BART GEURTS

DONKEY BUSINESS

ABSTRACT. In this paper I present experimental data showing that the interpretation of donkey sentences is influenced by certain aspects of world knowledge that seem to elude introspective observation, which I try to explain by reference to a scale ranging from prototypical individuals (like children) to quite marginal ones (such as railway lines). This ontological cline interacts with the semantics of donkey sentences: as suggested already by the anecdotal data on which much of the literature is based, the effect of world knowledge is by and large restricted to donkey sentences with non-intersective determiners. I outline a psychological model which incorporates both ontological and logical factors, and suggest that there may be something wrong with the standard assumption that a statement's receiving a truth value requires that it have a definite reading.

1. INTRODUCTION

What, if anything, do donkey sentences mean? Medieval logicians, according to Geach (1962), would have said that (1) has the same truth conditions as (2a). Geach argued, however, that these are too weak, and that (2b) gives a better paraphrase.[1]

(1) Every farmer who owns a donkey beats it.

(2)a. Every farmer who owns a donkey beats at least one of the donkeys he owns. (∃-reading)

b. Every farmer who owns a donkey beats all of the donkeys he owns. (∀-reading)

When other semanticists began discussing sentences like (1), it soon became clear that Geach's judgment is not universally shared. For, as Parsons (1978) may have been the first to note, when presented with situations satisfying (2a) but not (2b), quite a few informants hesitate to call (1) false. Furthermore, Parsons observed, although for many speakers (1) prefers a

[1] Actually, Geach's example has *any* instead of *every*, which doesn't seem to make much of difference, though. The terms '∃-reading' and '∀-reading' are due to Chierchia (1992, 1995).

Linguistics and Philosophy **25:** 129–156, 2002.

∀-reading, there are cases in which the weaker reading is clearly preferred; for example:

(3)a. Every guest who had a credit card used it to pay his hotel bill. (∃-reading preferred)

b. Every guest who had a credit card kept it in his wallet. (∀-reading preferred)

There are two factors that determine the preferred interpretation of a donkey sentence: the initial determiner and world knowledge; and these factors interact in ways that still remain to be clarified. Following Kanazawa (1994), I assume that the main generalizations are the following:

(4)a. Donkey sentences with weak determiners only have ∃-readings.

b. Donkey sentences with universal determiners prefer ∀-readings, but may have ∃-readings, too.

Since donkey sentences with *not every* and *not all* follow the same pattern as donkey sentences with *every* and *all*, it will be convenient to have one label subsuming them all; I use the term 'universal determiner' for this purpose, and that is how (4b) is to be understood. It has been argued by Krifka (1996) that *not every* and *not all* aren't determiners and that 'not every A B' is to be parsed as 'not (every A B)' rather than '(not every) A B'. I believe Krifka is right about this (and will provide supporting evidence in the postscript to this paper), but as for my current purposes nothing hinges on this point, I will for the most part pretend as if *not every* and *not all* are determiners.

Weak determiners are *some*, *a few*, *at least* n, *at most* n, (*exactly*) n, *no*, and perhaps *few* and *many*.[2] (4a) may be a bit too strong, because there are donkey sentences with weak determiners that seem to prefer ∀-readings:

(5)a. No man who had a credit card failed to use it. (Kanazawa 1994: 117)

b. At least one boy who had an apple for breakfast didn't give it to his best friend. (van der Does 1993, as cited by Chierchia 1995: 65)

[2] I am being a bit cautious here, because *few* and *many* are notoriously difficult (see Lappin 2000 for recent discussion). In donkey sentences they seem to pattern with the weak determiners (as they do in existential sentences, for example), but the analysis I will propose requires that weak determiners be intersective, and it is not clear that *few* and *many* have this property, although they appear to have it in some of their uses.

Such examples are hard to find, though, and always seem to involve some form of negation: typically, the main determiner is *no*, as in (5a), and if it isn't the predicate contains a negation operator, as in (5b). I don't know of any unmarked examples with *some* (say) which elicit ∀-readings as easily as universal donkey sentences elicit ∃-readings when the context is right (as witness (3a)).[3] So even if (4a) overstates the case, there is a distinct asymmetry between donkey sentences with weak determiners, on the one hand, and universal determiners, on the other. This asymmetry is also borne out by experimental evidence, as we will see below.

Between them, (4a) and (4b) cover most basic determiners, save for *most* (but see footnote 2). The reason for this is that intuitions about donkey sentences with *most* are especially problematic, as observed by Heim (1982), Rooth (1987), Kamp (1991), Kanazawa (1994), and others. It seems to me that donkey sentences with *most* tend towards ∀-readings, but even if this is so, the tendency is decidedly less pronounced than in the case of universal donkey sentences. By and large, Heim's (1982: 62) assessment that donkey sentences with *most* have 'rather nebulous truth conditions' remains valid, and therefore they will be left aside in the following.

The observations in (4) suggest that the interpretation of donkey sentences is a delicate interplay between meaning and subject matter, and as it seems unlikely that introspective data will afford us a clear view on this process, experimental data are called for. The purpose of this paper is to present and analyse the results of an experiment carried out some years ago (in the first half of 1997, to be more precise), with a view to clarifying the interaction between the semantic properties of determiners and world knowledge. As we will see, this experiment corroborates the observations in (4), but it also throws new light on the role played by world knowledge, in that it suggests that it is not just considerations of plausibility that affect the interpretation of donkey sentences (as one might expect on the basis of examples like (3a, b)): in addition world knowledge enters the picture in a rather more fundamental way, which doesn't seem to be open to introspection.

I am not the first to have gathered experimental evidence on the interpretation of donkey sentences. Previously, experiments were carried out by Yoon (1994) and Conway and Crain (1995). However, since the latter study is about children's understanding of donkey sentences, and my concern is with adult interpretations, I will only consider Yoon's contribution here (see also Krifka 1996 for discussion of this work). Yoon's study is based on the hypothesis that the interpretation of a donkey sentence may be affected by the content of the predicate, and she focuses her attention on two

[3] These remarks are expanded upon in the postscript to this paper.

factors in particular. One factor is the distinction between total and partial predicates. If a total predicate is applied to a collection of individuals there is a tendency to infer that the predicate applies to all members, whereas with a partial predicate, a weaker inference is preferred. The following examples illustrate this distinction:

(6)a. *total:* The glasses are clean ≫ All the glasses are clean.
b. *partial:* The glasses are dirty ≫ At least some of the glasses are dirty.

Yoon's expectation was that the interpretation of donkey sentences, too, is affected by the total/partial distinction. A further distinction she expected to have an effect is that between stative and episodic predicates, as illustrated by the following:

(7)a. *stative:* The children are five years old ≫ All the children are five years old.
b. *episodic:* The children are making a toy plane ≫ At least some of the children are making a toy plane.

