No	2	5	0	2	3	3
C-04	Yes	No	Yes	Yes	Yes	1	2	3	3	0	4
C-05	Yes	No	No	No	No	2	5	0	2	3	3
C-06	Yes	No	Yes	No	Yes	0	3	2	4	1	5
C-07	No	Yes	Yes	Yes	Yes	3	0	5	3	2	2
C-08	Yes	No	Yes	No	Yes	0	3	2	4	1	5
C-09	Yes	No	Yes	No	Yes	0	3	2	4	1	5
C-10	Yes	No	Yes	Yes	Yes	1	2	3	3	0	4

Subject	Test 1	Test 2	Test 3	Test 4	Test 5	all	none	some	some+	only	gleeb′
NC-01	No	Yes	Yes	No	No	3	2	3	3	4	1
NC-02	No	No	Yes	No	Yes	1	2	3	5	2	3
NC-03	Yes	No	Yes	No	Yes	0	3	2	4	1	4
NC-04	Yes	No	No	Yes	Yes	2	3	2	2	1	4
NC-05	No	No	Yes	No	Yes	1	2	3	5	2	3
NC-06	Yes	No	No	Yes	Yes	2	3	2	2	1	4
NC-07	No	Yes	Yes	Yes	Yes	3	0	5	3	2	3
NC-08	Yes	No	Yes	Yes	Yes	1	2	3	3	0	5
NC-09	No	Yes	Yes	Yes	Yes	3	0	5	3	2	3
NC-10	No	Yes	Yes	No	No	3	2	3	3	4	1

Table 3 Responses of each subject to each test card, with counts of the number of responses that are consistent with various potential meanings

consistent with 'gleeb' as a side effect. The participants for whom any non-'gleeb' interpretation had more matches than 'gleeb' itself (namely C-03, C-05 and C-07) were the three with the *lowest* scores with respect to 'gleeb' (three, three and two). So the higher scores of four and five with respect to 'gleeb' were not 'piggy-backing' on some other determiner with which the responses were actually more consistent. Because of the similarity between the target 'gleeb' and the 'some+' candidate (they differ only on one card, Test 1), high scores on 'gleeb' certainly do correlate with high scores on 'some+'; but on the one card where these hypotheses do differ, only one participant (C-07) sided with 'some+', and so this is the only participant where 'some+' has more matches (three) than 'gleeb' has (two).

This informal analysis can be verified by computing Bayes factors in order to identify the candidate interpretations that best fit the participants' response patterns. For each candidate interpretation *i*, let H_i be the hypothesis that interpretation *i* is adopted by the participant. We assume an 'error rate' p_e, such that if a participant adopts interpretation *i* then we assume that for each card the probability of sorting the card in accordance with *i* is actually only $(1 - p_e)$; without this assumption, $\Pr(D|H_i)$

would be zero for any response pattern D that contains even a single response that disagrees with i. Then for a particular response pattern D containing c_i responses in accordance with interpretation i, we can ask whether hypothesis H_i is 'substantially more supported' (Jeffreys 1961) than H_j by asking whether the Bayes factor $K_{ij} = \frac{\Pr(D|H_i)}{\Pr(D|H_j)}$ is greater than 3, where $\Pr(D|H_i) = (1 - p_e)^{c_i} p_e^{(5-c_i)}$. We can also compare hypothesis H_i with a random-guessing hypothesis H_{rand}, where $\Pr(D|H_{\text{rand}}) = (\frac{1}{2})^5$ for any D. We have no principled way of choosing p_e precisely, but for any $p_e < 0.25$ the result is that a hypothesis H_i is more substantially supported than any other (including random guessing) if and only if it is the only interpretation with $c_i = 5$. On this basis we would conclude that five participants correctly adopted 'gleeb' in the conservative condition and one participant correctly adopted 'gleeb′' in the non-conservative condition; three other participants (C-03, C-05 and C-07, already mentioned above) adopted other interpretations in the conservative condition, and four did so in the non-conservative condition. The responses of the remaining participants—two in the conservative condition, and five in the non-conservative condition—either support the random guessing hypothesis if $p_e \lesssim 0.01$, or remain unclassified otherwise. This classification of participants is illustrated in Figure 3.[12]

The results are perhaps even more telling when we look more closely at the responses of the one child who sorted all five test cards correctly in the non-conservative condition (NC-08). This child told the experimenters that the puppet was confused about which characters on the cards were boys and which were girls. Recall that in this condition the true criterion for the puppet to like a card was 'gleeb′ girls are on the beach', or equivalently 'not all beach-goers are girls'. But another statement equivalent to these is 'some boys are on the beach'. So if the child thought that the puppet intended the internal argument of the determiner in the crucial sentence to denote the set of boys, then she in fact learnt a *conservative* meaning for 'gleeb', with a meaning like 'some' has. One might even be tempted to suggest that she was led to believe that the puppet was confusing boys with girls *because of* a requirement that 'gleeb' be understood conservatively.

