# When being included is being excluded:
# A note on complement set focus
# and the inclusion relation

ANTHONY J. SANFORD, CLAIRE WILLIAMS, and NICOLAS FAY
*University of Glasgow, Glasgow, Scotland*

Some negative quantifiers lead to pronominal reference patterns that are different from those obtained with positive quantifiers (Moxey & Sanford, 1993). This has been interpreted as meaning that the negatives give rise to a focus on the complement set (Moxey & Sanford, 1987); so, given *few of the children enjoyed the trip*, focus is on those who did not enjoy the trip. To date, this interpretation has depended on subjective judgments as to which set an anaphoric plural pronoun is referring to, allowing other interpretations of the data to be given by discourse semanticists. In two studies, we use the attachment patterns associated with the expression *including*, thereby circumventing the judgment problem. We show that a case like *not many people enjoyed the race, including John* leads to a representation in which *John* maps into the set of individuals who did not enjoy the race. We test and support the earlier claim that complement set focus is driven by denials associated with some negative quantifiers.

In a series of studies, it has been shown that positive (e.g., *a few*) and negative (e.g., *few, not many*) quantifiers create different patterns of attentional focus (Moxey & Sanford, 1987; Moxey, Sanford, & Dawydiak, 2001; Paterson, Sanford, Moxey, & Dawydiak, 1998; Sanford, Moxey, & Paterson, 1996). Focus has been inferred by investigators through obtained patterns of preferred pronominal attachment. Thus, in Sentence 1′, *they* appears to refer to the set of fans who went to the match, whereas in Sentence 2′, *they* refers to those who did not:

(1) A few of the fans went to the football match.

(1′) They enjoyed the game thoroughly.

(2) Few of the fans went to the match.

(2′) They watched it on television instead.

Pairing (1′) with (2) and (2′) with (1) yields results that are less coherent. Thus, (2′) certainly cannot follow (1) without seeming to be infelicitous, and (1′) following (2) does not sound so good as does (2′). These intuitions are borne out by experiments showing that after certain negative quantifiers, participants write continuations starting with pronouns that they claim refer to the set of whom the predicate is false (the *complement set*), whereas after positive quantifiers, this does not happen. Rather, reference here is to the subset of which the predicate of the quantified sentence is true, or the *reference set* (Moxey & Sanford, 1987,

1993; Sanford et al., 1996). The present paper is concerned with the nature of this phenomenon.

There is a practical difficulty in establishing the referent of a plural pronoun and, hence, the focus pattern associated with a quantifier, in that it is necessary to rely on the judgments of either participants or judges in order to estimate which referent is intended. This creates problems. For instance, it has been suggested that the pronoun effect, like those in (1) and (2), results from negative quantifiers' facilitating reference to the whole set, or to the quantified individuals in general. So, in (2′), rather than *they* referring to the complement set (those fans who did not go to the match), the anaphoric sentence is a shortened version of

(3) In general they watched it on TV.

This is certainly plausible for continuations to (2), like

(4) They didn't want to stand out in all that rain.

Here, it seems obvious that they *all* did not want to stand in the rain (but nevertheless some did). The generalization position has been especially espoused by followers of Kamp and Reyle's (1993) discourse representation theory, in which it is claimed that the complement set is not part of a mental representation of a quantified sentence (Corblin, 1997; Guerts, 1997; Percus, Gibson, & Tunstall, 1997).

In the present paper, we explore another way of establishing which sets are made prominent as a result of using a quantified statement, circumventing the need for the subjective judgment of referents. It is based on the *including(x,y)* relation, which maps individuals into sets. For instance:

(5) Not many of the fans went to the match, and that includes John.

Correspondence concerning this article should be addressed to A. J. Sanford, Department of Psychology, University of Glasgow, Glasgow G12, Scotland (e-mail: tony@psy.gla.ac.uk).

Did John go to the match? Our intuitions suggest that he did not. That is, *John* is part of the complement set of individuals. Because there is no way in which the attachments can be made to the superset, any preference for the complement set pattern can be taken at face value.

Using the *include(x,y)* relation, we establish whether or not complement set mapping occurs. We also explore the conditions leading to complement patterns. Using an analysis of pronominal continuations, Moxey et al. (2001) have shown that the difference in focusing between negative and positive is a result of a difference between those quantifiers that lead to affirmations and those that lead to denials, rather than resulting from semantic negativity (downward monotonicity; Barwise & Cooper, 1981), as others have suggested (Kibble, 1997a, 1997b). The properties of denial and monotonicity are briefly spelled out in the Appendix. Some formally negative, or monotone-decreasing, quantifiers (such as *at most 10*) do not appear to be denials by the set of standard tests described in the Appendix; rather, they simply affirm a state of affairs. In the present context, our hypothesis is that only denials will give rise to inclusions in the complement set.

