A RELEVANCE THEORETIC ACCOUNT OF DEFINITE DESCRIPTIONS

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Philosophy

By

Diana C. Puglisi, B.A.

Washington, DC
April 11, 2014
Copyright 2014 by Diana C. Puglisi
All Rights Reserved
A RELEVANCE THEORETIC ACCOUNT OF DEFINITE DESCRIPTIONS

Diana C. Puglisi, B.A.

Thesis Advisor: Wayne Davis, Ph.D.

ABSTRACT

Definite descriptions have a number of different readings, the most prominent being the referential and attributive readings. On the referential reading, the description ‘The current President of the USA’ is interpreted as “Barack Obama”; read attributively, it is interpreted as “the current President of the USA, whoever it is”. Philosophical work on definite descriptions tends to examine the question of whether the referential/attributive distinction is a semantic or a pragmatic phenomenon. I use Sperber and Wilson’s relevance theory to provide a contextualist account of the English definite article, which views all of its uses as arising out of the interplay between semantics and pragmatics. I examine the strengths and weakness of previous contextualist approaches to definite descriptions, and I offer an account that treats the English definite article as encoding a procedural constraint on interpretation, where 'interpretation' is understood in terms of relevance theory. According to relevance theory, hearers select the interpretation which yields the greatest cognitive benefits for the least cognitive processing cost. The cost of processing, and the benefits derived from it, vary with the context. Thus, while the procedural constraint encoded by the definite article limits the range of potential interpretations, which interpretation is assigned to a definite description depends on the context.

In addition to the referential and attributive uses, this relevance theoretic view can also account for the other readings of definite descriptions. The view is subject to a serious objection, however, one which faces any relevance theoretic account of a communicative phenomenon.
Because it is not currently possible to quantify cognitive benefits and cognitive processing costs, it does not appear possible to show that the relevance theoretic account of descriptions predicts that a particular interpretation of a description would be the natural interpretation. This would render the account untestable. I argue that with some additions to relevance theory, it is possible to overcome this objection.
This work is dedicated to all the people whose support made it possible: my committee chair, Wayne Davis, my readers, Linda Wetzel, Steven Gross, and Mark Lance, my dear friends Tony Manela, Jenifer Palucki, and Dan Quattrone, and finally my parents Al and LeAnn Puglisi.

I could not have done it without you.

Many thanks,

Diana C. Puglisi
# Table Of Contents

Chapter 1: Introduction .................................................................................................................. 1
Chapter 2: The Neo-Strawsonian Account ....................................................................................... 10
  2.1 Strawson ............................................................................................................................ 10
  2.2 The Neo-Strawsonian Account .......................................................................................... 18
  2.3 Problems with the Neo-Strawsonian Account ..................................................................... 28
Chapter 3: Contextualist Accounts ................................................................................................. 33
  3.1 The Two-Dimensional Account ......................................................................................... 35
  3.2 Problems with the Two-Dimensional Account .................................................................... 42
  3.3 The Synecdoche Theory .................................................................................................... 59
  3.4 Critique of the Synecdoche Theory .................................................................................... 64
Chapter 4: Relevance Theoretic Accounts ..................................................................................... 82
  4.1 Relevance Theory .............................................................................................................. 83
  4.2 Bezuidenhout’s Account ..................................................................................................... 90
  4.3 Problems with Bezuidenhout’s Account ............................................................................ 97
  4.4 Powell’s Account ............................................................................................................... 107
Chapter 5: Relevance Theory ....................................................................................................... 116
  5.1 Ostensive Inferential Communication ................................................................................. 117
  5.2 The Principles of Relevance .............................................................................................. 119
  5.3 Cognitive Processing ......................................................................................................... 132
Chapter 6: An Account of Definite Descriptions ........................................................................... 155
  6.1 The Relevance Theoretic Account ...................................................................................... 155
  6.2 Individual Concepts .......................................................................................................... 165
  6.3 The Specific, Functional and Predicative Uses .................................................................... 180
  6.4 Other Individual Concepts ............................................................................................... 196
  6.5 The Problem of Testability ............................................................................................... 200
Chapter 7: Comparative Assessment of Relevance ....................................................................... 203
  7.1 A Basic Strategy ............................................................................................................... 203
  7.2 Contrived Cases and Interpreters ...................................................................................... 206
  7.3 Available Interpretations .................................................................................................. 211
  7.4 Justification and Cognitive Effects ..................................................................................... 229
  7.5 Comparing Processing Costs ............................................................................................. 237
Chapter 8: The Test Cases .............................................................................................................. 248
  8.1 Key Theoretical Assumptions ............................................................................................. 248
  8.2 The Attributive Case .......................................................................................................... 255
  8.3 The Referential Case ......................................................................................................... 279
  8.4 The Kind-Referring Case .................................................................................................. 286
  8.5 Conclusion ......................................................................................................................... 295
Bibliography ................................................................................................................................. 296
Consider an utterance of the following sentence:

(1.1)  The fairy in my pocket is frightened.

The expression ‘the fairy in my pocket’ in (1.1) is a definite description. In attempting to provide an analysis of (1.1), one’s first thought might be to treat it as a directly referential expression, much like a name, whose contribution to the proposition expressed by the utterance in which the expression occurs consists solely in the referent of the expression. It turns out, however, that treating definite descriptions in this way is problematic. If we take the contribution made by a description to the proposition expressed by the utterance in which the description occurs to be exhausted by the description’s referent, then (1.1) would appear to be meaningless, for there is no fairy in my pocket. The description in (1.1) lacks a referent in the actual world, and thus an utterance of (1.1) would seem to express but half a proposition. Intuitively, however, an utterance of (1.1) expresses a complete proposition which we are perfectly capable of understanding.

Bertrand Russell’s solution to this problem was to reject the notion that definite descriptions are directly referential in favor of the view that they are what he called “denoting phrases”.\(^1\) According to Russell, expressions of the form the $F$ are quantificational in nature, and the utterances containing them express general propositions. On Russell’s view, an utterance of the form the $F$ is $G$ should be given the interpretation in (1.2).\(^2\)

\[
(1.2) \quad \exists x (Fx \land \forall y (Fy \rightarrow x = y) \land Gx)
\]

---

\(^1\) Russell (1905), p. 479
\(^2\) Ibid., pp. 481-482
Essentially, according to Russell, an utterance of the form *the F is G* says that there is exactly one *F* and it is *G*. (1.1), then, on Russell’s view, says that there is exactly one fairy in my pocket and that fairy is frightened. Because Russell interprets the semantics of an utterance of (1.1) as quantificational in nature, we can view it as expressing a complete, though false, proposition despite the fact that there is no fairy in my pocket.

Russell’s account of definite descriptions remains highly influential a century later. It has not escaped criticism, however. Strawson, for example, famously objected to Russell’s claim that utterances like (1.1) are false, proposing instead to return to a view of definite descriptions as essentially directly referential and address the problem of nonexistent referents by appeal to pragmatic presuppositions.  

Perhaps even more troubling for Russell’s account of definite descriptions, however, is the case made by Keith Donnellan for the claim that there are at least two distinct uses of definite descriptions, one of which does not appear amenable to Russell’s original analysis.

Donnellan presents us with a number of different cases in an attempt to draw out the intuitive distinction between the different uses of definite descriptions. Take, for instance, the now famous case of Smith’s murderer.

(1.3) **Smith’s murderer is insane.**

We can imagine two different contexts in which (1.3) might be uttered, and in which the same description – ‘Smith’s murderer’ – is used in two distinct ways. In the first context, imagine we have just discovered Smith, who we know, murdered. Suppose further that neither of us knows

---

3 Strawson (1950)
4 Donnellan (1966), (1968)
5 Donnellan (1966), pp. 285-286
who could have done this foul deed. In reaction to the scene, I exclaim (1.3). Intuitively, the
description as it occurs in my utterance means something like that suggested by Russell. In
uttering (1.3) in this context I mean that the murderer of Smith, whoever it is, is insane. But now
imagine that we are in the courtroom watching the trial of Jones for Smith’s murder, and that we
both know that we share the belief that Jones is the killer. Suppose further that Jones has been
behaving in an extremely strange manner all throughout the trial. In response to some of this
behavior, I lean over to you and whisper (1.3). In this case, intuitively, I have used the
description in (1.3) to refer to Jones. In uttering (1.3) in this case, I am not saying that the
murderer of Smith, whoever it is, is insane; I am saying that Jones is insane. In the first case, my
utterance of (1.3) serves to communicate a general proposition. In the second case, it serves to
communicate a singular proposition about Jones.

On the basis examples like forgoing, Donnellan concludes that descriptions have two
distinct uses. The attributive use is that seen in the first scenario we examined, where the
description appears to operate in accordance with Russell’s view of descriptions, and, intuitively,
what is communicated is a general proposition. The referential use is that seen in the second
scenario, where the description would seem to behave much like a directly referential expression,
and, intuitively, what is communicated is a singular proposition. It seems clear that if Donnellan
is correct, and there are, in fact, two distinct uses of definite descriptions, then Russell’s account
alone is insufficient to account for definite descriptions. For while Russell’s view of descriptions
can easily handle the attributive use of descriptions, it is without the machinery to account for the
referential use.
It is now widely accepted that definite descriptions have both an attributive and a referential use. Indeed, since Donnellan identified the referential/attributive distinction, a significant portion of the philosophical work on definite descriptions has been dedicated to attempting to deal with it. Two main camps have long dominated the philosophical landscape surrounding the analysis of definite descriptions and the referential/attributive distinction. Following Neale, let us call these two main opposing viewpoints the referentialist and Russellian positions. The referentialist view is motivated largely by the intuition that the referential use of a definite description is semantically relevant. To put it another way, the referentialist view is motivated by the intuition that when ‘the F’ in an utterance of the form *The F is G* is used referentially, a singular proposition is explicitly expressed or, to put it in Grice’s terms, is part of what is said. To account for this intuition, the referentialist maintains that definite descriptions are semantically ambiguous between the attributive and referential readings.

In contrast, the Russellian attempts to preserve Russell’s account of the semantics of definite descriptions, claiming that an utterance containing a definite description always explicitly expresses a general proposition. Relying on Gricean or neo-Gricean principles, the Russellian holds that referential uses of descriptions arise in contexts in which an implicature is generated. In such cases, on this view, the utterance explicitly expresses a general proposition, but a singular proposition is implicated. To put it in Gricean terminology, the Russellian holds that when a description ‘The F’ in an utterance of the form *The F is G* is used referentially, a general proposition is what is said, while a singular proposition is what is meant.

A substantial number of arguments have been put forward attempting either to support or undermine both the Russellian and referentialist accounts of the referential/attributive
distinction. The Russellian’s main argument against the referentialist view, however, lies in the claim that the referentialist account violates the principle of theoretical parsimony. In his work on implicature, Grice suggested that when attempting to account for various semantic phenomena, we should adhere to a modified version of Occam’s razor – do not multiple senses beyond necessity. This principle, which is sometimes known as *Grice’s razor*, essentially tells us that we should not posit extra linguistic senses when it is possible to do without them. Since the Russellian offers an account of the referential use of descriptions which does not require positing an additional sense for definite descriptions, the argument goes, Grice’s razor tells us that the Russellian view should be preferred over that of the referentialist, which accounts for the referential/attributive distinction in terms of a semantic ambiguity.

In support of the view that the referential use of descriptions is semantically relevant, Donnellan appealed to cases of misdescription. Consider Donnellan’s example of an utterance of (1.4).

(1.4) The man with the martini is sad.

