4 More of English: Nonverbal Predicates, Modifiers, Definite Descriptions

We ultimately want to determine appropriate types of meanings for lexical items of all syntactic categories and to predict how these meanings are composed in all sorts of syntactic constructions. In this chapter, we take some further small steps towards this goal. We continue with the Fregean working hypothesis that there are two types of basic, or saturated, meanings – namely, individuals and truth-values – and that all other meanings are functions that are somehow constructed out of these.

We will run into important general issues like vagueness, context-dependency, and presupposition, but will not stop to consider them seriously and systematically. We will also make quite a few concrete decisions without defending them against alternatives. The primary purpose of this chapter is to build a reasonably broad base of simple constructions for which we have at least a preliminary treatment, so that we are not too constrained in our choice of examples in the subsequent chapters. A secondary purpose is to give an overview of some of the basic issues regarding modification and presupposition.

4.1 Semantically vacuous words

Some lexical items are widely held to make no semantic contribution to the structures in which they occur. The standard example is of certain occurrences of prepositions, such as "of" in "proud of John" or "father of John". Another plausible candidate is the copula "be" in predicative sentences like "John is rich". We will also assume, at least for the time being, that the indefinite article "a" is vacuous when it occurs in predicate nominals such as "a cat" in "Kaline is a cat".[1] We would want the following equalities, for example:

(1) $[\![\textbf{of John}]\!] = [\![\textbf{John}]\!]$
$[\![\textbf{be rich}]\!] = [\![\textbf{rich}]\!]$
$[\![\textbf{a cat}]\!] = [\![\textbf{cat}]\!]$

There are various ways of making this come out. One way is to list semantically vacuous items in the lexicon as denoting the identity function of the appropriate type, for instance:

(2) $[\![\textbf{of}]\!] = \lambda x \in D_e \,.\, x$

(3) $[\![\textbf{be}]\!] = \lambda f \in D_{<e,t>} \,.\, f$

(4) $[\![\textbf{a}]\!] = \lambda f \in D_{<e,t>} \,.\, f$

In words: $[\![\textbf{of}]\!]$ is that function which maps every individual in D_e to itself, and $[\![\textbf{be}]\!]$ (= $[\![\textbf{a}]\!]$) is that function which maps every function in $D_{<e,t>}$ to itself.

An even easier possibility is to assume that the semantic component simply "doesn't see" such items. In other words, a structure that is really binary-branching may be treated as non-branching in the semantics: a branch occupied only by a vacuous item doesn't count. The principle for nonbranching nodes then applies and passes up the meaning unchanged. Either way, we ensure the equalities in (1).

4.2 Nonverbal predicates

What we have assumed for verbs can be extended straightforwardly to adjectives, nouns, and prepositions. Just as intransitive verbs denote functions from individuals to truth-values, so do many nouns and adjectives, for instance:

(1) $[\![\textbf{cat}]\!] = \lambda x \in D_e \,.\, x$ is a cat

(2) $[\![\textbf{gray}]\!] = \lambda x \in D_e \,.\, x$ is gray

Among prepositions, intransitive (1-place, monadic) ones are the exception rather than the rule, but there are some candidates:

(3) $[\![\textbf{out}]\!] = \lambda x \in D_e \,.\, x$ is not in x's home

Each of these categories also has transitive (2-place, dyadic) members, whose extensions are just like those of transitive verbs – for example, "part" in "part of Europe" and "proud" in "proud of John" (note the vacuity of "of"), and all run-of-the-mill prepositions:

(4) $[\![\textbf{part}]\!] = \lambda x \in D_e \,.\, [\lambda y \in D_e \,.\, y \text{ is part of } x]$

(5) $[\![\textbf{fond}]\!] = \lambda x \in D_e \,.\, [\lambda y \in D_e \,.\, y \text{ is fond of } x]$

(6) $[\![\textbf{in}]\!] = \lambda x \in D_e \,.\, [\lambda y \in D_e \,.\, y \text{ is in } x]$

These lexical entries allow us to calculate appropriate denotations for the phrases "part of Europe", "fond of Joe", "in Texas", by means of our composition principle Functional Application (FA), for instance:

By FA: $[\![\textbf{in Texas}]\!] = [\![\textbf{in}]\!]([\![\textbf{Texas}]\!])$
By lexical entry for **Texas**: $[\![\textbf{Texas}]\!] =$ Texas
Hence: $[\![\textbf{in Texas}]\!] = [\![\textbf{in}]\!](\text{Texas})$
By lexical entry for **in**: $[\![\textbf{in}]\!](\text{Texas})$
$= [\lambda x \in D_e \,.\, [\lambda y \in D_e \,.\, y \text{ is in } x]](\text{Texas}) = \lambda y \in D_e \,.\, y \text{ is in Texas.}$
Hence: $[\![\textbf{in Texas}]\!] = \lambda y \in D_e \,.\, y \text{ is in Texas.}$

We will disregard the case of ditransitive (3-place, triadic) predicates, though there are presumably some analogs to verbs like "give" and "introduce" in other syntactic categories.

Exercise

Calculate the truth-conditions for at least one of the sentences "Joe is in Texas", "Joe is fond of Kaline", and "Kaline is a cat".

4.3 Predicates as restrictive modifiers

It has often been observed that prepositional phrases (PPs) may appear inside NPs in three distinct semantic roles: as arguments, as restrictive modifiers, or as nonrestrictive modifiers.[2] Typical examples of each type are the following:

(1) **a part *of Europe*** (argument)

(2) **a city *in Texas*** (restrictive modifier)

(3) **Susan, *from Nebraska*,** (nonrestrictive modifier)

We have already dealt with the first case: PPs that are arguments are headed by vacuous prepositions, thus they have the same denotations as the NPs they contain (individuals), and they are arguments of 2-place (relational) nouns such as "part".

About the third case, we have next to nothing to say, and we mentioned it only to guard against confusing nonrestrictive modifiers with restrictive ones. The basic intuition that most authors have expressed about the semantics of nonrestrictive modification[3] is that nonrestrictive modifiers are not semantically composed at all with the phrases they modify. Rather, they have the status of separate sentences which serve to make side-remarks of some kind. For example, the meaning of (4) is not unlike that of (5).

(4) **It is surprising that Susan, from Nebraska, finds it cold in here.**

(5) **It is surprising that Susan finds it cold in here. Note that she is from Nebraska.**

This makes it reasonable to assume that at the level at which our semantic rules apply, the nonrestrictive modifier isn't part of the structure at all, so the question of how its denotation should be composed with that of its modifier doesn't arise in the first place. This said, we will concentrate on restrictive modifiers, which are our primary topic in this section.

