Restrictions on the meanings of determiners: Typological generalisations and learnability

Tim Hunter, Jeffrey Lidz, Alexis Wellwood, and Anastasia Conroy

Introduction

We investigate children's abilities to learn various determiner meanings that are not attested in natural languages.

- If the typological absence of a determiner is due to an **inherent property of the language faculty**, we would predict that children will **fail to learn it**.
- If the determiner is absent for **other reasons**, we would predict that children will **succeed in learning it**.

The novel determiners under investigation vary along the following formal dimensions:

	First-order (Expt. 1)	Second-order/Proportional (Expt. 2)
Conservative	NALL	FOST
Non-Conservative	GRALL	GROST, GRFOST

All these determiners are easily representable in GQT [1]; if they are not all learnable, we require something more restrictive for our theory of the semantics of natural language determiners.

A determiner D is conservative iff: $D(X)(Y) \equiv D(X)(X \cap Y)$
It is well-known that all natural language determiners are conservative [2,3], but whether children entertain non-conservative determiner meanings as hypotheses during acquisition is an open question (despite some impressions to the contrary [4,5]).

The determiner we call FOST ("less than half") is unattested but conservative, and apparently minimally different from (attested, conservative) MOST. Whether FOST is learnable is also an open question.

Major findings

- **Experiment 1:** Children succeed at learning conservative NALL but fail at learning non-conservative GRALL
- **Experiment 2:** Early evidence suggests that conservativity continues to make the cut between learnable and unlearnable in the space of second-order determiners: children appear to entertain the possibility of FOST, but not GROST

These results strengthen the argument that we should strive for a theory of natural language semantics from which it follows that all determiners are conservative.

Methods

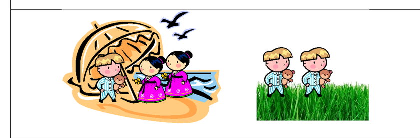
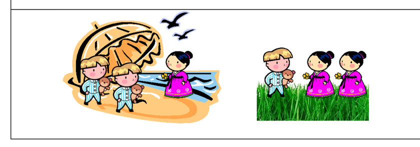
- 38 participants across 2 experiments, age 4;5-5;10; variant of the 'picky puppet' task [7]
- For both experiments, the puppet likes cards where **'gleeb girls are on the beach'**; the meaning of 'gleeb' varies
- Training cards presented where the target sentence is true (the puppet likes these cards) and false (the puppet doesn't like these cards)
- Participants attempt to sort test cards into like (true) and dislike (false) on the basis of target sentence
- Training cards: 5 for Expt.1, 6 for Expt.2
- Test cards: 5 for Expt.1, 12 for Expt.2

Experiment 1: First-Order

We compare the learnability of the following two determiners:

$$\text{NALL}(X)(Y) \equiv X \not\subseteq Y \qquad \textit{conservative}$$
$$\text{GRALL}(X)(Y) \equiv Y \not\subseteq X \qquad \textit{non-conservative}$$

Sample training cards (with truth values for DET(GIRL)(BEACH)):

	NALL GIRL $\not\subseteq$ BEACH	GRALL BEACH $\not\subseteq$ GIRL
	TRUE	FALSE
	FALSE	TRUE
	TRUE	TRUE

Participants: 20 children, ages 4;5–5;6 (mean 5;0)
Conservative condition: 10 children, ages 4;5–5;5 (mean 4;11)
Non-Conservative condition: 10 children, ages 4;11–5;3 (mean 5;1)

Experiment 1 Results

Results indicate a correlation between conservativity and success in learning ($p = 0.07$, Fisher's exact test).