Imagine (1.4) uttered at a party while the speaker gestures at a man who is holding a martini glass filled with clear liquid. Now suppose that the man in question is in fact sad, but unbeknownst to the speaker and hearer, the liquid in the glass is actually just water. The description in such an utterance of (1.4) would appear to be referentially used, and many have

---

6 The body of literature surrounding the debate between Russellsians and referentialists is quite large. Since that debate is not the main focus of this work, however, we do not have space to go into it in detail here. Instead, we will limit our focus to an overview of those aspects of the debate which will be relevant to our later discussion of contextualist and relevance theoretic accounts of definite descriptions. For arguments in favor of the Russellian view, see, e.g., Bach (1987), (1994: Ch. 6); Kripke (1977); Neale (1990: Ch. 3), (2004); Salmon (1982), (1991). For arguments in favor of the referentialist view, see, e.g., Barwise & Perry (1983); Devitt (1981), (2004); Wettstein (1981), (1983), (1986).

7 Donnellan (1966)
the intuition that the speaker has said something true. This intuition could be taken as evidence that the referential use of definite descriptions is semantically relevant. Not everyone shares the same intuition as to the truth value of such an utterance of (1.4), however. And even if we do take this sort of utterance of (1.4) to be saying something true, it also seems clear that the speaker has done something incorrect. Because the Russellian claims that utterances containing referentially used descriptions linguistically express a general proposition while implicating a singular one, the Russellian has a ready means of explaining this tension in our intuitions. In the cases of misdescription like that of (1.4), the Russellian can claim that the speaker has said something false, while communicating something true.

Referentialists also argue that the Russellian theory is incapable of handling incomplete descriptions — that is, descriptions whose content alone is insufficient to identify a referent. The description in (1.5) is an incomplete description.8

(1.5) The table is covered with books.

The description in an utterance of (1.5) would appear to be used referentially. Since the referentialist treats the referential sense of definite descriptions as functioning essentially as demonstratives do, he or she should have no problem accounting for the description in (1.5) or other incomplete descriptions like it. If the Russellian view of descriptions is correct, on the other hand, then an utterance of (1.5) says that there is exactly one table and it is covered with books. As there are many books in the universe, the Russellian account would seem to predict that an utterance of (1.5) is false. Intuitively, however, it is not. Russelians offer two different potential solutions to this problem. The first holds that incomplete descriptions are actually

8 The example was originally Strawson’s. See Strawson (1950).
complete descriptions in which some of the descriptive content has been elided. The second
holds that in the case of incomplete descriptions the domain of quantification is contextually
restricted such that there is a unique table in the restricted domain. Whether or not either of these
solutions to the problem actually works is a matter of considerable debate between Russellians
and referentialists.

In more recent work, however, an alternative to the referentialist and Russellian views of
descriptions has emerged – contextualism. This new approach to descriptions would allow us to
account for the referentialist’s intuition that the referential use of descriptions is semantically
relevant without running afoul of Grice’s razor. According to the contextualist, the linguistic
meaning of a definite description is neither referential nor attributive, and gives rise to different
explicitly communicated contents depending on how it is contextually supplemented. Thus, in
some contexts a general proposition is explicitly communicated, while in others a singular
proposition is explicitly communicated. Multiple contextualist accounts of definite descriptions
have been proposed over time, differing in their accounts of the linguistic meaning of definite
descriptions and how that meaning gives rise to different readings in different contexts. We will
look at some of these accounts in detail in later chapters. One contextualist view of definite
descriptions relies on the account of communication put forward by Sperber & Wilson –
relevance theory.

The overarching goal of this work is to examine the possibility of putting forward a
viable relevance theoretic account of English definite descriptions. The relevance theoretic view
of communication holds that in most cases, the linguistic meaning of an utterance alone is
insufficient to determine a proposition and must be filled in via pragmatic processing. The
account of definite descriptions which will be offered here treats the English definite article as just such a case, contending that the linguistic meaning of an utterance containing a definite description fails to determine a proposition. Instead, on this view, the definite article encodes procedural information which limits interpretation of the description to concepts with individual referents. Thus, the interpretation of the description – including the determination as to whether it is being used referentially or attributively – must be filled in during pragmatic processing.

This pragmatic processing, according to relevance theory, is guided by the search for optimal relevance, and the interpretation which will ultimately be assigned to an utterance is the optimally relevant one. On a relevance theoretic account of definite descriptions, then, a description will be given a referential reading in contexts in which that interpretation is optimally relevant and an attributive reading in contexts in which that interpretation is optimally relevant.

We will discuss relevance theory and the technical definitions of relevance and optimal relevance in much greater detail in chapter 5. Roughly, however, relevance is a ratio of cognitive benefits to cognitive costs. Unfortunately, as we will see, there is no way to quantify cognitive benefits or cognitive costs. As a result, it would seem that we as theorists would be unable to say that a particular interpretation of an utterance is the optimally relevant interpretation, as we would appear to be without the means of quantifying the relevance of any particular interpretation. If this is the case, however, then it will not be possible to test a relevance theoretic account of definite descriptions to determine whether or not it makes accurate predictions as to the natural interpretations of different descriptions in different contexts. Since relevance theory holds that the interpretation which will be selected in a context is the optimally relevant interpretation, if we are unable to show that a particular interpretation of a description is
optimally relevant, we will be unable to demonstrate that our relevance theoretic account of
descriptions predicts that a particular interpretation will be assigned to a description in a context.
This problem, which we will call the problem of testability, presents a serious difficulty which
any relevance theoretic account of definite descriptions would have to overcome. Chapters 7 and 8 are dedicated to attempting to address this problem. Before putting forward our own relevance theoretic account of descriptions in chapter 6, however, we will begin by looking at the strengths and weaknesses of previous contextualist accounts.
Chapter 2
The Neo-Strawsonian Account

Despite the fact that the discussion of definite descriptions has largely centered on the
debate between Russellians and referentialists, a number of alternative accounts have been put
forward. While not all of these accounts attempt to make sense of the referential/attributive
distinction, those that do share the claim that the distinction is an artifact of some level of context
dependence in the semantics of definite descriptions. And even those theories which do not pay
particular attention to the referential and attributive uses of descriptions hold that the
interpretation of utterances containing definite descriptions involves significant pragmatic
influences. Though, as we shall see, each of these accounts suffers from significant problems,
understanding the successes and failures of previous attempts to characterize definite
descriptions as context-dependent should prove useful in attempting to complete our overall goal
of constructing a viable relevance theoretic account of definite descriptions. With that in mind,
the next few chapters will examine several alternative accounts of definite descriptions, their
insights, and their difficulties. Chapter 3 examines the two-dimensional and synecdoche
accounts of definite descriptions, while chapter 4 discusses two differing attempts to apply
relevance theory to descriptions. This chapter, however, will assess the neo-Strawsonian account.

2.1 Strawson

Before Donellan had even made us aware of the distinct uses of definite descriptions, P.F.
Strawson leveled a number of objections against the Russellian picture of the semantics of
definite descriptions, and proposed a radically different view. The resulting debate between Strawson and Russell was not limited to the question of how to properly characterize definite descriptions, however; it was extensively entangled with broader theoretical issues about the overall understanding of reference, semantics and communication. As a result, many felt that once the overarching theoretical issues were addressed, and in particular, a distinction drawn between sentence meaning and speaker meaning (between semantics and pragmatics), the problems that Strawson found in the Russellian picture dissolved and the intuitions that motivated Strawson’s own account could be accounted for.

Murali Ramachandran has argued, however, that even with an understanding of the difference between semantics and pragmatics in place, Strawson’s objections to Russell still have a good deal of bite, and that a modified version of his theory is the best way to make sense of definite descriptions in light of these objections. As we shall see, Ramachandran’s proposed replacement of the Russelian characterization of definite descriptions suffers from a number of its own difficulties. Before we can properly discuss Ramachandran’s theory, however, we must first have a clear understanding of Strawson’s original proposal.

In “On Referring”, Strawson claims that Russell’s characterization of the semantics of definite descriptions leads to counterintuitive assessments of the truth values of utterances containing empty descriptions – that is, definite descriptions that are not satisfied by any object. To see this consider Strawson’s now famous example:

(2.1) The king of France is wise.

---

9 Ramachandran (2008)
10 Strawson (1950)
Is (2.1) true or false? Uttered during the reign of Louis XIV, (2.1) would have (presumably) been viewed as true. Uttered during the reign of Louis II, on the other hand, it would probably have been seen as false. This, of course, is no problem for the Russellian, since different individuals will be denoted by the description at different times, and hence (2.1) will express different propositions on different occasions of use. But what if (2.1) were uttered at some point during the 20th century? At no point during the 20th century has there been an individual who satisfies the description in (2.1). As a result, many have the intuition that such an utterance would be neither true nor false, that we are not in a position to assign a truth value to an utterance of (2.1) made in the 20th century. But this intuition conflicts with the Russellian’s analysis. If Russell and the Russellians are right, then the semantic content of any utterance of (2.1) is there is exactly one king of France and he is wise. Since at any point during the 20th century there was no king of France, then, a 20th century utterance of (2.1) would clearly be false.

It might seem, then, that Russell’s theory of definite descriptions is incorrect, and we ought to look for an alternative means of accounting for the semantics of definite descriptions. According to Strawson, Russell’s mistake lies in his failure to draw the distinction between an expression and the use of that expression, or the parallel distinction between a sentence and a use of that sentence. Once we draw these distinctions, Strawson argues, it becomes clear that the mechanism of reference is not the expression itself, but rather the use of that expression. While the sentence ‘The king of France is wise’ may have a particular semantic content, it is only partial. And hence the sentence on its own does not express a proposition. Only utterances – that is, sentences in context – can express propositions. And hence only utterances are true or false, and different utterances of the same sentence may have different truth values. The reason that an
utterance of (2.1) during the reign of Louis XIV would be true, while an utterance of the same sentence during the reign of Louis II would be false, then, is the fact that the expression ‘the king of France’ is being used in the two cases to refer to different individuals, and hence, the different utterances express different propositions.\(^{11}\) And since France had no king during the 20th century, the expression ‘the king of France’ could not properly be used to refer to anything during that century. As a result, a 20th century utterance of (2.1) fails to express a proposition or expresses a proposition which is neither true nor false.\(^{12}\)

Strawson’s evaluation of the phenomena of reference alone is not sufficient to make sense of definite descriptions, however. After all, part of what motivated the Russellian analysis was the fact that when we utter a sentence like (2.1), we do seem to be committing ourselves to the existence of the object identified by the description used. When I claimed at the beginning of this chapter that the third chapter of this work would examine the two-dimensional account of definite descriptions, that gave you every right to believe that a third chapter exists, and that there is, in fact, a two-dimensional account of definite descriptions. But if, as Strawson suggests, the actual device of reference is the use of an expression, there seems little reason to think that an utterance containing a definite description would justify the expectation that something satisfying the description actually exists.

\(^{11}\) Strawson himself does not explicitly make this claim. But if we assume, with Strawson, that the definite description is being used to refer to different individuals, it follows that the associated utterances express different propositions. One utterance expresses the proposition that Louis XIV is wise, the other than Louis II is wise.

\(^{12}\) Strawson (1950), p. 331
To account for this, Strawson suggests that utterances containing definite descriptions presuppose the existence of the object identified by the description.\(^{13}\) A presupposition is, essentially, a condition on the truth aptness of an utterance of a sentence. If Strawson is right, (2.1) presupposes the existence of a king of France. If there is a king of France, then an utterance of (2.1) is truth apt. It is true if the king of France is wise and false if he is not. If there is no king of France, however, then the presupposition is not satisfied, and (2.1) is not truth apt. Note, however, that presupposition is a pragmatic phenomenon. While the satisfaction of a presupposition associated with a sentence is essential if an utterance of that sentence is to express a true or false proposition, the presupposition itself is not part of what is asserted by an utterance of that sentence; it is not part of the proposition expressed by the utterance.\(^{14}\) Thus, because (2.1) presupposes the existence of the king of France, an utterance of (2.1) licenses the hearer to expect that there is indeed a king of France, but because that presupposition is not part of what is asserted by the utterance, the lack of a king of France does not result in a false statement, as it would if Russell’s semantic analysis were correct.

