

6 Quantifiers: Their Semantic Type

The only DPs treated so far (not counting predicate nominals) have been proper names, definite descriptions, pronouns, and traces.¹ We assumed that their denotations were individuals – that is, elements of D_e . There are many other kinds of DPs – for example, DPs made up with a variety of determiners ("this", "that", "a(n)", "every", "no", "many", "few", "most", and so forth) and yet other types, such as "only John". What denotations should we assign to all these DPs and to the determiners they contain? What types of denotations will be suitable in each case?

Before we address these questions directly, we will establish a couple of important negative points. First, we show that at least some DPs do not denote *individuals*. Second, we are going to see that *sets of individuals* (or their characteristic functions²) are not a suitable type of denotation either. We will consider more arguments than are strictly speaking needed to make these points. It may seem like overkill, but it serves the additional purpose of drawing up a checklist against which to test positive proposals, in particular the one adopted below.

6.1 Problems with individuals as DP-denotations

There are various types of DPs for which denotations of type e seem to work well. We had some success treating proper names in this way, and also definite descriptions. Pronouns ("he", "I", ...) and demonstratives ("this book", "that cat") might also be accommodated if we allowed that their extensions are not fixed by the semantics once and for all, but vary from one occasion of utterance to the next. "I", for instance, seems to denote on each occasion when it is uttered the individual who utters it. This individual then enters the calculation of the truth-value of the uttered sentence in the usual way. For instance, when Irene Heim utters "I am sleepy", this is true just in case Irene Heim is in the extension of "sleepy", and false otherwise. We will give more detailed

consideration to the proper treatment of such context dependency later on. Even if denotations may be fixed in part by the utterance occasion, however, individuals don't seem to be the right kind of denotation for many DPs. Let us see why not.

6.1.1 Predictions about truth-conditions and entailment patterns

Naive intuition is not a reliable guide as to which DPs denote individuals. If asked whom or what "no man" denotes, a lay person might say that it doesn't denote anything, but with "only John", they might say it refers to John.

As it turns out, the latter suggestion is immediately falsified by its predictions. If **[only John]** = John, then **[only John left]** = 1 iff John left. But if John and Sam both left, this sentence is intuitively false, not true as just predicted. Of course, this argument relies on a number of assumptions: in particular, assumptions about the constituent structure of the example and about the principles of semantic composition. It is thus not an indefeasible argument, but it is sound.

By comparison, the mere fact that it offends common sense to assign a referent to "no man", is at best a weak reason not to do so in our theory.³ More decisive, in this instance as well, will be considerations pertaining to predicted semantic judgments about sentences. Such considerations can be used to refute a *particular* assignment of denotation for a given DP, as in the case above, where we showed that **[only John]** cannot be John. And sometimes they can even be used to show that *no denotation of a certain type* will do. Let us give a few illustrations of this type of reasoning.

DPs that fail to validate subset-to-superset inferences

The following inference is intuitively valid:

- (1) **John came yesterday morning.**
 ∴ **John came yesterday.**

Not only is it intuitively valid, but it is *predicted* to be valid by any semantics that implies these three assumptions:

- (i) $\llbracket \text{John} \rrbracket \in D_e$
- (ii) $\llbracket \text{came yesterday morning} \rrbracket \subseteq \llbracket \text{came yesterday} \rrbracket^4$
- (iii) A sentence whose subject denotes an individual is true iff that individual is a member of the set denoted by the VP.

Proof: Suppose the premise is true. Then, by (iii), $\llbracket \text{John} \rrbracket \in \llbracket \text{came yesterday morning} \rrbracket$. Hence, by (ii) and set theory, $\llbracket \text{John} \rrbracket \in \llbracket \text{came yesterday} \rrbracket$. And, again by (iii), the conclusion is true. QED. Notice that no concrete assumption about **which** element of D_e **John** denotes was needed for this proof.

Now consider the parallel inference in (2).

- (2) **At most one letter came yesterday morning.**
 ∴ **At most one letter came yesterday.**

This one is intuitively invalid: It is easy to imagine a state of affairs where the premise is true and the conclusion is false. For example, only one letter comes in the morning but two more in the afternoon. If we want to maintain (ii) and (iii), it therefore follows that **[at most one letter]** $\notin D_e$, or else we could prove the validity of (2) exactly as we proved that of (1). Should we give up (ii) or (iii) to avoid this conclusion? (ii) seems well-motivated by other data: for example, the tautological status of "If John came yesterday morning, he came yesterday". And giving up (iii) would set our whole enterprise back to square one, with no plausible substitute in sight. These would be costly moves at best.

We can test other DPs in this inference schema:

- (3) **a came yesterday morning.**
 ∴ **a came yesterday.**

Among the substitutions for **a** that systematically *fail* to make this valid are DPs with the determiners "no" and "few", and with all determiners of the form "less than n", "at most n", "exactly n" (for some numeral "n"). For all such DPs, we thus have a strong reason to assume that their denotations are not of type e. (The reverse does not hold: there are many DPs that do validate (3) but still cannot have type e denotations, for other reasons, such as those considered next.)

DPs that fail the Law of Contradiction⁵

If you choose two VPs with disjoint extensions and combine first one, then the other, with a given proper name, you get two sentences that contradict each other. For example, (4) is contradictory.

- (4) **Mount Rainier is on this side of the border, and Mount Rainier is on the other side of the border.**

Again, this is a judgment predicted by any semantics committed to a few plausible assumptions. The following four suffice:

- (i) [Mount Rainier] $\in D_e$
- (ii) [be on this side of the border] n [be on the other side of the border]
 $= \emptyset$
- (iii) (composition rule for subject + VP, same as above)
- (iv) standard analysis of and.

We leave the proof to the reader.

The important point is again that this proof does not rely on a more specific assumption about Mount Rainier than what type of denotation it has. It will therefore generalize to analogous sentences with any other DP that denotes an individual. But many such sentences are not in fact contradictory, for example:

- (5) More than two mountains are on this side of the border, and more than two mountains are on the other side of the border.

So, unless we want to mess with (ii), (iii), or (iv), we have to conclude that more than two mountains does not denote anything in D_e .

Further DPs to which this argument extends are "a mountain", " n mountains" (for any numeral " n "), "no mountain", and lots of others.

DPs that fail the Law of Excluded Middle⁶

Again we form minimal pairs of sentences whose subjects are identical and whose VPs differ. This time we choose two VPs the union of whose extensions exhausts everything there is, and we coordinate the two sentences by "or". For example:

- (6) I am over **30** years old, or I am under **40** years old.

(6) is a tautology, which we can prove if our semantics implies this much:

- (i) $\llbracket I \rrbracket \in D_e$
- (ii) [be over **30** years old] \cup [be under **40** years old] = D
- (iii) (as above)
- (iv) standard analysis of or.

Our reasoning must be foreseeable at this point: Any other individual-denoting DP in place of "I" in (6) would likewise be predicted to yield a tautology. So if there are counterexamples, the DPs in them cannot denote individuals. Here is one:

- (7) Every woman in this room is over **30** years old, or every woman in this room is under 40 years old.

There are other such systematic differences in the entailments, tautologies, and contradictions that we get for proper names on the one hand and for the lot of nondefinite DPs on the other. They all yield potential arguments for a distinction in semantic type. But you have the idea, and we can turn to an argument of a somewhat different kind.

6.1.2 *Predictions about ambiguity and the effects of syntactic reorganization*

English syntax often allows us to construct different sentences out of more or less the same lexical items. Often we will get different meanings as a result. For instance, "John saw Mary" has different truth-conditions from "Mary saw John", and this is easily accounted for in our current theory. Other times, we get pairs of sentences whose truth-conditions coincide. This occurs with active–passive pairs, with pairs of a topicalized sentence and its plain counterpart, or with certain more stilted circumlocutions like the "such that" construction:

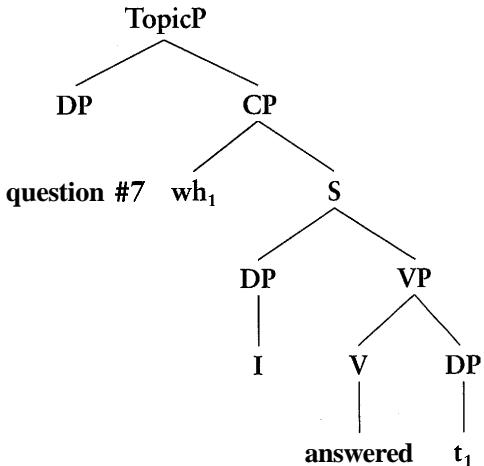
- (8a) I answered question #7.
- (8b) Question #7, I answered.

- (9a) John saw Mary.
- (9b) Mary is such that John saw her.
- (9c) John is such that Mary saw him.

Whatever the subtler meaning differences in each of these groups, we cannot imagine states of affairs in which one member of the pair or triple is true and the other(s) false. Our semantics (given suitable syntactic assumptions) predicts these equivalences. Let us briefly sketch how before we resume our main line of argument.

First, take the unlyrical (9b) and (9c). If you treat the pronouns "her" and "him" as variables and "such that" as signaling predicate abstraction, then [such that John saw her] = (the characteristic function of) the set {x : John saw x}. This is also the value of the whole VP, and (9b) is thus predicted to be true iff Mary is in this set – which means nothing more and nothing less than that John saw her. So we have proved (9b) equivalent to (9a). The same goes for (9c).

Second, consider the topicalization construction in (8b). This may be much like the "such that" construction. In the analysis of Chomsky's "On Wh-Movement,"⁷ the topicalized phrase in (8b) would be linked to a moved wh-phrase, and we might assume a structure of the following kind:



The CP here looks like just another predicate abstract, and indeed interpreting it this way predicts the right meaning and the equivalence with (8a).

Why did we bring up these equivalences? Because we now want you to look at analogous cases in which the names and first person pronouns of (8) and (9) have been replaced by some other kinds of DPs:

- (10a) **Almost everybody answered at least one question.**
- (10b) **At least one question, almost everybody answered.**

- (11a) **Nobody saw more than one policeman.**
- (11b) **More than one policeman is such that nobody saw him.**
- (11c) **Nobody is such that he or she saw more than one policeman.**

Suddenly the "transformations" affect truth-conditional meaning. (10a) can be true when no two people answered any of the same questions, whereas (10b) requires there to be some one question that was answered by almost all. For instance, imagine there were ten students and ten questions; student #1 answered just question #1, student #2 just question #2, and so on, except for student #10, who didn't answer any. (10a) is clearly a true description of this state of affairs, but (10b) is false of it. (Some speakers might decide upon reflection that they actually get two distinct readings for (10a), one of which is equivalent to (10b). But even so, the fact remains that (10b) cannot express the salient reading of (10a).)

Similar comments apply to (11a), (11b), and (11c). (11c) does not require for its truth that any policeman went completely unnoticed, whereas (11b) claims that two or more did. Conversely, (11b) is true and (11c) false when everybody saw several policemen and there were some additional unseen policemen hiding

in the shadows. (11a) seems to be equivalent to (11c), and hence distinct from (11b). Even if some speakers find (11a) ambiguous, (11b) clearly differs from (11c) on any reading of either, and this is enough for us to make our point.

The point is this: The truth-conditional effects that we have just seen to be associated with certain structural rearrangements are completely unexpected if the DPs in (10) and (11) are treated like those in (8) and (9). The basic semantic properties of the "such that" and topicalization constructions directly imply that shuffling around DPs with meanings of type e in these constructions cannot possibly change truth-values.⁸ So when we do observe such changes, as between (11b) and (11c), they are one more piece of evidence that certain DPs don't denote individuals.

A final problem with trying to assimilate the semantics of all DPs to that of proper names (closely related, as it turns out, to the preceding puzzle) is that such an approach does not anticipate certain judgments of *ambiguity*. For example, (13) is ambiguous in a way that (12) is not.

(12) It didn't snow on Christmas Day.

(13) It didn't snow on more than two of these days.

Suppose ten days are under consideration, and it snowed on exactly three of them. Is (13) true or false then? This depends on how you understand it. I could argue: "(13) is true. There was no snow on as many as *seven* days, and seven is more than two, so surely it didn't snow on *more than two* days." You might argue: "No, (13) is false. It snowed on *three* days, and three is more than two, so it *did* snow on more than two." We are not going to settle our argument, because each of us is reading the sentence in a different way, and both are apparently readings allowed by English grammar. (There may be differences in the ways the two readings are most naturally pronounced, and if so, these might be traceable to some subtle difference in syntactic structure. Should this turn out to be the case, we should really group this example with the ones in the previous section; that is, it would then be another instance of how changes in structure affect truth-conditions. For our purposes here, this wouldn't make an important difference, since one way or the other, the example highlights a fundamental semantic difference between "more than two of these days" and a proper name like "Christmas Day".)

