Distinct Verification Strategies for *Most* and *More Than Half*: Experimental Evidence for a Decompositional Analysis of Quantificational Determiners

Martin Hackl & Ben Acland
Pomona College
Martin.Hackl@pomona.edu
Ben.Acland@pomona.edu

1 Introduction

Formal semantic analyses aim to establish a systematic relation between the truth-conditional (TC) import of an expression and its syntactic/combinatorial properties. How (descriptions of) TCs are used by other systems of the mind – for instance in verification tasks – is typically not seen as something that formal semantics needs to account for or that could help distinguish between competing semantic analyses. An area where this lack of interest yields a particularly wide gap that more complete theories eventually will have to bridge is quantification.

This paper presents a novel experimental technique ("Self-paced Counting")that allows us to gather fine grained timing information about how subjects gather information incrementally in verification tasks that involve counting. We show that this technique can detect different verification profiles for semantically equivalent quantified statements and that evidence of this sort can help distinguish between competing analyses of quantifiers that are said to be indistinguishable in their TC import and their compositional commitments.

2 Limitations of Generalized Quantifier Theory

A mile stone in the development of the theory of quantification in natural language came from the study of proportional quantifiers (such as *most*, *more than half*, *two thirds*, etc.) which – as Barwise and Cooper (1981) showed for *most* – cannot be defined within first order predicate logic. They demand a strictly richer apparatus for their proper characterization and therefore set a bench mark for the expressive power found in natural language. To accommodate the expressive power displayed by proportional quantifiers Generalized Quantifier Theory (GQT) was adopted. In GQT the basic building blocks of quantificational expressions are determiners which denote relations between sets of individuals. To see the import of this move, compare the semantics of the universal quantifier *every* as given in familiar first order terms, (1)a, with its treatment in GQT where *every* denotes the subset relation, (1)b.

(1) a. $[\![\text{Every}]\!](A)(B) = 1$ iff $\forall x\, [A(x) \rightarrow B(x)]$
b. $[\![\text{Every}]\!](A)(B) = 1$ iff $A \subseteq B$

While both treatments seems equally feasible in the case of *every*, GQT proves its worth in the analysis of determiners like *most,* which cannot be defined in first order terms. I.e. there is no operator "Mx" that together with suitable Boolean operations over the two sets A and B would provide a definition for *most* in the sense of *more than half,* (2)a. GQT's assumption that determiners denote relations between sets – irrespective of whether they are morpho-syntactically simple or complex – provides an elegant solution to the puzzle, (2)b.

(2) a.*$[\![\text{most}]\!](A)(B) = 1$ iff $Mx[A(x)\ \{\&,\vee,\rightarrow,\neg\}\ B(x)]$
b. $[\![\text{most}]\!](A)(B) = 1$ iff $|A \cap B| > |A - B|$

An alternative approach, dismissed by Barwise & Cooper (1981) in passing and henceforth largely unexplored,[1] would be to include numbers and functions from (sets of) individuals to numbers among the basic building blocks for constructing natural language quantifiers. Prima facie appeal for such an approach lies in its promise to deliver a fully compositional account of the semantics of morpho-syntactically complex determiners like *more than half* which is not given in GQT.[2]

While GQT is rather successful in providing analyses of quantifiers in natural language, its success comes at a cost that is typical of systems that are too powerful. This paper focuses on one particular limitation of GQT: GQT is too coarse to distinguish between denotationally equivalent determiners even when their internal, morpho-syntactic make-up is quite different. The fact that GQT is too coarse can be seen language internally as well as through a study of verification strategies triggered by different but denotationally equivalent determiners. Both lines of argumentation converge and together argue quite forcefully for a theory of quantification that needs to derive determiner meanings based how they are internally structured.

3 Most versus More than Half

A case that illustrates GQT's insensitivity to the internal structure of determiners is provided by GQT's treatment of *most* and *more than half* as in given in (3)a-b.

[1] But see Nerbonne (1994), Krifka (1999), and Hackl (2000) among others.
[2] It is, of course, not inconceivable to develop such an account within GQT but it would have the status of an ad hoc amendment rather than that of a genuine extension since the relation between the form of a determiner and its denotation is entirely arbitrary according to GQT.