[12] Some more of the relevant thresholds for various settings of error rate are as follows. For $p_e < 0.25$, $K_{ij} > 3$ iff $c_i - c_j > 0$; for $0.25 \leq p_e \lesssim 0.366$, $K_{ij} > 3$ iff $c_i - c_j > 1$. For $p_e \lesssim 0.377$, H_i is substantially more supported than H_{rand} simply iff $c_i = 5$; for larger values of p_e, no H_i is ever substantially more supported than H_{rand}. In the other direction, H_{rand} is substantially more supported than H_i: (i) for $p_e \lesssim 0.011$, iff $c_i < 5$, i.e. iff any responses disagree with i; or (ii) for $0.011 \lesssim p_e \lesssim 0.125$, iff $c_i < 4$, i.e. iff more than one response disagrees with i. Note, however, that there is no participant for whom all $c_i < 4$, so H_{rand} can only be substantially more supported than all other hypotheses if $p_e \lesssim 0.011$.

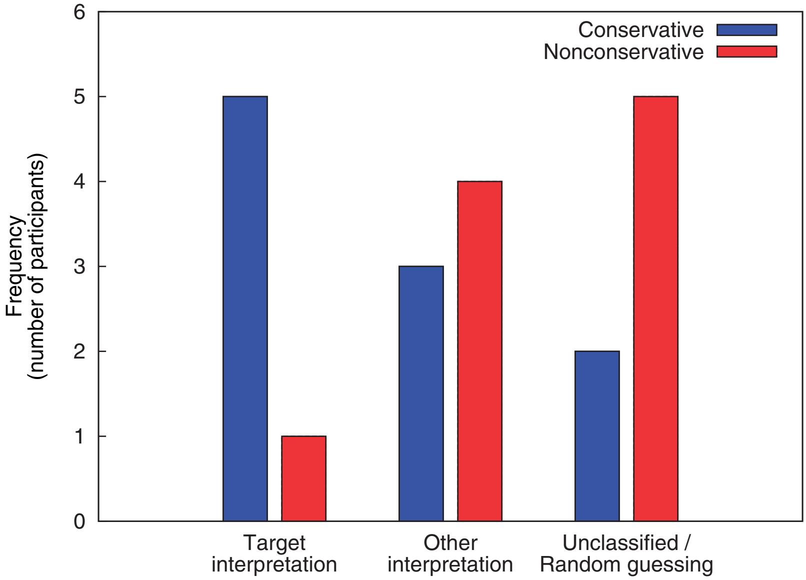


Figure 3 Classification of participants according to Bayes factors.

5.3 *Discussion and potential objections*

Here, we will consider some potential concerns and remaining open questions.

First, these results should of course only bear on the issue of determiner meanings to the extent that we are confident that the participants really did understand the relevant parts of the explanations in (11) to have the structure shown in (1). Nothing in the design of the experiment itself eliminates the possibility that the participants might have been trying to identify an interpretation for some different structural analysis of the crucial utterance, or for this utterance as an unanalysed whole. Had we found no difference between the conservative and non-conservative conditions, one might be hesitant to reject the hypothesis that determiners are restricted to conservative meanings, because of these possibilities. But it is unlikely that we would have found results consistent with the independently motivated restriction to conservative determiner meanings if participants had not been using determiner structures.

Second, taking the results to contribute to an explanation of the typology of determiners requires us to assume that the way children approached our word-learning task is relevantly similar to the way children naturally acquire the lexicon of their native language. We do not intend to claim that our participants came away from the experiment with the novel conservative determiner as a new fully fledged member of their mental lexicon, or that they could *never* learn to use the novel non-conservative determiner no matter how much training they received; and we cannot offer any explicit theory of exactly what relationship our task bears to 'natural' word-learning. Our conclusion that learnability plays a role in explaining the typological generalisation is based on the assumption that the asymmetry between children's responses to the two determiners we tested would carry over to situations of natural word-learning, but nothing in the methodology we adopted guarantees this.