While the distinction between arguments and restrictive modifiers is notoriously difficult to make in practice, the basic semantic intuition behind it is simple: Arguments reduce the adicity of the noun they combine with; modifiers leave it unchanged.

What it means to reduce the adicity of the noun is illustrated by the example we have treated: "part" is a 2-place predicate, while the result of combining it with "of Europe" to form "part of Europe" is a 1-place predicate. Restrictive modifiers, by contrast, are characterized by the fact that they leave the semantic type, including the adicity, of the modifier completely unchanged. Intuitively, "city in Texas" has the same kind of extension as "city": namely, the characteristic function of a set. More specifically, if $[\![\textbf{city in Texas}]\!]$ is the characteristic function of a set A and $[\![\textbf{city}]\!]$ is the characteristic function of a set B, then A is a subset of B: namely, that subset which results by intersecting B with the set of things in Texas. Consider the truth-conditions of "Lockhart is a city in Texas" to confirm this intuition.

Now we have already proposed an analysis of "in" above (motivated at the time by occurrences of this preposition in simple copula + PP phrases), under which "in Texas" has a denotation of type <e,t>. Given this decision and our current inventory of composition rules, we predict "city in Texas" to be uninterpretable. As a branching structure with daughters "city" and "in Texas", we should be interpreting it by applying either ⟦**city**⟧ to ⟦**in Texas**⟧ or vice versa. But neither is possible, as both are functions of type <e,t>.

What shall we do about this problem? There are two directions we could take: either revise out lexical semantics for (some of) the ingredients, or else stipulate a new composition rule. We will entertain both options, beginning with the second.

4.3.1 A new composition rule

Here is a composition principle which is tailored to the situation at hand.

(6) *Predicate Modification* (PM)[4]
If α is a branching node, {β, γ} is the set of α's daughters, and ⟦β⟧ and ⟦γ⟧ are both in $D_{<e,t>}$, then
⟦α⟧ = λx ∈ D_e . ⟦β⟧(x) = ⟦γ⟧(x) = 1.[5]

Applied to "city in Texas", (6) gives the desired result:

(7) ⟦**city in Texas**⟧
= λx ∈ D_e . ⟦**city**⟧(x) = ⟦**in Texas**⟧(x) = 1
= λx ∈ D_e . x is a city and x is in Texas.

(In the last step, we used the lexical entries and the result of a previous calculation from section 4.2 above.)

PM is general enough to cover not just PPs modifying nouns, but likewise adjective phrases (APs), whether to the right or the left of a noun, and also stacked modifiers in unlimited numbers. For instance, we can now predict correctly the truth-conditions of (8).

(8) **Kaline is a gray cat in Texas fond of Joe.**

To treat this example, we must, of course, impose some binary-branching hierarchy among the three modifiers "gray", "in Texas", and "fond of Joe". Given the nature of the semantic operation performed by PM, it so happens that all our different choices in this regard yield logically equivalent results. This is as it should be. The syntax of English (to the best of our knowledge) does not determine a unique parse for this sentence, but it is nevertheless perceived as truth-conditionally unambiguous.

Exercise

Calculate the truth-conditions for (8), given one possible syntactic parse.

The operation performed by PM has also been called "intersective modification", because if we look at the sets instead of at their characteristic functions, it amounts to set-theoretic intersection. "Conjunctive composition" would be another natural name, highlighting the connection with the semantics of "and".[6] (Notice that "city in Texas" receives exactly the meaning we derive for "city *and* in Texas", using the predicate-coordinating "and" we defined in section 3.4.)

Predicate modification is a genuinely new semantic composition principle on our list. It is obviously not functional application. If it is really needed, there is more to semantic composition than Frege's Conjecture. Are we forced to this conclusion?

4.3.2 *Modification as functional application*

As we have already mentioned, the alternative to PM is to explore revised lexical entries for the words that may head modifiers. Suppose we insist that [[**city**]] and [[**in Texas**]] combine by Functional Application after all. If we keep [[**city**]] as before, with type <e,t>, then [[**in Texas**]] will have to be of type <<e,t>,<e,t>>. This in turn requires a new semantics for at least one of "in" and "Texas". If we keep the assumption that [[**Texas**]] = Texas $\in D_e$, we must reinterpret "in". It must now denote a function of type <e,<<e,t>,<e,t>>>. Which such function? Well, we know that we want to be able to derive equations like these:

(9) [[**in**]](Texas)($\lambda x \in D_e$. x is a city)
= $\lambda x \in D_e$. x is a city and x is in Texas.
[[**in**]](Building 20)($\lambda x \in D_e$. x is a room)
= $\lambda x \in D_e$. x is a room and x is in Building 20.

The generalization appears to be that, for any individual $y \in D_e$ and any function $f \in D_{<e,t>}$, [[**in**]](y)(f) = $\lambda x \in D_e$. f(x) = 1 and x is in y. This determines directly the desired definition for the function [[**in**]]:

(10) [[**in**]] = $\lambda y \in D_e$. [$\lambda f \in D_{<e,t>}$. [$\lambda x \in D_e$. f(x) = 1 and x is in y]]

By similar reasoning, we can determine an entry for the adjective "gray" that permits phrases like "gray cat" to be interpreted by Functional Application:

(11) [[**gray**]] = $\lambda f \in D_{<e,t>}$. [$\lambda x \in D_e$. f(x) = 1 and x is gray]

Exercise

Calculate the truth-conditions for (8), given one possible syntactic parse. This time, use FA instead of PM.

By systematically revising the entries of all adjectives and prepositions, we are able to interpret all phrases containing a noun with one or more modifiers in them by means of Functional Application alone, and so we can eliminate Predicate Modification from the theory.

But there is a trade-off. What happens now when an AP or PP stands as a (maximal) predicate by itself, as in "Julius is gray" or "Julius is in Amherst"? If APs and PPs denote functions of type <<e,t>,<e,t>>, these sentences are *prima facie* uninterpretable. We could try to solve the problem by assigning a suitable denotation to the copula "be".

Exercise

Define such a denotation. There are two distinct solutions.

But then this same "be" would not be interpretable in sentences with a nominal predicate, for example, "Julius is a cat", since NPs like "a cat" are still of type <e,t>. So the copula would have to be ambiguous between vacuous and nonvacuous occurrences.

Another solution is to assume a systematic lexical ambiguity in all adjectives and prepositions. Each has both the initially assumed type <e,t> (or <e,<e,t>>) meaning and the new type <<e,t>,<et>> (or <e,<<e,t>,<et>>>) denotation. The syntax may freely generate both homonyms in all the same places, but the Principle of Interpretability will allow only one in any given environment.