Finally, earlier data showed an asymmetry in focus effects (e.g., Paterson et al., 1998; Sanford et al., 1996), such that positives and affirming negatives block complement reference, giving a strong reference set pattern, whereas denials give rise to a complement set preference but still allow the reference set pattern. Our final hypothesis is that this asymmetry should be reproduced with the *including(x,y)* relation.

In Experiment 1, participants saw a number of quantified clauses with the inclusion relationship and made direct judgments of whether or not the character presented in the inclusion relation should be assigned to the complement set group or to the reference set group. In Experiment 2, we investigated whether similar focus effects occur in self-paced reading.

## EXPERIMENT 1
## Direct Judgments

### Method

**Participants.** These were 20 unpaid volunteer individuals of varied age and occupation. None had taken part in a similar study.

**Materials and Design.** Six quantifiers were used, as is shown in Table 1, along with relevant information about their properties. The properties are monotonicity (a test of semantic negativity) and denial, expressed as a strength of denial index. The basis for these classifications is described in the Appendix. The denial index is an operational way of establishing the extent to which a quantified sentence (e.g., *not many people like marmalade*) is construed as introducing and denying the possibility that more than that number like marmalade. The quantifiers themselves were chosen first to represent common monotone-decreasing expressions that form denials in propositions (*not many*, *few*, and *no more than 10*). Second, a monotone-decreasing quantifier that forms assertions in propositions (*at most 10%*) was chosen. This expression forms a minimal contrast with *no more than 10%*, which is its semantic equivalent but is a denial. Third, two monotone-increasing quantifiers were used: *a few,* which forms a minimal pairing with *few*, and *no less than 10%*, which forms a pairing with *no more than 10%*. The set of expressions allows a test of the major hypotheses. A series of five statements had the following form:

> *Q* of the students had a car, and that includes Sophie.
> Did Sophie have a car?        YES        NO        DON'T KNOW

Here *Q* denotes a quantifier. The six quantifiers were paired with the five statements, creating 30 materials in all. These were arranged in four random orders and printed in four booklets.

**Procedure.** Each participant was given one of the four booklets and was invited to answer each question by circling "yes," "no," or "don't know."

### Results and Discussion

The mean value for each response assignment as a function of quantifier is shown in Table 1. Following an arcsin transformation on the scores expressed as proportions, the data for the incidence of reference set judgments were subjected to an analysis of variance (ANOVA). There was a reliable difference over conditions [$F(5,39) = 53.7$, $p < .0001$]. Specific contrasts were analyzed by $t$ tests. First, as we predicted, reference set focus occurred less often with denials (*few, not many, no more than 10%*) than with positives (*a few, at least 10%*): All six of these comparisons were reliable at $p < .001$ by $t$ tests. Second, as was predicted, denials produced fewer reference set references than did the negative yet affirming *at most 10%*: The three $t$ tests all were reliable at $p < .005$ or better.

First, participants do make the judgment that inclusion can be to the complement set. Second, negative quantifiers that form denials lead to the complement set focus pattern. Finally, there is an asymmetry in focus patterns: Denial-forming quantifiers do not rule out reference set patterns, whereas those that do not form denials almost block complement attachment. In Experiment 2, we investigated the

**Table 1**
**Quantifiers Used in Experiment 1 and Their Properties**

| Quantifier | Monotone | Deny/Affirm | Denial Index | Mean Refset | *SE* Refset |
|---|---|---|---|---|---|
| Not many | Dec | deny | .98 | 0.56 | 0.20 |
| Few | Dec | deny | .90 | 2.40 | 0.37 |
| No more than 10% | Dec | deny | .40 | 2.95 | 0.40 |
| At most 10% | Dec | affirm | .10 | 4.05 | 0.37 |
| A few | Inc | affirm | .04 | 4.65 | 0.18 |
| At least 10% | Inc | affirm | .02 | 4.90 | 0.65 |

Note—The rightmost columns show the results for Experiment 1. The mean proportion of refset references is out of a maximum of 5. The Appendix explains monotone increasing/decreasing, denial, and the calculation of the denial index. Dec, decreasing; Inc, increasing; Refset, reference set.

focus effects observed with the *including* relation by using a self-paced reading paradigm.