In order to test these predictions, subjects were presented with situation descriptions, and were subsequently asked to judge whether a given sentence correctly described the situation in question. For example, it was expected that subjects would tend to accept (6b) in a situation in which only some of the glasses are dirty, and that (8) would be accepted in a situation in which every boy who had a baseball card in his pocket soiled at least one (though not necessarily all) of his baseball cards (*soil* being a partial predicate):

(8) Every boy who had a baseball card in his pocket soiled it while playing in the mud.

The collection of materials used by Yoon is exemplified by the following:

(9) $\left\{\begin{array}{c}\text{Every}\\ \text{No}\end{array}\right\}$ boy who $\left\{\begin{array}{c}\text{had}\\ \text{has}\end{array}\right\}$ a toy car $\left\{\begin{array}{ll}\text{kept it in good order} & \text{(total)}\\ \text{broke it} & \text{(partial)}\\ \text{likes it} & \text{(stative)}\\ \text{let his friend play with it} & \text{(episodic)}\end{array}\right.$

Yoon's results are summarized in Table I. Note that the preferred interpretations of simple sentences with plural subjects vary with their predicates, as expected, and that donkey sentences with *every* follow the same

TABLE I

% ∀-readings for simple and quantified sentences (after Yoon 1994)

	Determiner		
Predicate	The	Every	No
Total	84	74	52
Partial	18	22	20
Stative	90	72	24
Episodic	28	16	22

pattern: ∀-readings are predominant with total and stative predicates, while ∃-readings are clearly preferred whenever the predicate is either partial or episodic. In donkey sentences with *no*, by contrast, the pattern of responses is quite different. With *no* there is a clear overall preference for ∃-readings, which is affected only by total predicates, and even in this case the effect is much weaker than in the other two sentence types.

These results should not be taken at face value, for two reasons. First, it may be that Yoon's stative/episodic distinction is not entirely unproblematic, as indeed certain remarks in Yoon (1996) confirm. For example, I should say that the stative predicate *clever* is compatible with a weak reading on at least some occasions, whereas episodic *ready* may favour a strong reading, as witness:

(10)a. The children are clever.
b. The children are ready.

Despite the fact that *clever* is stative, (10a) might be uttered truthfully even if not quite all children are clever, whereas (10b) suggests rather strongly that all the children are ready, although *ready* is episodic. This being said, it should be noted that, whatever Yoon's stative/episodic distinction comes down to, it only affected the interpretation of donkey sentences with *every*, and had no effect on donkey sentences with *no*.

A second, and more serious, problem with Yoon's results is that her data for *no*-sentences with total predicates may be an experimental artifact. To show this, I list here four samples of Yoon's materials (emphasis added):

(11) *claim:* No boy who has a toy car *likes* it. (stative)
situation: We have 5 boys. 3 of them each have one toy car and none of them likes it. The other 2 each have 4 toy cars, and each *likes* 2 of his 4 toy cars while not liking the other 2.

(12) *claim:* No boy who had a toy car *let his friend play with it.* (episodic)
situation: We have 5 boys. One of them had one toy car and didn't let his friend play with it. The other 4 each had 2 toy cars, and each *let his friend play with* one of his toy cars while not letting him play with the other toy car.

(13) *claim:* No boy who had a toy car *kept* it in good order. (total)
situation: We have 5 boys. 2 of them each had one toy car and each broke it. The other 3 each had 3 toy cars, and each *broke* 1 or 2 of his toy cars.

(14) *claim:* No boy who had a toy car *broke* it. (partial)
situation: We have 5 boys. 3 of them each had one toy car and none of them broke it. The other 2 each had 2 toy cars, and each *broke* one of them while not breaking the other.

The crucial observation is that (13) deviates from the others not only in that it features a total predicate, but also in that its description focuses on the cases that *confirm* the sentence; that there are disconfirming cases, as well, remains implicit, and must be inferred. In the other three tasks it is the other way round: in each of these the story puts the disconfirming cases into the foreground. Since all of Yoon's materials for *no* follow this pattern, it doesn't come as a surprise that subjects had greater difficulty seeing that the sentence with the total predicate is actually false. Hence, Yoon's results give no reason for doubting that the interpretation of *no*-sentences is essentially independent of the content of the predicate, or in other words, they don't contradict (4a).

If we want to probe native speakers' intuitions about the truth conditions of donkey sentences, we must present them with situations containing a mixture of positive and negative evidence. The foregoing observation suggests that it can matter a great deal *how* this evidence is presented. In Yoon's experiment negative evidence was inadvertently de-emphasized because it had to be inferred from the way the positive evidence was presented, and this may have had an effect on her data. This is a more general problem, of course, since the bulk of the literature on donkey

sentences is based on introspective verdicts on sentences about described situations, and describing a situation virtually entails putting emphasis on some aspects at the expense of others. Consider, for example, the following description:

> [Suppose] John owns ten donkeys and beats exactly nine of them, and that every other man beats every donkey he owns. Then is ['Every man who owns a donkey beats it'] true? (Rooth 1987: 253)

Rooth notes that informants tend to give 'varied and guarded judgements about this case.' (1987: 254) But what would happen if the situation description (i) actually *mentioned* the one donkey John owns but does not beat, or (ii) *stated* that John doesn't beat this donkey, or even (iii) *emphasized* that John doesn't beat this donkey? It is likely, I believe, that such changes would affect informants' verdicts, though strictly speaking we just don't know, because this possibility has never been investigated. At any rate, as long as the verbal medium is used for presenting states of affairs, such effects are difficult to rule out.

2. THE DONKEY EXPERIMENT

Considering that verbal descriptions may have unnoticeable effects on subjects' judgments, I decided to try an experiment which didn't describe but picture situations.[4] Otherwise my set-up was essentially the same as Yoon's. Twenty native speakers of Dutch were asked to judge whether or not donkey sentences correctly described pictured situations. Instructions urged subjects to answer either true or false, but they were also given the option of leaving the matter open in case they couldn't make up their minds.

Two of the tasks were the following:

(15) *situation type:* boy/girl
claim: Every boy that stands next to a girl holds her hand.
picture: Figure 1

[4] It would be naïve to presume that pictures are necessarily more objective than descriptions; indeed, it isn't hard to think of pictures that bias subjects towards a certain response. However, it is generally (though not always) possible to design pictures in such a way that this possibility can be safely disregarded.

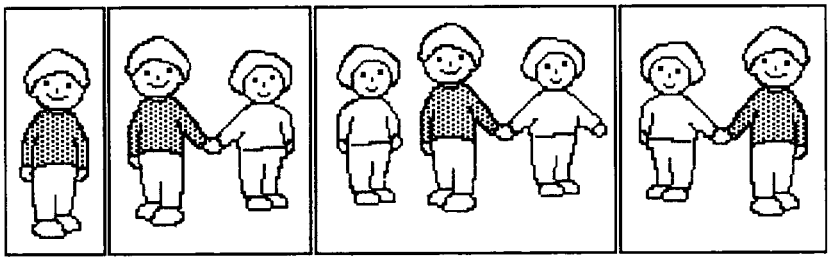

Figure 1. Every boy that stands next to a girl holds her hand.