Exercise

It would not be adequate to list the two readings of each preposition or adjective separately in the lexicon, as if they had to be learned individually. Evidently, there is a systematic relation between the two readings, which makes one predicable, given the other. So we would want to list only one in the

lexicon, and derive the other by means of a general "lexical rule". Spell out two versions of this proposal. For the first version, assume that the homonyms with the simpler types (<e,t> or <e,<e,t>>) are basic and listed in the lexicon. For the second version, assume that the more complicated types <<e,t>,<et>> or <e,<<e,t>,<et>>> are the ones of the basic, individually listed items. Your task is to formulate the appropriate lexical rules for either version. That is, you have to specify general recipes that map arbitrary denotations of the basic type to secondary denotations of the appropriate nonbasic type. Rules of this kind are called "type-shifting rules".[7]

Yet another solution is to posit a certain amount of non-overt structure in VPs of the surface form "be" + AP or "be" + PP. Perhaps these VPs contain an invisible predicate that the AP or PP modifies, something like a zero equivalent of a bland noun like "thing" or "individual".

We will not explore or try to evaluate these options further. All we wanted to show here is that the elimination of the Predicate Modification rule is not without its price. As matters stand, it does not look entirely unreasonable if we decide to adopt PM after all.

4.3.3 *Evidence from nonintersective adjectives?*

Both analyses of adjectives that we have entertained so far predict that the following pair of sentences are logically equivalent:

(12) (a) **Julius is a gray cat.**
(b) **Julius is gray and Julius is a cat.**

In the analysis that uses PM, the equivalence follows directly from the content of the rule. We can prove it without using any specific information about the meanings of the lexical items, except the information about their semantic type. In the analysis that relies on Functional Application alone, the equivalence follows from the lexical meaning of the adjective "gray". On this alternative, mere inspection of the *types* of the words and the applicable composition principles does *not* suffice to prove it.

Since the equivalence in (12) is indeed intuitively valid, both analyses make the correct prediction, albeit in different ways. But it has often been noted[8] that analogous equivalences do not obtain for many other adjectives, and it seems that this fact might have some bearing on the choice we have been contemplating.

Consider adjectives like "large" and "small". One might truthfully assert that a small elephant is still a very large animal. So it is intuitively possible for (13) to be true while (14) is false.

(13) **Jumbo is a small elephant.**

(14) **Jumbo is a small animal.**

This shows that (13) does not intuitively entail (14). But "Jumbo is an elephant" does entail "Jumbo is an animal", and given the meaning of "and", this implies that (15) entails (16).

(15) **Jumbo is small and Jumbo is an elephant.**

(16) **Jumbo is small and Jumbo is an animal.**

So (13) and (14) cannot be equivalent to (15) and (16) respectively. At least one, perhaps both, of these equivalences must be denied, or else we falsely predict that (13) entails (14).

Now we have seen that if adjectives are of type <e,t> and combine with their modifiees by PM, the equivalences in question follow regardless of specific lexical meaning. We have to conclude, therefore, that it is not possible to define an adequate type <e,t> meaning for "small".

But if the meaning of "small" is of type <<e,t>,<e,t>>, it seems that we have a chance of defining it appropriately, so that the inference from (13) to (14) does not go through. We do *not* want a lexical entry analogous to (11) above, of course (that is, we don't want to define ⟦**small**⟧ as $\lambda f \in D_{<e,t>}$. [λx . f(x) = 1 and x is small]), but fortunately there are other functions of type <<e,t>,<e,t>> to consider. Here is a proposal that seems to reflect the relevant intuition about why small elephants are not small animals.

(17) ⟦**small**⟧ =
$\lambda f \in D_{<e,t>}$. [$\lambda x \in D_e$. f(x) = 1 and the size of x is below the average size of the elements of {y : f(y) = 1}]

One might quibble with some of the details, but (17) is definitely on the right track towards an explanation of why (13) fails to imply (14). According to (17), (13) asserts that Jumbo is an elephant and Jumbo's size is below the average elephant size. The set of all animals contains mostly individuals that are smaller than any elephants, so the average animal size is much lower than the average elephant size, and it is easy for Jumbo's size to fall between the two. This is how (14) can be false when (13) is true.

So it seems that we should allow at least some adjectives to denote functions of type <<e,t>,<e,t>>. "Small", we have seen, cannot possibly be interpreted with the lower type <e,t>, and this is just one of many examples for which the same kind of argument can be made. These adjectives are often called

nonintersective. In the adjective inventories of natural languages, they are apparently in the majority. "Intersective" adjectives – that is, those which validate equivalences like (12), like our initial example "gray" – represent the exception rather than the rule.

Are there any implications of this discussion for the analysis of the intersective adjectives? Not as far as we have seen up to now. Adding type <<e,t>,<e,t>> adjectives like "small" to the lexicon does not seem to cost us anything extra, whether we add them to a grammar that also has type <e,t> adjectives and a rule of PM, or to one where *all* adjectives have type <<e,t>,<e,t>> and PM is absent. Notice that our theory does not place any premium *per se* on having a uniform semantic type for all members of a given syntactic category. So the fact that on the PM option, some adjectives have type <e,t> and others type <<e,t>,<e,t>> is not in itself a reason to disprefer it.

The picture changes if we recall some of the trade-offs we thought to be tied to the elimination of PM. For instance, we observed that type <<e,t>,<e,t>> meanings are less straightforwardly interpretable than type <e,t> meanings in predicative occurrences of APs, as in "Julius is gray". But it turns out that adjectives like "small", which we have seen *cannot* have lexical entries of type <e,t>, also occur as seemingly complete predicates:

(18) **Jumbo is small.**

What do we make of this? We seem to need one of the mechanisms that we considered in section 4.3.2 as a replacement for PM after all, whether we have PM or not. For instance, we may posit a zero modifiee in the syntax of (18), or assume a lexical rule that routinely produces secondary type <e,t> meanings from type <<e,t>,<e,t>> inputs. (See exercise above.) But this tips the balance, and a theory without PM begins to look more parsimonious over all.

Or does it? Let's consider the whole case based on "small" a bit more carefully before we jump to conclusions. Is a type <<e,t>,<e,t>> analysis as in the entry (17) really forced on us by the data we have considered?

An elementary observation about adjectives like "small" that we have so far left aside is that they are vague and heavily context-dependent. In suitable specific discourse settings, people may have truth-value judgments about utterances containing "small" which are firm and uniform across speakers. But if we try to generalize over all felicitous uses of a given "small" sentence, we find that objects of practically any size can count as "small" and also as "not small". It is important to note that this vagueness and context-dependency are not limited to predicative occurrences like the one in (18). It remains even when there is a noun that "small" modifies. Consider Jumbo again. When you first read sentence (13), which described him as "a small elephant", in the context in which we presented it above, you spontaneously interpreted it in the manner we have

described: namely, as true only if Jumbo is smaller than the average elephant. But imagine we had first introduced a scenario populated with an army of monsters like King Kong. We might then have said something like: "Jumbo doesn't have a chance; he's only a small elephant", and this could have been true even if Jumbo were as large as or even larger than most other elephants.