There is no ambiguity of this kind in (12), and there can't really be, given that "Christmas Day" denotes a day.⁹ The three ingredients of this sentence – the name Christmas Day, the predicate it snows on, and the negation – can only be put together to produce a truth-value in one way, so there is simply no room for any possible ambiguity. If DPs like more than two days likewise denoted individuals, there ought to be no more room for ambiguity in (13). (The argument as stated

is a bit simple-minded, we admit: one can think of more than one way to semantically compose the verb snow, the temporal preposition on, a time-name, and a negation. But since all plausible alternatives come to the same thing in the case of (12) – they had better, or else we would predict an ambiguity that doesn't in fact exist! – it's hard to see how they could wind up with different truth-values for (B).)

6.2 Problems with having DPs denote sets of individuals

The view that all DPs denote individuals may have been a dead horse as soon as anyone gave it conscious thought. But another similarly problematic idea still comes up every now and then. P. T. Geach introduces and warns against it as follows:

Another bad habit often picked up from the same training is the way of thinking that Frege called mechanical or quantificatiou思 thinking: mechanische oder *quantifizierende Auffassung*. I have used a rude made-up word "quantificatiou思" because Frege was being rude; "quantificational" and "quantifying" are innocent descriptive terms of modern logic, but they are innocent only because they are mere labels and have no longer any suggestion of quantity. But people who think quantificatiou思ly do take seriously the idea that words like "all", "some", "most", "none", tell us how much, how large a part, of a class is being considered. "All men" would refer to the whole of the class men; "most men", to the greater part of the class; "some men" to some part of the class men (better not ask which part!); "no man", finally, to a null or empty class which contains no men. One can indeed legitimately get to the concept of a null class – but not this way.

I have not time to bring out in detail how destructive of logical insight this quantificatiou思 way of thinking is.¹⁰

Exercise

Consider the following "quantificatiou思" fragment of English. Assume our customary composition principles. As before, the basic domains are $D (= \mathbb{D})$ and $\{0, 1\}$, but the denotations of lexical items are now different.

[Ann] = {Ann}

[Jacob] = {Jacob}

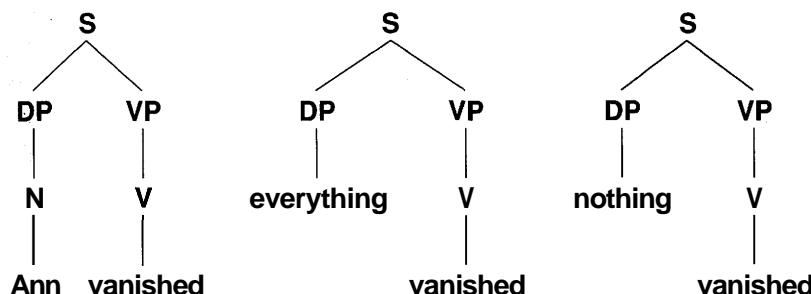
[everything] = D

[nothing] = O

[vanished] = $\lambda X \in \text{Pow}(D) . X \subseteq \{y \in D : y \text{ vanished}\}.$ "

[reappeared] = $\lambda X \in \text{Pow}(D) . X \subseteq \{y \in D : y \text{ reappeared}\}.$

- (a) Discuss the adequacy of this proposal with respect to the following types of structures:



Are the correct truth-conditions predicted in each of the three cases?
Can the proposal be extended to other quantifier phrases?

- (b) Which of the following inferences are predicted to be valid, given a "quantificational" semantics of the sort presented above? Which predictions are correct?

- (i) **Ann vanished fast.**
Ergo: **Ann vanished.**
- (ii) **Everything vanished fast.**
Ergo: **Everything vanished.**
- (iii) **Nothing vanished fast.**
Ergo: **Nothing vanished.**

For this exercise, add an unanalyzed predicate **vanished fast** to our fragment, and assume the following relationship:

$$\{x \in D : x \text{ vanished fast}\} \subseteq \{x \in D : x \text{ vanished}\}.$$

Given this assumption, determine whether you are able to prove that the premise logically implies the conclusion in each of the above examples. Note that, strictly speaking, the premises and conclusions are phrase structure trees, not just strings of words.

- (c) Consider now the following passage from Geach:

In a reputable textbook of modern logic I once came across a shocking specimen of quantificational thinking. Before presenting it to you, I must supply some background. In ordinary affairs we quite often need to talk about kinds of things that do not exist or about which we do not yet know whether they exist or not; and this applies to ordinary scientific discourse as well – I once saw a lengthy chemical treatise with the title "Nonexistent Compounds." Accordingly, logicians need to lay down rules for propositions with empty subject-terms. The convention generally adopted is that when the subject-term is empty, ostensibly contrary categorical propositions are taken to be both true; for example, if there are no dragons, "All dragons are blue" and "No dragons are blue" are both true. This convention may surprise you, but there is nothing really against it; there are other equally consistent conventions for construing such propositions, but no consistent convention can avoid some surprising and even startling results.

Now my author was trying to show the soundness of this convention, and to secure that, came out with the following argument. . . . "If there are no dragons, the phrases 'all dragons' and 'no dragons' both refer to one and the same class – a null or empty class. Therefore 'All dragons are blue' and 'No dragons are blue' say the same thing about the same class; so if one is true, the other is true. But if there are no dragons to be blue, 'No dragons are blue' is true; therefore, 'All dragons are blue' is also true." I know the argument sounds like bosh; but don't you be fooled – it is bosh.¹²

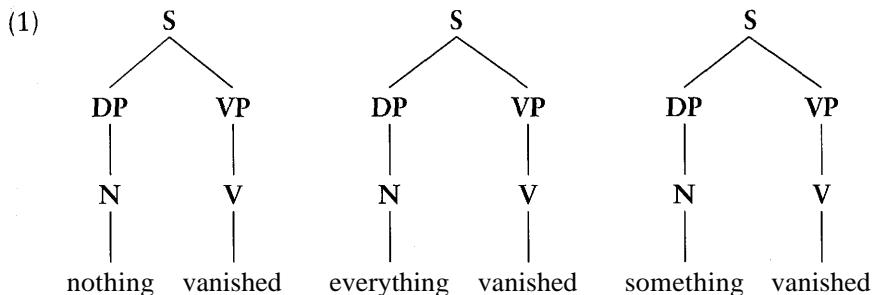
Question: What is wrong with the logician's argument?

6.3 The solution: generalized quantifiers

6.3.1 “*Something*”, “*nothing*”, “*everything*”

We have seen that quantificational DPs like "something", "everything", and "nothing" are not proper names. Within our current framework, this means that their denotations are not in D_e . That is, quantificational DPs are not of semantic type e . We have also seen that the denotations of quantificational DPs like "something", "everything", and "nothing" are not sets of individuals. Hence they cannot be of semantic type $\langle e, t \rangle$. What, then, is the semantic type of quantificational DPs? The Fregean reasoning that we have been following so far

gives a clear answer, given (a)the usual kind of phrase structure, as in (1)below, (b)our previous assumptions about the denotations for VPs and sentences, and (c)our goal to reduce semantic composition to functional application whenever possible.



Our rules of semantic composition determine that the denotations of the VP-nodes in the above structures coincide with the denotation of *vanish*, hence are elements of $D_{\langle e, t \rangle}$. Since the denotations of the sentences must be members of D_e , we now conclude that the semantic type of quantificational DPs must be $\langle \langle e, t \rangle, t \rangle$, provided that the mode of composition is indeed functional application, and given that type e is excluded.

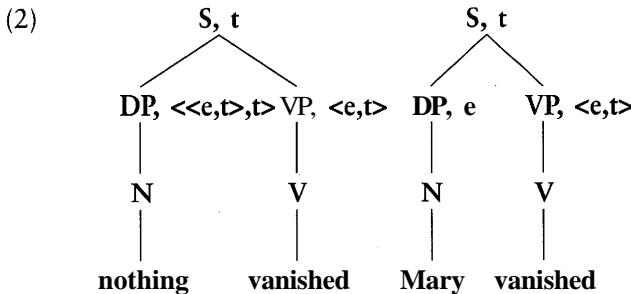
Here is a more intuitive way of thinking about the semantic contribution of "nothing", "everything", and "something" in subject position. Rather than denoting an individual or a set of individuals, "nothing" in "nothing vanished" says something about the denotation of the predicate "vanished". It states that there is no individual of which the predicate is true; that is, there is no individual that vanished. If we replace "nothing" with "everything", the claim is that the predicate is true of all individuals. Finally, sticking "something" in the subject position of a predicate leads to the statement that there is at least one individual of which the predicate is true. Quantificational DPs, then, denote functions whose arguments are characteristic functions of sets, and whose values are truth-values. Such functions are sometimes called "second-order properties", first-order properties being functions of type $\langle e, t \rangle$. The second-order property $[[\text{nothing}]]$, for example, applies to the first-order property $[[\text{vanished}]]$, and yields truth just in case $[[\text{vanished}]]$ does not apply truly to any individual. More recently, second-order properties are commonly referred to as "generalized quantifiers".¹³ The reasoning we just went through leads to the following lexical entries for "nothing", "everything", and "something":

$[\text{nothing}] = \lambda f \in D_{\langle e, t \rangle} . \text{there is no } x \in D_e \text{ such that } f(x) = 1$.

$[\text{everything}] = \lambda f \in D_{\langle e, t \rangle} . \text{for all } x \in D_e, f(x) = 1$.

$[\text{something}] = \lambda f \in D_{\langle e, t \rangle} . \text{there is some } x \in D_e \text{ such that } f(x) = 1$.

Our semantic composition rules guarantee the right modes of combination for subject DPs with their VPs without any further machinery. If the DP is a proper name, the denotation of the VP applies to the denotation of the DP. If the DP is a quantifier phrase, the denotation of the DP applies to the denotation of the VP. The difference between the two modes of semantic composition can be illustrated by trees with semantic type annotations as in (2):



In both structures, the rule for Non-Branching Nodes (NN) determines that the denotation of the lexical item occupying the subject position is passed up to the DP-node. The difference between the denotations of the lexical items "nothing" and "John", then, brings about a difference in the way the respective DP-nodes are semantically combined with their VPs.

Exercise

Calculate the truth-conditions for the trees in (1) above.

6.3.2 Problems avoided

Before we adopt the new higher type for quantificational DPs, we want to be sure that it will indeed avoid the problems we noted for the simpler types *e* and *<e,t>*. Let us return to the inference patterns we looked at in section 6.1.1.

Subset to superset

When $[\alpha]$ is of type *e*, and β necessarily denotes a subset of γ , then a β entails a γ . But At **most one letter came yesterday morning** does not entail At **most one letter came yesterday**. So we concluded that [at most one letter] cannot be in D_{α} .

Now we must show that $D_{\langle\langle e,t\rangle,t\rangle}$ contains a denotation that will predict the invalidity of this inference. That is, we must show that there exists a function $f \in D_{\langle\langle e,t\rangle,t\rangle}$ such that it is possible that $f(\llbracket \text{came yesterday morning} \rrbracket) = 1$, but $f(\llbracket \text{came yesterday} \rrbracket) = 0$. This is quite evident. We only need to point out that it is possible for the actual facts to be such that $\llbracket \text{came yesterday morning} \rrbracket \neq \llbracket \text{came yesterday} \rrbracket$. There are as many different functions in $D_{\langle\langle e,t\rangle,t\rangle}$ as there are ways of mapping the elements of $D_{\langle e,t\rangle}$ to $\{0, 1\}$. So for each given pair of distinct elements of $D_{\langle e,t\rangle}$, there are lots of functions in $D_{\langle\langle e,t\rangle,t\rangle}$ that map the first to 1 and the second to 0.

Law of Contradiction

When $\llbracket \alpha \rrbracket$ is of type e, and β and γ denote necessarily disjoint sets, then " $\alpha \beta$ and $\alpha \gamma$ " is a contradiction. But **More than two cats are indoors and more than two cats are outdoors** is not a contradiction. So [**more than two cats**] cannot be in D_e .

Now we must show that $D_{\langle\langle e,t\rangle,t\rangle}$ contains a denotation that will predict the possible truth of this conjunction. That is, we must show that there exists a function $f \in D_{\langle\langle e,t\rangle,t\rangle}$ such that it is possible that $f(\llbracket \text{indoors} \rrbracket) = 1$ and $f(\llbracket \text{outdoors} \rrbracket) = 1$. Obviously, there exist plenty such functions.

Law of Excluded Middle

When $\llbracket \alpha \rrbracket$ is of type e, and β and γ denote sets whose union is necessarily all of D, then " $\alpha \beta$ or $\alpha \gamma$ " is a tautology. But **Everybody here is over 30 or everybody here is under 40** is not a tautology. So [**everybody**] here cannot be in D_e . The proof that a suitable denotation exists in $D_{\langle\langle e,t\rangle,t\rangle}$ is similarly trivial.

In sum, once we allow DPs to have meanings of type $\langle\langle e,t\rangle,t\rangle$, the problematic predictions of a theory that would only allow type e (or types e and $\langle e,t\rangle$) are no longer derived. This in itself is not a tremendously impressive accomplishment, of course. There are always many boring theories that avoid making bad predictions about logical relations by avoiding making any at all.