(16) *situation type:* railway/road
claim: Every railway line that crosses a road goes over it.
picture: Figure 2

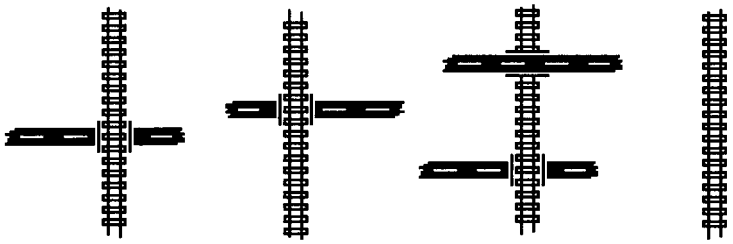

Figure 2. Every railway line that crosses a road goes over it.

Experimental materials contained sentences with four different determiners, corresponding to English *every, not every, some*, and *no*. These sentences were paired with six different types of situations.[5] With four

[5] The determiners used were *iedere* 'every', *niet iedere* 'not every', *enkele* 'some', and *geen* 'no'. As I see no reason to suppose that there are relevant differences between English and Dutch, I will refer to English translations throughout. Apart from the two situation types illustrated in the text, there were four more:

situation type: line/square
claim: Every line that goes through a square divides it into two equal halves.

situation type: child/balloon
claim: Every child that has a balloon is holding it in his right hand.

situation type: O/K
claim: Every O that is adjacent to a K is separated from it by a line.

situation type: girl/dog
claim: Every girl that stands next to a dog keeps it on a leash.

TABLE II

Results for weak determiners. T = true; F = false; ? = no answer

Determiner	Situation type	Answers (%) T	F	?
Some	Railway/road	80	15	5
	Line/square	90	10	
	O/K	100		
	Child/balloon	85	10	5
	Girl/dog	100		
	Boy/girl	90	10	
No	Railway/road		95	5
	Line/square	10	90	
	O/K		100	
	Child/balloon		100	
	Girl/dog		100	
	Boy/girl		100	

determiners and six situation types, we had 24 test items. These were interspersed with filler items so that there were 75 tasks altogether, which were presented in the same order to all subjects.

The main results are summarized in Tables II and III. Let us begin with the results for the weak determiners, which almost speak for themselves (Table II). In nearly all cases subjects judged that the sentences with *some* were true, and sentences with *no* false, of their respective situations. Hence, subjects obtain ∃-readings for both types of sentence, and these interpretations are not affected by differences in content. Thus weakly quantified donkey sentences conform to the generalization in (4a). Note that the number of cases in which subjects couldn't decide between true and false is almost negligible; this held not only for sentences with weak determiners, but across the board.

The observation in (4b) leads us to expect that the results for the universal determiners are less uniform, and this is what we find (Table III). It may seem at first that these results are simply random, but this initial impression is misleading; for it turns out that the six situation types line up along a gradient. To explain this, consider the responses to sentences with *every*. With the railway/road situation type (illustrated in (16)), subjects' judgments are practically unanimous: nine out of ten deem the sentence

TABLE III

Results for universal determiners. T = true; F = false; ? = no answer

Determiner	Situation type	Answers (%)		
		T	F	?
Every	Railway/road	5	90	5
	Line/square	20	80	
	O/K	30	65	5
	Child/balloon	45	50	5
	Girl/dog	60	40	
	Boy/girl	65	35	
Not every	Railway/road	85	10	5
	Line/square	85	15	
	O/K	55	40	5
	Child/balloon	20	70	10
	Girl/dog	55	40	5
	Boy/girl	55	45	

false, which implies that they assign it a ∀-reading. With the boy/girl situation type (illustrated in (15)), the balance has shifted considerably, but not entirely: most subjects get a ∃-reading in this case, but there is still a substantial minority who prefer a ∀-reading. The other situation types lie between these two opposites. This in itself is not a particularly remarkable finding. However, to begin with, this scale is mirrored almost perfectly in the experimental results for *not every*. Here, too, subjects are practically unanimous about the railway/road example: 85% say that it is true, and therefore prefer a ∀-reading. Again we see that opinions are divided about the boy/girl case, with a sizeable 45% preferring the ∀-reading. And again the other situation types line up between these two opposites in the same order as with *every*, with one exception only, namely the child/balloon item. (I don't have a good explanation for this deviation.)

There is a further observation which confirms that this ordering of situation types is not accidental: it is reflected in the response patterns of individual subjects. The data in Table III show that, as a group, subjects have a preference for ∀-readings with universal quantifiers, which is manifested in two ways. On the one hand, there is an overall preference for ∀-readings as opposed to ∃-readings (60% vs. 38%). On the other hand, whereas there are situation types for which virtually all subjects get a ∀-reading, there is

never a consensus that the ∃-reading is the one. This being so, we might suspect that a response for a item lower on the scale (i.e., closer to the boy/girl situation type) is predictive of the response that will be given, by the same subject, for a higher item. This conjecture turns out to be justified. Leaving out of account the aberrant child/balloon situation type, we find that, if a subject gets a ∀-reading with a given situation type, the odds are 4 to 1 that he will get the same reading with the next-higher situation type on the scale.

Summing up our findings, we can conclude that the results of this experiment confirm the observations stated in (4). Whereas donkey sentences with *some* and *no* get ∃-readings, and are impervious to changes in subject matter, sentences with *every* and *not every* merely prefer ∀-readings, and may shift towards ∃-readings depending on what is in the domain of quantification. Our results bring up a new question, as well. We have seen that the interpretation of universal donkey sentences may be affected by plausibility judgments. However, it is doubtful that the content effect we found in our experiment has anything to do with plausibility. We saw that a majority of subjects ticked the 'true' box in task (15), whereas nearly all subjects ticked 'false' in task (16). This shift is evidently related to the difference in situation type between the two tasks, but it doesn't seem right to say that a ∀-reading is less plausible in the former case than in the latter. So how can we account for these shifting responses?

3. Outline of an Explanation

The explanation I want to propose is partly semantic and partly ontological. The semantic part is that weak but not universal determiners are intersective. The ontological part is that less typical individuals (e.g., railway lines) are readily viewed as several 'cases', whereas it is much harder to view prototypical individuals (e.g., boys) this way. Between them, these two factors explain our data; or so I will argue.

Before I present my analysis, a caveat is in order. In the following my main concern will be to explain the data presented above, and it remains to be seen how these data relate to the introspective evidence on which the bulk of the literature is based. In my discussion of Yoon's experiments I emphasized the potential differences between describing and displaying target situations. It may make a great deal of difference whether a subject actually sees four boys or merely hears or reads about them.[6] However, by

[6] Further possible complications arise when we take into account the extent of the situation. The situations used in the experiment were kept as small as possible, so as to

the same token, explanations that apply in the former case may not carry over to the latter, for example because descriptions may cause response biases where pictures are more neutral. None of this is to imply that my experimental data are entirely unrelated to the anecdotal evidence on which previous accounts are based. It is merely to say that the connection need not be a straightforward one.