So the meaning for "small" we codified in (17) represents at best a sort of default that applies when phrases of the form "small N" are interpreted more or less out of context. The contribution of the modified noun's meaning to the meaning of the whole phrase is not in general this mechanical. Perhaps the basic generalization about "small" is that it means "of a size less than the contextually salient standard". How the contextually salient standard is established for each given utterance of the word "small" is a complex affair. Previous discourse and the nonlinguistic circumstances of the utterance play a role. The mention of the word "elephant" in the immediate vicinity of the adjective draws attention to the elephant stereotype, including the stereotypical elephant size. Perhaps this is the whole reason why the average size of elephants happens to be the most salient standard in most situations where somebody utters the phrase "small elephant". In other words, the contribution of the modified noun may be rather indirect and mediated by the context.

If this picture is correct, a type <e,t> interpretation for "small" may be viable after all. The lexical entry might say essentially the following:

(19) $[\![\textbf{small}]\!] = \lambda x \in D_e$. x's size is below c, where c is the size standard made salient by the utterance context.

"Small elephant", as it occurs in (13), could then be interpreted by PM after all. It would receive the truth-conditions in (17) just in case the context of utterance does not supply a more salient size standard than the average size of elephants. To explain the intuition that (13) does not entail (14), we would assume that utterance contexts change quickly. We must acknowledge that it is possible, in fact highly natural, for an utterance of (14) which follows right after (13) to change the prevailing size standard from average elephant size to average animal size. We do predict that (13) entails (14) *if the context for both is the same.* But this prediction may be compatible with the evidence, if we can tell a plausible story about why the context will automatically change whenever an utterance of (14) follows one of (13).

We have evidently scratched only the surface of a complex of important issues here. Vagueness and context-dependency have been studied quite carefully by philosophers and linguists within the general framework of natural language semantics that we are presenting here. But the results of their work, both substantive and technical, are largely beyond the scope of this text.[9] Our present conclusion must therefore remain open for reconsideration. But for the purposes

of this book, we assume that type <e,t> entries for vague adjectives like "small" are viable, and we continue to work with the Predicate Modification rule.

We should briefly mention that there are a few other types of apparently nonintersective adjectives, with a behavior rather different from "small". One group is represented by "former" and "alleged". Clearly, if John is a former teacher, it does not follow that John is former and John is a teacher. The second conjunct is clearly false in this case, and the first is not even grammatical. The latter fact suggests that "former" does not have a meaning of type <e,t>, and moreover that it is not, after all, advisable to make a zero modifiee available in the syntax of all VPs of the form "be" + AP. At least, these are the obvious conclusions to draw if we want to predict that "John is former" is plainly uninterpretable. In this respect, "former", unlike "small", shows just the distribution that we would expect of an adjective of type <<e,t>,<e,t>>. But unfortunately its meaning can be shown not to be of this type. The reasoning goes as follows: if [[**former**]] were a function of type <<e,t>,<e,t>>, then for any two nouns α and β, the following would hold:

(20) If $[\![\alpha]\!] = [\![\beta]\!]$, then $[\![\textbf{former } \alpha]\!] = [\![\textbf{former } \beta]\!]$.

The reason for (20) is that "former α" (and likewise "former β") could only be interpreted by Functional Application: that is, $[\![\textbf{former } \alpha]\!] = [\![\textbf{former}]\!]([\![\alpha]\!])$. Since $[\![\alpha]\!]$ is by assumption the same as $[\![\beta]\!]$, we must have $[\![\textbf{former}]\!]([\![\alpha]\!]) = [\![\textbf{former}]\!]([\![\beta]\!])$. But (20) implies counterintuitive predictions. Suppose Bill's lovers happen to be exactly the tenants of 13 Green Street. So for any $x \in D$, $[\![\textbf{lover of Bill's}]\!](x) = 1$ iff $[\![\textbf{tenant of 13 Green St}]\!](x) = 1$. By the mathematical definition of a function, this means that $[\![\textbf{lover of Bill's}]\!] = [\![\textbf{tenant of 13 Green St}]\!]$. With (20), it then follows that, if "John is a former lover of Bill's" is true, then so is "John is a former tenant of 13 Green Street". But intuitively, the situation described is entirely compatible with the first of these being true and the second false. So we conclude that [[**former**]] is not of type <<e,t>,<e,t>>.

Apart from these negative conclusions, our present framework doesn't enable us to say anything precise. Very roughly, a successful analysis of "former" presupposes a general account of the time parameter in predicates. We must first adapt our semantics to the elementary fact that, for example, "John is a teacher" may be true in 1970 and false in 1980. Our current entries for predicates only make sense if we either disregard any change over time, or tacitly agree that we are considering a certain fixed point in time. The proper treatment of implicit time reference and temporal quantification requires an intensional semantics. An intensional semantics will be introduced in chapter 12, but even there, we will neglect temporal dependencies. Accordingly, we cannot fully answer the question of what the existence of adjectives like "former" implies for the treatment of adjectives and modification in general.

Many other loose ends have been left in this brief introduction to modifiers. The astute reader may have noticed, for instance, that we have said nothing about PPs modifying verbs rather than nouns. A few cautiously chosen examples of this type happen to mean what we predict them to mean: for example, "Julius is sleeping on the couch" does seem to be true just in case Julius is sleeping and is on the couch. (*Exercise*: Show that this is what we predict.) Many other examples that readily come to mind, however, do not have the predicted meanings at all. Consider "John wrote on the blackboard", "Mary put the book on the table", or "Max tossed the salad in the bowl", to name just a few. There are plenty of good grounds here to suspect that our present account of modifiers is very far from how modification really works in natural language. On the other hand, it might turn out that the problems lie elsewhere – for instance, in our simplistic assumptions about verb meaning and VP structure. Indeed, current research overwhelmingly points to the latter conclusion, and many *prima facie* counterexamples emerge as cases of intersective modification after all.[10]

4.4 The definite article

We have proposed that common nouns like "cat" denote the characteristic functions of sets of individuals. What does this imply for the semantic analysis of determiners? We will defer the general version of this question to chapter 6. Right here, we will look at only one determiner: namely, the definite article.[11]

4.4.1 A lexical entry inspired by Frege

The basic intuition about phrases of the form "the NP" is that they denote individuals, just like proper names. Had it not been for Bertrand Russell's famous claim to the contrary, few people would think otherwise. Frege, for one, thought it obvious: "let us start, e.g., with the expression 'the capital of the German Empire.' This obviously takes the place of a proper name, and has as its reference an object."[12] Hence his practice of referring to definite descriptions as "compound proper names".