Truth-conditional effects of syntactic reorganization

It is more challenging to reflect on our second argument against the uniform type e analysis of DPs. We observed that certain syntactic operations which systematically preserve truth-conditions when they affect names (for example, topicalization, passive) sometimes alter them when they affect quantifiers. Does the type $\langle\langle e,t\rangle,t\rangle$ analysis help us predict *this* fact, and if so, why?

For reasons that we will attend to in the next chapter, the English examples we used to illustrate this issue in section 6.1.2 involve complications that we cannot yet handle. For instance, we are not ready to treat both sentences in pairs like (3a) and (3b).

- (3) (a) Everybody answered many questions correctly.
 (b) Many questions, everybody answered correctly.

But we can construct a hypothetical case that shows how it is possible in principle to alter truth-conditions by the topicalization of a phrase whose meaning is of type $\langle\langle e, t \rangle, t \rangle$.

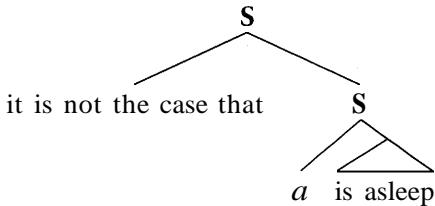
Suppose we topicalized the embedded subject a in a structure of the form (4a), yielding (4b).

- (4) (a) It is not the case that a is asleep.
 (b) a , it is not the case that t is asleep.

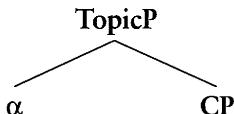
Never mind that structures like (4b) are syntactically ill-formed. We only want to show here that (4a) and (4b) can differ in truth-conditions when $[\alpha] \in D_{\langle\langle e, t \rangle, t \rangle}$, which they couldn't when $[\alpha] \in D$.

Here are two possible structures (see section 6.1.2).

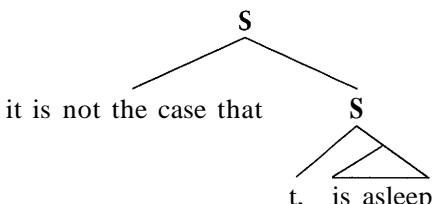
(4a')



(4b')



wh_1



If $\llbracket \alpha \rrbracket \in D$, the two structures are provably equivalent. (Exercise: Give the proof.) But what happens if $\llbracket a \rrbracket \in D_{\langle\langle e, t\rangle, t\rangle}$? We will show that in this case their truth-conditions may differ. Suppose, for instance, that $a = \text{everything}$, with the lexical entry given above. Then we calculate as follows:

$$(5) \quad \llbracket (4a) \rrbracket = 1 \\ \text{iff} \\ \llbracket \text{everything is asleep} \rrbracket = 0 \\ \text{iff} \\ \llbracket \text{everything} \rrbracket (\llbracket \text{asleep} \rrbracket) = 0 \\ \text{iff} \\ \llbracket \text{asleep} \rrbracket(x) = 0 \text{ for some } x \in D.$$

$$(6) \quad \llbracket (4b) \rrbracket = 1 \\ \text{iff} \\ \llbracket \text{everything} \rrbracket (\llbracket \text{wh}_1 \text{ it is not the case that } t_1 \text{ is asleep} \rrbracket) = 1 \\ \text{iff} \\ \llbracket \text{wh}_1 \text{ it is not the case that } t_1 \text{ is asleep} \rrbracket(x) = 1 \text{ for all } x \in D \\ \text{iff} \\ \llbracket \text{it is not the case that } t, \text{ is asleep} \rrbracket^{[1 \rightarrow x]} = 1 \text{ for all } x \in D \\ \text{iff} \\ \llbracket t_1 \text{ is asleep} \rrbracket^{[1 \rightarrow x]} = 0 \text{ for all } x \in D \\ \text{iff} \\ \llbracket \text{asleep} \rrbracket(x) = 0 \text{ for all } x \in D.$$

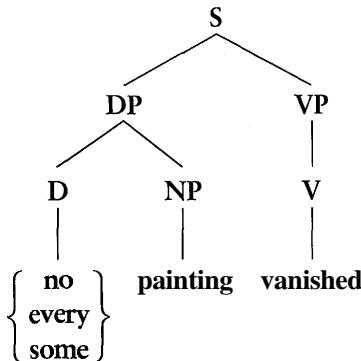
The last lines of (5) and (6) express clearly distinct conditions. It is easy to imagine that the actual facts might verify the former but falsify the latter: this happens whenever some, but not all, individuals sleep.

So we have shown that topicalization of a phrase with a type $\langle\langle e, t\rangle, t\rangle$ meaning can affect truth-conditions. The example was a hypothetical one, but we will soon be able to analyze real-life examples of the same kind.

6.4 Quantifying determiners

Having decided on the denotations for quantifying DPs like "nothing", "everything", or "something", we are now ready to find denotations for the quantifying determiners "no", "every", "some", and what have you. Consider the following structure:

(1)

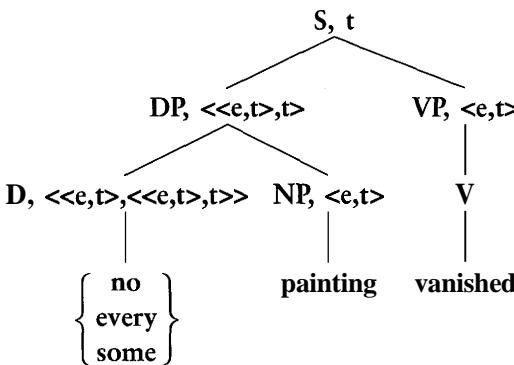


We reason as before. Assume that

- (a) The phrase structures for phrases containing quantifying determiners are as given above.
- (b) The semantic type of common nouns is $\langle e, t \rangle$.
- (c) The semantic type of quantificational DPs is $\langle \langle e, t \rangle, t \rangle$.
- (d) Determiners and NPs semantically combine via functional application.

It now follows that the semantic type of quantifying determiners is $\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$. The annotated tree (2) illustrates the composition process. (2) has two binary-branching nodes, and in each case, the mode of composition is Functional Application.

(2)



As for the lexical entries, we have the following:

$$\begin{aligned} [\text{every}] &= \lambda f \in D_{\langle e, t \rangle} . [hg \in D_{\langle e, t \rangle} . \text{for all } x \in D_e \text{ such that } f(x) = 1, g(x) = 1] \\ [\text{no}] &= hf \in D_{\langle e, t \rangle} . [\lambda g \in D_{\langle e, t \rangle} . \text{there is no } x \in D_e \text{ such that } f(x) = 1 \text{ and } g(x) = 1] \end{aligned}$$

$$[\text{some}] = \lambda f \in D_{\langle e, t \rangle} . [hg \in D_{\langle e, t \rangle} . \text{there is some } x \in D_e \text{ such that } f(x) = 1 \text{ and } g(x) = 1]$$

Exercise

Calculate the truth-conditions for (1) step by step.

Our analysis of quantifying determiners with denotations of type $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t \rangle\rangle$, is essentially the one first proposed by David Lewis in "General Semantics."¹⁴ A similar proposal for the treatment of quantification in natural languages was independently made by Richard Montague, and was taken up in subsequent work by Cresswell, Barwise and Cooper, and Keenan and Stavi.¹⁵

6.5 Quantifier meanings and relations between sets

6.5.1 A little history

The semantics for quantifying determiners that we arrived at in the last section is a variant of the oldest known view of quantification, the so-called relational view, which can be traced back to Aristotle. About Aristotelian logic (syllogistics), the dominant paradigm up to the nineteenth century, Dag Westerståhl says the following:

The syllogistics is basically a theory of inference patterns among quantified sentences. Here a quantified sentence has the form

(1) QXY

where X, Y are universal terms (roughly 1-place predicates) and Q is one of the quantifiers all, some, no, **not all**. In practice, Aristotle treated these quantifiers as relations between terms. Aristotle chose to study a particular type of inference pattern with sentences of the form (1), the syllogisms. A syllogism has two premisses, one conclusion, and three universal terms (variables). Every sentence has two different terms, all three terms occur in the premisses, and one term, the "middle" one, occurs in both premisses but not in the conclusion. It follows that the syllogisms can be grouped into four different "figures", according to the possible configurations of variables:

Q_1ZY	Q_1YZ	Q_1ZY	Q_1YZ
Q_2XZ	Q_2XZ	Q_2ZX	Q_2ZX
<hr/>	<hr/>	<hr/>	<hr/>
Q_3XY	Q_3XY	Q_3XY	Q_3XY

Here the Q_i can be chosen among the above quantifiers, so there are $4^4 = 256$ syllogisms.

Now the question Aristotle posed – and, in essence, completely answered – can be formulated as follows:

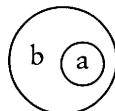
For what choices of quantifiers are the above figures valid?

For example, if we in the first figure let $Q_1 = Q_2 = Q_3 = \text{all}$, a valid syllogism results ("Barbara" in the medieval mnemonic); likewise if $Q_1 = Q_3 = \text{no}$ and $Q_2 = \text{all}$ ("Celarent"). Note that Aristotle's notion of validity is essentially the modern one: a syllogism is valid if every instantiation of X, Y, Z verifying the premisses also verifies the conclusion.¹⁶

In their history of logic, William and Martha Kneale describe a contribution by Leibniz as popularized by the eighteenth-century mathematician L. Euler:¹⁷

Leonhard Euler's "Lettres a une Princesse d'Allemagne" (which were written in 1761 and published in St. Petersburg in 1768) must be mentioned among works of the eighteenth century that contributed something to mathematical logic. Those letters which deal with logic contain no attempt to work out a calculus, though Euler was a great mathematician; but they popularized Leibniz's device of illustrating logical relations by geometrical analogies, and this had some influence on thinkers in the next century. In particular it directed attention to the extensional or class interpretation of general statements; for Euler represented (or rather illustrated) the four Aristotelian forms of statements by three relations of closed figures according to the following scheme ...

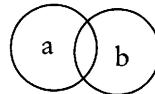
Every a is b



No a is b



Some a is b



And Frege (though he may be more famous for his invention of predicate logic, with its 1-place quantifiers) also endorsed the relational view of quantifiers in various places: "the words all, every, no, some combine with concept words [our NPs or VPs]. In universal and particular affirmative and negative statements we express relations between concepts and indicate the specific nature of these relations by means of those words."¹⁸

6.5.2 Relational and Schonfinkled denotations for determiners

On the relational theory of quantification, quantifiers denote relations between sets. For instance, "every" denotes the subset relation, and "some" denotes the relation of non-disjointness. A sentence like "Every goat is a mutt" is understood as stating that the set of goats is a subset of the set of mutts. And a sentence like "Some goat is a mutt" is understood as stating that the set of all goats is not disjoint from the set of all mutts.

If we take "relation" in its exact mathematical sense (a set of ordered pairs), then our semantics for quantifying determiners from the previous section is not literally an instance of the relational theory. But there is a very straightforward connection between our determiner denotations and relations between sets.

Here is a sample of the relations that common quantifying determiners would express on a relational theory in the strictest sense of the word:

(1) For any $A \subseteq D$ and any $B \subseteq D$:

- (a) $\langle A, B \rangle \in R_{\text{every}}$ iff $A \subseteq B$
- (b) $\langle A, B \rangle \in R_{\text{some}}$ iff $A \cap B \neq \emptyset$
- (c) $\langle A, B \rangle \in R_{\text{no}}$ iff $A \cap B = \emptyset$
- (d) $\langle A, B \rangle \in R_{\text{at least two}}$ iff $|A \cap B| \geq 2^{19}$
- (e) $\langle A, B \rangle \in R_{\text{most}}$ iff $|A \cap B| > |A - B|$

It is clear why we are not assuming in this book that $[\text{every}] = R_{\text{every}}$, $[\text{some}] = R_{\text{some}}$, etcetera. This would not be consistent with our goal of minimizing the number of principles of semantic composition. (If, for example, **every** literally were R_{every} , we could not interpret **Every painting vanished** by functional application, but would need special new rules.) Nevertheless, just as it is sometimes convenient to pretend that VPs denote sets (though they really denote functions), it can be convenient to talk as if determiners denoted relations between sets. We just have to understand clearly how such talk translates back into our official theory.

To see the connection between the relations defined in (1) above and our determiner meanings of type $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t \rangle\rangle$, it is helpful to appreciate an important analogy between quantificational determiners and transitive verbs: Both can be viewed as expressing 2-place relations, the only difference being that the latter are first-order relations (they relate individuals), while the former are second-order relations (they relate sets of individuals or characteristic functions thereof). In the section on Schonfinkelization, we saw how to construct functions in $D_{\langle\langle e, \langle\langle e, t \rangle \rangle \rangle}$ out of first-order 2-place relations. By analogous operations, we can get from second-order 2-place relations to functions in $D_{\langle\langle e, \langle\langle e, t \rangle \rangle \rangle}$.