Weak determiners have an important property that universal determiners lack: they are intersective, which is to say that the truth of a sentence 'D A B', where D is a weak determiner, can always be determined by inspecting just A ∩ B. In other words, weak determiners allow us to concentrate on positive evidence, and ignore all else. To illustrate how this property may affect the way a situation is 'taken in', consider how one might ascertain if the sentences in (17) are true:

(17) *situation:* 14S 9H 7X 21K 15B 3X 6P 23C 12T

a. Some prime numbers are followed by an X.

b. Every prime number is followed by an X.

Verifying (17a) is comparatively easy, because it can be seen at once that there are only two X's, hence only two numbers that are followed by an X, both of which turn out to be prime. If someone adopts this verification strategy, which is sanctioned by the truth-conditional meaning of *some*, he effectively ignores all numbers that don't have the property of being followed by an X. This he cannot do, obviously, if he has to verify (17b) in the same situation, and therefore it takes a bit longer to see that this sentence is false. This example should not be taken to imply that people have one strategy for verifying universal sentences and another one for verifying existential sentences. Rather, it is just to illustrate how the meaning of a determiner *can* have an influence on how a scene is interpreted.[7]

allow subjects to take them in at a glance, and there is no guarantee that the same results would be obtained if materials were scaled up to pictures with 15 or 20 individuals. This limits the experimental set-up, of course. For example, it is partly for this reason that I refrained from using sentences with *most*, which would have required larger situations.

[7] Such effects have been experimentally demonstrated for negative expressions and determiners like *few* and *many*, for example (see Just and Carpenter 1971, Clark 1974). That existential sentences are easier to verify than universal ones has been shown by Meyer (1970) and Just (1974).

Now consider the following task:

(18) *situation type:* boy/girl
claim: Some of the boys that stand next to a girl hold her hand.
picture: We have 4 boys altogether; 1 boy is standing alone; 1 boy is standing next to 1 girl and not holding her hand; 1 boy is standing next to 1 girl and holding her hand; 1 boy ('Fred', to give him a name) is standing between 2 girls, holding hands with 1 of them.

The potential troublemaker in this task is Fred. The evidence Fred brings to bear on the target statement is ambivalent, because he is holding hands with only one of the girls flanking him. However, since *some* is intersective, we may ignore all negative evidence and confine our attention to the boys and girls holding hands; this leaves us with two boys holding hands with a girl, and therefore the claim is judged true.[8] Furthermore, this outcome doesn't depend upon the subject matter of (18): if we replaced the boys and girls with railway lines and roads, say, the same conclusion would follow. So if this is how, in our experiment, sentences with *some* were evaluated, we should expect them to always come out true, regardless the situation type, which is what we found. (The same reasoning applies, *mutatis mutandis*, for *no*.)

Now for donkey sentences with universal determiners: how might (16) be handled? Since *every* is not intersective, here the fact must be faced that there is one railway line that passes over one road and underneath another. This circumstance creates an unusual dilemma, because a single individual presents positive and negative evidence at the same time. My suggestion is that this dilemma can be solved by counting the same individual twice. In the next section I will discuss at some length what *that* might mean, but the basic idea is that it is possible to divide a railway line (in one's mind, of course), and parse it as two separate individuals, one crossing a road passing through a tunnel, while the other crosses the same road over a bridge. But if this is what is counted, the claim in (16) is false.

Unlike roads or railway lines, boys are individuals *par excellence*: cleanly delineated, medium-sized objects that can be relocated without

[8] I am not saying that this judgment follows *logically* from what the situation is and what the statement in (18) means. That would imply that the statement has a definite meaning, which I doubt. What I want to suggest is merely that it is *natural*, psychologically speaking, to use the intersection property as an excuse for dismissing part of the evidence the situation provides. What we would call 'the interpretation' of the statement is the result of this decision. This way of looking at the interpretation of donkey sentences is akin to Kanazawa's (1994); I will return to it below.

their identity being compromised. Railway lines are not much like boys at all. Consider the rail connection between Amsterdam and Brussels, and that between Brussels and Paris. Are they two railway lines or merely parts of one connecting Amsterdam with Paris? The answer, I suppose, is that it depends how you look at it: there is no canonical way of individuating railway lines. There is a canon of individuating persons, of course, and this is what makes it more difficult than it is with railway lines to view the same boy as two individuals – more difficult, that is, to count the same boy twice. Hence, the interpretative dilemma that was solved so easily in the railway/road situation resists a similar treatment in (15). Counting Fred twice is not so easy, and someone who doesn't manage to do this will have to decide if Fred-the-boy, with one girl on his right and another on his left, counts as a boy standing next to a girl whose hand he is holding, or as a boy standing next to a girl whose hand he is *not* holding. In the former case, the sentence will be judged true, in the latter, false.

If this reasoning is sound, there are three groups of subjects:

(A) those who, ontological predicaments notwithstanding, manage to count Fred twice;

(B^+) those who count Fred once and view him as (or focus on the fact that he is) a boy standing next to a girl whose hand he is holding; and

(B^-) those who count Fred once and view him as (or focus on the fact that he is) a boy standing next to a girl whose hand he is not holding.

Subjects in groups A and B^- will reject the sentence, hence obtain the $\forall$-reading, whereas subjects in group B^+ will accept it, preferring the $\exists$-reading. Assuming that the A-strategy is preferred, *ceteris paribus*, and that the tendency to resort to the B-strategies mounts as we quantify over 'better' individuals, it should be the case, given our experimental results (cf. Table III), that children are more prototypical individuals than letters, which in turn should be better individuals than railway lines or geometrical lines. This, it seems to me, is a plausible result.

4. CASES VS. CHARACTERS

What does it mean to say that the same individual is counted twice? It is tempting to explicate this locution in terms of quantification over cases, but that wouldn't be quite right. The word *case* connotes such entities

as events, facts, situations, incidents, and states of affairs, and intuitively these are not the kind of things we quantify over when we count the same individual twice.

(19) Whenever Barney is in Paris he visits the Eifel tower.

This is evidently a statement about a certain class of events or cases (or whatever you choose to call them), and the entities (19) quantifies over *involve* Barney, but we wouldn't want to say that Barney is considered as often as he is in Paris; it is his visits that are under consideration. Now compare (19) with the following examples:

(20)a. National Airlines served at least 2 million passengers in 1975. (Gupta 1980: 23)

b. Four thousand ships passed through the lock last year. (Krifka 1990: 487)

Although every passenger is a person, (20a) does not entail that National Airlines served at least 2 million distinct persons; passengers are not counted like persons. Krifka's example, (20b), is similar in that its truth doesn't seem to require four thousand ships, as long the occasions on which a ship passed through the lock amount to that number; but the example also illustrates that it doesn't require a relational noun to induce such a reading.