Strawson’s analysis is not without its problems, however. First, despite the fact that the semantics of definite descriptions proposed by Russell yields a counterintuitive assessment of the truth value of a 20\(^{th}\) century utterance of (2.1), the Russellian would have a prima facie reason for maintaining his analysis. For while, intuitively, a 20\(^{th}\) century utterance of (2.1) lacks a truth value (or, at the very least, we are disinclined to call such an utterance an outright falsehood),

\(^{13}\) It is still an open question whether it is utterances or speakers that presuppose. In what follows, I may vacillate between attributing presuppositions to utterances and attributing them to speakers. This imprecision should have no bearing on my arguments, however, since the feature of presupposition that is crucial for our purposes is the failure of an utterance to express a true or false proposition when its associated presupposition fails.

\(^{14}\) Strawson equates asserted content and proposition expressed. See Strawson (1950) p. 326
Strawson leaves us with no explanation as to how a perfectly meaningful utterance of an indicative statement (for so he would characterize a 20th century utterance of (2.1)\textsuperscript{15}) can be without a truth value. To put it another way, Strawson has separated meaning from truth aptness. But, intuitively, these two things are closely connected. Given the various motivations for accepting Russell’s semantics discussed in chapter one, then, it might be argued that rather than abandoning Russell’s analysis, we should instead retain it and look for another explanation of our intuitions with respect to utterances of sentences containing empty descriptions. And in fact, Strawson’s distinction between the sentence and a use of that sentence provides the Russellian with theoretical machinery that could easily be used to provide just the sort of explanation he needs.\textsuperscript{16} Indeed, the Russellian already makes use of this same distinction in explaining referential uses of descriptions when he claims that referentially used descriptions result in implicated singular propositions. It might very well be, then, that our hesitance to assign a truth-value to a 20th century utterance of (2.1) is similarly the result of factors surrounding the use of the utterance.

For the sake of simplifying the discussion, let us assume a Gricean or Neo-Gricean theory of conversational implicature. Now, if the Russellian is correct and, regardless of the circumstances of use, (2.1) semantically expresses the proposition that there is exactly one king of France who is wise, a 20th century utterance of (2.1) would be in violation of the maxim of

\textsuperscript{15} Strawson (1950) p. 330-331
\textsuperscript{16} I have deliberately left the contours of this distinction vague. Strawson’s distinction is, in essence, the distinction between semantics and pragmatics. While there is wide agreement about the existence of such a distinction, the proper characterization of the two domains and the boundary between them has been the focus of protracted debate. The bare acknowledgement of the distinction, however, should be sufficient for any Russellian to explain our intuitions in the case of empty descriptions.
quantity unless it were also being used to express something further.\footnote{The quantity maxim at issue need not be Grice’s quantity maxim. Neo-Gricean versions of the maxim would seem equally well suited to the task. See, e.g. Horn (1972), (1989); Gazdar (1979); Levinson (2000).} That is, to use Grice’s terms, in order to view a 20th century utterance of (2.1) as in line with the cooperative principle, what was said by the utterance could not be what was meant. However, because the context in which our hypothetical utterance is embedded is so scanty (we know only that the utterance is made at some point during the 20th century), we, as interpreters of the utterance, lack a sufficient basis for generating hypotheses about what was meant. We have no foundation from which we might calculate an implicature. So, essentially, we are left with a situation in which we have good reason to think that the proposition expressed by the use of (2.1) is not the same as that expressed by (2.1) by itself, but we lack the means of discerning the content of the proposition expressed by the use of (2.1) in question. To put it in Gricean terms, we know what was said by the 20th century utterance of (2.1) and we know that something different was meant, but we don’t have any clue as to what that could be. If we don’t know what is being expressed by the utterance, though, then it is no surprise that we are unwilling to assign it a truth-value.

It would seem, then, that the Russellian can make sense of Strawson’s problematic empty descriptions without abandoning Russell’s original semantic analysis. Since the Russellian can also account for something that Strawson cannot – our intuition that utterances containing empty descriptions are not semantically defective – Strawson’s theory lacks the explanatory scope of its competitor.

This is not the only difficulty facing Strawson’s proposal, however. Consider modern day utterances of (2.2) – (2.6)

(2.2) The king of France does not exist.
If Strawson is right in his claim that definite descriptions trigger a presupposition of the existence of the object described, then each of the foregoing utterances should carry with it such a presupposition. And, if the failure of a presupposition associated with an utterance prevents that utterance from expressing a proposition, then we should be hesitant to assign truth values to (2.2)-(2.6) in the same way we are hesitant to assign a truth value to (2.1). But this is not the case. Intuitively, modern day utterances of (2.2)-(2.4) are true, while modern day utterances of (2.5) and (2.6) would be false. These utterances, then, would appear to be counterexamples to Strawson’s proposed account of definite descriptions.

Of course, it may be (and in fact has been) argued that (2.2), (2.3) and (2.5) do not presuppose the existence of the king of France. If this is in fact the case, however, the Strawsonian is still responsible for providing an account of why the presupposition is present in one case, but not another. This is an instance of what is known as the projection problem for presupposition – the problem of determining when a complex sentence will inherit the presuppositions of its simple components and explaining why presuppositions are inherited in one case but not another. Thus, the Strawsonian might put forward any of a number of proffered accounts of the projection problem in order to explain our willingness to assign truth-values to

---

18 These examples are taken from Ramachandran (2008). (2.6) is somewhat contentious. Though many have the intuition that an utterance of (2.6) is false, since the king of France is not Carlos Slim, Bill Gates or Warren Buffet, others have the intuition that an utterance of (2.6) is just as truth valueless as a 20th century utterance of (2.1). Still, those intuitions running counter to Strawson’s view, even if mistaken, deserve an explanation.
(2.2), (2.3) and (2.5).\textsuperscript{19} Of course, this still leaves the Strawsonian with the difficulty of demonstrating that the solution they give is a viable solution. And even if the projection problem actually admits of a solution, this still leaves (2.4) and (2.6) to be explained.\textsuperscript{20}

In light of the various difficulties facing Strawson’s analysis, along with the fact that it seems possible for a slightly modified Russellian account to deal with Strawson’s main worry – utterances involving empty descriptions – Strawson’s account has largely fallen out of favor over the years. Recently, however, Murali Ramachandran has attempted to revive Strawson’s position by offering a neo-Strawsonian account of definite descriptions.\textsuperscript{21} It is to this revised version of Strawson’s theory that we now turn.

### 2.2 The Neo-Strawsonian Account

Ramachandran’s revival of Strawson’s account begins with a recognition of the reasons why the theory has been given so little credence. While he finds the Russellian’s attempt to account for our intuitions with respect to utterances containing empty descriptions to be unconvincing, Ramachandran concedes that utterances containing empty descriptions provide only questionable motivation for accepting Strawson’s thesis. And since Strawson relies on such utterances as the sole support in favor of the theory, we have been left with very little reason to take it seriously. The situation is made even worse when we take into consideration the significant threat to Strawson’s view posed by apparent counterexamples like (2.4) and (2.6). In

\textsuperscript{19} For discussion of the projection problem see Karttunen (1973), (1974); Soames (1982); Heim (1990).
\textsuperscript{20} Despite the fact that (2.4) would appear to have a negation in it, it does not fall in the class of examples which might be attributed to the projection problem. First, it is not entirely clear that there really is a negation, as we would normally think of it, in this sentence. And second, even if there is a negation, the syntax of the sentence places the presupposition trigger in this case – the description ‘the king of France’ – outside its scope.
\textsuperscript{21} Ramachandran (2008)
light of all this, Ramachandran seeks to construct a modified version of Strawson’s account of
definite descriptions, one which avoids the problems plaguing the original theory and which
enjoys independent support beyond that of intuitions about utterances containing empty
descriptions.

According to Ramachandran, we can find much better support for a Strawsonian view by
shifting our attention away from empty descriptions and towards incomplete descriptions. Recall
that an incomplete description is a description that does not identify a unique object. ‘The current
President of the United States of America’ is a complete description – it identifies one and only
one object. The description ‘the cat’, on the other hand, is incomplete; it could be applied to any
one of millions of past or present felines. The Russellian accounts for their incompleteness by
appealing to the context, either as the source of the complete, but elided, semantic content of the
description, or as the means of restricting the domain of quantification. It is not the issue of how
to addresses the incompleteness of such descriptions that is of interest to Ramachandran,
however, but rather the unintelligibility of utterances containing such descriptions in certain
circumstances.

Ramachandran asks us to consider a room containing many different tables. Some of
these tables are covered with books, while others are not, and no table is more salient than any
other. Call this the many-table scenario. Now, imagine someone uttering (T) in the many-table
scenario.

(T) The table is covered with books.

Now, our first thought upon hearing such an utterance would be to ask the speaker which table
he was talking about. But suppose he responds that he had no particular table in mind. There is
no doubt that we would view the speaker as most odd for producing such an utterance. But is the utterance true or false? Intuitively, no truth value can be assigned in this case. Indeed, it is not clear what we should take such an utterance to even mean. In Ramachandran’s words, “we would be at a loss as to [(T)’s] particular truth conditions”. \(^{22}\) An utterance of (T) in the many-table scenario is unintelligible.

Now imagine an utterance of (T*) in the many-table scenario.

\(\text{(T*)} \) There is exactly one table and it is covered with books.

Is this utterance true or false? Intuitively, it is false. There is more than one table. Moreover, there is no question as to the truth conditions of the particular utterance of (T*) – it is true if there is exactly one table in the relevant context which is covered with books. Unlike an utterance of (T) in the many-table scenario, an utterance of (T*) would be completely intelligible. Given that the Russellian holds that (T*) is a translation of (T), that (T*) and (T) express the same proposition, he owes us an explanation of both the unintelligibility of a many-table utterance of (T) and the disparity between the intelligibility of utterances of (T) and (T*). Given that the Russellian holds that (T*) characterizes the content of (T), it seems unlikely these phenomena admit of an explanation which is consistent with the Russellian account of the semantics of definite descriptions. This is especially true given that the pragmatic factors surrounding the two utterances are nearly identical. As a result, if the Russellian were to appeal to pragmatic considerations to explain the unintelligibility of (T), those same pragmatic considerations would likely be operative in the case of (T*), meaning that (T*) should also be unintelligible.

\(^{22}\) Ramachandran (2008)
If we reject the Russellian identification of the content of \((T)\) and \((T^*)\), however, Ramachandran suggests that we may be able to explain the many table scenario in a way which, while differing from Strawson’s original view, is still very much in line with it. Instead working with presuppositions, as Strawson does, we could hold that the existence of a unique table is a \textit{precondition} on the expression of a proposition in the case of \((T)\). Since the precondition in not met, an utterance of \((T)\) in the many-table scenario fails to express a proposition, which is why it is neither intelligible nor truth-evaluable. \((T^*)\), on the other hand, succeeds in expressing a proposition, and is hence intelligible and truth-evaluable.

There is, however, one way that the Russellian might make sense of the difference in intelligibility between utterances of \((T)\) and \((T^*)\) in a many-table scenario. The idea, originally put forward by Ganeri,\textsuperscript{23} is that both \((T)\) and \((T^*)\), as uttered in the many-table scenario, fail to express propositions. According to this line of thought, the context fails to determine a domain of quantification for either \((T)\) or \((T^*)\), and without such a determination, neither utterance is capable of expressing a proposition. The reason we find the utterance of \((T^*)\) intelligible is that the use of the quantificational expression leads us to arbitrarily select a restricted domain of quantification. But because the context itself does not determine a domain, a complete proposition is not actually expressed.