(3) a. [[more than half]] (A)(B) = 1 iff $|A \cap B| > \frac{1}{2} |A|$
b. [[most]] (A)(B) = 1 iff $|A \cap B| > |A - B|$

Under standard assumption about counting the descriptions of the TC import of the two determiners in (3) characterize exactly the same set of models. I.e. *most* and *more than half* are denotataionally equivalent. Since they are also said to be equivalent in terms of their compositional properties, they combine with two (descriptions of) sets, they are indistinguishable for GQT. This means, in particular, that the right hand sides in (3) are interchangeable for GQT even though it seems quite plausible that the TC import of *more than half* should be given as in (3)a – given the close parallel between the internal make-up of *more than half* and the functors used in (3)a – while the right hand side of (3)b is a better approxiamtion of the TC import of *most* under the assumption that *most* is the superlative of *many* (e.g. Yabushita 1998).

3.1 *Most = many+est*

Support for the idea that *most* should be analyzed as a superlative of *many* (*most = many+est*[3]) comes, inter alia, from the fact that *most* is subject to the same constraints that govern the interpretation of superlatives in general. More specifically, *most* can in principle be either understood as superlative of *many* or as a proportional determiner. However, as we show, environments that disallow the so called relative (or comparative) reading of superlatives only allow the proportional interpretation of *most*. This suggests that *most* is treated as a superlative by the grammar and calls for an analysis of proportional *most* as a special case of a superlative interpretation. We show that proportional *most* can be see as a default interpretation of the superlative of *many* that arises when *–est* compares the cardinalities of two disjoint subsets of the NP denotation that together exhaust the NP denotation. This analysis of *most* clearly favors (3)b as the better approxiamtion for the TC import of *most* over (3)a.

3.2 Distinct Verification Strategies for *Most* and *More than Half*

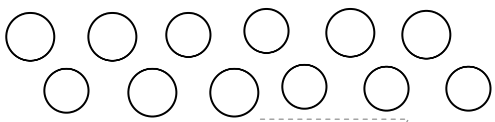
While the two descriptions of the TC import in (3)a and (3)b are equivalent it seems plausible that they are cognitively distinct. In particular, (3)a, taken literally, calls for determining half the total number of As while (3)b doesn't. Instead, (3)b requires the comparison of the number of As that are Bs to the number of As that are not Bs. Hence, assuming that these TC descriptions inform verification procedures we would expect to see different verification profiles if *most* and *more than half* are associated with different TC descriptions. More specifically, a natural verification strategy for *most* would seem to be a form of vote-counting where subjects simply keep track

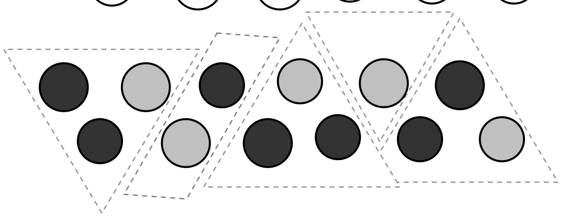
[3] Cf. Bresnan (1973)

whether for each A that is a B there is also an A that is not a B. *More than half* on the other hand should trigger a profile that is akin to checking whether the number of As that are Bs is bigger than some criterion n which represents (an estimate of) half the total number of As.

To see whether such a difference in verification strategy exists, we developed a new experimental paradigm ("Self-paced Counting“ – SPC). In a typical SPC trial subjects hear a sentence such as (4)a or b. Then they see two scattered rows of dots that are at first only outlined, (5). As subjects press the space bar the dots are uncovered in increments of 1, 2, or 3 – as indicated in (6) – while previously seen dots are masked. We hypothesize that the speed with which subjects move from one frame to the next indicates how fast they integrate the new information relative to a given verification criterion.

(4) a. More than half of the dots are blue.
b. Most of the dots are blue.