To consider a concrete case, how do we get from R_{every} to [every]? Our starting point is (1a) (repeated from above).

$$(1a) \quad R_{\text{every}} = \{<A, B> \in \text{Pow}(D) \times \text{Pow}(D) : A \subseteq B\}^{20}$$

The set R_{every} has a characteristic function, which we'll call " F_{every} ":

$$(2) \quad F_{\text{every}} = \lambda <A, B> \in \text{Pow}(D) \times \text{Pow}(D) . A \subseteq B.^{21}$$

F_{every} is a 2-place function that maps pairs of sets of individuals into truth-values. It can now be Schonfinkled in two ways. We Schönfinkel it from left to right and call the result " f_{every} ":

$$(3) \quad f_{\text{every}} = \lambda A \in \text{Pow}(D) . [\lambda B \in \text{Pow}(D) . A \subseteq B].$$

We are almost there now: f_{every} just about could serve as the meaning of **every** if it weren't for the fact that NPs and VPs denote functions, not sets. To correct this, we must replace the sets A and B by functions f and g , and accordingly rewrite the subset condition to the right of "iff".

$$(4) \quad [\![\text{every}]\!] = \lambda f \in D_{\ll e, t \rr} . [\lambda g \in D_{\ll e, t \rr} . \{x \in D : f(x) = 1\} \subseteq \{x \in D : g(x) = 1\}].$$

This is the lexical entry for **every** as we determined it in section 6.4 in a slightly different, but equivalent, formulation. In definition (4), we just used a bit more set-theoretic notation for the formulation of the value description of the h-term.

Exercise

Every one of R_{every} , F_{every} , f_{every} , and our [every] is in principle a candidate for the denotation of the English determiner **every** provided that one assumes suitable composition principles to fit one's choice. And these four are not the only possibilities. Due to the systematic relations between sets and their characteristic functions, and between 2-place functions and their various Schönfinkelizations, there are lots of additional variants. Most of them are of no particular interest and don't occur in the literature; others happen to be common. In the influential paper by Barwise and Cooper (see n. 13), determiners were treated as denoting *functions from Pow(D) into Pow(Pow(D))*. Your task in this exercise is to spell out this option.

-
- (a) Give examples of Barwise and Cooper-style lexical entries for a couple of run-of-the-mill determiners.
 - (b) Specify the composition rules that are needed in conjunction with these lexical entries.
 - (c) Show that there is a one-to-one correspondence between functions from $\text{Pow}(D)$ into $\text{Pow}(\text{Pow}(D))$ and relations between subsets of D .
-

6.6 Formal properties of relational determiner meanings

If quantifiers correspond to binary relations, we may investigate whether they do or do not have standard properties of binary relations like symmetry, reflexivity, and so on. Aristotle was already interested in questions of this kind. A substantial part of the *Prior Analytics* is dedicated to investigating whether the two terms A and B in a statement of the form QAB are "convertible"; that is, whether the quantifier involved expresses a symmetric relation. By way of illustration, let Q be the determiner "some", and A and B the predicates "US citizen" and "native speaker of Spanish" respectively. Since "Some US citizens are native speakers of Spanish" and "Some native speakers of Spanish are US citizens" are logically equivalent (that is, they have the same truth-conditions), we have grounds to believe that "some" is convertible. It all depends, of course, on whether the logical equivalence is unaffected by the particular choice of predicates.

Here is a list of some potentially interesting mathematical properties of relational determiner denotations. In every definition, δ stands for a determiner, and A, B, C range over subsets of the domain D .²²

Defniendum

- δ is reflexive
- δ is irreflexive²³
- δ is symmetric
- δ is antisymmetric

- δ is transitive

- δ is conservative

- δ is left upward monotone²⁴

- δ is left downward monotone

Definiens

- for all $A : \langle A, A \rangle \in R_\delta$
- for all $A : \langle A, A \rangle \notin R_\delta$
- for all $A, B : \text{if } \langle A, B \rangle \in R_\delta, \text{ then } \langle B, A \rangle \in R_\delta$
- for all $A, B : \text{if } \langle A, B \rangle \in R_\delta \text{ and } \langle B, A \rangle \in R_\delta, \text{ then } A = B$
- for all $A, B, C : \text{if } \langle A, B \rangle \in R_\delta \text{ and } \langle B, C \rangle \in R_\delta, \text{ then } \langle A, C \rangle \in R_\delta$
- for all $A, B : \langle A, B \rangle \in R_\delta \text{ iff } \langle A, A \cap B \rangle \in R_\delta$
- for all $A, B, C : \text{if } A \subseteq B \text{ and } \langle A, C \rangle \in R_\delta, \text{ then } \langle B, C \rangle \in R_\delta$
- for all $A, B, C : \text{if } A \subseteq B \text{ and } \langle B, C \rangle \in R_\delta, \text{ then } \langle A, C \rangle \in R_\delta$

δ is right upward monotone	for all A, B, C : if $A \subseteq B$ and $\langle C, A \rangle \in R_\delta$, then $\langle C, B \rangle \in R_\delta$
δ is right downward monotone	for all A, B, C : if $A \subseteq B$ and $\langle C, B \rangle \in R_\delta$, then $\langle C, A \rangle \in R_\delta$

Exercise

For every property defined above, try to find determiners whose denotations have it, as well as determiners whose denotations don't.

Why have linguists been interested in classifying determiner meanings according to such mathematical properties? Is this just a formal game, or does it throw some light on the workings of natural language? Modern research within the generalized quantifier tradition²¹ has shown that some of those mathematical properties may help formulate constraints for possible determiner meanings. Keenan and Stavi, for example, have proposed that all natural language determiners are conservative.²⁶ In order to see what a non-conservative determiner would look like, imagine that English "only" was a determiner rather than an adverb. The non-equivalence of "only children cry" and "only children are children that cry" would now establish that "only" is a determiner that is not conservative. Conservativity, then, is a non-trivial potential semantic universal.

Other mathematical properties have been argued to pick out linguistically significant classes of DPs. The following two exercises will give you some taste of this influential line of research. When you work on those exercises, be warned that you may not be able to come up with completely satisfactory answers. Try your best, and note any open problems. If you want to delve deeper into those areas, consult the pioneering work of Milsark, Fauconnier, and Ladusaw,²⁷ and the handbook articles mentioned in note 25 at the end of the chapter for further directions.

Exercise on "there"-insertion

It has often been observed that not all kinds of NPs are allowed in "there"-insertion constructions. Here are two examples:

- (i) There are some apples in my pocket.
- (ii) *There is every apple in my pocket.

Test a number of quantifiers as to their behavior in "there"-insertion constructions, and try to characterize the class of quantifiers that are permitted in this environment with the help of some formal property of determiner denotations. Consider the mathematical properties defined above.

Exercise on negative polarity

The adverb "ever" is an example of a so-called *negative polarity item* (NPI), so called because it seems to require a negative environment:

- (i) I haven't ever visited the Big Bend National Park.
- (ii) *I have ever visited the Big Bend National Park.

However, there needn't always be a "not" to license "ever". For instance, the following examples are also grammatical.

- (iii) Very few people ever made it across the Cisnos range.
- (iv) Every friend of mine who ever visited Big Bend loved it.

Try out other sentences like (iii) and like (iv) with different determiners in place of "very few" and "every". Which property of these determiners seems to correlate with the distribution of "ever"? Again, consider the properties defined above.

6.7 Presuppositional quantifier phrases

Determiners denote functions of type $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t \rangle \rangle$. Total or partial functions? So far we have tacitly assumed the former. The lexical entries we have given for every, some, no, more than two, etcetera, all define total functions. They thus guarantee that every a , some a , no a , ... always have a semantic value, regardless of the facts (provided that a itself has a semantic value²⁸). In other words, quantifying determiners, as we have treated them so far, never give rise to presuppositions.

But we have no good reason to assume that this is generally correct for all quantifying determiners of natural languages. Indeed, there are some persuasive examples of determiners which seem to denote partial functions.

6.7.1 “Both” and “neither”

Consider the determiners both and neither, as in both cats, neither cat.²⁹ What are the intuitive truth-conditions of a sentence like (1), for instance?

(1) Neither cat has stripes.

If there are exactly two cats and neither has stripes, (1) is clearly true. If there are exactly two cats and one or both of them have stripes, (1) is clearly false. But what if there aren't exactly two cats? For example, suppose there is just one cat and it doesn't have stripes. Or suppose there are three cats (all equally relevant and salient) and none of them has stripes. In such circumstances, we are reluctant to judge (1) either true or false; rather, it seems inappropriate in much the same way as an utterance of the cat when there isn't a unique (relevant and salient) cat. This suggests the following lexical entry:³⁰

(2) $f_{\text{neither}} = \lambda A : A \in \text{Pow}(D) \& |A| = 2 . [\lambda B \in \text{Pow}(D) . A \cap B = \emptyset]$

Together with the definition of "presupposition" in chapter 4, (2) predicts that (1) presupposes there to be exactly two cats.

Similar judgments apply to sentences like Both cats have stripes, suggesting an analogous lexical entry for both:

(3) $f_{\text{both}} = \lambda A : A \in \text{Pow}(D) \& |A| = 2 . [\lambda B \in \text{Pow}(D) . A \subseteq B]$.

Exercise

Give a precise characterization of the relation between [neither] and [no] and the relation between [both] and [every].

6.7.2 Presuppositionality and the relational theory

The existence of presuppositional determiners like neither and both is actually incompatible with a strictly relational theory of quantifiers. A relation, as you recall, is a set of ordered pairs. A given ordered pair is either an element of a given relation, or else it is not. There is no third possibility. Suppose, for instance, that R is some relation between sets of individuals; that is, R is some subset of $\text{Pow}(D) \times \text{Pow}(D)$. Then for any arbitrary pair of sets $\langle A, B \rangle \in \text{Pow}(D) \times \text{Pow}(D)$, we either have $\langle A, B \rangle \in R$ or $\langle A, B \rangle \notin R$. The characteristic function of R is a total function with domain $\text{Pow}(D) \times \text{Pow}(D)$, and any Schonfinkelizations of it are total functions with domain $\text{Pow}(D)$.

This means that the procedure for constructing determiner meanings from relations which we gave in section 6.5.2 always produces total functions, hence non-presuppositional meanings. In practice, this means that we cannot obtain the entries for both and neither by this construction. Put differently, we cannot fully describe the meanings of these determiners by sets of ordered pairs of sets. If we try, for example, for both, the best we can come up with is (4).

$$(4) \quad R_{\text{both}} =^? \{<A, B> \in \text{Pow}(D) \times \text{Pow}(D) : A \subseteq B \ \& \ |A| = 2\}$$

(4) correctly characterizes the conditions under which a sentence of the form both $a \beta$ is true, but it fails to distinguish the conditions under which it is false from those where it has no truth-value. If we based our lexical entry for both on (4), we would therefore predict, for example, that (5) is true (!) in a situation in which there is only one cat and I saw it.

$$(5) \quad \text{I didn't see both cats.}$$

This is undesirable. Speakers' actual judgments about (5) in this situation fit much better with the predictions of the presuppositional analysis we adopted earlier.

While a strictly relational theory cannot describe the meanings of both and neither, an almost relational theory, on which determiner meanings are potentially partial functions from $\text{Pow}(D) \times \text{Pow}(D)$ to $\{0, 1\}$, would work fine. On such a theory, both and no, for instance, would denote the functions defined in (6).

- $$(6) \quad \begin{aligned} (a) \quad F_{\text{both}} &= \lambda<A, B> : A \subseteq D \ \& \ B \subseteq D \ \& \ |A| = 2 . \quad A \subseteq B. \\ (b) \quad F_{\text{no}} &= \lambda<A, B> : A \subseteq D \ \& \ B \subseteq D . \quad A \neq B = O. \end{aligned}$$

In the case of a nonpresuppositional determiner 6, F_δ happens to be the characteristic function of some relation on $\text{Pow}(D)$, but not when δ is presuppositional. In practice, the label "relational theory" is also applied to such an almost relational theory.

The fact that presuppositional determiners do not correspond to relations in the strictest sense gives rise to some indeterminacy as to how standard properties of relations apply to them. Consider, for example, the notion of irreflexivity. As a property of relations between subsets of D, it may be defined as in (7a) or (7b).

- $$(7) \quad \begin{aligned} (7a) \quad R \text{ is irreflexive} &\quad \text{iff for all } A \subseteq D, <A, A> \notin R. \\ (7b) \quad &\quad \text{... iff for no } A \subseteq D, <A, A> \in R. \end{aligned}$$

(7a) is the definition we employed above; (7b) would have been a fully equivalent choice, and no less natural. Suppose we now ask ourselves whether neither is an irreflexive determiner. The initial answer that we come up with is that this

is not a well-defined question. The only definition that we have for "irreflexivity" as a property of determiners is the one in section 6.6. But this cannot be applied to neither, since there is no such thing as R_{neither} . We just learned that it is not possible to define such a relation.