Counting passengers is rather like counting characters in a play or film. Just as an actor can play several characters (even at the same time), for the purpose of evaluating (20a) everyone is counted as often as he flew with National Airlines. Let us say, then, that (20a) quantifies over a certain type of character, i.e., the National-Airlines-passenger type. Characters are individuals as seen under a given aspect; they might therefore be modeled as pairs of individuals and properties.[9] It should be noted, however, that individuals are always seen from a vantage point, even if some vantage points are so common that they tend to go unnoticed. Viewing Barney as a person is quite different from, and more natural than, viewing him

[9] Fine (1982) introduces such pairs under the name of 'qua individuals', and uses them to capture the distinction between an individual and the matter it consists of: though a statue is, in a sense, identical with the bronze it is made of, the two may diverge over time or across worlds; they are different characters ('qua individuals') of the same individual. As Fine points out, characters are very much like Aristotle's aggregates of matter and form. Partly for the kinds of reasons adduced by Fine, I prefer to think of characters in intensional terms, but for many purposes an extensional treatment will do just fine. Krifka (1990) argues against such treatments as applied to examples like (20a, b), but his objections are surmountable, as Barker (1999) shows.

as a National Airlines passenger, but of course we view him in a certain way in either case. If we count the same person as often as he flew with National Airlines in 1975, we are negotiating between two standards of individuation. Counting always is counting characters.

I am making a point of stressing the fundamental similarity between viewing someone as a person and viewing him as a passenger, because we need to explain how speakers manage to view the 'same' domain of quantification in varying terms. I have claimed that certain tasks in our experiment prompted subjects to count the same individual twice. If we explicate this as a transition from quantification over individuals to quantification over cases, it doesn't seem to make much sense. For why should subjects suddenly begin to count entities that are so different from the ones they are supposed to count, grammatically speaking? But if the shift is merely from one cast of characters to another, from Fred-as-a-person to Fred-as-a-boy-standing-next-to-Betty, say, there is no mystery to be reasoned away.

So what happened in our experiment was that subjects tried to cope with universal donkey sentences by shifting from one cast of characters (the standard one) to another. It is essential to my purposes that some casts of characters are more readily available than others. While viewing people as (well) people is the most natural thing to do, viewing someone in his capacity of standing next to someone else demands considerable mental effort. But viewing a railway line in its capacity of crossing a certain road is easy again. There are more or less obvious reasons why such differences should exist, but the important thing is *that* they exist.[10]

5. What do Donkey Sentences Mean?

Thus far I have made only minimal assumptions about the meaning of donkey sentences. As it stands my proposal merely requires the premise that *some* and *no* are intersective determiners while *every* and *not every* are not; the main burden of explanation is carried by extra-semantic mechanisms. However, the story is as yet incomplete, because it doesn't explain why such mechanisms should enter the fray, to begin with. I have suggested that intersective quantifiers *allow* subjects to ignore or background certain aspects of a given situation, but why should subjects *want* to do that? I

[10] This is another way in which quantifying over characters differs from quantifying over cases (or events). Cases can be constructed almost *ad libitum*. It is not particularly difficult to hypothesize the case of Pedro beating Chiquita, for example. But viewing Pedro in his capacity of being a Chiquita beater is quite a different matter.

have suggested that, and tried to explain why, recasting the quantificational domain is required by non-intersective determiners only, but why should it happen in the first place? In order to answer these questions we must have a closer look at the semantics and pragmatics of donkey sentences, and especially at donkey anaphora.

It has often been claimed that donkey sentences carry an implication of uniqueness (e.g., Parsons 1978, Cooper 1979, Heim 1990, Kadmon 1990). The idea is that, in the first instance at least, a sentence like (21) implies that every farmer owns at most one donkey.

(21) Every farmer who owns a donkey beats it. (=(1))

This uniqueness implication is usually attributed to the pronoun. Despite various cogent objections against this idea, I believe that it is essentially correct, and that it is the key to the problems posed in the last paragraph. But the key will not fit unless it is appreciated that this uniqueness implication has some rather peculiar properties, which have to do with the fact that it has not one but two sources. To see this, consider the following turkey sentence, which Krifka (1996) attributes to C. L. Baker:

(22)a. ?Most Iowa farmers that raise a turkey sell it for Thanksgiving.
b. Most Iowa farmers that raise turkeys sell them for Thanksgiving.

If it is agreed that it would be highly unusual for an Iowa farmer to raise one turkey at a time, then (22a) becomes infelicitous, and (22b) must be used instead. It is hard to see how this fact can be accounted for unless it is assumed that (22a) conveys a uniqueness assumption, of whatever nature, to the effect that an Iowa farmers normally raise one turkey at a time. Hence this observation corroborates theories proposing that the singular pronoun comes with a uniqueness requirement. However, even if this is the case, the pronoun is not the only source of the uniqueness assumption conveyed by the sentence, as the following examples show:

(23)a. ?Most Iowa farmers that raise a turkey are well off.
b. Most Iowa farmers that raise turkeys are well off.

(23a) is just as infelicitous as (22a) is, and evidently for the same reason, but now there is no pronoun that may be held responsible for the uniqueness implication the sentence gives rise to. So the source of this implication

can only be the singular indefinite *a turkey*, and it is presumably due to a conversational implicature.[11,12]

In (22a) the uniqueness implicature triggered by the indefinite NP *a turkey* is exploited and reinforced by the use of the singular pronoun *it*. By this I don't mean to suggest that the pronoun comes with a uniqueness implication of its own (in this respect I dissociate myself from the analyses endorsed by Parsons, Cooper, Heim, Kadmon, and others). Being a definite expression, the job of a (singular) pronominal expression is to single out one and only one individual, given the context in which the expression occurs. But as Strawson pointed out half a century ago, a definite NP may be used (and often is used) for singling out an individual, with respect to a given context, regardless whether its descriptive content is satisfied by one individual or by several. And if this is true of definite NP's it holds *a fortiori* of pronominal definites. So the right way of looking at (22a) is the following. The singular indefinite *a turkey* introduces an individual of the order Galliformes within the scope of the quantifier *most*, implicating that there is at most one specimen per Iowan farmer, and this relative uniqueness enables the use of the pronoun, whose purpose is to single out one individual matching its meagre descriptive content, and is able to do so because, within the local context in which it is deployed, there is only one candidate to be considered.

In order to show how this view on donkey pronouns helps to explain the peculiarities of the sentences in which they occur, let us consider the following example:

(24) Every O that is adjacent to an M stands on the left of it.

[11] One of the referees discerns a clear difference between (22a) and (23a), the uniqueness implication being weaker in the latter case than it is in the former. I am not sure that I share this intuition, but if it is correct it corroborates my analysis, according to which the pronoun in (22a) reaffirms, so to speak, the uniqueness implication associated with the indefinite.