If you read on in that particular passage, Frege imposes a rather odd syntactic analysis on his example. Instead of dividing it into the constituents "the" and "capital of the German Empire", he splits it up into "the capital of" and "the German Empire". He then proceeds to analyze "the capital of" as denoting a function from objects to objects (our type <e,e>). Elsewhere, however, he treats a similar example as follows:

> "the negative square root of 4". We have here a case in which out of a concept-expression a compound proper name is formed with the help of the definite article in the singular, which is at any rate permissible when one and only one object falls under the concept.[13]

We follow Frege in this second syntactic and semantic analysis. Notice that by "concept-expression", he means an expression whose meaning is of type <e,t>. In his example, that's the NP "negative square root of 4", which indeed receives a meaning of that type if we analyze it along the lines of the previous sections of this chapter. ("Square root", it seems, is a transitive noun like "part", with a meaning of type <e,<e,t>>. "Of" is vacuous, **[[square root]]** applies to 4 by Functional Application, and the result of that composes with **[[negative]]** under Predicate Modification.)

The determiner "the", then, denotes a function with arguments in $D_{<e,t>}$ and values in D_e. For instance, **[[the]]** applied to the function **[[negative square root of 4]]** yields the number –2. **[[the]]** applied to **[[president of the USA]]** yields Clinton at the time of writing. **[[the]]** applied to **[[opera by Beethoven]]** yields *Fidelio*. The generalization that emerges is (1).

(1) For any $f \in D_{<e,t>}$ such that there is exactly one x for which $f(x) = 1$, **[[the]]**(f) = the unique x for which $f(x) = 1$.

What about functions f which do *not* map exactly one individual to 1? What is **[[the]]**(f) for one of those?

Let's examine our intuitions on this matter. What are the objects denoted by the following definites?

(2) **the escalator in South College**

(3) **the stairway in South College**

You should know that South College has no escalator and more than one stairway. Once we are aware of this, we are hard pressed to say which objects (2) and (3) denote. The only natural answer is that neither of these phrases denotes any object at all. Let's go ahead and implement precisely this simple-minded intuition in our lexical entry for "the".

What we are aiming to predict is that (2) and (3) have no semantic value. In other words, there is no such thing as **[[the escalator in South College]]** or **[[the stairway in South College]]**. The reason has to be that the functions **[[escalator in South College]]** and **[[stairway in South College]]** are *not in the domain of* **[[the]]**. If they are not in the domain of **[[the]]**, then **[[the]]** can't apply to them, and this means that we cannot apply FA to calculate a semantic value for the DP-nodes in (2) or (3). The generalization that emerges regarding the domain of **[[the]]** is this:

(4) The domain of [[**the**]] contains just those functions $f \in D_{<e,t>}$ which satisfy the condition that there is exactly one x for which $f(x) = 1$.

Putting (1) and (4) together, we can now formulate our lexical entry for "the":

(5) [[**the**]] =
$\lambda f : f \in D_{<e,t>}$ and there is exactly one x such that $f(x) = 1$.
the unique y such that $f(y) = 1$.

This is a bit of an unwieldy λ-term,[14] but if you apply it to the examples above, you can see that it describes the function we were trying to define.

Before we end this subsection, let's dispose of a technical matter. What is the semantic type of [[**the**]]? In the strict sense of our definitions so far, it actually has none. To say that its type is <<e,t>,e> would mean that it is a function from $D_{<e,t>}$ to D_e. But "from $D_{<e,t>}$" means "with domain $D_{<e,t>}$", and we have just seen that the domain of [[**the**]] is not $D_{<e,t>}$ but only a subset thereof. At this point, we find ourselves with an inconvenient terminology, and we will simply change it. We henceforth define $D_{<\sigma,\tau>}$ (for any types σ, τ) as the set of all *partial* functions from D_σ to D_τ. "Partial function from" is defined as follows:

(6) A *partial* function *from* A *to* B is a function from a subset of A to B.

(When we emphatically mean "function from" rather than "partial function from", we will sometimes say "*total* function from".) With these new definitions, we can now say that [[**the**]] is in $D_{<<e,t>e>}$, or that its type is <<e,t>,e>.

4.4.2 *Partial denotations and the distinction between presupposition and assertion*

When a tree contains a lexical item that denotes a partial function, this may cause the tree to wind up without a semantic value. We have already seen this happen in (2) and (3) above. In larger structures, there are repercussions all the way up the tree. For instance, the following sentences are predicted not to have any semantic values, neither 1 nor 0 nor anything else.

(7) **The stairway in South College is dirty.**

(8) **John is on the escalator in South College.**

This is a direct consequence of our set of composition principles. The only principle that could potentially provide a semantic value for the branching node

above the definite description (that is, the S-node in (7) and the PP-node in (8)) is Functional Application, repeated here from chapter 3.

(9) *Functional Application* (FA)
If α is a branching node and $\{\beta, \gamma\}$ is the set of α's daughters, then α is in the domain of $[\![\ ]\!]$ if both β and γ are and $[\![\gamma]\!]$ is in the domain of $[\![\beta]\!]$. In this case, $[\![\alpha]\!] = [\![\beta]\!]([\![\gamma]\!])$.

But, as this formulation makes plain, you can't apply the denotation of one daughter to that of the other unless both daughters *have* denotations. So FA can't apply (and no other composition principle even came close). By the same reasoning, no semantic values can be obtained for any higher nodes that indirectly dominate a denotationless definite description like (2) or (3).

Exercise

Consider an example with one definite description embedded in another:

(i) The killer of the black cat escaped.

(a) Draw an interpretable syntactic structure for (i).

(b) Describe three possible states of affairs:

one where (i) is false;
another one where (i) lacks a truth-value because "the black cat" has no extension;
a third one where (i) is also without a truth-value, but this time because "the killer of the black cat" lacks a denotation. (Assume for this third scenario that "the black cat" does have a denotation.)

Are the empirical predictions that are implied by our current semantic component correct? Consider what we predict about (8): If you know English, and if you furthermore know that there is no escalator in South College, then you know that:

(a) the sentence "John is on the escalator in South College" is not true, and
(b) the sentence "John is on the escalator in South College" is not false.

Part (a) of this prediction is unobjectionable. But in apparent disagreement with (b), many informants will spontaneously classify the assertion "John is on the escalator in South College" as false.