# EXPERIMENT 2
## Reading Time Measures

## Method

**Participants.** There were 48 participants who were paid £4 for a session that lasted approximately 20 min. They were undergraduates of the University of Glasgow and were naive with respect to the aims of the experiment. They were randomly assigned to three groups of 16 for the three subexperiments making up Experiment 2.

**Design and Materials.** A total of 32 passages was created, based on earlier materials from Sanford et al. (1996, Experiment 3). An example material is shown in (6):

(6)  Not many of the MPs attended the meeting, and that includes John. His presence/absence helped the meeting to run smoothly. (TARGET) Did the meeting run smoothly?

For each passage, the six quantifiers were used as options, with two target conditions: reference set focus (e.g., *presence*, referring to the set for which the predicate of the first sentence holds) or complement set focus (e.g., *absence*, referring to the set for which the predicate of the preceding sentence does not hold). Each material was followed by a simple comprehension question. In order to prevent the requirement of many more materials and to guarantee that the participants would not be exposed to too many trials of the same type, the experiment was broken down into three subexperiments, A, B and C, which were run with independent groups of participants. In 2A, the quantifiers *few* and *a few* were used, providing a test of negative denial versus positive. In 2B, *at most 10* and *not many* were compared, contrasting negative affirmation with negative denial. In 2C, *no more than 10* and *at least 10* compared denial versus positive affirmation.

For a given subexperiment, four materials files were created, with each material appearing only once in a file. There were 8 materials in each of the four (2 quantifiers × 2 target type) conditions. By rotation over the four files, each material appeared in each of the four conditions. Any participant saw only one of the files. In addition to the 32 experimental materials in a file, there were 32 fillers unre-

lated to the experiment. Materials within a file were presented in quasirandom order.

**Procedure.** The participants were allocated to Subexperiment A, B, or C and were presented with one of the four files. Materials were presented a sentence at a time on a VDU. In order to see each sentence, the participants pressed the center button on a button box. They were instructed to read each sentence at their own pace, but basically as quickly as they could, compatible with comprehension.

## Results

**Error rates on questions.** The error rate on questions was 8% overall, evenly split over conditions, indicating that the participants were performing the reading task adequately.

**Experiment 2A.** Mean reading times are shown in Figure 1A. For *a few*, complement targets were read much more slowly than reference set targets. There was a very weak trend in the opposite direction for *few*. ANOVAs showed that there was a main effect of continuation, with reference set targets being read faster than complement set targets [$F_1(1,15) = 17.28, p < .0008; F_2(1,31) = 11.05, p < .0023$]. More important, and as predicted, there was reliable interaction between target type and quantifier [$F_1(1,15) = 24.74, p < .0002; F_2(1,31) = 34.00, p < .0001$]. Simple main effects tests showed that the target effect for *a few* was reliable at .0001 for both $F$s but that, for *few*, it was not. We may conclude that there is a focusing difference between the quantifiers: *A few* puts focus clearly on the reference set, whereas focus is diffuse between complement and reference sets for *few*. This is consistent with what we expected, showing an asymmetry in the extent of focusing between the two quantifiers.

**Experiment 2B.** Figure 1B shows a predicted crossover interaction, with *not many* giving a clear reading time advantage for complement targets and *at most 10%* giving an equally strong reading time advantage for the reference
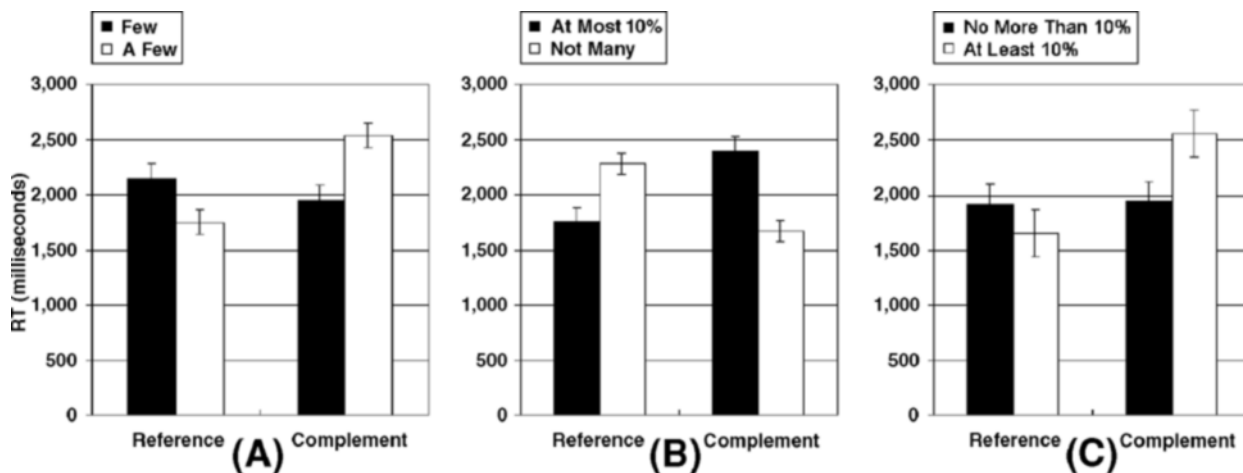