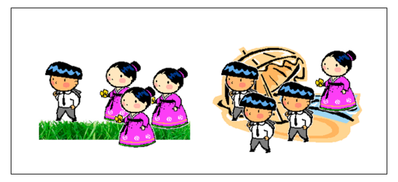
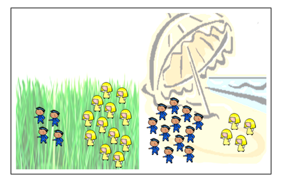
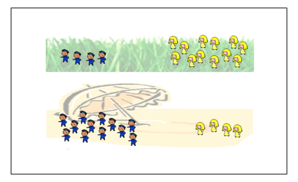
	Conservative NALL	Non-Conservative GRALL
Cards correctly sorted (out of 5)	mean 4.1 **greater than chance** (2.5) ($p < 0.001$)	mean 3.1 **not different from chance** (2.5) ($p > 0.17$)
Subjects with "perfect" accuracy	50%	10%

Furthermore: the one participant (10%) who responded with perfect accuracy in the non-conservative condition told us that the puppet had confused the girls and the boys, so this participant was evaluating the truth of '*some boys* are on the beach' (with a conservative determiner).

Adapting the Task

To ease the added burden on the participant of detecting a second-order generalisation over the training cards in Expt.2, we explored the effects of different card layouts.

- increasing the numbers of characters present, from 1, 2 or 3 (for each of the four sets) to 4, 8, or 12
- making the set of boys and the set of girls more visually salient through colour choices
- rearranging the beach/grass so that comparison between girls-on-the-beach and girls-on-the-grass is as visually salient as comparison between girls-on-the-beach and boys-on-the-beach

As these manipulations have not had any significant effects, below we report results collapsed across all three.

Experiment 2: Second-Order/Proportional

We would like to investigate the learnability of the following unattested variants of the naturally-occuring determiner 'most' [8]

	$\cap > -$	$\cap < -$
conservative	MOST$(X)(Y)$ $\lvert X \cap Y\rvert > \lvert X - Y\rvert$	FOST$(X)(Y)$ $\lvert X \cap Y\rvert < \lvert X - Y\rvert$
non-conservative	GROST$(X)(Y)$ $\lvert Y \cap X\rvert > \lvert Y - X\rvert$	GRFOST$(X)(Y)$ $\lvert Y \cap X\rvert < \lvert Y - X\rvert$

Sample training cards (with truth values for DET(GIRL)(BEACH)):

	MOST	FOST	GROST	GRFOST
	FALSE	TRUE	FALSE	TRUE
	FALSE	TRUE	FALSE	TRUE
	FALSE	TRUE	TRUE	FALSE

Participants (all in the FOST condition, so far): 18 children, ages 4;6–5;10 (mean 5;2)

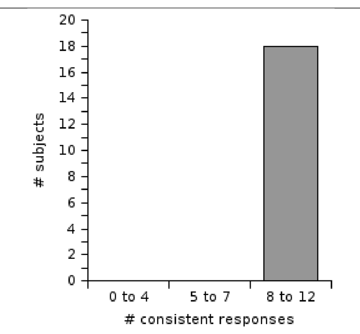
Experiment 2 Results/Discussion

Participants are not successfully learning FOST: they sort an average of 5.7 out of 12 test cards correctly (i.e. at chance).

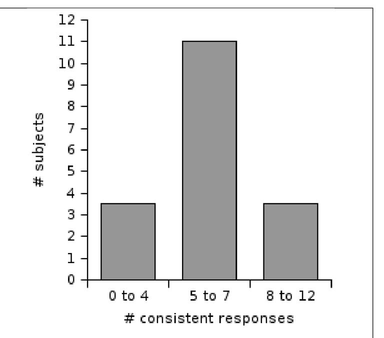
But closer analysis of their responses suggests that while participants do not detect the correct generalisation over the training cards, FOST is a hypothesised meaning they entertain — whereas GROST and GRFOST are not.