Third, one might object to our inferring, from an asymmetry between 'gleeb' and 'gleeb′', that there is a general asymmetry between the class of conservative determiners and the class of non-conservative determiners. In other words, perhaps it is not the conservative/non-conservative distinction that is the underlying cause of the asymmetry we discovered between 'gleeb' and 'gleeb′', but rather some other distinction between these two novel determiners.[13] Because the two determiners were pronounced identically, an alternative would necessarily need to refer to their semantics, possibly in interaction somehow with the determiners' arguments 'girl' and '(are) on the beach'. One such alternative explanation is that participants in the conservative condition succeeded not by recognising that the puppet likes cards where GIRL$\nsubseteq$BEACH, but rather by recognising that the puppet *dislikes* cards where GIRL $\subseteq$ BEACH. In order for this idea to account for the asymmetry we found between the two conditions, there would need to be reason to believe that the non-conservative condition made it less feasible to adopt the equivalent strategy, namely recognising that the puppet dislikes cards where BEACH $\subseteq$ GIRL. Since the conditions are equally mathematically complex, an alternative explanation unrelated to conservativity would need to suppose that the difference between these two 'disliking' criteria stems from somewhere else, perhaps from the

[13] To test this possibility one would need to run some variation of our experiment with another pair of well-matched determiners, one conservative and one non-conservative. Our determiners were chosen to be the (intuitively) simplest possible: 'gleeb' is the only determiner in the 'square of opposition' (arguably the four 'simplest' determiners) that does not exist as a lexical item in English, and so experiments with other determiners would likely require significantly more training items (to allow participants to identify the intended meaning) and significantly more test items (to assess participants' conclusions). Even determiners of the form 'at least *n*' or 'exactly *n*', arguably the next simplest, will unfortunately not be suitable, since there is no non-conservative 'mirror image' of these determiners: in these cases, $\mathcal{R}(X)(Y)$ and $\mathcal{R}(Y)(X)$ are equivalent. One candidate we identified was the determiner meaning 'less than half' and its non-conservative mirror-image—but these did not give meaningful results in pilot studies we ran, presumably because of the significantly increased complexity.

ways in which *they* can be expressed in English. More specifically, one might look for an independently plausible reason for 'all girls are on the beach' (an expression of the disliking criterion in the conservative condition) to be more accessible than 'only girls are on the beach' (the disliking criterion in the non-conservative condition). In principle one might attribute this to either: (i) the fact that the 'all' sentence better matches the intended determiner syntax of the 'gleeb' sentence, since 'all' (but not 'only') is a determiner; or (ii) a simple asymmetry in these children's knowledge of the two words 'all' and 'only'.

To repeat, however, recall that any of these alternative explanations of the asymmetry between 'gleeb' and 'gleeb′' in our experiment will leave open the existing typological question of why non-conservative determiners are unattested. Results that failed to distinguish between 'gleeb' and 'gleeb′' in an experimental setting might not, one could argue, have told strongly *against* the learnability hypothesis, because of the kinds of concerns just discussed. But with the observations from our experiment and the typological generalisation both at hand, the hypothesis that conservative and non-conservative determiners differ in learnability seems appealing. Additionally, the asymmetry between children's acquiring 'gleeb' and 'gleeb′' in our experiment does not demonstrate that 'gleeb′' is completely unlearnable. Rather what we see here is simply an advantage for learning the conservative 'gleeb' over the non-conservative 'gleeb′'. It may be that conservative relations have an advantage (e.g. a higher prior probability of being a determiner meaning) without there being an absolute prohibition on non-conservative relations as determiner meanings. If this were the case, then perhaps the lack of non-conservative determiners in natural languages derives from their relative likelihood (as compared to their conservative counterparts) and not from an absolute prohibition in the formal system underlying determiner meanings. In either case, however, there would be a critical link between learnability and typology, whether that link is absolute or gradient.

Finally, we have mentioned that the typological generalisation alone arguably provides only very weak evidence for an asymmetry in the learnability status of conservative and non-conservative determiners; hence the significance of work that tries to investigate learnability more directly. The distinction between the role of typological evidence and more direct 'learnability evidence' in reaching conclusions about the language learner's hypothesis space would be particularly clearly brought out if we could identify clear cases of both (i) typologically unattested patterns that, evidence suggests, have an explanation in learnability asymmetries, and (ii) typologically unattested patterns that appear to have no such explanation. If our conclusions here are correct, then non-conservative determiners constitute an instance of the first pattern. Other recent work with children suggests a possible instance of the second pattern, also in the domain of determiner meanings: Halberda et al. (submitted for publication) report results suggesting that some children assign the meaning 'less than half' to the determiner pronounced 'most' in English. This meaning is conservative but nonetheless unattested, and is at least considered to be a possible meaning for 'most' in a particular experimental context, suggesting that it is available as a possible determiner meaning, even if not the correct meaning for the word pronounced 'most'. If correct, this would mean that the typological absence of this determiner would need some other sort of explanation, perhaps relating to pragmatic or functional pressures. We mention this here as an indicator of possible directions for future work following on from the experiment reported here, and to caution against the temptation to take the learnability asymmetry between conservative and non-conservative determiners as a foregone conclusion on the basis of the typological facts alone.