**Figure 1. Mean reading times (RTs, in milliseconds) for target sentences in Experiment 2, with standard error bars. Target sentences were consistent with either the reference set or the complement set. A, B, and C refer to the corresponding subexperiments.**

set. The interaction is reliable [$F_1(1,15) = 32.37, p < .0001$; $F_2(1,31) = 60.27, p < .0001$].

**Experiment 2C.** Figure 1C shows that the positive quantifier *at least 10%* gives a clear reading time advantage for reference set targets over complement targets, whereas the negative quantifier *no more than 10%* produces no advantage for one over the other. ANOVAs show a reliable interaction between quantifier and target type [$F_1(1,15) = 19.14, p < .0005$; $F_2(1,31) = 12.37, p < .001$]. This pattern is similar to that in Experiment 2A.

Additional simple main effects analyses show that the reading times for targets following *no more than 10%* lie toward the middle of the range of times obtained for reference and complement set targets for *at least 10%*. Thus, the time for reading a reference set target in Experiment 2C for *at least 10%* was faster than it was for targets following *no more than 10%*, a difference that is nearly reliable by participants [$F_1(1,15) = 4.442, p < .052$] and is reliable by materials [$F_2(1,31) = 9.67, p < .004$]. The reading time for a complement set target for *at least 10%* was slower than it was for *no more than 10%* [$F_1(1,15) = 23.68, p < .0002$; $F_2(1,15) = 23.68, p < .0001$]. The pattern of these results match those found in Experiment 2A.

## GENERAL DISCUSSION

The *including(x,y)* relation gave rise to expected patterns of attachment. There was clear evidence for complement set attachment and for negative expressions that form denials giving rise to less reference set attachments and more complement set attachments than did either positives or negatives that were affirmations (Experiment 1). Furthermore, in Experiment 2, these attachments were found to influence the processing time for subsequent targets that took attached individuals and presented information that portrayed them as being in either the reference set or the complement set. Positives, and negatives that formed affirmations, produced attachments that were consonant with focus on the reference set, whereas negatives that formed denials either showed a preference for the complement set pattern (with *not many*) or were indifferent as to which set was preferred (with *few* and *no more than 10%*). Because there is no way in which the attachments can be made to the superset (an individual has to be in either the complement set or the reference set), the results support the view that complement set reference occurs as a result of negative denials. The studies overcome the problems of subjectivity in the classification of referents to plural continuations, used in earlier work (e.g., Sanford et al., 1996). They suggest that complement reference is a reality and that such reference cannot be attributed to reference to the superset (e.g., Corblin, 1997). Although it was shown in earlier work that focus patterns influence self-paced reading times (Sanford et al., 1996, Experiment 3) and eyetracking patterns (Paterson et al., 1998), the effects could still have been due to reference to the superset (Percus et al., 1997). The present experiments circumvented these problems.

The present data are broadly consistent with the inference theory of complement focus (Moxey et al., 2001; Sanford et al., 1996) that attributes complement focus patterns to denial. However, the differences between quantifiers in the denial group require further comment. They did not all give rise to equally strong complement reference focus. In terms of both direct judgments (Experiment 1) and reading times (Experiment 2), *not many* is the strongest complement set focuser. *Few* and *no more than 10* are less strong and appear to be equivalent in their effects and almost indifferent in their preference toward complement or reference focus patterns. Note that this ordering shows some correspondence to the denial index data, insofar as *not many* leads to the strongest denial index and also shows the strongest mapping effects. (However, the larger difference in denial index between *few* and *no more than 10* is not reflected in the attachment patterns.) At present, there is no explanation for the differences, which constitute simple empirical observations. However, the rough correspondence with denial promises the possibility of internal consistency for the claim that it is denial that controls complement focus.