Does this mean that our Fregean semantics for the definite article has proved empirically inadequate and must be abandoned? Many philosophers and linguists have drawn this conclusion.[15] It is a reasonable conclusion, but it is not inescapable. An alternative response is to reconsider the straightforward identification which we have assumed so far between the semantic values 1 and 0 and the pre-theoretical notions of truth and falsity. Might we perhaps reconcile the present semantic analysis with the empirical evidence if we posit a somewhat more indirect correspondence between the truth-values of our semantic theory and the intuitions that people report in truth-value judgment tasks?

As a first step in this direction, let us propose that the colloquial term "false" covers both truth-value-less sentences and those that are false in the technical sense of denoting 0. In other words, the technical terms of our theory translate into pre-theoretical terms as follows:

$[\![\phi]\!] = 1$	ϕ is true
ϕ has no semantic value $[\![\phi]\!] = 0$	ϕ is false

This stipulation makes the predictions of our semantics consistent with the data we reported above: for example, with the fact that informants who are told that South College contains no escalator and are asked to decide whether (8) is true or false will choose "false".

Mere compatibility with the data is not all we are aiming for, of course. To justify our choice over competing theories, in particular those that make only a 2-way distinction between true and false, we have to show that the additional distinction between two sources of intuitive falsity does some useful work. For instance, we might argue that it helps us to explain certain other manifestations of semantic competence, which can be observed when we move beyond simple truth-value judgment tasks and elicit subtler intuitions.

Indeed, this kind of argument has been offered.[16] Specifically, it has been argued that the technical distinction between lacking a value and denoting 0 can be systematically related to an intuitive distinction: namely, the distinction between what is *asserted* and what is *presupposed*. For an illustration, consider the following three sentences.

(10) (a) **John is absent again today.**
(b) **Today is not the first time that John is absent.**
(c) **John is absent today, and that has happened before.**

All three of these sentences somehow express the speaker's belief that John is absent today and has been absent at least once before. But they are not simply interchangeable. If you are talking to somebody with whom you already share the information that John has been absent in the past, but who doesn't yet know about today, (10a) is a natural choice, but (10b) is not. If you are talking to somebody who already knows that John is absent today, but knows nothing about his past history, then (10b) is natural, whereas (10a) is not. And if your audience knows nothing at all about John's past or present attendance, the most natural choice is (10c). We accordingly say that (10a) *presupposes* that John was absent before and *asserts* that he is absent today. With (10b), it's the other way round: this sentence presupposes that John is absent today and asserts that he was absent before. Finally, (10c) asserts that John both is absent today and was absent before, without presupposing anything.

The ability to discriminate between (10a), (10b), and (10c) and decide which is most appropriate for an audience with a given state of information is clearly part of understanding English. We would therefore like to capture it somehow in our semantics. The hypothesis presently under consideration is that a semantic theory equipped to distinguish two kinds of non-true sentences is better suited to accomplish this than one that isn't. The concrete proposal is that ϕ having no value represents the case that ϕ has a false presupposition, and $[\![\phi]\!] = 0$ means that ϕ does not presuppose anything false but makes a false assertion. (The third case, $[\![\phi]\!] = 1$, thus has to mean that both what ϕ presupposes and what it asserts are true.)

Let's return to the definite article. The analysis we presented above (following Frege) may be called a "presuppositional" analysis of "the". In light of what we have said in the present subsection, it predicts that a sentence like (8) ("John is on the escalator in South College") would be used most naturally by a speaker who assumes that her audience knows that there is a unique escalator in South College, but doesn't know about John's whereabouts. This seems basically right. Minimal pairs like (11a, b) below point in the same direction.

(11) (a) **There will be one mid-term, which will be on November 21st.**
(b) **The mid-term will be on November 21st.**

If the topic of mid-term exams for this course hasn't come up yet at all, (11a) is the natural choice; (11b) is fully appropriate only when the audience is already aware that there will be one mid-term. Neither sentence can be true unless there will be a unique mid-term, but in (11a) this is part of the assertion, while in (11b) it is presupposed. Our semantic analysis predicts this: If there isn't a unique mid-term, then (11b) has no truth-value at all. (In contrast with (11a), which should denote 0 in this case. We say "should", because we have yet to develop an analysis of the words and constructions in (11a) that actually predicts this.)

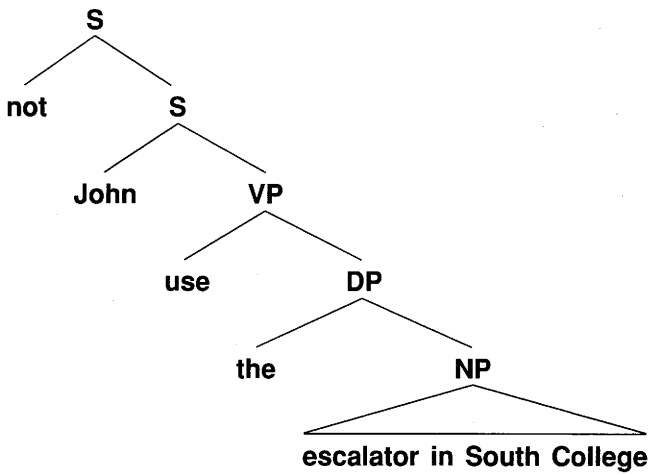
We have made an attempt here to provide some initial evidence in favor of our presuppositional analysis. This is not meant to be a compelling justification of this analysis, however. There are many well-known objections which we have not even mentioned, much less addressed. For the time being (especially for the next chapter), it is convenient to be able to assume *some* concrete analysis of the definite article, even if it turns out to be only a crude approximation. This will allow us to get our analysis of other parts of English off the ground.

Exercise 1

Assume that sentence (i) has the structure (i′) at the level at which it is interpreted.

(i) **John doesn't use the escalator in South College.**

(i′)



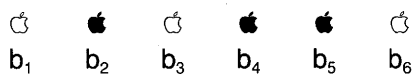
In (i′), the determiner "the" is assumed to be a D (determiner, not to be confused with the domain of individuals D, also referred to as "D_e") heading a DP (determiner phrase).[17] The structure (i′) differs from the surface structure of (i) with respect to the position of the subject "John". We might assume that (i′) is the Deep Structure of (i), where the subject appears below the negation before movement to a higher position. Or else we could assume that (i′) is a Logical Form representation that results from reconstruction of the raised subject into its original position.

What does our current theory predict about this sentence? Are the predictions empirically correct? Construct suitable examples and scenarios that help

sharpen intuitions and promote discussion. It might also be useful to go through Bertrand Russell's "On Denoting" (see n. 15) and see which of his examples would argue against the analysis of the definite article proposed in this chapter.