This strategy faces a number of different difficulties, not the least of which is the plausibility of the claim that we arbitrarily select a restricted domain of quantification in the absence of contextual determination in the case of \((T^*)\), but not the in case of \((T)\). Even if we set

\textsuperscript{23} Ganeri (1995)
such worries aside, however, Ramachandran argues that the Russellian who adopts this strategy will quickly end up committed to a Strawsonian picture of descriptions.

According to Ramachandran, the Russellian who attempts to explain the unintelligibility of a many-table utterance of (T) by appealing to the absence of a contextually determined domain of quantification is committed to two principles, the Precondition Principle (PP) and the Content Principle (CP)

(PP) It is a precondition of expressing a proposition with a description sentence The $F$ is $G$ where ‘the $F$’ is incomplete that there be a contextually determined predicate or domain, $C$.

(CP) When a proposition is expressed by an utterance of such a sentence, it is the proposition: there is exactly one $F$ that is [in] $C$, and it is $G$.\textsuperscript{24}

(CP) is the principle that makes this a Russellian account, while (PP) is the basis for the Russellian’s explanation of (T) and (T*). But, as Ramachandran points out, (PP) is insufficient.\textsuperscript{25} A contextually determined domain of quantification is not enough to guarantee that a proposition is expressed by an utterance containing an incomplete definite description. Consider, again, the many-table scenario, and suppose that (2.7) is uttered in this scenario.

(2.7) The table in this room is covered with books.

(2.7) explicitly indicates a domain of quantification, and yet it would appear to be unintelligible in much the same way as the many-table utterance of (T). So what is the difficulty? The problem appears to be that there is more than one table in the domain in question. It would seem, then, that in order for a proposition to be expressed by an utterance containing an incomplete description, there must be at most one object which satisfies the description in the domain in

\textsuperscript{24} Ramachandran (2008). Emphasis in the original
\textsuperscript{25} The examples that follow are from Ramachandran (2008).
question. Even with this modification, however, the principle is still faulty. For suppose that the speaker points to an empty corner and utters (T).

(T) The table is covered with books.

This utterance, too, seems unintelligible. But why? The reason, presumably, is that the description is empty in this case. There is no table in the contextually determined domain (the empty corner). According to Ramachandran, then, we must add to (PP) the condition that there be at least one object satisfying the description in the contextually determined domain. With these modifications, (PP) becomes (PP*).

(PP*) It is a precondition of expressing a proposition with a description sentence \( \text{The } F \text{ is } G \) where ‘the F’ is incomplete that there be a contextually determined predicate or domain, C, such that there is exactly one \( F \) which is \([\text{in}]\) C.

Recall, however, that the Russellian holds that the content of a description sentence of the form \( \text{The } F \text{ is } G \) is \textit{there is exactly one }F\textit{ in }C\textit{ and it is }G.\ But why should this be both a precondition for the expression of a proposition with a description sentence \textit{and} part of the content of the proposition expressed? Why should this content be part of the proposition expressed by a description sentence when it is already a precondition for the very expression of the proposition in the first place? Since (PP*) is necessary for the explanation of the unintelligibility of the many-table utterance of (T) and utterances like it, Ramachandran concludes that (CP) should be abandoned. Once we do that, however, the account is no long Russellian, but instead Strawsonian.

---

We have thus reached the core of Ramachandran’s account. Like Strawson, Ramachandran holds that descriptions are devices for referring. But his account deviates from Strawson’s in that, rather than maintaining that utterances containing definite descriptions presuppose the existence of an object (within the contextually identified domain) which uniquely satisfies the description, he instead holds that the existence of such an object is a precondition which must be met in order for such utterances to express propositions. And such preconditions are not part of what is asserted by an utterance containing a description.

According to Ramachandran, however, (PP*) is only half of the story. So far, we have only been looking at utterances in which the description is being used to pick out an individual. Ramachandran calls this the referring use of definite descriptions. But in some cases, descriptions are used to specify a property that is being predicated of an individual.

Ramachandran calls such descriptions predicative. The distinction can readily be seen in (2.8) and (2.9).

(2.8) The greatest playwright of all time lived in London.
(2.9) Shakespeare was the greatest playwright of all time.

The description ‘the greatest playwright of all time’ is being used to pick out an individual – Shakespeare – in (2.8). This is a referring use. In (2.9), however, the description is being used to indicate a property which the sentence ascribes to Shakespeare. This is a predicative use. It is

---

27 As we shall see shortly, Ramachandran does not hold that all definite descriptions are devices for referring. The main point here, however, is that Ramachandran rejects the Russellian picture of definite descriptions as quantificational expressions.
28 ‘Asserted content’, for Ramachandran, appears to be the same as ‘semantic content’ or ‘the proposition expressed’.
29 This is not to say that (2.9) could not be interpreted as indicating an identity between Shakespeare and the greatest playwright of all time. Ramachandran admits that it can be given such a reading. What is important here is that it can also be heard as attributing a property to Shakespeare.
important to note that this distinction is separate from that between referentially and attributively used descriptions. According to Ramachandran, the referential use and the attributive use are both cases of referring use.

Now, the preconditions for the predicative use are different from those for the referring use. (PPR) gives the preconditions for the referring use (which we have already seen in (PP*)). The preconditions for the predicative use are described in (PPP).

(PPR) An utterance involving the referring use of a description ‘the F’ will express a proposition if and only if there is a contextually determined domain that contains exactly one F.

(PPP) An utterance involving the predicative use of a description ‘the F’ will express a proposition if and only if the following conditional holds: *Had things been slightly different, there might have been an F (in the circumstances), but no more than one.*

(PPR) and (PPP) described the conditions that must be met for utterances containing the referring and predicative uses of definite descriptions to successfully express propositions. But what propositions would such utterances express? Ramachandran is rather vague in his answer to this question. He explicitly states that the preconditions for expressing a proposition are *not* part of what is asserted by an utterance containing a definite description. And, as we noted above, he states that definite descriptions – or rather, referring uses of definite descriptions - are referring expressions, but it is not entirely clear what he means by ‘referring expression’. One might be inclined to assume that in calling definite descriptions ‘referring expressions’, Ramachandran is treating them as directly referential, that is, as contributing only their referents to the proposition expressed. But this cannot be what he means, as he classifies both referentially and attributively used descriptions as referring uses, and consequently claims that

\[\text{Ramachandran (2008)}\]
Obviously, given this take on referring, I must deny some familiar theses – e.g. (a) that one must already be acquainted (in some sense) with an object in order to refer to it; and also (b) that one cannot understand an utterance involving reference without knowing which object is being referred to.31

He further suggests that an attributively used description refers to an individual “by way of the descriptive content”. It would seem, then, that Ramachandran does not view definite descriptions as directly referential expressions. Unfortunately, however, we are given little further clue as to what he takes to be the contribution to the proposition expressed made by a referring use of a definite description, leaving a significant gap in the overall theory.

Ramachandran attempts to use the theory to make sense of utterances like (2.2) – (2.6) which proved problematic for Strawson’s original account.

(2.2) The king of France does not exist.
(2.3) P.F. Strawson is not the king of France.
(2.4) No one had lunch with the king of France yesterday.
(2.5) Necessarily, the king of France is wise.
(2.6) The king of France is the world’s richest man

Recall that in order to make sense of them, the original Strawsonian account had to treat (2.2), (2.3) and (2.5) as instances of the projection problem for presupposition. This maneuver is not available to Ramachandran, however, since the neo-Strawsonian account has shifted from talk of presupposition to talk of preconditions. Instead, Ramachandran attempts to handle these problematic sentences by appealing to the predicative use of definite descriptions. (2.3), for example, is taken to express something like the proposition that P.F. Strawson does not possess the property of being the king of France. Since this is clearly true, it is no wonder that we intuitively view a modern day utterance of (2.3) as true.

The other problematic utterances are given a different treatment. Since the descriptions in these utterances are being used referringly, the neo-Strawsonian is committed to claiming that no proposition is expressed. The predicative condition, according to Ramachandran, is met in these cases - that is, if things had been slightly different, there might have been a king of France. As a result, Ramachandran claims, since no proposition is literally expressed, we hear the speaker as communicating a slightly different proposition. In the case of (2.2), for example, we understand the speaker to be communicating something like (2.10), and an utterance of (2.4) communicates something like (2.11).

(2.10) No one is the king of France.
(2.11) No one had lunch with someone who is the king of France.

This strategy has a rather unfortunate consequence, however. Recall Strawson’s original example utterance, (2.1).

(2.1) The king of France is wise.

We are intuitively disinclined to assign a truth value to a modern day utterance of (2.1). But if Ramachandran is correct that utterances containing empty referringly used descriptions communicate further propositions when the predicative condition is met, then (2.1), while failing to express a proposition, should communicate something like the proposition that someone is the king of France and wise. Since that proposition is false, we should expect to intuitively view (2.1) as false. But we do not. Ramachandran admits to being without any explanation of this inconsistency, but does not view this as a serious problem, since the Russellian is in much the
same situation with respect to empty descriptions like that in (2.1).32 Neither the Russellian nor the neo-Strawsonian has a satisfactory explanation of utterances containing such descriptions.

Of course, the claim that empty descriptions cannot be held against the neo-Strawsonian since the Russellian cannot handle them either only holds water if the neo-Strawsonian account and the Russellian view are the only theories available. They are not, however; hence, other things being equal, any view which could adequately deal with utterances containing empty descriptions would have an advantage over both the Russellian and the neo-Strawsonian.

2.3 Problems with the Neo-Strawsonian Account

We have already seen a number of problems with the neo-Strawsonian account in the previous section. We found that the neo-Strawsonian appears to have difficulty accounting for utterances containing empty descriptions. Our discussion also raised some question as to the viability of Ramachandran’s suggestion that the preconditions of an utterance are represented in the logical form of that utterance despite not being part of the semantic content. And, finally, we saw that the account leaves us with a largely incomplete picture of the semantic content of utterances containing definite descriptions. These are not the only problems facing the neo-Strawsonian, however. In what follows, we will examine a number of potential difficulties facing the account.

Before turning to new problems for the neo-Strawsonian, it might be helpful to get a more detailed understanding of the implications of the incomplete nature of the neo-Strawsonian account of the semantics of definite descriptions. As we have already noted, Ramachandran

32 Ramachandran (2008)
leaves the neo-Strawsonian view of the semantics of utterances containing definite descriptions rather vague and incomplete. This is in itself problematic, as any gap in a theory’s explanation of its topic is a weakness in that theory. But the failure to give a definite semantics for utterances containing definite descriptions leads to other problems as well. As we have seen, there are a number of different ways in which definite descriptions can be used. Of particular interest to many, including us, is the distinction between referentially and attributively used descriptions. Without a definite picture of the semantics of descriptions, however, it is unclear how the neo-Strawsonian’s account might handle this distinction.

We might make a few conjectures. Ramachandran does claim that both referentially and attributively used descriptions are instances of the referring use – that is, they are devices for referring. And attributively used descriptions, on his view, refer through descriptive content. Presumably, then, referentially used descriptions refer without the mediation of descriptive content. This would seem to indicate that he takes the distinction to bear on the semantics of description utterances – attributively used descriptions contribute descriptive content (along with referent) to the proposition expressed while referentially used descriptions do not contribute descriptive content. But, then, what explains this difference? What is meant by descriptive content? And what about the generic or functional uses of definite descriptions? We cannot even begin to provide answers to these questions without a more detailed picture of the neo-Strawsonian view of the semantic content of description utterances. Thus, in the absence of a more explicit account, the neo-Strawsonian theory faces a serious theoretical deficiency.