[12] If the uniqueness implication carried by (23a) is in fact a conversational implicature, then why isn't it cancelled? The reason for this, as argued to some length in Geurts (1999), is that implicatures can be cancelled under rather special circumstances only. If a speaker uses an expression α that gives rise to an implicature φ, and then proceeds to deny that φ is true, he must have a good reason for using α in the first place. And if there is an alternative expression α' which would have done just as well without implicating φ, there is generally no good reason for using α rather than α'. So (23a) is awkward because an alternative wording is readily available, i.e., (23b), which doesn't implicate that Iowa farmers raise at most one turkey at a time.

On the account I have sketched an unwary informant will take this to mean that every O is adjacent to at most one M. But now the following scene is sprung upon our victim:

(25) OM MOM POM EOG

This is the situation that (24) is intended to describe. True or false? This puts our informant in something of a bind, because he has accepted (24) as a well-formed English sentence, he has assumed that the uniqueness implication carried by this sentence is satisfied, but is now forced to concede that the second O in (25) counts as an O that is adjacent to an M despite the fact that it violates the uniqueness implication. So the uniqueness requirement turns out to be problematic, and it cannot be set aside without further ado because it is exploited by the pronoun: if the second O counts as an O that is adjacent to an M, which of its two neighbours, if any, does the neuter pronoun refer to?

There are several ways out of this interpretative crisis, some of which I have discussed in the foregoing (the following list is not meant to be complete):

(i) Our informant may dig in, and declare the statement infelicitous under the circumstances (or something to that effect). In our experiment this didn't happen very often, but there were a number of subjects who took this way out. This solution is in line with the familiar observation that speakers' intuitions about the truth conditions of donkey sentences aren't always clear.

(ii) Our informant may choose to shift his perspective on the quantificational domain, and try to come up with a cast of characters that allows him to save the uniqueness assumption. This is the option that was discussed at some length in the foregoing.

(iii) Another possible strategy is to home in on one of the individuals the pronoun might denote, and try to forget about the other(s). This move may be sanctioned in a number of ways. I have argued that the meaning of a weak determiner allows the hearer to ignore aspects of the situation that cause conflicts of evidence in the case of non-intersective determiners. This is not the only way to justify an unequal treatment of individuals, however. Another way may be by appeal to considerations of plausibility. We have seen that plausibility judgments may bias the hearer towards an otherwise dispreferred ∃-reading for universal donkey sentences; for example:

(26) Every guest who had a credit card used it to pay his hotel bill. (=(3a))

For most people, this is true in a situation that satisfies the ∃-reading though not the ∀-reading, which means that excess credit cards are simply ignored, presumably on the strength of the argument that it would be highly unusual for someone to use more than one credit card for paying a single bill.

(iv) Finally, it is possible that the hearer decides to waive the uniqueness requirement. In her treatment of Heim's (1982) sage plant example, Kadmon (1990: 317) proposes that 'speakers accept this example because it can't make any difference to truth conditions which sage plant *it* stands for, out of all the sage plants that a buyer *x* bought (for each buyer *x*)'.

(27) Every woman who bought a sage plant here bought eight others along with it.

According to Kadmon, her proposal actually allows her to save the assumption that the pronoun carries a uniqueness requirement, even in the face of examples like (27): 'We might say that the speaker does invariably assume that the sage plant is unique per choice of buyer, but leaves the choice of unique plant undetermined, since it makes no difference to truth conditions.' (ibid.) We might say this, of course, but I don't find it very plausible, and prefer to say instead that the speaker gives up the uniqueness requirement for the very reason Kadmon mentions. Heim (1990) also points out that Rooth's (1987) example (28) presents problems for Kadmon's uniqueness analysis:

(28) No parent with a son still in high school has ever lent him the car on a weeknight.

This sentence is falsified by any parent *a* who has a son in high school and has lent him the car on a weeknight, even if *a* has another son who was never even allowed near the car. This contradicts Kadmon's analysis, since it does matter in this case which of *a*'s sons the pronoun is taken to pick up. However, the contradiction disappears once it is seen that the Kadmon strategy is only one way of coping with the uniqueness requirement.

6. READINGS FOR ALL OCCASIONS?

The following sentences are at least similar in form, so it is natural to expect them to have similar logical properties:

(29)a. No farmer owns a donkey.
b. No farmer who owns a donkey beats it.

The main finding of Kanazawa (1994) is that this expectation is warranted only if donkey sentences take certain readings. For example, the following inference is not valid unless donkey sentences with *no* select ∃-readings (Kanazawa 1994: 131):

(30) No farmer who owns a donkey beats it.
No farmer who owns a female donkey beats it.

More generally, Kanazawa shows that the preferences summarized in (4) can be accounted for on the plausible assumption that basic patterns of inference which are valid for other sentences hold for donkey sentences, as well. (The inference patterns Kanazawa focuses upon are based on monotonicity, intersection, and the square of opposition.)

Kanazawa considers two ways of looking at this result, without committing himself one way or the other. On the one hand, he suggests that the semantics of natural language may be 'designed' to be logically homogeneous in the sense that basic inferences are valid throughout the system. This is what Kanazawa calls a 'functional explanation' of why donkey sentences prefer certain readings over others. On the other hand, the issue may be approached in psychological terms, that is, by considering how speakers actually process donkey sentences. Kanazawa outlines a model of interpretation based on the assumption that the interpretation of a donkey sentence is underspecified by the grammar, and need not assign a truth value in situations failing the uniqueness requirement associated with the donkey pronoun. In general, this interpretation demands further specification when the uniqueness requirement is not met, and Kanazawa's proposal is that this specification is constrained by the tendency to preserve certain inference patterns.

Even without going into the details of Kanazawa's analysis, it will be clear that there are close affinities between his ideas and mine. To begin with, we agree on what the main question is, namely: Why is it that donkey sentences are interpreted the way they are? This problem had barely been addressed in the literature before Kanazawa, and has been taken up only sporadically since, laudable exceptions being Yoon (1994) and Krifka

(1996). I deplore this lack of interest, because I believe that this question is the most urgent, and the most promising route towards a better understanding of donkey sentences. Be that as it may, apart from the fact that we address the same problem, my approach is partly in the same spirit as Kanazawa's, in that I have argued, as he has, that the logical properties of their determiners are crucial to explain why some donkey sentences prefer ∀-readings while others must have ∃-readings. Furthermore, I entirely agree with his suggestion that such preferences may be accounted for by way of a processing model (although Kanazawa is more careful than I am to leave room for alternative explanations).