Turning to the empirical observations themselves, the finding that the quantifiers *few* and *no more then 10* gave rise to attachment patterns that did not show a particular bias toward complement set or reference set may be explained in (at least) two quite different ways. First, it is possible that, on some occasions, they cause an attachment in one direction and, on other occasions, a bias in the other direction. This should lead to the larger variances of reading times in Experiment 2 for the two quantifiers in question than for the others, which have more clearly marked focus patterns. Examination of the variances suggests that this is not the case (see the standard errors in Figures 1A, 1B, and 1C). A second possibility is that these two quantifiers do not give rise to fully instantiated (realized) mappings but that the mapping onto sets is underspecified until it is needed (either as a judgment, as in Experiment 1, or for the purposes of integration, as in Experiment 2). The results of Experiment 2 fit this, in that the times for targets following *few* and *no more than 10%* are slower than the times for targets consistent with the most clearly focusing expressions but are faster than the times for targets that are inconsistent with the patterns of the most focusing expressions. This suggests that they occupy a middle ground in terms of focusing and will accommodate reference set and complement set focusing with equal ease or difficulty.

Dowty (1994) has suggested that negative quantifiers do not give rise to classical anaphoric possibilities. The results with *few* and *no more than 10%* would appear to support such a view, especially if the underspecified attachment argument given above is correct. However, *not many* shows a very clear complement set bias. In fact, the denial model is consistent with both of these patterns and does not disconfirm Dowty's position. The results with *at most 10%*, which shows an overwhelming reference set preference, is more of a problem if it is downward monotonicity that

is supposed to lead to a lack of anaphoric reference. Here, it is not possible to invoke denial as a mechanism because *at most 10%* is an affirmation. Together, the results are most consistent with the claim that if monotone-decreasing expressions are also denials, they create either a complement set focus or an underspecified pattern of attachment that readily accommodates complement attachment. It is not monotonicity alone that determines either reference pattern (as was suggested by Kibble, 1997a, 1997b) or lack of preferences for anaphoric patterns (Dowty, 1994).

The *including(x,y)* relation is a viable way of examining the prominence of sets during the processing of quantified sentences. The experiments show that complement set attachment occurs readily and, in the case of *not many*, very markedly. The results cannot be explained in terms of general or superset attachment. However, in the cases of *few* and *no more than 10%*, both complement and reference attachment may be enabled as possibilities, but a definite attachment not be made until needed.

## REFERENCES

BARWISE, J., & COOPER, R. (1981). Generalised quantifiers and natural language. *Linguistics & Philosophy*, **4**, 159-219.

CLARK, H. H. (1976). *Semantics and comprehension*. The Hague: Mouton.

CORBLIN, F. (1997). Quantification et anaphore discursive: La reference aux complimentaires. *Languages*, **123**, 51-74.

DAWYDIAK, E. (2001). *The inference model of quantifier focus*. Unpublished doctoral dissertation, University of Glasgow.

DOWTY, D. (1994). The role of negative polarity and concord marking in natural language reasoning. In M. Harvey & L. Santelmann (Eds.), *Proceedings of the 4th Annual Conference on Semantics and Linguistic Theory* (Vol. 4, pp. 114-144). Ithaca, NY: Cornell University, Department of Linguistics.

GUERTS, B. (1997). Review of L. M. Moxey and A. J. Sanford, (1993): *Communicating quantities. Journal of Semantics*, **18**, 87-94.

HORN, L. R. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

KAMP, H., & REYLE, U. (1993). *From discourse to logic: Introduction to model theoretic semantics of natural language, formal logic, and discourse representation theory*. Dordrecht: Kluwer.

KIBBLE, R. (1997a). Complement anaphora and dynamic binding. In *Proceedings of the 7th conference of SALT (semantics and linguistic theory)*. Ithaca, NY: Cornell University, Department of Modern Languages and Linguistics.

KIBBLE, R. (1997b). Complement anaphora and monotonicity. In G. J. M. Kruijff, G. V. Morrill, & R. T. Oehrle (Eds.), *Proceedings of Formal Grammar Conference* (pp. 125-136). Aix-en-Provence.

KLIMA, E. S. (1964). Negation in English. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language* (pp. 246-323). Englewood Cliffs, NJ: Prentice-Hall.