We might stop right here and agree henceforth that the concept of irreflexivity is applicable only to nonpresuppositional quantifiers. But this may not be desirable if this concept plays some role in our semantic theory. For example, Barwise and Cooper (see note 13) have claimed that it plays a central role in the analysis of the English "there"-construction. If we are interested in this sort of claim, we have an incentive to look for a natural extension of the definition of irreflexivity that will allow it to apply to both total and partial determiner meanings.

So let's replace the "R" in (7) by an "F", which stands for a possibly partial function from $\text{Pow}(D) \times \text{Pow}(D)$ into $\{0, 1\}$. How should we rewrite the rest? There are, in principle, two possibilities:

(8) F is irreflexive

- (a) ... iff for all $A \subseteq D$, $F(A, A) = 0$.
- (b) ... iff for no $A \subseteq D$, $F(A, A) = 1$.

When F happens to be total – that is, when F is the characteristic function of some relation – then (8a) and (8b) are equivalent, and they amount to exactly the same thing as (7). More precisely, if R is irreflexive in the sense of (7), then char_R^{31} is irreflexive in the sense of both (8a) and (8b); and if char_R is irreflexive in the sense of either (8a) or (8b), then R is irreflexive in the sense of (7). (Exercise: Prove this.) This being so, both (8a) and (8b) qualify as natural extensions of the basic concept of irreflexivity. However, the two are not equivalent! They do coincide for total functions F, but they diverge for partial ones. To see this, consider the function F_{neither} :

(9) $F_{\text{neither}} = \lambda \langle A, B \rangle : A \subseteq D \ \& \ B \subseteq D \ \& \ |A| = 2 \ . \ A \cap B = \emptyset$.

If we adopt (8b), F_{neither} qualifies as irreflexive, but if we adopt (8a), it doesn't. (Exercise: Prove this.)

So we have to make a choice between (8a) and (8b). From a purely formal point of view, the choice is entirely arbitrary. The standard choice in the linguistic literature is (8b), for reasons that have to do with the intended empirical applications (see exercise below). (8b) is the more liberal definition of the two, in the sense that it makes it easier to qualify as irreflexive: Any F that is irreflexive according to (8a) is also irreflexive under (8b), but not vice versa. This fact becomes more transparent if we reformulate (8b) as follows:

(10) F is irreflexive

- iff for all $A \subseteq D$ such that $\langle A, A \rangle \in \text{dom}(F) : F(A, A) = 0$.

Exercise

Prove that (10) is equivalent to (8b).

Similar choices arise when other mathematical properties are extended from relations between sets to potentially partial determiner meanings. In many such cases, one of the *prima facie* reasonable definitions has been chosen as the standard definition in the linguistic literature. For instance, here are the official extended versions for selected concepts from our list in section 6.6. As before, "A", "B", and "C" range over subsets of D, and δ stands for a determiner.

Definiendum

δ is reflexive

δ is irreflexive

δ is symmetric

δ is conservative

δ is left upward monotoneⁱ²

δ is left downward monotone

δ is right upward monotone

δ is right downward monotone

Definiens

for all A such that $\langle A, A \rangle \in \text{dom}(F_\delta)$:
 $F_\delta(A, A) = 1$

for all A such that $\langle A, A \rangle \in \text{dom}(F_\delta)$:
 $F_\delta(A, A) = 0$

for all A, B : $\langle A, B \rangle \in \text{dom}(F_\delta)$ iff $\langle B, A \rangle \in \text{dom}(F_\delta)$, and $F_\delta(A, B) = 1$ iff $F_\delta(B, A) = 1$
 for all A, B : $\langle A, B \rangle \in \text{dom}(F_\delta)$ iff $\langle A, A \sqcap B \rangle \in \text{dom}(F_\delta)$, and $F_\delta(A, B) = 1$ iff $F_\delta(A, A \sqcap B) = 1$

for all A, B, C : if $A \subseteq B$, $F_\delta(A, C) = 1$, and
 $\langle B, C \rangle \in \text{dom}(F_\delta)$, then $F_\delta(B, C) = 1$

for all A, B, C : if $A \subseteq B$, $F_\delta(B, C) = 1$, and
 $\langle A, C \rangle \in \text{dom}(F_\delta)$, then $F_\delta(A, C) = 1$

for all A, B, C : if $A \subseteq B$, $F_\delta(C, A) = 1$, and
 $\langle C, B \rangle \in \text{dom}(F_\delta)$, then $F_\delta(C, B) = 1$.

for all A, B, C : if $A \subseteq B$, $F_\delta(C, B) = 1$, and
 $\langle C, A \rangle \in \text{dom}(F_\delta)$, then $F_\delta(C, A) = 1$.

Exercise

Go back to the exercises on "there-sentences and negative polarity in section 6.6, and reconsider your answers in the light of the present discussion of presuppositional determiners.

6.7.3 Other examples of presupposing DPs

Our observations about the presupposition of both cats generalize to DPs of the form the two cats, the three cats, the four cats, etcetera. A sentence like

The twenty-five cats are in the kitchen is felt to presuppose that there are twenty-five (relevant)cats, and to assert that all (relevant)cats are in the kitchen. Our semantics should thus predict that, for any numeral **n** and any NP α :

- (11) **the n a** has a semantic value only if $|[\alpha]| = n$.

Where defined, $[\text{the } n \alpha] = \lambda A . [\alpha] \subseteq A$.

How do we ensure this prediction by appropriate lexical entries in conjunction with the usual composition principles?

The answer to this question depends on what we take to be the syntactic structure of DPs of this form. If we could assume that their constituent structure is $[\text{DP} [D \text{ the } n] \text{ NP}]$, and could treat **the n** as a lexical unit, it would be straightforward. We would only have to write lexical entries like those in (12), then.

- (12) $F_{\text{the three}} = \lambda A, B : |A| = 3 . A \subseteq B$

$F_{\text{the four}} = \lambda A, B : |A| = 4 . A \subseteq B$

etc.

This analysis is quite clearly not right. The definite article and the numeral are evidently lexical items in their own right, and we would like to explain how their separate meanings contribute to the meaning of the DP as a whole. Moreover, the constituent structure of such DPs seems to be $[\text{the} [\text{n NP}]]$ rather than $[[\text{the n} \text{ NP}]]$. Unfortunately, we cannot attend to these facts, since we will not be able to go into the semantics for plural NPs in this book.³³ Well aware of its limitations, then, we will assume the ad hoc analysis in (12).

Exercise

Suppose we treated **the cat** and **the cats** as elliptical surface variants of **the one cat** and **the at least two cats** respectively. How does this treatment compare to the semantics for singular definites that we gave in chapter 4?

Another type of DP that gives rise to presuppositions is the so-called partitive construction, exemplified by **two of the five cats**, **none of the (at least two) dogs**, etcetera. Assuming again a very ad hoc syntactic analysis, we could capture their presuppositions by lexical entries for the complex determiners along the lines of (13).

- (13) $F_{\text{none of the (at least two)}} = \lambda A, B : |A| \geq 2 . A \cap B = \emptyset$

Exercise

Return to the exercises on there-sentences and NPIs in section 6.6. What predictions do the solutions you proposed imply for DPs of the forms considered in this section, e.g., **the seven cats, none of the cats?**

6.8 Presuppositional quantifier phrases: controversial cases

In the previous section, we considered some types of DPs which carry presuppositions about the cardinality of their restrictors. For most of them, we did not provide a serious compositional analysis to pinpoint the exact lexical sources of these presuppositions. But whatever their sources may turn out to be, the fact that the DPs as a whole carry the presuppositions in question is rather apparent and uncontroversial.

With these observations in the background, we now turn to a family of much more controversial claims about the presuppositions of quantified structures in natural languages.

6.8.1 Strawson's reconstruction of Aristotelian logic

The meanings we have been assuming throughout this chapter for sentences with every, some (**a**), and no are the ones that were promoted by the founders of modern logic. We repeat once more the definitions of the relevant second-order relations:

- (1) (a) $R_{\text{every}} = \{<A, B> : A \subseteq B\}$
- (b) $R_{\text{some}} = \{<A, B> : A \cap B \neq \emptyset\}$
- (c) $R_{\text{no}} = \{<A, B> : A \cap B = \emptyset\}$

As has been noted by many commentators on the history of logic,³⁴ the definitions in (1) are not consistent with certain assumptions about the semantics for every, some, no that were part of (at least some versions of) Aristotelian logic. In many works in this tradition, generalizations such as the ones in (2) were considered valid.

- (2) For any predicates α , β :
- every $\alpha \beta$ and no $\alpha \beta$ is a contradiction.
 - some $\alpha \beta$ or some α not β is a tautology.
 - every $\alpha \beta$ entails some $\alpha \beta$.
 - no $\alpha \beta$ entails some α not β .

We give a concrete English instance for every of the sentence schemata in (2(i))–(2(iv)):

- (3) (a) Every first-year student in this class did well and no first-year student in this class did well.
 (b) Some cousin of mine smokes, or some cousin of mine doesn't smoke.
 (c) Every professor at the meeting was against the proposal.
 ∴ Some professor at the meeting was against the proposal.
 (d) No student presentation today was longer than an hour.
 ∴ Some student presentation today wasn't longer than an hour.

According to the Aristotelian laws in (2), (3a) is contradictory, (3b) is a tautology, and (3c) and (3d) are valid inferences. The predictions of the modern (“classical”³⁵) analysis in (1), are otherwise: (a) is contingent; specifically, it is true when there are no first-year students in this class. (b) is likewise contingent; it is false when the speaker has no cousins. The premise in (c) doesn't entail the conclusion: the former is true but the latter false when there were no professors at the meeting. Likewise for (d): its premise is true, but its conclusion false if there were no student presentations today.

Which of the two sets of predictions is correct for English? As we will see, the empirical evidence bearing on this question is surprisingly difficult to assess. Before we take a closer look at it, let's present a concrete semantic analysis of the determiners every and some that predicts the validity of the Aristotelian laws in (2). Such an analysis is suggested in the following passage from Strawson:³⁶

Suppose someone says “All John's children are asleep”. Obviously he will not normally, or properly, say this, unless he believes that John has children (who are asleep). But suppose he is mistaken. Suppose John has no children. Then is it true or false that all John's children are asleep? Either answer would seem to be misleading. But we are not compelled to give either answer. We can, and normally should, say that, since John has no children, the question does not arise. . . .

. . . The more realistic view seems to be that the existence of children of John's is a necessary precondition not merely of the truth of what is said, but of its being *either* true *or* false. . . .

... What I am proposing, then, is this. There are many ordinary sentences beginning with such phrases as "All . . .", "All the . . .", "No . . .", "None of the . . .", "Some . . .", "Some of the . . .", "At least one . . .", "At least one of the . . ." which exhibit, in their standard employment, parallel characteristics to those I have just described in the case of a representative "All . . ." sentence. That is to say, the existence of members of the subject-class³⁷ is to be regarded as presupposed (in the special sense described) by statements made by the use of these sentences; to be regarded as a necessary condition, not of the truth simply, but of the truth or falsity, of such statements. I am proposing that the four Aristotelian forms [i.e., "all a β ", "no a β ", "some a β ", "some a not β ", as they are interpreted in Aristotelian logic] should be interpreted as forms of statements of this kind. Will the adoption of this proposal protect the system from the charge of being inconsistent when interpreted? Obviously it will. For every case of invalidity, of breakdown in the laws [of Aristotelian logic], arose from the non-existence of members of some subject-class being incompatible with either the truth or the falsity of some statement of one of the four forms. So our proposal, which makes the non-existence of members of the subject-class incompatible with either the truth or the falsity of any statement of these forms, will cure all these troubles at one stroke. We are to imagine that every logical rule in the system, when expressed in terms of truth and falsity, is preceded by the phrase "Assuming that the statements concerned are either true or false, then . . ." Thus . . . the rule that [all a β] entails [some a β] states that, if corresponding statements of these forms have truth-values, then if the statement of the [form all a β] is true, the statement of the [form some a β] must be true; and so on.

At the beginning of this quote, it is not immediately apparent which set of English determiners Strawson means to make claims about. His primary example all John's children is probably among those less controversial candidates for a presuppositional analysis which we already surveyed in the previous section. Possessive DPs like John's children are standardly analyzed as covert definite descriptions (with structures essentially of the form the children (of)John), and all + definite is taken to be a partitive. (So all John's children is a surface variant of all of John's children, the optionality of of being an idiosyncratic property of all.) As far as this particular example goes, then, Strawson's claims may not go beyond what is commonplace in contemporary formal semantics, and the same goes for none of the, some of the, and at least one of the, which he lists a few paragraphs down. But in this same list he also includes plain all, no, some, and at least one. So it is clear that he means his proposal to extend beyond the partitives, and we concentrate here on its application to the simple determiners he mentions, plus every (which we assume Strawson would not distinguish from all in any respect relevant here).