In addition to significantly limiting the explanatory breadth of the theory, the lack of a comprehensive account of the semantics of descriptions makes it difficult to evaluate the theory.
If we do not know what proposition the neo-Strawsonian takes to be expressed by a description utterance, we will be unable, in many cases, to test the predictions of the theory against our intuitions about the natural interpretations of utterances. This is a serious weakness for any theory.

Even if the neo-Strawsonian were to give a complete, explicit account of the semantics of description utterances, the theory may have other failings, particularly with respect to the predicative condition. Recall that the predicative condition is satisfied with respect to a description if the predicative conditional obtains.

**Predicative Conditional (PC):**

Had things been slightly different, there might have been a contextually relevant object satisfying the description, but no more than one.

It is not clear, however, just what it would mean for this conditional to be true. For instance, how different from the actual circumstances can things be before they no longer count as being only ‘slightly different’? Consider, for example, an utterance of (2.12).

(2.12) Billy is not the most eloquent unicorn politician.

On its most natural interpretation, (2.12) is true. According to Ramachandran, the reason that utterances like (2.12) express complete, truth evaluable propositions is that the predicative condition is satisfied. But it seems difficult to hold that the predicative condition is satisfied in this case. The circumstances in which there are not only unicorns, but unicorns who use language and have a political system, is not just *slightly* different from the actual world, it is *drastically* different. If on the other hand, we were to take ‘slightly different’ as it occurs in PC to mean something like “in the closest possible world”, then the condition would not be satisfied in the
case of (2.12), and we should expect the utterance to fail to express any proposition. It would seem, then, that the predicative condition is unable to handle utterances like (2.12).

The neo-Strawsonian might respond to this objection by claiming that the predicative condition is indeed satisfied for (2.12) and that the objection is based on a misinterpretation of PC. The objection to the predicative condition assumes (2.13) is the proper way of understanding PC with respect to (2.12).

\[(2.13)\text{ If things were slightly different, then there would be a most eloquent unicorn politician, but not more than one.}\]

But a close look at the consequent of PC reveals that (2.14) is, perhaps, the correct way of understanding PC as it pertains to (2.12).

\[(2.14)\text{ If things were slightly different, then it would be possible for there to be a most eloquent unicorn politician, but not more than one.}\]

Since (2.14) is true, the predicative condition is satisfied for (2.12), thus making it unsurprising that we intuitively view it as being truth evaluable.

Even if (2.14) is the proper interpretation, PC is still not capable of dealing with all predicative descriptions. Consider, for example, (2.15).

\[(2.15)\text{ Victor is not the owner of a round square.}\]

On its most natural interpretation, (2.15) is true. Thus, if the neo-Strawsonian is correct, the predicative condition should be satisfied. Assuming that PC should be interpreted along the line indicated by (2.14), for the condition to be satisfied with respect to (2.15), the conditional in (2.16) would have to obtain.

\[(2.16)\text{ If things were slightly different, then it would be possible for there to be an owner of a round square, but not more than one.}\]
It doesn’t matter how different circumstances are from those in the actual world, however, we will not encounter a world at which the existence of an owner of a round square is possible. (2.16) is false, and the predicative condition with respect to (2.15) is not satisfied. And it seems that any predicatively used description that describes a logically impossible property would pose the same problem for when PC is interpreted along the lines of (2.14). It would appear, then, that the neo-Strawsonian’s account of predicative descriptions is flawed.

As we have seen, the neo-Strawsonian account faces a number of significant difficulties. Ramachandran’s attempt to revive Strawson’s view has provided us with one important insight into our subject of inquiry, however. As Ramachandran argues, certain uses of definite descriptions are unintelligible. And these instances of unintelligibility will have to be explained by any theory of the semantics of definite descriptions. We will need to remember this as we evaluate other theories of descriptions and as we construct our own account.
Chapter 3  
Contextualist Accounts

At this point, one might think that we have exhausted all the possible accounts of the referential/attributive distinction. The distinction brings to light two possible uses of any definite description – the referential use and the attributive use, and the literature makes clear that there are two levels of meaning that might be appealed to in making sense of the differing uses of descriptions – the semantic level and the pragmatic level. The referentialist contends that both uses should be seen as occurring at the level of semantics. The Russellian claims that the attributive use is yielded by nothing more than the standard semantics of definite descriptions, while the referential use arises at the level of pragmatics. And the neo-Strawsonian claims that the referential use occurs at the semantic level, while the attributive use may be accounted for by appealing to the level of pragmatics. Since it seems implausible to hold that both uses are pragmatic phenomena, given that, intuitively, one or the other of these uses is truth conditionally relevant to any utterance in which a definite description occurs, the three accounts we currently have on offer would seem to cover all the plausible explanatory avenues available.

We are only limited to the explanations offered by the referentialist, the Russellian, and the neo-Strawsonian, however, if we accept the claim that there are only two levels of meaning to which we may appeal when accounting for communicative phenomena like the referential/attributive distinction. Certain accounts of the referential/attributive distinction and the semantics of definite descriptions reject this claim and assume instead that there are three levels of meaning to any utterance. Let us call such accounts contextualist.
The claim that a proper analysis of communicated content requires a more nuanced
division than the simple distinction between semantics and pragmatics is easily justified when
we examine utterances containing indexicals. Consider, for example, Bertrand’s utterance in the
following dialogue. Assume that the conversation is taking place on June 15th.

(3.1)  Alfred: I’m seeing a movie tonight. Do you want to come?
       Bertrand: I have a paper due tomorrow.

On the assumption that communicated content is either semantic or pragmatic, the natural
analysis of the content communicated by Bertrand’s utterance would be something like that
described in (3.2).

(3.2)  Semantics: Bertrand has a paper due on June 16th.
       Pragmatics: Bertrand cannot go to a movie with Alfred during the evening
                   of June 15th.

But while the description of the content of Bertrand’s utterance given by (3.2) is intuitively
accurate, it is clearly incomplete. There are important elements of the semantics of Bertrand’s
utterance that (3.2) fails to capture. For example, Bertrand did not use his name when uttering
(3.2); he used the linguistic expression ‘I’. But ‘I’ and ‘Bertrand’ are not synonyms. Nor are they
universally co-referential. The expression picks out Bertrand as it occurs in his utterance, but it
picks out Alfred as it occurs in Alfred’s utterance. Nonetheless, our intuitions tell us that the
expression ‘I’ means the same thing in both Alfred’s utterance and Bertrand’s utterance, at least
in some sense. What this shows is that the division of communicated content into semantics and
pragmatics is not sufficient for completely characterizing communicative phenomena. Rather,
there must be a third level of meaning at which the content associated with ‘I’ in Bertrand’s
utterance is represented as being the same as the content associated with the expression as it occurs in Alfred’s utterance.

Of course, when we say that there is a third level of meaning, this does not rule out the possibility that this level could be a pragmatic level of meaning of a different sort from the standard understanding of pragmatics as involving just implicatures or a semantic level of a different sort from our standard understanding of semantics. In fact, if we take semantic content to be encoded content and pragmatic content to be inferred content, then this third level of meaning would appear to be a bit of both, since it involves both types of content.

On the basis of the assumption that the content associated with an utterance is properly analyzed in terms of three levels of meaning, Francois Recanati has offered two different contextualist accounts of the semantics of definite descriptions and, in particular, the referential/attributive distinction. We will examine each of these accounts in turn.

3.1 The Two-Dimensional Account

Recanati’s initial approach to the referential/attributive distinction accounts for it in terms of a two-dimensional approach to semantics.33 Following Kaplan and Stalnaker34, Recanati views the meaning of a sentence $S$ as at least partially consisting in the conditions under which any utterance of $S$ would express a true proposition. It is important to note that this view does not equate sentence meaning with truth conditions. The conditions in question are not those under which $S$ (or, more precisely, a particular utterance of $S$) is true; they are the conditions that must

---

33 Recanati (1989)
34 Kaplan (1977), (1979); Stalnaker (1970), (1978)
be met for any utterance of $S$ to express a true proposition. The difference may be clarified by an example. Consider the sentence (3.3).

\begin{equation}
(3.3) \quad \text{Willard arrives in Santa Fe tomorrow.}
\end{equation}

What are the truth conditions for (3.3)? That is, what conditions must be met for (3.3) to be true? Independent of a context of utterance, this question appears to be unanswerable, and its answer will vary across different contexts of utterance. If (3.3) is uttered on November 10$^{th}$, 2020, then it is true if and only if Willard arrives in Santa Fe on November 11$^{th}$, 2020. If it is uttered on February 4$^{th}$, 1954, however, then it is true if and only if Willard arrives in Santa Fe on February 5$^{th}$, 1954. Truth-conditions, it seems, are by and large specifiable only for utterances, not for sentences independent of context.

What about the conditions under which any utterance of (3.3) expresses a true proposition? Can we specify these conditions independent of any context of utterance? Consider (3.4).

\begin{equation}
(3.4) \quad \text{An utterance of (3.3), } u, \text{ expresses a true proposition if and only if Willard arrives in Santa Fe on the day after } u \text{ is uttered.}
\end{equation}

The right-hand side of the biconditional in (3.4) gives a rough specification of the conditions under which an utterance of (3.3) expresses a true proposition. According to Recanati, these conditions are essentially what is specified by what he calls the diagonal proposition. The diagonal proposition associated with an utterance $u$ is such that for any world $w$, the proposition is true with respect to $w$ if and only if $u$ as uttered in $w$ would express a true proposition in $w$.\(^{35}\)

Now, suppose (3.3) is uttered by Donald on February 4$^{th}$, 1954. Clearly, the diagonal proposition associated with (3.3) is not what Donald said – that is, it is not what Donald explicitly expressed.

\(^{35}\) Recanati (1989), p. 235
Donald said that Willard arrives in Santa Fe on February 5th, 1954. The diagonal proposition does not give us what Donald’s utterance literally meant, so it clearly does not exhaust the semantic content of the utterance. What it does give us, according to Recanati, is the linguistic meaning of Donald’s utterance. Thus, there are three levels of meaning associated with any particular utterance: (1) The linguistic meaning, which is given by the diagonal proposition associated with the sentence uttered. This is the level at which lexical ambiguity occurs. In the case of Donald’s utterance, this is the proposition that Willard arrives in Santa Fe on the day following the utterance. (2) What is said, that is, the proposition explicitly expressed by the utterance. In the case of Donald’s utterance, this is the proposition that Willard arrives in Santa Fe on February 5th, 1954. (3) What is meant, that is, any implicatures, presuppositions or speech acts associated with the utterance. In terms of the division between semantics and pragmatics, we might roughly characterize levels (1) and (2) as the semantics of the utterance and level (3) as the pragmatics. I call the characterization rough, however, because there is an important difference between levels (1) and (2) that would lead some to view level (2) as capturing parts of both the semantics and the pragmatics of an utterance. The linguistic meaning of an utterance – the first level of meaning - is, on this view, contextually invariant. It is the meaning of the sentence that is uttered, independent of any context of utterance. By contrast, the second level of meaning – what is said – is sensitive to context. It is the meaning of the sentence in context. In fact, the distinction between linguistic meaning and what is said bears remarkable similarity to the distinction between sentence meaning and utterance meaning.

Using this distinction between linguistic meaning and what is said, Recanati suggests a new means of explaining the referential/attributive distinction. The Russelian, Recanati claims,
is correct to hold that Russell’s account of definite descriptions properly describes their semantics. But since the Russellian fails to draw the distinction between linguistic meaning and what is said, he incorrectly assumes that the Russellian analysis applies to the whole of the semantics of any utterance containing a definite description. According to Recanati, what Russell’s analysis gets right is the linguistic meaning of utterances containing definite descriptions. In other words, the diagonal proposition associated with an utterance of the form \( \text{The } F \text{ is } G \) is a general proposition – that is, a proposition of the form \( \text{there exists an } x \text{ such that } x \text{ is uniquely } F \text{ and } x \text{ is } G \). At the second level of meaning, however, what is said by an utterance of the form \( \text{The } F \text{ is } G \) could be either a general proposition or a singular proposition.\(^{36}\) In other words, the linguistic meaning of a description is always attributive, but at the level of what is said, the description could be interpreted as either attributive or referential.