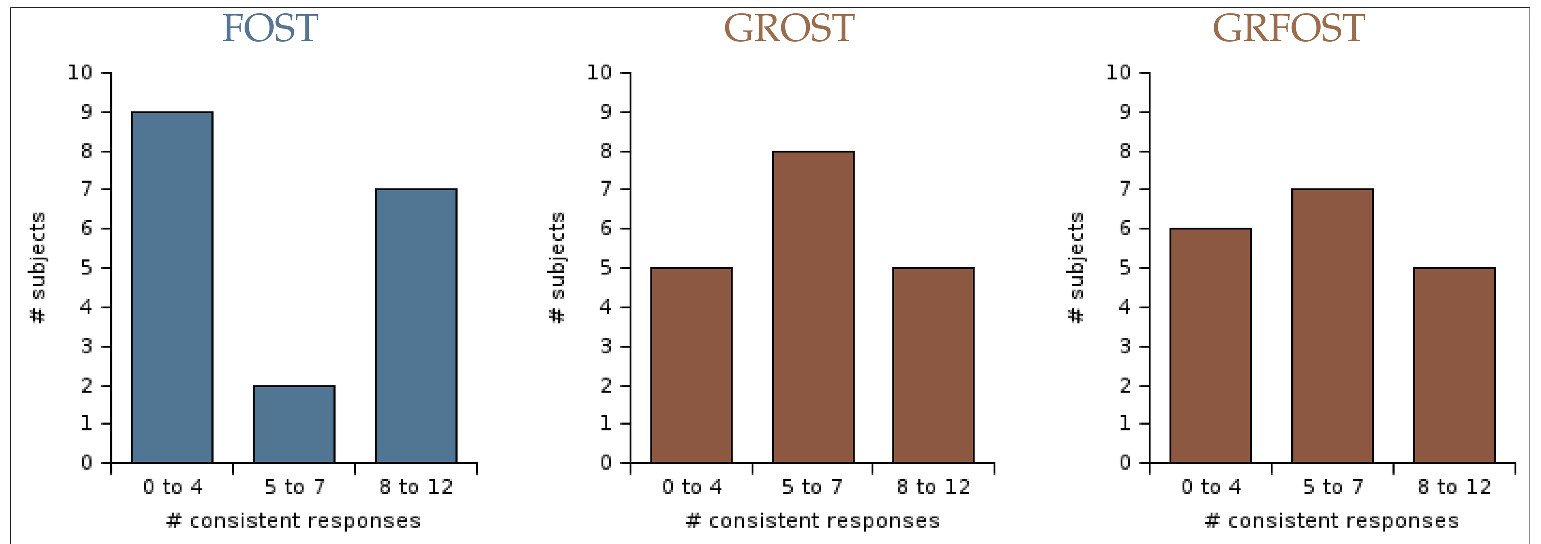
Specifically, we ask: For each participant, which determiner (if any) is his/her set of responses consistent with? Then, how many participants responded in a way consistent with FOST? . . . with GROST? . . . with GRFOST?

Predicted pattern for a determiner D if all participants respond as if 'gleeb' means D



Predicted pattern for a determiner D if all participants respond in some way unrelated to D





$\chi^2 = 21.6, df = 2, p < .0001$ $\qquad$ $\chi^2 = 2.4, df = 2, p > .30$ $\qquad$ $\chi^2 = 3.88, df = 2, p > .14$

The difference between these graphs needs *some* explanation. The following speculative story is, at least, a candidate:

- Participants are unable to determine which proportional determiner captures the correct generalisation about the training cards; but they choose some proportional determiner "at random" to sort cards on during the testing phase
- MOST and FOST are possible meanings, so are chosen relatively frequently
- GROST and GRFOST are not possible meanings, so are rarely/never chosen

References

[1] Mostowski. (1957). On a generalization of quantifiers. *Fundamenta mathematicae, 44*:12–36. [2] Barwise & Cooper. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy, 4*:159–219. [3] Higginbotham & May. (1981). Questions, quantifiers and crossing. *The Linguistic Review*, 1(1):41–80. [4] Inhelder & Piaget. (1964). *The Early Growth of Logic in the Child*. Routledge & Kegan Paul, London. [5] Crain, Thornton, Boster, Conway, Lillo-Martin, & Woodams. (1996). Quantification without qualification. *Language Acquisition, 5(2)*:83–153. [6] Hackl. (To appear). On the Grammar and Processing of Proportional Quantifiers: 'Most' versus 'more than half'. To appear in *Natural Language Semantics*. [7] Waxman & Gelman. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development*, 1:139–156. [8] Halberda, Taing, & Lidz. (2008). The age of onset of 'most' comprehension and its potential dependence on counting ability in preschoolers. *Language Learning and Development, 4(2)*: 99–121.