According to Strawson then, at least some occurrences of English every, some, etcetera behave as if their lexical entries were not as in (1) above, but rather as below:

- (4) (a) $F_{\text{every}} = \lambda A, B : A \neq O . A \subseteq B.$
- (b) $F_{\text{no}} = \lambda A, B : A \neq O . A \cap B = O.$
- (c) $F_{\text{some}} = \lambda A, B : A \neq O . A \cap B \neq O.$

These entries validate the Aristotelian generalizations in (2), provided, as Strawson notes, that there are suitable definitions of the basic semantic properties. Here is a proposal:

Basic semantic properties

ϕ is a *tautology* iff the semantic rules alone establish that, if ϕ is in the domain of $\llbracket \cdot \rrbracket$, then $\llbracket \phi \rrbracket = 1$.

ϕ is a *contradiction* iff the semantic rules alone establish that, if ϕ is in the domain of $\llbracket \cdot \rrbracket$, then $\llbracket \phi \rrbracket = 0$.

ϕ *entails* ψ iff the semantic rules alone establish that, if ϕ and ψ are both in the domain of $\llbracket \cdot \rrbracket$ and $\llbracket \phi \rrbracket = 1$, then $\llbracket \psi \rrbracket = 1$.

So it is reasonable to hypothesize that, to the extent that native speakers' judgments about examples like (3) conform to the predictions in (2), this is due to their using the entries in (4) rather than those in (1).

6.8.2 Are all determiners presuppositional?

Strawson was not directly engaged in natural language semantics, and it is impossible to attribute to him any very specific claim in that domain. He did say that at least some uses of simple English determiners carried existence presuppositions, but this certainly doesn't imply that (4a), (4b), and (4c) are the lexical entries for English every, no, some. At best it implies that some English utterances are correctly translated into a symbolic language whose determiners have the semantics of (4). As regards English itself, this leaves many possibilities open. Perhaps (4) represents certain readings, among others, of lexically ambiguous items of English. Or perhaps it's not the lexical denotations of the determiners at all that are responsible for the relevant presuppositions, but other ingredients of the structures in which they occur. Linguists inspired by Strawson's discussion have explored various concrete options in this regard. We will begin here with a hypothesis that is simpler and more radical than most of the other proposals that are out there. By considering some standard objections to it, we will develop an appreciation for a variety of hypotheses that are currently under debate.

Following up on Strawson's remarks, let us look at the following Presuppositionality Hypothesis, versions of which have been argued for by James McCawley, and more recently by Molly Diesing.³⁸

(5) Presuppositionality Hypothesis

In natural languages, all lexical items with denotations of type $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t \rangle \rangle$ are presuppositional, in the sense of the following mathematical definition (where δ is a lexical item of the appropriate semantic type, such as a determiner):

δ is presuppositional iff for all $A \subseteq D, B \subseteq D : \text{if } A = O, \text{ then } \langle A, B \rangle \notin \text{dom}(F).$

According to this hypothesis, determiner denotations like those defined in (1) are not possible in natural languages at all, and a fortiori cannot be the denotations of English every, all, some, at least one, a, no on any of their readings. The closest allowable denotations are the minimally different presuppositional ones defined in (4). Similarly, the denotations of most other determiners that we have considered in this chapter must be slightly different from what we have assumed. The following would be some revised lexical entries that conform to the Presuppositionality Hypothesis:

- (6) (a) $F_{\text{few}} = \lambda \langle A, B \rangle : A \neq O . |A \cap B| \text{ is small.}$
- (b) $F_{\text{most}} = \lambda \langle A, B \rangle : A \neq O . |A \cap B| > \frac{1}{2}|A|.$
- (c) $F_{\text{at least three}} = \lambda \langle A, B \rangle : A \neq O . |A \cap B| \geq 3.$
- (d) $F_{\text{at most three}} = \lambda \langle A, B \rangle : A \neq O . |A \cap B| \leq 3.$

How does this hypothesis fare with respect to the linguistic facts?

There are some observations that appear to support it, as even its harshest critics concede. Lappin and Reinhart,³⁹ for instance, report that their informants judge 7(a) below to be a presupposition failure rather than a true statement. (The informants were all aware, of course, that America has never had kings, and made their judgment on the basis of this piece of factual knowledge.) McCawley⁴⁰ already reported similar intuitions about 7(b):

- (7) (a) All/every American king(s) lived in New York.
- (b) All unicorns have accounts at the Chase Manhattan Bank.

These judgments are predicted by the lexical entry for every in (4) (and a parallel entry for all), whereas the one in (1) would predict judgments of "true". So we have observations here which, ceteris paribus, favor (4) over (1), and thus support the Presuppositionality Hypothesis. But for other examples, the predictions seem not to be borne out. Here are some more data from Reinhart:⁴¹

- (8) (a) No American king lived in New York.
 (b) Two American kings lived in New York.

Regarding (8a) and (8b), only about half of Reinhart's informants judged them presupposition failures on a par with (7). The other half judged (8a) true and (8b) false. So this looks as if some people employed the presuppositional entries in (4) and (6), whereas others employed the standard ones from (1) and previous sections. Reinhart also contrasts (7) with (9).

- (9) (a) Every unicorn has exactly one horn.
 (b) Every unicorn is a unicorn.

Her informants judged (9a) and (9b) true without hesitation, as if in this case, unlike with (7), they employed the standard nonpresuppositional entry of every.

So the evidence seems to be mixed. Can we detect some systematic pattern? Lappin and Reinhart (following earlier authors, especially Barwise and Cooper and de Jong and Verkuyl⁴²) endorse two descriptive generalizations. The relevant difference between (7) and (9), they maintain, is that (9) need not be taken as a description of the actual world, whereas (7) cannot naturally be taken any other way. They formulate this first generalization roughly as follows:

- (10) In non-contingent contexts, speakers' judgments about presupposition failure and truth-value conform to the standard (nonpresuppositional) analyses of determiners.

The notion of a "non-contingent context", of course, cries out for further clarification, and we will return to this shortly.

If (10) succeeds in distinguishing (7) from (9), the difference between (7) and (8) must lie elsewhere, both presumably being understood as "contingent" statements in the relevant sense. Apparently, what is decisive here is the choice of determiner. (7) and (8) are completely alike up to their determiners, and by testing a larger sample of additional determiners in the same kind of context, it emerges that the dividing line is the same as that which determines the grammaticality of there be sentences. Determiners disallowed in the there construction (every, almost, every, not every, most, and, unsurprisingly, uncontroversially presuppositional ones like both and neither) give rise to presupposition failure judgments in examples like (7). The ones that are alright in there sentences (no, numerals, **few**, **many**), when placed in the same sentence frame, evoke the mixed reactions found with (8). If (following Milsark) we define a "strong" determiner as one that is barred from there sentences, and a "weak" determiner as one that's allowed there, we can state Lappin and Reinhart's second descriptive generalization as follows:

- (11) In contingent contexts, strong determiners evoke judgments that conform to the presuppositional analysis, whereas weak determiners give rise to mixed judgments that conform sometimes to the presuppositional and sometimes to the standard analysis.

We will assume that both of these empirical generalizations are at least roughly on the right track. Either one of them seems to threaten the Presuppositionality Hypothesis by outlining a set of *prima facie* counterexamples. Suppose we nevertheless wanted to defend this hypothesis. Let us take a closer look, first at (10) and then at (11), to see whether such a defense is possible and what it would commit us to.

6.8.3 Nonextensional interpretation

The fact that (9a) and (9b) are spontaneously judged true is unexpected if **every** carries a Strawsonian existence presupposition. Similar facts were acknowledged by previous proponents of a presuppositional semantics for **every**: in particular, Diesing, de Jong and Verkuyl, and Strawson himself. Let's look at some of their examples and how they responded to the challenge. We quote from de Jong and Verkuyl:⁴³

... we claim that the standard interpretation of universal quantification is not based upon the most regular use of *all* in natural language, but rather upon the marked use of this expression: its conditional use. Due to its conditional structure,

- (12) $\forall x (Px \rightarrow Qx)$

expresses a specific relation between P and Q. Such an interpretation of

- (13) All ravens are black

is favored by the fact that the set of all ravens is a subset of the set of black entities, *which is not based on observation, but on induction or hypothesis. Blackness is taken as a property inherent to ravens*, as long as no counterexample shows up. Examples such as (13) must be treated as marked cases in comparison with *contingent sentences* such as (14) and (15).

- (14) All seats are taken.

- (15) All men are ill.

We use the term "marked" here in the linguistic sense. Sentence (13) is a clear example of a statement having *the status of a law* – or a hypothesis,

an opinion or a belief – that is firmly settled in science, in biological theory and also in our everyday naive physics. . . . In general, sentences like (14) and (15) are fully nonsensical if there are no men or seats in the context of use. This seems to be due to the fact that there is *no inherent relation* between seats and the property "to be taken", or between men and the property "being ill". As a consequence (12) cannot serve as a basis for this interpretation of (14) and (15). However, suppose that (ultrafeminist) science discovers that (15) is a law of nature. In that case, the interpretation of (15) is on a par with (13). So it depends on whether a certain sentence functions as a lawlike statement in a theory (or a more or less consistent set of everyday assumptions), when the conditional use gets the upper hand. . . . We regard the lawlike use of sentences as marked because we do not think it is a property of natural language that there are theories, whether scientific or embedded in our everyday opinions. Summarizing, all can be used in lawlike sentences as well as in contingent statements. Both contexts impose different interpretations on all. Only in hypothetical contexts can all be interpreted without presuppositions on the size of $\llbracket N \rrbracket$. In contingent statements the use of all requires a non-empty noun denotation.

De Jong and Verkuyl and Lappin and Reinhart, though they come down on opposite sides about the two competing lexical entries for all (every), agree essentially on the terms of the debate. Not only do they give similar descriptions of the relevant intuitions, they also assume that, first, the judgments about (9a), (9b), and (13) are incompatible with a (unambiguously) presuppositional analysis of **every/all**, and that, second, they support the standard nonpresuppositional analysis. But are these two assumptions so evidently correct? Strawson, it turns out, argued long ago that the second was quite mistaken: "

... There are, in fact, many differences among general sentences. Some of these differences have been exploited in support of the claim that there are at least some general sentences to which the negatively existential analysis (" $(x)(fx \supset gx)$ ") is applicable.⁴⁵ For example, it may be said that every one of the following sentences viz.,

- (16) All twenty-side rectilinear plane figures have the sum of their angles equal to 2×18 right angles
- (17) All trespassers on this land will be prosecuted
- (18) All moving bodies not acted upon by external forces continue in a state of uniform motion in a straight line

might be truthfully uttered; but in no case is it a necessary condition of their truthful utterance that their subject-class should have members. Nor can it be said that the question of whether or not they are truthfully uttered is one that arises only if their subject-class has members. . . . These facts, however, are very inadequate to support the proposed analysis. If the proposed analysis were correct, it would be a *sufficient condition* of the truthful utterance of these sentences that their subject-classes had no members; for " $\sim(\exists x)(fx)$ " entails " $(x)(fx \supset gx)$ ". But this is very far from being the case for these, or for any other, general sentences.

Let us consider this important point more carefully. If Strawson is right, it was no more than a coincidence that the "true"-judgments which Reinhart found with (9a) and (9b) conformed to the standard analysis of every. If we vary the predicate, we will find just as many "false"-judgments:

- (9) (c) Every unicorn has exactly two horns.
- (d) Every unicorn fails to be a unicorn.⁴⁶

Similarly, if we make suitable alterations in the predicates of Strawson's examples, their intuitive truth-values go from true to false:

- (18') All moving bodies not acted upon by external forces decelerate at a rate of 2.5 m/sec².

The same knowledge of physics that makes us assent to (18), regardless of whether we believe that there actually exist any bodies not acted upon by external forces, will make us dissent from (18').

The same point can be made by minimally varying the quantifier, say from all/every to no:

- (9) (e) No unicorn has exactly one horn.
- (18'') No moving body not acted upon by external forces continues in a state of uniform motion in a straight line.

If the "true"-judgments on (9a) and (18) are claimed as evidence for the standard analysis of **every/all**, shouldn't the "false"-judgments on (9e) and (18'') be counted as evidence *against* the standard analysis of no?

Summarizing Strawson's point, the judgments that speakers have about truth and falsity of lawlike quantificational statements do not support the standard analysis of quantifiers any more than the presuppositional one. Both analyses

seem to make systematically wrong predictions in this domain, albeit different ones. The presuppositional analysis errs by predicting that emptiness of the restrictor's extension suffices to render all such statements truth-value-less. The standard analysis errs by predicting that emptiness of the restrictor's extension suffices to verify all of them. It looks like both analyses miss the point of what really determines the intuitive truth-values of these statements.