(5)


(6)


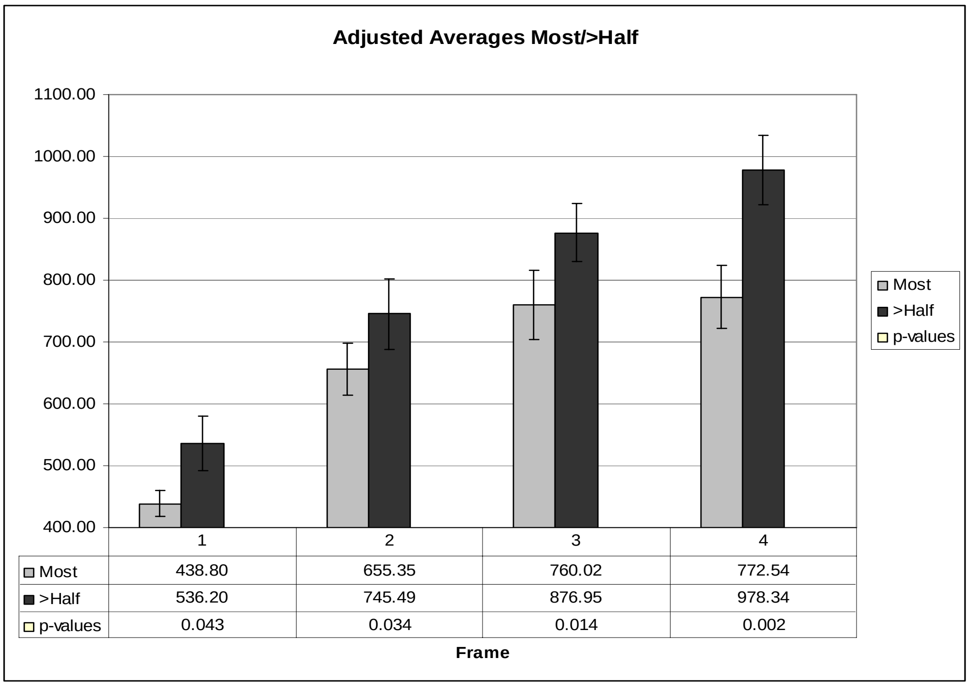
Frame 1 2 3 4 5

To show reliability of the procedure we show that the cumulative RTs as well as the RTs for moving one frame forward into the array are linear functions of n for verifying statements of the form *more than n* A B.

In the *most/more than half* experiment we varied pseudo-randomly the total number of dots across target and filler items to avoid that subjects know before hand what half of the number of dots is. To discourage strategies that rely on estimating the total amount of e.g. blue stuff vs. red stuff, we varied the size of dots in a particular color systematically so that subjects encountered on average as much blue stuff as non-blue stuff in the region of interest (up to and including frame 3).

The graph in **Error! Reference source not found.** shows the results (n = 6) up to frame 3 where it is not yet known whether the sentence is in fact true. While RTs for both *most* and *more than half* increase linearly, RTs for *most* are consistently below RTs for *more than half* and increase less. This, we argue, indicates a different verification profile from the one we see for *more than half* which is more similar to the pattern of counting to a criterion triggered by *more than n.* It is consistent with a strategy that simply performs a comparison operation at each step and keeps track of the fact whether the target color leads. I.e. each frame is evaluated in terms of whether there are more dots in the target color than there are dots that aren't and whether overall the target color leads.

(7)



References

Barwise, J. and Cooper, R. (1981). Generalized Quantifiers and Natural Language. Linguistics and Philosophy, **4**: 159-219.

Bresnan, J. (1973). Syntax of the Comparative Clause Construction in English. Linguistic Inquiry **4,3**: 275-344.

Hackl, M. (2000). Comparative Quantifiers. PhD dissertation MIT, MITWPL, Cambridge, MA.

Krifka, M. (1999). At least some Determiners aren't Determiners. In K. Turner, ed., The Semantics/Pragmatics Interface from Different Points of View. (= Current Research in the Semantics/Pragmatics Interface Vol. 1). Elsevier Science B.V., pp. 257-291.

Nerbonne, J. (1994). A Semantics for Nominal Comparatives. In Dekker, P. and M. Stockhof, eds., Proceedings of the 9th Amsterdam Colloquium, Amsterdam: ILLG, pp. 487-506.

Yabushita, K. 1998. The Unified Semantics of Mosts. Proceedings of WCCFL **18**: 320-334.