The obvious question at this point is why would both the attributive and referential readings of a description be available at the level of what is said when only the attributive interpretation is available at the level of linguistic meaning? How did this extra potential interpretation creep in between the meaning of the sentence and the meaning of the utterance? According to Recanati, there are two interpretations of an utterance of the form \( \text{The } F \text{ is } G \) at the level of what is said because there are two ways of interpreting the conditions specified in the diagonal proposition at the level of linguistic meaning.

To see this, consider the sentence in (3.5) and its associated diagonal proposition given in (3.6).

\(^{36}\) Recanati does not give an explicit definition of a singular proposition. Rather, he simply characterizes singular propositions as “the type of proposition expressed by means of sentences with a proper name or a demonstrative expression as subject-term.” See Recanati (1989), p. 217.
(3.5) I am a philosopher
(3.6) There is an $x$ such that $x$ utters (3.5) and $x$ is a philosopher.

On the view we are currently considering, the diagonal proposition given by (3.6) is a specification of the conditions that must be met for (3.5) to express a true proposition. We can distinguish two different conditions given by (3.6). The first condition is that there be an $x$ such that $x$ utters (3.5). The second condition is that $x$ is a philosopher. Recanati calls these, respectively, the *identifying condition* and the *predicative condition*. These two conditions are of fundamentally different types, however. The predicative condition is a truth condition. An utterance of (3.5) by me, Diana Puglisi, is true if and only if Diana Puglisi is a philosopher. An utterance of (3.5) by Greta Garbo is true if and only if she is a philosopher. The identifying condition – the condition that there be an $x$ that utters (3.5) – is not a truth condition, however. Consider a world $\alpha$ where I am a philosopher, but I refuse to call myself one. If I *were* to utter (3.5), the proposition expressed would be true in $\alpha$, despite the fact that I did not utter (3.5) in $\alpha$. Rather than being a condition on the truth of an utterance of (3.5), the identifying condition is a condition on the expression of a proposition by (3.5). The sentence ‘I am a philosopher’, by itself, does not express a proposition.\(^{37}\) In order for the sentence to express a proposition, there must be some individual who utters it. Recanati calls conditions that must be met in order for a sentence to express a proposition *contextual conditions*.\(^{38}\)

\(^{37}\) One might be inclined to object on the following grounds: Truth-conditions belong to utterances – they are the conditions under which an utterance is true. Utterances are only truth-apt (that is, only capable of being true or false) if they express definite propositions. Since the condition that there be an $x$ that utters (3.5) must be met for (3.5) to express a definite proposition, it is a condition under which an utterance of (3.5) is true, and is hence a truth-condition. Such an objection need not worry us here, however. For even if we accept the objector’s conclusion – that both conditions are truth conditions – his own argument highlights the difference between the condition that there is an $x$ that utters (3.5) and the condition that $x$ is a philosopher. The latter bears directly on the truth of the utterance, while the former bears on it only indirectly. This difference is all that is really needed.

\(^{38}\) Recanati (1989), p. 237
While it is evident how each of the conditions specified in (3.6) should be understood, things are not so clear cut when it comes to diagonal propositions associated with utterances of the form *The F is G*. Consider, for example, an utterance of (3.7) and its associated diagonal proposition, given in (3.8).

(3.7) The man in Armani is handsome.
(3.8) There exists an $x$ such that $x$ is uniquely a man wearing Armani and $x$ is handsome.\(^{39}\)

(3.8), again, specifies two conditions: (1) the identifying condition: there is an $x$ such that $x$ is uniquely wearing Armani, and (2) the predicative condition: $x$ is handsome. Clearly, the condition that $x$ is handsome is a truth-condition. Any utterance of (3.7) will be true if and only if the individual indicated is handsome. But what about the identifying condition? It looks like the identifying condition could safely be interpreted as either a contextual condition or a truth condition.

Suppose I utter (3.7) at a party while nodding toward Jeffery Donovan. We already know that my utterance is true if and only if Donovan is handsome. But what about his being a man wearing Armani? If he is wearing Armani, does that fact make it possible for my utterance to express a definite proposition which is made true by Donovan’s good looks? Or does my utterance express a definite proposition that is made true both by Jeffrey Donovan’s good looks and by his wearing Armani? There are no intuitively correct answers to these questions. And the protracted arguments over cases of misdescription indicate that we’ll get no productive intuition pumps out of assuming that the condition in question is unsatisfied. If we assume that Donovan is not wearing Armani when I nod toward him and utter (3.7), some will have the intuition that I

\(^{39}\) The uniqueness condition here should be taken as relative to a contextually restricted domain.
have failed to express a proposition, some will have the intuition that I have expressed a false proposition, and still others will have the intuition that I have expressed a true proposition, albeit with a faulty indicator of the subject of my utterance.

If the identifying condition in (3.8) may be interpreted as either a contextual condition or a truth condition, then, according to Recanati, there are two ways in which the diagonal proposition given by (3.8) might interact with the context of utterance, yielding two possible interpretations of what is said by an utterance of (3.7). If the identifying condition is a truth condition then (3.8) expresses a general proposition – one which is true if and only if there is an $x$ such that $x$ is a man wearing Armani and $x$ is handsome. If, on the other hand, the identifying condition is a contextual condition, then (3.7) expresses a singular proposition – one which is true if and only if Jeffrey Donovan is handsome (given that Donovan is the only man in the context wearing Armani). Interpreting the identifying condition in the diagonal proposition as a truth condition yields an attributive reading of the description in (3.7); interpreting it as a contextual condition yields a referential reading. And according to Recanati, which interpretation of the diagonal proposition is preferable will depend on the context of utterance.

This two-dimensionalist account of the referential/attributive distinction has the great advantage of being capable of accounting for the intuitions behind both the Russellian and referentialist viewpoints. The view delivers a univocal Russellian semantics for definite descriptions (at the level of linguistic meaning) while simultaneously making sense of the intuition that the referential/attributive distinction is semantically significant (at the level of what is said). Since it does not characterize the distinction as being pragmatic, the view does not carry with it the burden of explaining how one of the two uses of descriptions is conveyed at the level
of pragmatics. The two-dimensional contextualist need not justify claims about presuppositions or demonstrate the presence of implicatures. Moreover, since the view does not posit a semantic ambiguity, it cannot be charged with multiplying senses beyond necessity, and hence can escape Grice’s razor.

### 3.2 Problems with the Two-Dimensional Account

Despite its benefits, the two-dimensional view is not beyond criticism. First, it relies heavily on the notion of a diagonal proposition, which may result in problems. The diagonal proposition associated with an utterance $u$, according to Recanati, is that proposition such that, for any world $w$, the proposition is true with respect to $w$ if and only if were $u$ uttered in $w$, it would express a true proposition with respect to $w$. Now, an utterance is a particular production of a sentence in a particular context. And the world in which an utterance is produced is part of the context of utterance. Given this, it is unclear how an utterance produced in one world could be uttered in another. An utterance produced in world $\alpha$ must be a distinct utterance from one produced in world $\beta$, for the identity of an utterance is tied to the context in which it is produced, and the world of production is a part of that context. And yet, Recanati’s characterization of the diagonal proposition seems to require at least the possibility of this occurrence. Insofar as it does, it would appear to be incoherent.

There may be a way to fix this problem, however. While a single utterance cannot made in two distinct worlds, a single sentence can easily be uttered in different worlds. Given this fact, the following altered definition of the diagonal proposition may allow the two dimensional theorist to circumvent the objection:
The diagonal proposition associated with an utterance \( u \) of a sentence \( s \) is the proposition such that, for any world \( w \) the proposition is true with respect to \( w \) if and only if were \( s \) uttered in \( w \) that utterance would express a true proposition in \( w \).

These alterations to Recanati’s original definition are not likely to generate problems with the rest of the account, either. According to Recanati, remember, the diagonal proposition associated with an utterance is the linguistic meaning of that utterance. Presumably, the linguistic meaning of an utterance is the linguistic meaning of the sentence uttered.

According to Recanati, remember, (3.6) provides the diagonal proposition associated with an utterance of (3.5).

\[
\begin{align*}
(3.5) & \quad \text{I am a philosopher} \\
(3.6) & \quad \text{There is an } x \text{ such that } x \text{ utters (3.5) and } x \text{ is a philosopher.}
\end{align*}
\]

Note how the indexical expression ‘I’ is cashed out in (3.6) – it is analyzed in precisely the same way we would normally interpret an attributive use of the description ‘the utterer of (3.5)’. Since Recanati’s two-dimensional account assumes that the linguistic meaning of an utterance is the diagonal proposition associated with that utterance, accepting the view would seem to require committing oneself to a descriptive view of the semantics of indexicals. Such a commitment may not be wholly unreasonable, but there are two problems with this version of the descriptive theory of indexicals, however, that might make some hesitant to accept it.

First, while many find the analysis of, say, ‘I’ as the speaker intuitively compelling, when the analysis is furthered by giving the Russellian interpretation of the description, the result runs counter to our intuitions, or at the very least fails to perfectly mesh with them. Intuitively, when I say “I’m tired”, I do not mean that there is an \( x \) such that \( x \) utters “I’m tired” and \( x \) is tired. The existential claim is not, intuitively, part of what I mean. Of course, the two-dimensionalist could
easily respond that my intuitions are only attentive to the level of what is said, and not the level of linguistic meaning, but then why would we have a positive intuitive reaction to the notion that ‘I’ means the speaker? After all, at the level of what is said, the content associated with my use of ‘I’, at least according to the two-dimensionalist view on offer here, is the referent of the term – that is, Diana Puglisi. One might think that the problem lies in giving the diagonal proposition in (3.6) a reading on which the condition that there is an utterer is a truth condition rather than a contextual condition. This would seem likely to result in a counterintuitive analysis, since, as we saw earlier, the condition that there be an utterer is, in the case of indexicals, a contextual condition. But this way of approaching the problem comes up short. Whether we read the condition as a truth condition or a contextual condition, the problem still remains. Our counterintuitive reaction to (3.6) as an analysis of (3.5) isn’t the result of assuming that the part of the linguistic meaning specifying that there is an utterer of (3.5) bears on the truth conditions when intuitively it does not; it results from our intuition that the notion that there is an utterer is just not part of the meaning of ‘I’.

Another problem with the descriptive theory of indexicals lies in its prima facie inconsistency with widely held notions about rigid designation. A rigid designator is an expression that picks out the same object in all possible worlds where that object exists. It is generally accepted that indexicals, like proper names, are rigid designators, while definite descriptions do not rigidly designate. Given this assumption, it does not seem possible to give a coherent account of indexicals in terms of definite descriptions. If indexicals have a semantically

---

40 See Kripke (1980)
significant property that definite descriptions do not, then definite descriptions cannot provide an adequate analysis of the semantics of indexicals.

Recanati may have a way around this problem, however, though it was originally developed to deal with a different, but related problem. Intuitively, there is a semantic difference between my utterance of “I am a brunette” and my utterance of “The present speaker is a brunette”, where the description is used referentially. But the two-dimensional account described thus far is incapable of explaining this difference. Both utterances express singular propositions about me at the level of what is said. And both have the linguistic meaning there is an x such that x utters U and x is a brunette. So, as far as the theory on offer is concerned, there should be no difference whatsoever in the semantic content of the two utterances.