We might leave it at this and simply set the semantics of lawlike quantificational sentences aside. What data we have looked at concerning their truth-conditions turned out to be simply irrelevant to the question we set out to answer: namely, whether natural language exhibits a universal constraint against non-presuppositional determiners. But this is not entirely satisfactory. If none of the semantic treatments of quantifiers that we are currently entertaining is applicable to lawlike statements, then shouldn't we abandon them all, instead of wasting our time comparing them to each other?⁴⁷

Fortunately, we can do better. Diesing maintains her version of the Presuppositionality Hypothesis in spite of the apparently conflicting data in connection with lawlike statements that she is well aware of. Here is a possible story that is in the spirit of Diesing's reaction to the challenge.⁴⁸

So far, we have only seen negative characterizations of what people do when they decide that, say, (9a) is true and (9c) is false. They don't, we saw, treat the fact that there are no actual unicorns as evidence one way or the other. What *do* they treat as evidence, then? It's not really so hard to give at least the rough outlines of a positive answer. There is a certain body of mythology that we have acquired together with the word unicorn. This mythology specifies a set of possible worlds in which there exist unicorns and in which these unicorns have certain properties and not others. All unicorns in the worlds that instantiate the myth have one horn, for instance, and none of them in any of these worlds have two horns. This, it seems, is the intuitive reason why (9a) is true and (9c) is false. If we consider not the set of *actual* unicorns (which is empty), but rather the set of *mythologically possible* unicorns (which consists of all the unicorns in the possible worlds which are truly described by the relevant myth), then it turns out that this set is (i) nonempty, (ii) a subset of the set of possible individuals with exactly one horn, and (iii) disjoint from the set of possible individuals with two horns.

This account suggests that the interpretation of (9a) and (9c) does, after all, involve a run-of-the-mill interpretation of every: namely, the one in (1) or (4). Either makes the correct prediction that (9a) is true and (9c) false, once we assume that the quantifier does not quantify over actual unicorns, but rather over mythologically possible unicorns. We do not have to commit ourselves to a particular technical realization of this idea here. Diesing assumes that lawlike statements are implicitly modalized. Quantificational DPs in lawlike statements, then, would be under the scope of a modal operator, and this is why quantification

is over possible individuals. Suppose some such analysis is on the right track. It appears, then, that what's special about the class of statements that Lappin and Reinhart call "non-contingent" and de Jong and Verkuyl "lawlike" is not a special interpretation of the *quantifier* at all. Rather, it is a special property of the environment the quantifier finds itself in.

Analogous stories can be told about de Jong and Verkuyl's biological law (13) or Strawson's sentences (16)–(18). The point about (18), we might say, is that it quantifies over all *physically possible* bodies not acted on by external forces, not just the actual ones. And (17) quantifies over all possible trespassers in the worlds that conform to the intentions of the land-owner at least as well as any world with trespassers in it can conform to them. In every one of these examples, we make our truth-value judgments by considering such sets of (wholly or partially) non-actual individuals.

At this point, it looks as if the truth-conditions of lawlike statements such as (9a)–(9e), (13), (16)–(18), and (18'), (18'') are, after all, consistent with both the presuppositional or the standard analysis of **every/all** and **no**. The existence presuppositions of the former are automatically fulfilled once the restrictor is interpreted in a suitably large domain of possible individuals, and then both analyses make the same, correct, predictions. It seems, then, that the data about lawlike quantificational statements that we have considered so far have no bearing one way or the other on the decision between presuppositional and standard lexical entries for the quantifiers. But this time around, the conclusion does not come packaged with reasons to suspect that both are wrong. Rather, both remain in the running, and we are ready to continue our search for decisive evidence.

Diesing reports an argument due to Kratzer⁴⁹ that intends to show that the presuppositional analysis is not just compatible with, but even necessary for, an adequate treatment of quantifying determiners in lawlike and other modalized statements. The bottom line of the argument is that unless we adopt a presuppositional analysis, quantifying determiners within the scope of modal operators might give rise to the Samaritan Paradox, a paradox that is well known to scholars of modal logic and conditionals.⁵⁰ Although we cannot go into the technical details of the argument here, we think an informal sketch may be useful, even in the absence of the necessary background from modal logic.⁵¹

The Samaritan Paradox comes up with sentences like (19a)–(19c):

- (19) (a) The town regulations require that there be no trespassing.
- (b) The town regulations require that all trespassers be fined.
- (c) The town regulations require that no trespassers be fined.

Suppose we live in a world in which (19a) and (19b) are true, but (19c) is false. Intuitively, there is nothing wrong with such an assumption. Yet any theory

that combines the standard modal analysis for verbs like "require" with the nonpresuppositional analysis of "all trespassers" and "no trespassers" would predict otherwise. (19a) says that there is no trespassing in any possible worlds that are compatible with the actual town regulations (these are the possible worlds in which no actual town regulation is violated). If "all trespassers" and "no trespassers" are nonpresuppositional, (19b) and (19c) would both be true. (19b) is true just in case all trespassers are fined in all possible worlds that conform to the actual town regulations. Since there is no trespassing in any of those worlds, there are no trespassers, and (19b) comes out true. And so does (19c). Consequently, we can't draw the desired distinction between (19b) and (19c).

The way we seem to understand (19b) and (19c) is that we temporarily suspend the regulation that there be no trespassing. But how come we suspend this regulation? Because we are temporarily assuming that there are trespassers. Where does this assumption come from? If "all trespassers" and "no trespassers" are presuppositional, we have an answer. The presupposition that there are trespassers could play a systematic role in picking out the set of possible worlds we are considering. For this to be an acceptable answer, however, we have to give a good account of the difference between the nonpresuppositional "no trespassing" in (19a) and the presuppositional "no trespassers" in (19c). The next section will look into this issue.

6.8.4 Nonpresuppositional behavior in weak determiners

We now turn to the challenge that the Presuppositionality Hypothesis faces from the behavior of weak determiners. Recall Reinhart and Lappin's generalization (11):

- (11) In contingent contexts, strong determiners evoke judgments that conform to the presuppositional analysis, whereas weak determiners give rise to mixed judgments that conform sometimes to the presuppositional and sometimes to the standard analysis.

The salience of the standard (nonpresuppositional)interpretation is known to be affected by pragmatic and grammatical factors, and in some examples it is entirely natural.⁵² Consider these:

- (20) (a) No phonologists with psychology degrees applied for our job.
 (b) Two UFOs landed in my backyard yesterday.
 (c) At most, twenty local calls from this number were recorded.

These sentences might very well be used in conversations where speaker and hearer are fully conscious of the possibility that there may be nothing that satisfies the restrictor. For instance, if I believe that there just aren't any phonologists with psychology degrees and am trying to convince you of this, I might use (20a) to cite circumstantial evidence for my conviction. If it later turns out that indeed there are no phonologists with psychology degrees, I will not feel any pressure to rephrase my statement. Quite to the contrary, I may then reiterate: "That's what I thought: there aren't any phonologists with psychology degrees. No wonder that none applied."

The situation with (20b) is a bit different. Unlike the previous sentence, this one cannot be used by a sincere speaker who believes that the restrictor is empty. After all, if there are no UFOs, then (20b) cannot be true. But it seems quite clear intuitively that it can be *false* in this case. Imagine that you and I have an ongoing disagreement about the existence of UFOs: I believe they exist, you do not. (20b) could be uttered in the course of one of our arguments about this matter: I might use it as evidence for my position. If you want to defend yours then, you will argue that (20b) is false. So in this case as well, the non-emptiness of the restrictor seems not to be a presupposition of (20b).

The case of (20c) is more similar to that of (20a), in that the non-existence of any local calls from this number would make it true (rather than false). Imagine I just got the phone bill and there is no extra charge for local calls. According to my contract with the phone company, the first twenty calls every month are covered by the basic rate. If I utter (20c) in this situation, I will not be taken as prejudging the question of whether there were any local calls from this number at all.

In every one of these cases, the standard, nonpresuppositional entries for the determiners seems to fit our intuitions much better than their presuppositional alternatives, and the Presuppositionality Hypothesis might accordingly look undesirable. This conclusion is further reinforced when we broaden the scope of our examples to include quantifiers in positions other than subject position.⁵³ An especially natural environment for nonpresuppositional interpretations is the post-copular position of **there** sentences. Reinhart,⁵⁴ for instance, reports that even those informants who perceived (8a) and (8b) as presupposition failures had no hesitation in judging (21a) and (21b) as respectively true and false.

- (8) (a) No American king lived in New York.
(b) Two American kings lived in New York.

- (21) (a) There were no American kings in New York.
(b) There were two American kings in New York.

In this particular construction, strong determiners are not grammatical in the first place. But in non-subject positions where both weak and strong determiners can occur, we see quite clearly that there is indeed a minimal contrast between weak and strong determiners. Strong determiners receive presuppositional readings regardless of position, whereas weak ones need not. Zucchi cites the following paradigm from Lumsden:⁵⁵

- (22) If you find every mistake(s), I'll give you a fine reward.
 most
 many
 a
 no
 three
 . . .

The examples with the strong determiners (every, most) convey that the speaker assumes there to be mistakes, whereas the ones with weak determiners (many, a, no, less than 2) sound neutral in this regard.

What do we conclude from this (cursory) survey of data? An obvious possibility is to maintain the Presuppositionality Hypothesis in its radical form and explore the possibility that weak determiners might be affected by a type ambiguity, as considered by Partee.⁵⁶ In chapter 4, we looked at predicative uses of indefinites like a cat in **Julius is** a cat, and concluded that when so used, indefinites are of semantic type $\langle e, t \rangle$. Maybe all nonpresuppositional uses of weak DPs can be assimilated to the predicative use. This is the line taken by Diesing,⁵⁷ who builds on insights from Discourse Representation Theory.⁵⁸ Weak DPs may or may not be interpreted as generalized quantifiers. Only if they are, are they presuppositional. If weak DPs are ambiguous, we now need to say something about the observed distribution of possible readings. Diesing invokes a special principle, her Mapping Hypothesis, to this effect.

Much current research explores ways of doing away with principles like the Mapping Principle. De Hoop⁵⁹ follows Diesing in assuming a type ambiguity for weak DPs, but links the different semantic types to different cases (in the syntactic sense). She then proposes to derive the distribution of the two types of weak DPs from syntactic principles governing the two types of cases, plus independent principles governing the interpretation of topic and focus. Other recent attempts in the literature reject the idea that weak DPs are ambiguous as to their semantic type, and argue that topicality and focus all by themselves are able to account for the presuppositionality facts.⁶⁰ As things stand, there is no consensus yet.

Notes

- 1 Remember that we are assuming (following Abney) that phrases like "the mouse" are DPs (determiner phrases) headed by a determiner whose sister node is an NP. Since proper names, pronouns, and traces behave syntactically like phrases headed by an overt determiner, they are classified as DPs as well. Our DPs correspond to Montague's "term phrases". Our NPs correspond to Montague's "common noun phrases".
- 2 In this chapter, we will indulge in a lot of set talk that you should understand as a sloppy substitute for the function talk that we would need to use to be fully accurate.
- 3 Russell wouldn't disagree here, we presume, even as he criticizes Meinong for his "failure of that feeling for reality which ought to be preserved even in the most abstract studies," and continues: "Logic, I should maintain, must no more admit a unicorn than zoology can" ("Descriptions," in A. P. Martinich (ed.), *The Philosophy of Language*, 2nd edn (New York and Oxford, Oxford University Press, 1990), pp. 212–18, at p. 213). Such pronouncements can sound dogmatic out of context, but they don't, of course, carry the burden of Russell's argument.
- 4 Once again, we pretend that VPs denote sets when they really denote the corresponding characteristic functions.
- 5 The Law of Contradiction states the validity of "not (p and not p)".
- 6 The Law of Excluded Middle states the validity of "p or not p".
- 7 Noam Chomsky, "On Wh-Movement," in P. Culicover (ed.), *Formal Syntax* (New York, Academic Press, 1977), pp. 71–132.
- 8 This statement would have to be qualified if we considered DPs that contain variables that are free in them. In that case, movement may affect variable binding relations, and thus affect truth-conditions, even if the moving phrase is of type e. We may safely disregard this qualification here, since it seems obvious that the DPs under consideration ("at least one question", "more than one policeman", etc.) have no free variables in them.
- 9 Days are, of course, included among the elements of D. They are more abstract entities than chairs and policemen, but objects nonetheless.
- 10 P. T. Geach, *Logic Matters* (Berkeley and Los Angeles, University of California Press, 1972), pp. 56ff.
- 11 Recall that $\text{Pow}(D)$, the power set of D, is defined as $\{X : X \subseteq D\}$.
- 12 Geach, *Logic Matters*, pp. 58f.
- 13 The modern notion of a generalized quantifier was introduced by A. Mostowski, "On a Generalization of Quantifiers," *Fundamenta Mathematica*, 44 (1957), pp. 12–36. The standard reference for linguists is J. Barwise and R. Cooper, "Generalized Quantifiers and Natural Language," *Linguistics and Philosophy*, 4 (1981), pp. 159–219.
- 14 D. Lewis, "General Semantics," in D. Davidson and G. Harman (eds), *Semantics of Natural Languages* (Dordrecht, Reidel, 1972), pp. 169–218. Lewis has intensional denotations instead of our extensional ones, but this has little impact on the treatment of quantifier phrases. Lewis's paper was first presented at the third La Jolla Conference on Linguistic Theory in 1969.
- 15 R. Montague, *Formal Philosophy*, ed. R. M. Thomason (New Haven, Yale University Press, 1974); M. J. Cresswell, *Logics and Languages* (London, Methuen, 1973); Barwise and Cooper, "Generalized Quantifiers"; E. L. Keenan and J. Stavi,