There are differences, as well. In the first place, whereas Kanazawa acknowledges the importance of world knowledge, his focus is on the logical properties of determiners. For me it is the other way round: my main interest lies in the role world knowledge plays in the interpretation of donkey sentences. Secondly, and relatedly, Kanazawa's ideas about the contribution of world knowledge are sketchy at best. Why is that world knowledge affects the interpretation of some donkey sentences only? How is it possible for world knowledge to have an effect at all? Kanazawa has very little to say about such questions, and what he has to say amounts to one paragraph (the last one of his paper), the gist of which is as follows:

> If the speaker's meaning is clear from the beginning, then the person processing the sentence does not have to try to 'figure out' what is meant by the sentence, and consequently he or she will not go into the trouble of invoking inference. (Kanazawa 1994: 155)

This is not terribly convincing, for a number of reasons. First, even if it is granted that the intended interpretation may determine how a sentence is interpreted (provided the speaker's meaning is sufficiently obvious), it still remains to be explained why this happens in some cases only. For example, even if it is obvious that I want to say that Jack is stupid, my uttering 'Jill has beautiful nails' is unlikely to be heard as *meaning* that Jack is stupid. So why is the speaker's intended meaning sufficiently powerful in donkey sentences but not in others? Or rather: why is it sufficiently powerful in *some* donkey sentences but not in others? Kanazawa doesn't say.

In his brief remarks about the effects of world knowledge, it is plain that Kanazawa is thinking of one particular kind, i.e., plausibility judgments. His idea is that certain readings are preferred because they are more plausible than others. However, although many of the examples discussed in the literature fit this description (as witness the contrast between (3a) and (3b)), not all of them do. It is doubtful, for example, that the effects found by Yoon (1994) are adequately described in these terms. As we have

seen, one of Yoon's findings was that (31a) is much more likely to have a ∀-reading than (31b):

(31)a. Every boy who has a toy car likes it.
b. Every boy who had a toy car let his friend play with it.

It is anything but obvious, however, that the ∃-interpretation of (31b) is more plausible than its ∀-interpretation. Or in other words, it would be far-fetched to assume that in such cases 'the speaker's meaning is clear from the beginning.' Similar remarks apply to the experimental results reported above, which show that speakers' preferences are affected by the kind of objects a donkey sentence quantifies over. Again, it doesn't seem right to say that the hearer's ontology leads him to expect a particular meaning before he interprets the sentence.[13]

Perhaps the most important difference between Kanazawa's views and mine is that I am drawn towards a sceptical conclusion which he steers clear of, namely that it may be sometimes be futile if not wrong to suppose that donkey sentences must have a definite reading. Kanazawa comes close to this position at one point, but doesn't embrace it, as I am tempted to do. Let me explain. Having shown that donkey sentences tend to be interpreted in such a way that certain basic forms of inference are preserved, Kanazawa goes on to make the following remark:

> In fact, it would seem that the intuitions about the validity of these inferences with donkey sentences – the feeling that they must be valid, or the inclination to draw such inferences – are stronger than the intuitions about the truth conditions of individual donkey sentences. (Kanazawa 1994: 149–150)

In my opinion, the importance of this observation can hardly be exaggerated, because it raises doubts about a quite fundamental presupposition concerning the interpretation not only of donkey sentences but of linguistic expressions in general. For how is it possible for our intuitions about the meaning of a donkey sentence φ to be so 'nebulous', to use Heim's (1982: 62) term, whilst our intuitions about basic inferences involving φ are so robust, given that these inferences are valid on one particular reading only? Could it be the case that speakers use these inferences not to arrive at a definite meaning but merely to determine a truth value? I suspect that it could be, and that what I have been calling '∀-readings' and '∃-readings' may not really be readings, after all.

If this suspicion turns out to be justified, there would be something seriously amiss with the standard view on the relation between meaning and

[13] This is confirmed by the protocols of the think-aloud experiments to be discussed below.

truth. For it is usually supposed without argument that in order to determine if a sentence is true we have to settle its meaning first and subsequently confront that with the facts. Though this idea has been questioned by philosophers like Quine and Davidson, my worry is not a philosophical but an empirical one. According to the model outlined in the foregoing, informants have various strategies for dealing with the task at hand, which is *not* to determine what a sentence means but merely to decide whether it is true or false. The received view entails that the latter presupposes the former, but Kanazawa's observation and my experimental findings suggest otherwise. If our intuitions about the logical properties of donkey sentences are stronger than our intuitions about their truth conditions, then perhaps the latter are less explicit than we would like to admit. And if truth-value judgments vary with qualitative features of the domain of quantification in anything like the way I have suggested, then at least from a psychological point of view it is dubious that donkey sentences always have a definite interpretation.

7. Postscript: Processing Negatives

It has been known for quite some time that negative expressions are more difficult than their positive counterparts (see Horn (1989: chapter 3) for an overview of the literature). Negative expressions take longer to process, cause more errors, and are harder to retain. This holds for explicit negatives (e.g., sentences with *not* or *no*(*ne*)) as well as for implicit ones (e.g., *fail* or *lack*). It also holds for expressions that are traditionally described as being negative in nature; for example, *few* is negative in this sense, whereas *many* is positive.

It is also well established that, in many cases, speakers employ special procedures for dealing with negatives.

(32)a. The first number on the list is not odd.
b. The first number on the list is odd.
c. The first number on the list is even.

When he has to evaluate (32a), an informant may decide to remove the negation particle, evaluate (32b), and finally flip the truth value obtained in the second step. Alternatively, he may decide to mentally rewrite (32a) as (32c), and evaluate *that*. These two strategies were already described by Wason (1961); they are incorporated in Clark's (1974) theory, which is confirmed by a broad range of experimental data, and remains one of the

most successful psychological accounts of the interpretation of negation to date. Quantitative evidence aside, the availability of these strategies has been confirmed by introspective reports collected in many experiments over the years.

Given that these strategies are employed already in the interpretation of comparatively simple sentences like (32a), it is *a priori* likely that they are also used in the interpretation of donkey sentences containing negative expressions. But I have less indirect evidence, as well. In one of the experiments conducted in preparation of the 'official' experiment discussed above, subjects were asked to think aloud while they evaluated donkey sentences (see Ericsson and Simon 1984 for extensive discussion of this method). The resulting monologues were recorded, transcribed, and analysed to perfect the design of the main experiment. They also provide valuable clues as to how people actually tackle the problem of evaluating donkey sentences.

These think-aloud sessions prove that at least one of the procedures for dealing with negatives is employed in the interpretation of donkey sentences, too. The following extracts exemplify the evidence. In the first example, the subject is working on the line/square task illustrated in footnote 5; in the other examples, subjects begin by reading out the target statement (these tasks weren't used in the final experiment):

(33)a. not every line; forget about 'not' for a moment; every line that goes through a square divides it into two equal halves; well that's false; so the rest is true

b. not every woman who has a briefcase is holding it in her right hand; 'not' to the side; every woman who has a briefcase is holding it in her right hand; that's true so this is false

c. in none of the piles that contain a white draughtsman there's a black one on top of it; leave out 'none' for a moment; in the piles that contain a white draughtsman there's a black one on top of it; that's false; so this is true

In these cases the procedure is to take out the negative element, obtain a truth value for the remainder of the sentence, and turn that around. Note that, despite the fact that it is blatantly illicit in the case of *no*, the same procedure is applied in (33c), too.