Exercise 2

Look at the following scenario:

(← LEFT)  (RIGHT →)
b_1 b_2 b_3 b_4 b_5 b_6

and consider the following three definite descriptions:

(i) the leftmost apple in the row
(ii) the leftmost dark apple in the row
(iii) the apple that is both leftmost in the row and dark

(a) In your intuitive judgment, which individual, if any, does each of these definite descriptions refer to in this situation?
(b) What predictions concerning the denotations of the definite descriptions (i), (ii), and (iii) would follow from an analysis that treats adjectives like "leftmost" as 1-place predicates?
(c) Specify a more adequate denotation for attributive uses of "leftmost".
(d) In a compositional fashion, compute the denotation of the definite description "the leftmost dark apple in the row", given the above scenario. For the purpose of this computation, take "apple in the row" as an unanalyzed predicate. That is, you don't have to worry about the PP "in the row".

4.4.3 Uniqueness and utterance context

Frege's uniqueness presupposition has often been objected to as an idealization that does not really fit the definite singular article in English. We frequently say things like "the door is locked" or "the cat wants to be fed", yet we don't believe that there is just one door and just one cat in the world, and nobody that hears us speak this way will attribute such beliefs to us either.

There are a number of different responses to this objection, but this is not the occasion to give them serious consideration. Somehow or other, we have to

concede that "the cat" doesn't denote the unique cat that there is in the whole world, but rather denotes, on each occasion on which it is uttered, the unique cat among those individuals that are under consideration on this utterance occasion. Our best attempt for now at making this explicit is in the following revised lexical entry:

(5′) $[\![\textbf{the}]\!] =$
$\lambda f : f \in D_{<e,t>}$ and there is exactly one $x \in C$ such that $f(x) = 1$.
the unique $y \in C$ such that $f(y) = 1$,
where C is a contextually salient subset of D.

Once again, our lack of a serious account of context-dependency prevents us from stating this more precisely. Below, we will assume something like (5′) in informal discussion, but will abstract away from context-dependency and use (5) in our calculations.

4.4.4 *Presupposition failure versus uninterpretability*

In chapter 3, we talked about cases of a rather different sort in which a linguistic expression fails to have a semantic value. Recall our treatment of so-called Θ-Criterion violations like "Ann laughed Jan". We observed that this sentence is not in the domain of the $[\![\ ]\!]$ function as defined by our semantic theory. We called such structures "uninterpretable", and we proposed that the uninterpretability of "Ann laughed Jan" accounted for the ungrammaticality judgment represented by the asterisk.

In the present chapter, however, we have just suggested that sentences which lack a semantic value are intuitively judged as presupposition failures. So it seems that we have not been consistent. Is it a kind of falsity or a kind of ungrammaticality that we want our theory to capture when it provides no denotation for a given structure? The two are obviously quite different intuitively, and by simply conflating them we would be missing a systematic fact about people's linguistic intuitions. We might try to draw the intended distinction as follows:

(12) If α is *uninterpretable*, then it can be proved from the semantics alone that α is outside the domain of $[\![\ ]\!]$.

(13) If it is a contingent matter of fact that α is outside the domain of $[\![\ ]\!]$, then α is a *presupposition failure*.

(12) and (13) correctly distinguish between, say, "Ann laughed Jan" on the one hand and "The escalator in South College is moving" on the other. In the former

case, we need not assume anything about the world to show it lacks a denotation. In the latter case, we need to invoke physical facts to show this, and we can easily imagine counterfactual states of affairs in which that sentence *would* have a truth-value.

But if we tried to turn (12) and (13) into biconditionals that could stand as *definitions* of uninterpretability and presupposition failure, we would face an objection: namely, that there are sentences which intuitively are "necessary presupposition failures" – for example, "John met the man who died and didn't die". By the criteria given in (12), (13), this is indistinguishable from "Ann laughed Jan": we only need the semantics of English in order to infer that it has no semantic value. But its intuitive status is different, and it should be classified as a presupposition failure rather than as uninterpretable.

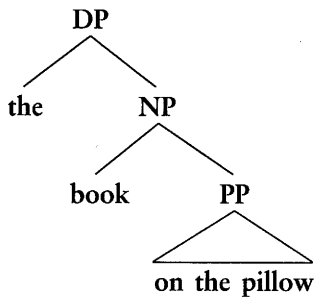
So the distinction we are after cannot simply be identified with the difference between necessary and contingent lack of denotation. If we want to characterize it in precise terms, we have to be more specific. In the case of an uninterpretable structure, information about the *type* of each subtree is sufficient to decide that the structure receives no denotation. To detect presupposition failure, by contrast, we must know more about the denotations of certain subtrees than their mere semantic types.

4.5 Modifiers in definite descriptions

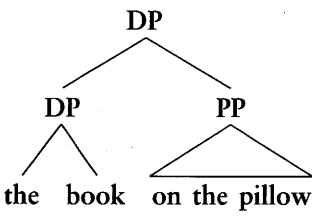
We conclude this chapter by highlighting some predictions that our current semantics makes about the interaction of modifiers and the definite article when these co-occur.

As was already implicit in our treatment of examples in section 4.4, we are assuming a syntax according to which restrictive modifiers within DPs (determiner phrases) form a constituent with the head noun to the exclusion of the determiner.[18] That is, the bracketing is as in (1), not as in (2).

(1)



(2)



Accordingly, our semantics composes the values of the noun and the modifier before composing the result of this with the value of the determiner.

Syntacticians have frequently entertained the opposite hierarchical organization (2). Suppose this were the structure of the object in the sentence "Ann dislikes the book on the pillow". What would happen if we attempted to interpret this sentence by our principles of semantic composition?

The first answer to this question is that the alternative structure in (2) leads to uninterpretability within our framework – not within (2) itself, but in whatever larger structure (2) is part of. What happens is that the lower DP, "the book", denotes the unique (contextually relevant) book (if any), and the higher DP then denotes a truth-value: namely, 1 if the unique book is on the pillow, and 0 if it is not. So any attempt to interpret the next higher node (say, the S-node if (2) is a subject, or the VP-node if it is the object of a transitive verb) will fail due to type-mismatch.[19]

In a historic debate, Barbara Partee[20] took the uninterpretability of structures like (2) (given certain plausible assumptions) to be an argument against syntactic analyses that imply the bracketing [[determiner noun]modifier] and in favor of those that assume [determiner [noun modifier]]. Noam Chomsky[21] pointed out in response that there is undeniable evidence for such alleged "uninterpretable" surface bracketings in at least some constructions in some languages. (Directly relevant here, for instance, are Scandinavian languages, where the definite article is realized as a suffix on the noun and intervenes in the linear surface order between it and the PP.) The proper conclusion to draw seems to be that surface structure need not always be the appropriate input to semantic interpretation. Rather, the level that is interpreted is a more abstract one. Our hypothesis then must be that even languages that exhibit the bracketing [[determiner noun] modifier] on the surface have [determiner [noun modifier]] at some other level of representation.