To get around this difficulty, Recanti suggests that all directly referential expressions – that is, expressions which contribute nothing more than their referents to the proposition explicitly expressed (in Recanati’s terms, the ‘what is said’) - have a property he calls REF. REF is a component of the linguistic meaning of directly referential terms which indicates that the identifying condition in the diagonal proposition associated with the utterance in which the term occurs is a contextual condition. In others words, referential expressions, including indexicals, have an additional aspect of their linguistic meaning via which they indicate that the utterances in which they occur in subject-position express singular propositions.\(^ {41} \) Since definite descriptions are not referential expressions, they do not have the REF component in their linguistic meaning. According to Recanati, it is this difference between the linguistic meaning of indexicals and definite descriptions that we are picking up one when we intuitively distinguish

\[^{41}\text{Recanati (1989), pp. 240-241}\]
the semantics of my utterance of “I am a brunette” from my utterance of “The present speaker is a brunette”.

Additionally, the notion that indexicals have the property REF while descriptions do not might allow us to explain the why indexicals are clearly rigid designators while descriptions appear not to be. Since expressions with the property REF indicate that the identifying condition is a contextual condition and not a truth condition, the identifying condition is used solely to pick out the object being referred to so that the utterance in question expresses a definite proposition. Once a definite proposition is successfully expressed, the identifying condition has no further role to play. Thus, when we evaluate an utterance containing a referential expression with respect to worlds other than the world in which it was uttered, the referential expression will always pick out the same object. Since the identifying condition, when functioning as a contextual condition, has no role beyond the contextual identification of the referent of the expression, it is no wonder that the reference of referential terms is invariant across possible worlds. There is no means by which the referent of such a term, once contextually fixed, could be changed.

In contrast, since the linguistic meaning of a definite description does not indicate that the identifying condition of utterances in which the description occurs in subject-position should be regarded as a contextual condition, a definite description will not necessarily pick out the same object across different possible worlds. If the identifying condition is functioning as a truth condition, the denotation of the associated description will vary as states of affairs vary, since which object satisfies the identifying condition depends on the state of affairs that obtains. As a result, definite descriptions (or, more accurately, attributively used definite descriptions) are not rigid designators.
Since the version of the descriptive theory of indexicals Recanati offers views the linguistic meaning of an indexical as, essentially, a definite description plus REF, the fact that definite descriptions are not rigid designators while indexicals are does not prevent him from accounting for the linguistic meaning of indexicals in terms of descriptions. As long as the two-dimensional theorist maintains that the linguistic meaning of indexicals contains the extra component REF, he can explain the different designating properties of indexicals and descriptions while simultaneous holding that they share the same linguistic meaning.

If the foregoing suggestions are correct, then the fact that the two-dimensional theory necessitates acceptance of the descriptive theory of indexicals may not be all that problematic. One should note, however, that by positing REF, the two-dimensional theorist has gained one advantage only at the cost of another. There are a number of thinkers who have argued that indexicals can be used attributively. If REF truly is a feature of all indexicals, however, then indexicals should not be capable of yielding attributive readings. Recanati is not persuaded by the arguments for the attributive uses of indexicals and so would not consider this a serious loss. If there truly are attributive uses of indexicals, however, the two-dimensional theory would be at a significant disadvantage against any theory that could explain them.

Attributively used indexicals may not be the only example of an explanatory burden that the two-dimensional theory fails to meet. The theory tells us that referential expressions possess a property called REF which, as a component of their linguistic meaning, indicates that the utterances in which these expressions occur as subjects express singular propositions. REF was postulated as a solution to a problem, but no independent justification is given for the idea that

---

42 Nunberg (1993), Schiffer (1995). We will discuss attributive uses of indexicals in a bit more detail in the next chapter.
such a property exists, aside from the fact that referential terms all seem to behave in the same way. Furthermore, while we are given clear descriptions of the diagonal propositions that constitute the linguistic meanings of various utterances, no characterization of how REF fits in with the diagonal propositions is given, nor are we provided with any substantial description of REF from either a cognitive standpoint or a formal semantics perspective.

The mystery surrounding REF is not the only weakness of Recanati’s two-dimensional account. The view is also plagued by two further gaps in its explanatory scope, both of which are rather significant. The first of these is the inability of the theory to account for the functional use of definite descriptions, an example of which is given by (3.9).

(3.9)  The President of the U.S. changes every 5 years.

The description ‘the President of the U.S.’ as it occurs in (3.9) is not being used either referentially or attributively. An utterer of (3.9) in, say, 2011, does not mean to say that Barack Obama changes every 5 years. Nor does he mean that whichever individual is currently the president of the United States changes every five years. But these two uses are all that the two-dimensional account makes room for. The linguistic meaning of (3.9), according to the two-dimensionalist, consists of the identifying and predicative conditions, in this case, the condition that there is an \(x\) such that \(x\) is the President of the U.S., and the condition that \(x\) changes every 5 years, respectively. If the identifying condition is taken as a truth condition, then the description is given an attributive interpretation, and a general proposition is expressed. If, on the other hand, the identifying condition is taken to be a contextual condition, the description is interpreted as being referentially used and a singular proposition is expressed. There is no room in the theory, as it is described, for a description such as that in (3.9) which is neither referentially nor
attributively used. As such, then, it would also seem that the two-dimensional theory is incapable of accounting for generically used descriptions, such as that in (3.10) or what Ramachandran refers to as predicatively used descriptions, such as that in (3.11).

(3.10) The dodo is extinct.
(3.11) Dr. Joseph Bell was the inspiration for Sherlock Holmes.

The description ‘the dodo’ as it occurs in (3.10) is being used neither referentially nor attributively. The typical utterance of (3.10) does not mean that a particular dodo is extinct, nor does it express the proposition that there is a contextually salient dodo which is extinct. It tells us that the species is extinct. And while we could certainly read (3.11) as indicating an identity between Dr. Joseph Bell and the object which served Conan Doyle’s inspiration for the character of Sherlock Holmes, there is also a reading of (3.11) on which it attributes the property of being the inspiration for Sherlock Holmes to Dr. Bell. But this reading cannot be captured by giving either a referential or an attributive interpretation to the description ‘the inspiration for Sherlock Holmes’, as both interpretations serve to pick out the object satisfying the description.

The two-dimensional theorist might argue that his theory can, in fact, handle predicative uses of definite descriptions. Note, after all, that the predicatively used description in (3.11) does not occur as a grammatical subject, but rather as a grammatical predicate. The theory provided by Recanati was offered solely as an account of description utterances of the form \( \text{The } F \text{ is } G \). Our example of a predicatively used description, however, like all those given by Ramachandran, is an utterance of the form \( A \text{ is the } F \). According to Recanati, in the case of utterances of the form \( \text{The } F \text{ is } G \), the predicative condition given by \( G \) can only be understood as a truth condition. Presumably, however, things are different in the case of utterances such as ‘\( A \text{ is the } F \)’. The
problem, the two-dimensional theorist might argue, is not that the two-dimensional theory cannot account for predicatively used descriptions, but rather that Recanati’s portrayal of it unnecessarily limited it to description utterances of the form *The F is G*. Such a limitation would certainly put the two-dimensional theory on poor footing against any other theory of descriptions capable of accounting for descriptions occurring in the grammatical predicate of an utterance, including predicatively used descriptions. It is not uncommon for descriptions to occur as part of the grammatical predicate of an utterance, and such descriptions can not only be used predicatively, but also referentially, attributively, or generically. Consider, for example, an utterance of (3.12).

(3.12) Her new lover is the man in the yellow hat.

Suppose that Chris and Alex arrive at a party, having heard that Helen’s new lover will be present, and they are both rather interested in laying eyes on him (perhaps they are rivals for Helen’s love who wish to assess the new competition). Suppose further that, sometime after arriving, Alex is told by an acquaintance that Helen’s new lover is the only man at the party wearing a yellow hat. Not immediately seeing anyone wearing a yellow hat, Alex finds Chris and utters (3.12). In such a case, the description ‘the man in the yellow hat’ appears to be attributively used. Now suppose instead that Alex has identified Helen’s new lover, and Chris asks which man he is. Pointing toward the man in question, Alex utters (3.12). In this case, the description is referentially used.

Now consider a schoolteacher’s utterance of (3.13) as part of a class lecture.

(3.13) The largest mammal is the whale.
In this case, the description ‘the whale’ would seem to be generically used. Yet, just like the description ‘the man in the yellow hat’ in the previous cases, the description occurs in the grammatical predicate of the utterance.

If we extend Recanati’s basic view beyond utterances of the form *The F is G* to utterances in which definite descriptions occur in the grammatical predicate, we might be able to handle examples like (3.11) and (3.12). Take (3.12), for example. As we noted, the description ‘the man in the yellow hat’ as it occurs there can be used either referentially or attributively. Given the assumption that natural languages are compositional, it seems reasonable to hold that the description has the same linguistic meaning when it occurs in the grammatical predicate as it does when it occurs in the grammatical subject. Thus, if the two-dimensional theory is correct, the linguistic meaning of the description ‘the man in the yellow hat’ consists in an identifying condition which can be interpreted as either a contextual condition or a truth condition. When interpreted as a full-fledged truth condition, the description is used attributively; when interpreted as a contextual condition, the description is used referentially.

In keeping with Recanati’s original understanding of what is contributed to the diagonal proposition associated with an utterance by the expressions comprising the utterance, the description ‘her lover’ is an identifying condition, while the expression ‘is the man in the yellow hat’ is a predicative condition. The identifying condition which makes up the linguistic meaning of the description ‘the man in the yellow hat’, then, is embedded in the predicative condition. This does not prevent us from holding that the identifying condition associated with ‘the man in

---

43 Arguably, the description, ‘the largest mammal’ is also being used generically. The point the example is intended to demonstrate, however, is that the generic use of definite descriptions can be found in the grammatical predicate of utterances.
the yellow hat’ can be interpreted as either a contextual condition or a truth condition. It simply means that the precise truth condition associated with the predicate of (3.12) will vary with context. Assume, for the sake of simplicity, that the description ‘her lover’ is used attributively. Further, let \( a \) name the man in the yellow hat, and let \( C \) stand for the context of discourse. Now, when the identifying condition associated with ‘the man in the yellow hat’ is taken to be a truth condition, an utterance of (3.12) will express the proposition described in (3.14); when the identifying condition is taken to be a contextual condition, however, an utterance of (3.12) will express the proposition described in (3.15). The part of the proposition corresponding to (3.12)’s predicative condition is marked in bold in each case.

(3.14) There exists an \( x \) such that, in \( C \), \( x \) is uniquely lover of Helen and there exists a \( y \) such that, in \( C \), \( y \) is uniquely male and wearing a yellow hat, and \( x \) is \( y \).

(3.15) There exists an \( x \) such that, in \( C \), \( x \) is uniquely lover of Helen and \( x \) is \( a \).

On this modified version of the two-dimensional theory is correct, then, the predicative condition is still viewed as a truth condition. When it has the identifying condition associated with a definite description embedded within it, however, its content becomes sensitive to context, and which condition must be satisfied for the utterance to be true depends on how the embedded identifying condition is understood.

It appears, then, that the two-dimensional theory can rather easily account for referentially or attributively used descriptions which occur in the grammatical predicate of an utterance. But what about so-called predicatively used descriptions, like that occurring in one reading of (3.11)?
(3.11) Dr. Joseph Bell was the inspiration for Sherlock Holmes

The foregoing modification of Recanati’s account will not be sufficient to handle the predicative reading of (3.11), since the description ‘the inspiration for Sherlock Holmes’ on that reading does not pick out an object, but rather describes a property. To deal with this, the two-dimensional theorist might be tempted to claim that predicatively used definite descriptions like that in (3.11) have an entirely different linguistic meaning from those descriptions which are used to pick out individuals. This move is unadvisable, however, for it would commit the two-dimensional theorist to the claim that definite descriptions are semantically ambiguous. And the main motivation behind the two-dimensional theory, remember, was to account for the apparent semantic import of the referential/attributive distinction without appeal to semantic ambiguity.