These findings are instructive in several ways. First, they support Krifka's (1996) claim that *not every* is not a determiner, and that sentences of the form 'not every A B' are better seen as instances of sentence negation. Secondly, they bear on the status of (4a), which states that donkey sentences with weak determiners only have ∃-readings. I conceded that

this may be too strong in view of examples like (5a, b), repeated here for convenience:

(34)a. No man who had a credit card failed to use it.
b. At least one boy who had an apple for breakfast didn't give it to his best friends.

I also noted, however, that it isn't clear how much weight such examples carry, because they always seem to involve some form of negation. The foregoing observations lend support to this point. This is not to say that they prove that (4a) holds without exception, but merely that apparent counterexamples like (34a, b) should not be taken at face value.

Finally, the examples in (33) illustrate the point I made at the end of my paper. The subject whose stream of consciousness is reflected in (33b), say, does not seem to be preoccupied with the meaning of the statement at issue. Rather, it appears that his sole concern is with the sentence's truth value, and that he is content to let its meaning fall where it may.

ACKNOWLEDGMENTS

This is a project I have been engaged in, on and off, since the summer of 1996, partly supported by a grant from the Deutsche Forschungsgemeinschaft (DFG). I have benefited from discussions with Herb Clark, Ileana Comorovski, Günther Gediga, Joyce Karreman, Manfred Krifka, Leo Noordman, Kai von Fintel, Wietske Vonk, and Young-Eun Yoon. I am especially indebted to Wietske Vonk and Joyce Karreman for helping me to set up my experiments and make sense of the results, and to the referees for *L & P*, without whose incisive criticisms the paper would have been rather different from what it is now.

REFERENCES

Barker, C.: 1999, 'Individuation and Quantification', *Linguistic Inquiry* **30**, 683–691.
Chierchia, G.: 1992, 'Anaphora and Dynamic Binding', *Linguistics and Philosophy* **15**, 111–183.
Chierchia, G.: 1995, *Dynamics of Meaning*, University of Chicago Press.
Clark, H. H.: 1974, 'Semantics and Comprehension', in T. Sebeok (ed.), *Current Trends in Linguistics*, **12**. Mouton, The Hague, pp. 1291–1428.
Conway, L. and S. Crain: 1995, 'Donkey Anaphora in Child Grammar', in J. N. Beckman (ed.), *NELS 25*, University of Massachusetts, Amherst.

Cooper, R.: 1979, 'The Interpretation of Pronouns', in F. Heny and H. Schnelle (eds.), *Selections from the Third Groningen Round Table*, Academic Press, New York, pp. 61–92.

Ericsson, K. A. and H. A. Simon: 1984, *Protocol Analysis: Verbal Reports as Data.* MIT Press, Cambridge, Mass.

Fine, K.: 1982, 'Acts, Events and Things', in W. Leinfellner, E. Kraemer, and J. Schank (eds.), *Language and Ontology: Proceedings of the Sixth International Wittgenstein Symposium*, Vienna, pp. 97–105.

Geach, P. T.: 1962, *Reference and Generality*, Cornell University Press, Ithaca, NY.

Geurts, B.: 1999, *Presuppositions and Pronouns*, Elsevier, Oxford.

Gupta, A. K.: 1980, *The Logic of Common Nouns*, Yale University Press, New Haven.

Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. thesis, University of Massachusetts, Amherst.

Heim, I.: 1990, 'E-type Pronouns and Donkey Anaphora', *Linguistics and Philosophy* **13**, 137–178.

Horn, L. R.: 1989, *A Natural History of Negation*, Chicago University Press.

Just, M. A. and P. A. Carpenter: 1971, 'Comprehension of Negation with Quantification', *Journal of Verbal Learning and Verbal Behavior* **10**, 244–253.

Just, M. A.: 1974, 'Comprehending Quantified Sentences: The Relation between Sentence-picture and Semantic Memory Verification', *Cognitive Psychology* **6**, 216–236.

Kadmon, N.: 1990, 'Uniqueness', *Linguistics and Philosophy* **13**, 273–324.

Kamp, H.: 1981, 'A Theory of Truth and Semantic Representation', in J. A. G. Groenendijk, T. M. V. Janssen, and M. B. J. Stokhof (eds.), *Formal Methods in the Study of Language*, Mathematical Centre Tracts 135, Amsterdam, pp. 277–322.

Kamp, H.: 1991, 'Uniqueness Presuppositions and Plural Anaphora in DTT and DRT', in M. Stokhof, J. Groenendijk, and D. Beaver (eds.), *Quantification and Anaphora, Part 1*, DYANA deliverable R2.2A, ILLC, University of Amsterdam.

Kanazawa, M.: 1994, 'Weak vs. Strong Readings of Donkey Sentences and Monotonicity Inference in a Dynamic Setting', *Linguistics and Philosophy* **17**, 109–158.

Krifka, M.: 1990, 'Four Thousand Ships Passed Through the Lock: Object-induced Measure Functions on Events', *Linguistics and Philosophy* **13**, 487–520.

Krifka, M.: 1996, 'Pragmatic Strengthening in Plural Predications and Donkey Sentences', in T. Galloway and J. Spence (eds.), *Proceedings from Semantics and Linguistic Theory (SALT) VI*, Cornell University, Ithaca, NY, pp. 136–153.

Lappin, S.: 2000, 'An Intensional Parametric Semantics for Vague Quantifiers', *Linguistics and Philosophy* **23**, 599–620.

Meyer, D. E.: 1970, 'On the Representation and Retrieval of Stored Semantic Information', *Cognitive Psychology* **1**, 242–300.

Parsons, T.: 1978, 'Pronouns as Paraphrases', Ms., University of Massachusetts at Amherst.

Rooth, M.: 1987, 'Noun Phrase Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics', in P. Gärdenfors (ed.), *Generalized Quantifiers: Linguistic and Logical Approaches*, Reidel, Dordrecht, pp. 237–268.

Strawson, P.: 1950, 'On referring', *Mind* **59**, 320–344.

Van der Does, J.: 1993, 'The Dynamics of Sophisticated Laziness', in J. Groenendijk (ed.), *DYANA Deliverable R22A*, ILLC, University of Amsterdam. [Reference as given by Chierchia (1995).]

Wason, P. C.: 1961, 'Response to Affirmative and Negative Binary Statements', *British Journal of Psychology* **52**, 133–142.

Yoon, Y.-E.: 1994, *Weak and Strong Interpretations of Quantifiers and Definite NPs in English and Korean*, Ph.D. Dissertation, University of Texas at Austin.
Yoon, Y.-E.: 1996, 'Total and Partial Predicates and the Weak and Strong Interpretations', *Natural Language Semantics* **4**, 217–236.

Humboldt University, Berlin &
University of Nijmegen
Department of Philosophy
P.O. Box 9103
6500 HD Nijmegen
The Netherlands
E-mail: bart.geurts@phil.kun.nl