- "A Semantic Characterization of Natural Language Determiners," *Linguistics and Philosophy*, 9 (1986), pp. 253–326.
- 16 D. Westerståhl, "Quantifiers in Formal and Natural Languages," in D. Gabbay and F. Guenther (eds), *Handbook of Philosophical Logic*, vol. 4: *Topics in the Philosophy of Language* (Dordrecht, Reidel, 1989), pp. 6f., our emphasis.
- 17 W. Kneale and M. Kneale, *The Development of Logic* (Oxford, Clarendon Press, 1962), p. 349, our emphasis.
- 18 This quote is from G. Frege, "Über Begriff und Gegenstand" ["On concept and object"], *Vierteljahrsschrift für wissenschaftliche Philosophie*, 16 (1892); our emphasis. See also *Grundgesetze der Arithmetik*.
- 19 For any set A, $|A|$ is the cardinality of A, i.e., the number of members of A.
- 20 For any set A, $\text{Pow}(A)$ (the power set of A) is the set of all subsets of A.
- 21 For any sets A and B, $A \times B$ (the Cartesian Product of A and B) is defined as $\{<x, y> : x \in A \text{ and } y \in B\}$.
- 22 These definitions don't directly mention the quantifier denotation $[\delta]$, but only the corresponding relation R_δ . Given the previous section, however, it is a routine exercise to translate every *definiens* into equivalent conditions on F_δ , f_δ , and $[\delta]$.
- 23 Note that "irreflexive" doesn't mean "nonreflexive".
- 24 For this and the following three monotonicity properties, a number of other terms are also in common use. "Upward monotone" is also called "upward-entailing" or "monotone increasing", and "downward monotone" is called "downward-entailing" or "monotone decreasing". "Left upward monotone" is called "persistent", and "left-downward monotone", "anti-persistent". Authors who use these terms "persistent" and "anti-persistent" tend to use "monotone increasing" or "upward monotone" in the narrower sense of "right upward monotone" (and similarly for "monotone decreasing" and "downward monotone").
- 25 Useful overview articles are Westerståhl, "Quantifiers"; J. van Eijck, "Quantification," in A. von Stechow and D. Wunderlich (eds), *Semantik/Semantics. An International Handbook of Contemporary Research* (Berlin and New York, de Gruyter, 1991), pp. 459–87; E. L. Keenan, "The Semantics of Determiners," in S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory* (Oxford, Basil Blackwell, 1996), pp. 41–63. E. L. Keenan and D. Westerståhl, "Generalized Quantifiers in Linguistics and Logic," in J. van Benthem and A. ter Meulen (eds), *Handbook of Language and Logic* (Amsterdam, Elsevier, forthcoming).
- 26 Keenan and Stavi, "Semantic Characterization." Conservativity is Barwise and Cooper's live-on property. Barwise and Cooper, "Generalized Quantifiers" conjecture that every natural language has conservative determiners.
- 27 G. Milsark, "Existential Sentences in English" (Ph.D. dissertation, MIT, 1974); G. Fauconnier, "Polarity and the Scale Principle," in R. E. Grossman, L. J. San, and T. J. Vance (eds), *Proceedings of the 11th Meeting of the Chicago Linguistic Society* (University of Chicago, 1975), pp. 188–99; W. A. Ladusaw, "Polarity Sensitivity and Inherent Scope Relations" (Ph.D. dissertation, University of Texas, 1979).
- 28 If the restrictor happens to lack a semantic value, as in **every** woman on the escalator in **South** College, then of course the whole DP doesn't get one either. This is evidently not due to any lexical property of the *determiner* (here **every**).
- 29 We are not talking here about "both" and "neither" when they occur as part of the discontinuous coordinators "both ... and ...", "neither ... nor ..." (as in "Both John and Bill left", "Neither John nor Bill left"). These are not determiners and do not interest us here.

- 30 Strictly speaking, the lexical entry has to define **[neither]**, of course. But you know how to get **[neither]** from f_{neither} .
- 31 Recall that char_i is the characteristic function of R_i .
- 32 Regarding the four monotonicity concepts, there is less of a consensus about how to extend them to partial determiners. For instance, an alternative to the present definition of "left upward monotone" would be the following:
- δ left upward monotone for all $A, B, C : \text{if } A \subseteq B \text{ and } F_\delta(A, C) = 1, \text{ then } \langle B, C \rangle \in \text{dom}(F_\delta) \text{ and } F_\delta(B, C) = 1$
- Analogous alternative definitions may be entertained for the other three monotonicity properties. As you may discover in the exercise below, it is not completely evident which choice is preferable.
- 33 The classical reference for linguists is G. Link, "The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretic Approach," in R. Bauerle et al. (eds), *Meaning, Use, and Interpretation* (Berlin, Walter de Gruyter, 1983), pp. 302–23. Useful handbook articles are G. Link, "Plural," in von Stechow and Wunderlich (eds), *Semantik*, pp. 418–40, and F. Landman, "Plurality," in S. Lappin (ed.), *Handbook*, pp. 425–57. See also R. Schwarzschild, *Pluralities* (Dordrecht, Kluwer Academic Publishers, 1996), and P. Lasersohn, *Plurality, Conjunction and Events* (Dordrecht, Kluwer Academic Publishers, 1995).
- 34 Our primary source here is P. F. Strawson, *Introduction to Logical Theory* (London, Methuen, 1952), who in turn refers to Miller, *The Structure of Aristotelian Logic*. According to Horn, however, Aristotle's own writings are not explicit enough to endorse all the "laws of the traditional system" as listed by Strawson (*Introduction*, pp. 156–63), and Aristotle's commentators have fleshed out the system in partially disagreeing directions. See L. Horn, *A Natural History of Negation* (Chicago, University of Chicago Press, 1989), esp. ch. 1, sect. 1.1.3: "Existential Import and the Square of Opposition," pp. 23–30. We do not want to get into questions of exegesis and history here. Strawson's proposal is interesting to us in its own right, even if (as Horn argues) it contradicts explicit statements by Aristotle and the majority of his medieval followers.
- 35 Don't be confused by the fact that "the classical analysis" refers to the modern one, not to any of those that date from antiquity. This is the customary terminology in the contemporary literature. The same perspective is reflected in the labels "standard" or "orthodox", which also mean the modern analysis.
- 36 Strawson, *Introduction*, pp. 173–7.
- 37 By "subject-class", Strawson means the set denoted by the restrictor (i.e., a in the schemata in (2)).
- 38 J. D. McCawley, "A Program for Logic," in Davidson and Harman (eds), *Semantics*, pp. 498–544; M. Diesing, "The Syntactic Roots of Semantic Partition" (Ph.D. dissertation, University of Massachusetts, Amherst, 1990); idem, *Indefinites* (Cambridge, Mass., MIT Press, 1992). McCawley associates all quantifying determiners other than "any" with existential presuppositions. Diesing ("Syntactic Roots" and *Indefinites*) connects the presuppositionality of quantifying DPs to the presence of a syntactically represented restrictive clause and the ability to undergo the syntactic operation of Quantifier Raising (see our chapters 7 and 8 for a discussion of quantifier raising). In M. Diesing and E. Jelinek, "Distributing Arguments," *Natural Language Semantics*, 312 (1996), pp. 123–76, the connection with type theory and generalized quantifiers is made explicit. Combining the two works by Diesing with

that by Diesing and Jelinek, we conclude, then, that we are justified in ascribing to Diesing the view that all DPs that express generalized quantifiers are presuppositional in the sense under discussion here. Be this as it may, the arguments for or against the presuppositionality hypothesis that we are about to discuss apply directly to Diesing's proposals concerning the presuppositionality of quantifying DPs, even under the syntactic characterization of Diesing 1990 and 1992.

- 39 S. Lappin and T. Reinhart, "Presuppositional Effects of Strong Determiners: A Processing Account," *Linguistics*, 26 (1988), pp. 1022–37; T. Reinhart, *Interface Strategies*, OTS Working Paper TL-95-002 (Utrecht University, 1995), ch. 4: "Topics and the Conceptual Interface."
- 40 McCawley, "Program for Logic."
- 41 Reinhart, *Interface Strategies*.
- 42 Barwise and Cooper, "Generalized Quantifiers"; F. de Jong and H. Verkuyl, "Generalized Quantifiers: The Properness of their Strength," in J. van Benthem and A. ter Meulen, *Generalized Quantifiers in Natural Language* (Dordrecht, Foris, 1985), pp. 21–43.
- 43 De Jong and Verkuyl, "Generalized Quantifiers," pp. 29f., emphasis added and examples renumbered.
- 44 Strawson, *Introduction*, pp. 195ff. (emphasis in original, examples renumbered). Reinhart, *Interface Strategies*, cites the examples from this passage.
- 45 Strawson's label "negatively existential" for the standard analysis of **every** derives from the fact that " $\forall x (fx \rightarrow gx)$ " is equivalent to " $\neg\exists x (fx \& \neg gx)$ ". Notice also that Strawson uses the so-called Principia *Mathematica* notation for predicate logic, where "(x)" is the symbol for " $\forall x$ ", and " \supset " for " \rightarrow ".
- 46 Lappin and Reinhart, "Presuppositional Effects," and Reinhart, *Interface Strategies*, actually cite (i):

- (i) Every unicorn is not a unicorn.

They report that their informants judge (i) false and imply that this judgment conforms to the predictions of the standard analysis. But this sentence is predicted false by the standard analysis only if the negation takes scope over the subject quantifier. On the other scope-order, it is predicted true. In (i), the latter scope-order may be excluded or dispreferred for independent reasons (see A. S. Kroch, "The Semantics of Scope in English" (Ph.D. dissertation, MIT, 1974). Our (9d) is constructed so as to avoid the potential scope ambiguity.

- 47 We might avoid this conclusion if we agreed with de Jong and Verkuyl that the lawlike use of quantified sentences is somehow "marked", i.e., beyond the proper domain of linguistic theory. But this does not seem right to us (and here we agree with Lappin and Reinhart). It is true that it is not "a property of natural language that there are theories," but – more to the point – it does appear to be a property of natural language that theories can be expressed in it.
- 48 Diesing, *Indefinites*, pp. 95ff.
- 49 The argument was developed by Kratzer in class lectures. See Diesing, *Indefinites*, p. 96.
- 50 For an overview and further references consult L. Aqvist, "Deontic Logic," in Gabbay and Guenther, *Handbook*, pp. 605–714. See also A. Kratzer, "Modality," in von Stechow and Wunderlich, *Semantik*, pp. 639–50.
- 51 See also our chapter 12, where an intensional semantics is introduced, and the semantics of attitude verbs is discussed.

- 52 The observations we are sketching here are found in a large number of works, many of which build on the work of Milsark. See Milsark, "Existential Sentences in English"; and idem, "Toward an Explanation of Certain Peculiarities of the Existential Construction in English," *Linguistic Analysis*, 311 (1977), pp. 1–29. (*Linguistic Analysis* misspells his name as "Milwark.")
- 53 The following discussion includes examples with quantifying DPs in object position that we cannot yet interpret in a compositional way. Quantifying DPs in object position are the topic of the next chapter. We are confident, however, that the point of the examples can still be appreciated.
- 54 Reinhart, *Interface Strategies*.
- 55 A. Zucchi, "The Ingredients of Definiteness and the Definiteness Effect," in *Natural Language Semantics*, 311 (1994), pp. 33–78. M. Lumsden, *Existential Sentences* (London, Croom Helm, 1988).
- 56 B. Partee, "Noun Phrase Interpretation and Type Shifting Principles," in J. Groenendijk, D. de Jongh, and M. Stokhof (eds), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers* (Dordrecht, Foris, 1987), pp. 115–43.
- 57 Diesing, "Syntactic Roots"; idem, *Indefinites*.
- 58 J. W. A. Kamp, "A Theory of Truth and Semantic Representation," in J. Groenendijk, T. Janssen, and M. Stokhof (eds), *Formal Methods in the Study of Language* (Amsterdam, Mathematical Centre, 1981), pp. 277–321; I. Heim, "The Semantics of Definite and Indefinite Noun Phrases" (Ph.D. dissertation, University of Massachusetts, Amherst, 1982).
- 59 H. de Hoop, "Case Configuration and Noun Phrase Interpretation" (Ph.D. dissertation, Rijksuniversiteit Groningen, 1992).
- 60 For attempts along those lines, see Reinhart, *Interface Strategies*, and D. Büring, "A Weak Theory of Strong Readings," in T. Galloway and J. Spence (eds), *Proceedings from SALT (Semantics and Linguistic Theory) VI* (Cornell University, Ithaca, NY, 1996), pp. 17–34.