MOXEY, L. M., & SANFORD, A. J. (1987). Quantifiers and focus. *Journal of Semantics*, **5**, 189-206.

MOXEY, L. M., & SANFORD, A. J. (1993). *Communicating quantities*. Hove, U.K.: Erlbaum.

MOXEY, L. M., SANFORD, A. J., & DAWYDIAK, E. (2001). Denials as controllers of negative quantifier focus. *Journal of Memory & Language*, **44**, 427-442.

PATERSON, K. B., SANFORD, A. J., MOXEY, L. M., & DAWYDIAK, E. J. (1998). Quantifier polarity and referential focus during reading. *Journal of Memory & Language*, **39**, 290-306.

PERCUS, O., GIBSON, T., & TUNSTALL, S. (1997, March). *Antecedenthood and the evaluation of quantifiers*. Poster presented at the 10th CUNY conference, Santa Monica, CA.

SANFORD, A. J., MOXEY, L. M., & PATERSON, K. B. (1996). Attentional focusing with quantifiers in production and comprehension. *Memory & Cognition*, **24**, 144-155.

ZWARTS, F. (1996). Facets of negation. In J. van der Does & J. van Eijk (Eds.), *Quantifiers, logic and language* (pp. 39-68). Stanford University, Center for the Study of Language and Information.

ZWARTS, F. (1998). Three types of polarity. In E. Hinrichs & F. Hamm (Eds.), *Plural quantification* (pp. 177-238). Dordrecht: Kluwer.

## APPENDIX
## Background to the Quantifier Classification

The experiments depend on a dissociation of two aspects of negativity: downward monotonicity, a formal, logical property of what it takes to make an expression minimally negative (Zwarts, 1996), and denial, a property associated with negativity in the monotonicity sense, but not necessary for it.

*Monotonicity*. Quantifiers are weakly negative if they are monotone decreasing in their left arguments (Zwarts, 1996; see Barwise & Cooper, 1981, for a definition). One test used for downward monotonicity is to put the quantifiers into the following frame: *If Q of the class came to the party, then q of the class came to the party early*. If this seems to be true, the quantifier is monotone decreasing. The positive expressions are monotone increasing. That is, they come out as producing true statements of the following form: *If Q of the class came to the party early, then Q of the class came to the party*. Since it is (analytically) true that *if at most 10% of the class came to the party, then at most 10% of the students came to the party early*, we may conclude that *at most 10%(x)* is monotone decreasing. A further test is based on whether declarative sentences containing a quantifier allow the inclusion of a negative polarity item, such as *anymore* or *give a damn* (Zwarts, 1996, 1998). For instance, *few of the students care anymore* is acceptable, whereas *a few of the students care anymore* is not. On this test, *at most 10% of students care anymore* is acceptable; thus, *at most 10% of X* is monotone decreasing by this test too.

*Denial*. Denial is feature of negation (Clark, 1976; Horn, 1989). Thus, *not many children like math* denies the possibility that more might like math. Not all quantifiers that are monotone decreasing lead to denials in propositions. Clark suggested that denials might be tested for, using a series of tests that Klima (1964) had identified as relating to negation. Moxey et al. (2001) used these tests as diagnostics when they examined the idea that denials are behind complement set focus. In this way of thinking, denial is indicated by the acceptability of a positive tag question, as in (A1), but affirmation goes with a negative tag question, as in (A2):

(A1) Not many people like cheese, do they?/ *don't they?
(A2) At most 10 people like the soup, don't they?/*do they?

## APPENDIX (Continued)

The second example is the *either* versus *too* tag:

(A3) Not many men are happy and not many women are happy either/ *too.

(A4) Many men are happy and many women are happy too/ *either.

The use of *either* connects two denial statements, whereas *too* connects two affirmations. A final test is the *neither*-versus-*so* case. *So* connects affirmations, whereas *neither* connects denials:

(A5) Many men are happy, and so/*neither are the women in the group (so, *many* does not form a denial).

(A6) Not many men are happy, and neither/ *so are the women in the group (so, *not many* forms a denial).

Dawydiak (2001) tested respondents̓ judgments on these tags for all of the quantifiers used in the present experiments. The denial index of each quantifier in Table 1 is a sum of all of the denial judgments divided by the total number of judgments for each quantifier and provides an empirical basis for the division of expressions into denials and affirmations. On this basis, *at most 10%* is monotone decreasing, but not a denial, whereas *not many* is monotone decreasing and *is* a denial.