6 CONCLUSION

We have examined the relationship between learnability and typology in determiner meanings. We presented an experiment that revealed no evidence of participants successfully learning a non-conservative determiner meaning, indicating that the typological generalisation concerning conservativity derives (at least in part) from an asymmetry in learnability. This in turn gives us reason to prefer theories of natural language semantics that rule out non-conservative relations as possible determiner meanings.

Acknowledgements

Thanks to Anastasia Conroy, Larry Horn, Paul Pietroski, Alexis Wellwood and Alexander Williams for their help with this work, and to the parents and children at the Center for Young Children at the University of Maryland.

TIM HUNTER
370 Temple St
New Haven, CT 06520-08366
USA
e-mail: tim.hunter@yale.edu

JEFFREY LIDZ
1401 Marie Mount Hall
College Park, MD 20742-7505
USA
e-mail: jlidz@umd.edu

REFERENCES

Barwise, Jon & Robin Cooper. (1981), 'Generalized quantifiers and natural language'. *Linguistics and Philosophy* **4:** 159–219.

Bhatt, Rajesh & Roumyana Pancheva. (2007), 'Degree quantifiers, position of merger effects with their restrictors, and conservativity'. In Barker, Chris & Pauline Jacobson (eds.), *Direct Compositionality*. Oxford University Press. Oxford. 102–31.

Crain, Stephen, Andrea Gualmini & Paul Pietroski. (2005), 'Brass tacks in linguistic theory'. In Carruthers, Peter, Stephen, Laurence & Stephen Stich (eds.), *The Innate Mind: Structure and Contents*. Oxford University Press. New York. 175–97.

Drozd, Kenneth. (2005/2006), 'Comments from our guest editor'. *Language Acquisition* **13:**169–75.

Fox, Danny. (2002), 'Antecedent contained deletion and the copy theory of movement'. *Linguistic Inquiry* **33:**63–96.

Geurts, Bart. (2003), 'Quantifying kids'. *Language Acquisition* **11:**197–218.

Gualmini, Andrea. (2003), *The Ups and Downs of Child Language*. Ph.D. dissertation, University of Maryland.

Halberda, Justin, Tim Hunter, Paul Pietroski & Jeffrey Lidz (submitted), The language-number interface: Evidence from the acquisition of 'most'.

Heim, Irene & Angelika Kratzer. (1998), *Semantics in Generative Grammar*. Blackwell. Oxford.

Herburger, Elena. (1997), 'Focus and weak noun phrases'. *Natural Language Semantics* **5:**53–78.

Higginbotham, James & Robert May. (1981), 'Questions, quantifiers and crossing'. *The Linguistic Review* **1:**41–80.

Inhelder, Barbel & Jean Piaget. (1964), *The Early Growth of Logic in the Child*. Routledge & Kegan Paul. London.

Jeffreys, Harold. (1961), *The Theory of Probability*. Oxford University Press. New York.

Keenan, Edward L. & Jonathan Stavi. (1986), 'A semantic characterization of natural language determiners'. *Linguistics and Philosophy* **9:**253–326.

Larson, Richard & Gabriel Segal. (1995), *Knowledge of Meaning*. MIT Press. Cambridge, MA.

Mostowski, Andrzej. (1957), 'On a generalization of quantifiers'. *Fundamenta mathematicae* **44:**12–36.

Musolino, Julien & Jeffrey Lidz. (2006), 'Why children aren't universally successful with quantification'. *Linguistics* **44:**817–52.

Papafragou, Anna & Julien Musolino. (2003), 'Scalar implicatures: experiments at the semantics–pragmatics interface'. *Cognition* **86:**253–82.

Philip, William. (1991), 'Spreading in the acquisition of universal quantifiers'. In Bates Dawn (ed.), In *Proceedings of the Tenth West Coast Conference on Formal Linguistics*. CSLI Publications. Stanford, CA. 359–73.

Philip, William. (1995), *Event Quantification in the Acquisition of Universal Quantification*. University of Massachusetts. Ph.D. dissertation, Amherst, MA.

Pietroski, Paul M. (2005), *Events and Semantic Architecture*. Oxford University Press. Oxford.

Waxman, Sandra R. & Rochel Gelman. (1986), 'Preschoolers' use of superordinate relations in classification and language'. *Cognitive Development* **1:** 139–56.