Notes

1 For a more systematic treatment of the indefinite article, see ch. 6.

2 See especially R. Jackendoff, *X′ Syntax* (Cambridge, Mass., MIT Press, 1977).

3 Ibid., chs 4 and 7.

4 A more explicit formulation along the lines of section 3.3 would be (4′).

(4′) If α is a branching node and $\{\beta, \gamma\}$ is the set of α's daughters, then α is in the domain of $[\![\ ]\!]$ if both β and γ are, and $[\![\beta]\!]$ and $[\![\gamma]\!]$ are both in $D_{<e,t>}$. In this case, $[\![\alpha]\!] = \lambda x \in D_e\ .\ [\![\beta]\!](x) = [\![\gamma]\!](x) = 1$.

5 Remember that we couldn't have written "$[\![\alpha]\!] = \lambda x \in D_e\ .\ [\![\beta]\!](x)$ and $[\![\gamma]\!](x)$" in the last line of (6). $[\![\beta]\!](x)$ and $[\![\gamma]\!](x)$ are truth-values. A value description of the

form "[[β]](x) and [[γ]](x)", then, is as ill-formed as a value description of the form "1 and 0".

6 Higginbotham calls this operation "theta identification": J. Higginbotham, "On Semantics," *Linguistic Inquiry*, 16 (1985), pp. 547–93, at p. 564.

7 Two influential papers concerning type shifting and flexible types are B. Partee and M. Rooth, "Generalized Conjunction and Type Ambiguity," in R. Bäuerle, C. Schwarze, and A. von Stechow (eds), *Meaning, Use and Interpretation of Language* (Berlin, de Gruyter, 1983), pp. 361–83, and B. H. 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.

8 For an informative overview and many references, consult C. Hamann, "Adjectival Semantics," in A. von Stechow and D. Wunderlich (eds), *Semantik/Semantics. An International Handbook of Contemporary Research* (Berlin and New York, de Gruyter, 1991), pp. 657–73. Another useful overview article (mainly discussing the semantics of adjectives) is B. H. Partee, "Lexical Semantics and Compositionality," in L. R. Gleitman and M. Liberman (eds), *Language. An Invitation to Cognitive Science*, 2nd edn, vol. 1 (Cambridge, Mass., MIT Press, 1995), pp. 311–60.

9 The best overview article on context dependency (unfortunately in German) is T. E. Zimmermann, "Kontexttheorie," in von Stechow and Wunderlich (eds), *Semantik/Semantics*, pp. 156–229. The references given in Zimmermann's article are a good starting point even for readers who cannot read the article itself. A classic reference is D. Kaplan, "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and other Indexicals," in J. Almog, J. Perry, and H. K. Wettstein (eds), *Themes from Kaplan* (New York, Oxford University Press, 1989), pp. 481–563. As for formal accounts of vagueness, we recommend M. Pinkal, *Logic and Lexicon: The Semantics of the Indefinite* (Dordrecht, Kluwer Academic Publishers, 1995). The classic reference for an account of the vagueness of adjectives is H. Kamp, "Two Theories about Adjectives," in E. Keenan (ed.), *Formal Semantics of Natural Language* (Cambridge, Cambridge University Press, 1975), pp. 123–55. See also E. Klein, "A Semantics for Positive and Comparative Adjectives," *Linguistics and Philosophy*, 4 (1981), pp. 1–45.

10 The classic reference is D. Davidson, "The Logical Form of Action Sentences," in N. Rescher (ed.), *The Logic of Decision and Action* (Pittsburgh, University of Pittsburgh Press, 1967), pp. 81–95. See also T. Parsons, *Events in the Semantics of English. A Study in Subatomic Semantics* (Cambridge, Mass., MIT Press, 1990).

11 And we won't even consider the full range of occurrences of this. Since all the common nouns we use in the book are singular count nouns, we say nothing about definite mass and plural terms. We confine our attention to what philosophers call "definite descriptions": viz. singular terms in which "the" combines with an NP headed by a count noun.

12 Frege, "Function and Concept" (1891), trans. in M. Black and P. Geach, *Translations from the Philosophical Writings of Gottlob Frege* (Oxford, Basil Blackwell, 1960), pp. 56–78.

13 Quote from Frege, "Über Sinn und Bedeutung," trans. as "On Sense and Nomination," in A. P. Martinich (ed.), *The Philosophy of Language*, 2nd edn (New York and Oxford, Oxford University Press, 1990), pp. 190–202, at pp. 197–8.

14 If we help ourselves to a little bit of mathematical and logical notation, we can make it shorter:

$\lambda f : f \in D_{<e,t>} \;\&\; \exists!x[f(x) = 1] \;.\; \iota y[f(y) = 1]$.

The abbreviatory conventions used here are the following:

"$\exists!x[\phi]$" abbreviates "there is exactly one x such that ϕ".
"$\iota x[\phi]$" abbreviates "the unique x such that ϕ".

(The first symbol in the last line is the Greek letter iota.)

15 Most notably B. Russell, in "On Denoting," *Mind*, 14 (1905), pp. 479–93, who proposed an alternative analysis which predicts that **[[John is on the escalator in South College]]** = 0 (given that there is no escalator in South College).

16 The classical reference here is once again Frege. See "Über Sinn und Bedeutung," pp. 196–7.

17 See S. Abney, "The English Noun Phrase in its Sentential Aspect" (Ph.D. dissertation, MIT, 1987).

18 The issue did not really arise before section 4.4, since in section 4.3 we only used the determiner "a". Since that was taken to be vacuous, nothing depended on its exact place in the phrase structure hierarchy.

19 We predict, thus, that if structures like (2) are generated by the syntax of English at all, they can only be used unembedded to make statements, or embedded in positions where truth-values are selected (e.g., as complements to truth-functional connectives). This possibility seems not to be realized, a fact for which we must seek some syntactic explanation.

20 B. H. Partee, "Some Transformational Extensions of Montague Grammar," in B. Partee (ed.), *Montague Grammar* (New York, Academic Press, 1976), pp. 51–76, esp. pp. 53ff. Actually, Partee makes her point about restrictive relative clauses rather than restrictive PP modifiers. But the internal semantics of relative clauses (see next chapter) is not relevant to it, and it carries over to any kind of postnominal restrictive modifier.

21 N. Chomsky, "Questions of Form and Interpretation," *Linguistic Analysis*, 1/1 (1975), pp. 75–109, esp. pp. 96ff.