The two-dimensional theorist might be able to account for the predicative use of definite descriptions without appeal to semantic ambiguity, however. With a few modifications of Recanati’s original understanding of the semantic content contributed by the various components of utterances containing definite descriptions, the predicative use could be treated in the same manner as the referential and attributive uses. Recall that the description in (3.11) can be given both an attributive reading and a predicative reading. If our analysis of those attributively used descriptions located in the grammatical predicate works, then the attributive reading of (3.11)

---

44 The account presented in what follows attempts to handle predicatively used descriptions in much the same way as referentially and attributively used descriptions are dealt with by the two-dimensional account. If viable, this strategy for dealing with predicatively used descriptions would make for the most elegant account overall. I do not here consider an account of the predicative use in terms of implicature because, on the face of things, there does not appear to be any clear violation of conversational cooperation which would necessitate interpreting the speaker as having communicated a predicatively used description as opposed to, say, an attributively used description. Indeed, that an appeal to implicature is not likely to yield an adequate explanation of the predicative use is made clear by the fact that we can hear both the attributive and predicative readings of (3.11) without being provided with any contextual background.
yields the proposition described in (3.16), where \( j \) names Dr. Joseph Bell, and \( C \) stands for the context of discourse.

(3.16) There exists an \( x \) such that, in \( C \), \( x \) is uniquely inspiration for Sherlock Holmes and \( j \) is \( x \).

Note that the predicative condition in this case is satisfied if and only if there is an object answering to the description and that object is identical to Dr. Joseph Bell. The ‘is’ in ‘\( j \) is \( x \)’ is the ‘is’ of identity. If we look closely at the two possible interpretations of the predicative condition expressed by ‘is the man in the yellow hat’ which are given in (3.14) and (3.15), we find the same ‘is’ of identity. It would seem that when a referentially or attributively used description is present in the grammatical predicate of an utterance of the form ‘A is the F’, the utterance is true when the object identified in the subject is identical with the object picked out by the description in the predicate.\(^{45}\)

Utterances of the form \( A \) is the \( F \) where ‘the \( F \)’ is predicatively used, on the other hand, do not express statements of identity, for, in contrast to referential or attributive uses, predicatively used descriptions occur only in the grammatical predicate and serve only to describe a property which applies to only one individual. In fact, these are, according to

\(^{45}\) Note that we have limited the present discussion to utterances of the form \( A \) is the \( F \) (which I take to include utterances of the form \( The \ F \ is \ the \ G \)). One might worry that we are ignoring a large class of utterances with definite descriptions in the grammatical predicate which do not constitute identity statements. Utterances of sentences like ‘Sally dried the last of the dinner dishes’, ‘Kyle didn’t finish the last problem’ or ‘Sam accidentally kicked the dog’, for example, do not serve to express propositions identifying the referent of the description in the predicate with the grammatical subject, or with anything else, for that matter. Given that predicatively used descriptions are used only in attributions of a property to an individual, however, they could not possibly occur in utterances like those above. We need only concern ourselves with those description utterances which could feasibly have predicative interpretations. This is not to say that utterances like ‘Sally dried the last of the dinner dishes’ have no place in the two-dimensional theory. On the contrary, it seems likely that the identifying conditions associated with the descriptions in these utterances are uniformly understood in terms what is below called the object interpretation.
Ramachandran, the defining characteristics of the predicative use.\textsuperscript{46} Since, unlike their referential and attributive counterparts, predicatively used descriptions do not identify individuals, they cannot stand on either side of an identity statement.

So what sort of proposition is expressed by utterances containing predicatively used descriptions? Consider an utterance of (3.11) interpreted predicatively. What would make such an utterance true? Intuitively, it is true when Dr. Joseph Bell instantiates the property \textit{being the inspiration for Sherlock Holmes}. To put it another way, it is true when there is a property \textit{uniquely inspiring Sherlock Holmes} and this property is possessed by Dr. Bell. If this is, in fact, what would make such an utterance of (3.11) true when the description is predicatively used, then (3.16) can be seen as describing both the proposition expressed by an attributive reading of an utterance of (3.11) and the proposition expressed by a predicative reading of such an utterance.

\begin{equation}
(3.16) \text{There exists an } x \text{ such that, in } C, x \text{ is uniquely inspiration for Sherlock Holmes, and } j \text{ is } x.
\end{equation}

The difference lies in how the predicative condition of (3.11) is interpreted, and this, in turn, depends on how the identifying condition embedded within the predicative condition is interpreted. Just like the referential and attributive uses, the two-dimensional theorist might argue, the predicative use results from a variation in the interpretation of the identifying condition. For both the referential and attributive uses, the variable in the identifying condition is understood as ranging over \textit{individuals}. As a result, the predicative condition is interpreted as identifying the grammatical subject with the individual supplying the value of the variable. On this interpretation, then, we understand (3.16) as saying that there is an object which, in the

\textsuperscript{46} Ramachandran (2008), p. 71
context, is uniquely inspiration for Sherlock Holmes, and we understand the ‘is’ in ‘j is x’ as the ‘is’ of identity. Call this the object interpretation of the identifying condition. In the case of a predicative use, however, the two-dimensional theorist might argue that the variable in the identifying condition is instead interpreted as ranging over properties. This, in turn, results in an interpretation of predicative condition as attributing a property to the grammatical subject. Thus, on this interpretation, we understand (3.16) as saying that, relative to the context, there is a property unique inspiration for Sherlock Holmes, and we interpret the ‘is’ in ‘j is x’ as the ‘is’ of predication. Call this the property interpretation of the identifying condition.

Given these two potential interpretations of the identifying condition associated with the description in utterances of the form ‘A is the F’, it would appear possible for the two-dimensional theory to account for the predicative use of definite descriptions. While such an account fills in one significant explanatory gap, the theory is still open to charges of insufficient theoretical breadth. For despite being capable of explaining three of the uses of definite descriptions – the referential, attributive and predicative uses – the two-dimensional theory still lacks any account of the generic and functional uses of definite descriptions.

In addition to worries about theoretical breadth, the two-dimensional theory is also open to the objection that it is deficient in explanatory depth. Indeed, there is one rather crucial aspect of the theory that is given virtually no detailed discussion, and, as a result, the theory as a whole may be untestable. Remember that according to the two-dimensional theory, definite descriptions have a univocal semantics – all uses of a definite description contribute the same identifying condition to the diagonal proposition associated with the utterance in which they occur. The different uses of definite descriptions result from different interpretations of this identifying
condition. Let us, for the moment, restrict our discussion to utterances of the form *The F is G*. The description, ‘the F’ contributes to the diagonal proposition associated with the utterance the following identifying condition: there exists an \( x \) such that, in the context, \( x \) is uniquely \( F \). If this condition is interpreted as a contextual condition – that is, as a condition which must be satisfied for a proposition to be expressed – then the description is being used referentially, and the utterance in which it occurs expresses a singular proposition. If, on the other hand, the identifying condition is interpreted as truth condition, then the description is being used attributively, and the utterance in which it occurs expresses a general proposition. According to Recanati, which interpretation is given to the identifying condition is determined contextually.\(^{47}\) In some contexts, the identifying condition will be interpreted as a contextual condition, and in others, it will be interpreted as a truth condition.

Since the variation in interpretation of the identifying condition is the source of the different interpretations of definite descriptions, an adequate explanation of the different uses of definite descriptions would seem to require at least some account of how the conversational context determines a particular understanding of the identifying condition. In the absence of such an account, the two-dimensional theory’s main claim – that definite descriptions have a single linguistic meaning which results in different interpretations in different contexts – amounts to little more than stipulation. The two-dimensional theorist must substantiate this claim by giving us at least a sketch of which factors of the conversational context serve to determine the interpretation of the identifying condition and indicating which contexts will result in which interpretation. Unfortunately, Recanati fails to give us even a brief description of how the

\(^{47}\) Recanati (1989), pp. 237-238
conversational context determines the interpretation of the identifying condition, and he provides us with no insight as to what determines an interpretation of the condition as a contextual condition as opposed to a truth condition.

The two-dimensional theorist claims that the referential and attributive uses of a definite description should be understood as different interpretations of the identifying condition which comprises the linguistic meaning of that definite description. Part of evaluation this claim will lie in examining concrete examples of uses of definite descriptions and determining whether or not the predictions of the two-dimensional theory are borne out in each case. This is how we go about testing theories of this sort against the relevant data. If the two-dimensional theory is correct, then we should expect that in any given case of, say, a referentially used description, the conversational context will determine an interpretation of the associated identifying condition as a contextual condition. Cases in which the expectation is met serve to increase the degree of confirmation associated with the theory. Cases in which the expectation is not met are counterexamples, and serve to disconfirm the theory. But this method of testing the theory cannot be employed without an account of how the conversational context determines one or the other of the potential interpretations of the identifying condition. In the absence of such an account, we have no basis for making the claim that a context determines this or that particular interpretation of an identifying condition. As a result, we cannot determine whether or not the claims of the two-dimensional theory are met with respect to any given example of a use of a definite description. It would seem, then, that as it stands, the two-dimensional account of definite descriptions cannot be tested. This is a serious failing in any theory.
It would seem, then, that while the two-dimensional account has its benefits, it also suffers from a number of substantial weaknesses. Perhaps the best feature of the theory is that it provides an account which maintains a univocal semantics for definite descriptions while simultaneously taking into account the apparent semantic import of the referential/attributive distinction. As a result, it multiplies neither senses nor implicatures beyond necessity. Despite this, however, it seems to be deficient in both explanatory breadth and depth. It provides explanations of the referential, attributive and predicative uses of definite descriptions, but fails to account for the generic and functional uses. In attempting to explain the difference between the semantics of referentially used descriptions and indexical expressions, it introduces a semantic feature – REF – but leaves much about REF a mystery. And perhaps most troubling of all, the theory accounts for different uses of definite descriptions in terms of contextually determined interpretations of their linguistic meaning, but it fails to provide any explanation of how the context determines an interpretation or which contexts determine which interpretations. This not only leaves a large explanatory gap in the theory, but also makes it impossible to test the theory’s predictions with respect to concrete examples of uses of definite descriptions.

The two-dimensional theory is not the only contextualist theory available, however. In fact, in light of the two-dimensional theory’s inability to account for functionally used definite descriptions, Recanati abandoned the view, and proposed an alternative – the synecdoche theory.

### 3.3 The Synecdoche Theory

Like the two-dimensional theory, the synecdoche theory is a contextualist theory – it holds that all uses of a definite description have the same linguistic meaning, but that in different
contexts of utterance, this linguistic meaning will lead to the expression of different propositions. While the synecdoche theory and the two-dimensional theory share the same basic strategy, they depend on slightly different sets of background assumptions. Since the differences in these background assumptions are responsible for a number of the significant differences between the synecdoche theory and the two-dimensional theory, it might be useful to briefly examine them.

The two-dimensional theory seems to work, for the most part, on the assumption that the proposition expressed by an utterance in a context is roughly equivalent to the truth conditions for that utterance, while the linguistic meaning of an utterance is given by a statement of the conditions under which any such utterance, in any context (or rather, in any world) would be true. In contrast, the framework Recanati places the synecdoche theory in does not identify the proposition expressed by an utterance with the truth conditions of that utterance, and it offers an entirely different understanding of the linguistic meaning of an utterance. On this new view, the linguistic meaning of an utterance is, essentially, its illocutionary act potential.\(^\text{48}\) In other words, the linguistic meaning of an utterance is an indication of both the sort of illocutionary act the utterance can be used to perform (command, assertion, request, etc.) and the conditions under which the speech act is satisfied (conditions for obeying the command, truth conditions, conditions on fulfillment of a request, etc.). The linguistic meaning of (3.17), for example, would include an indication that (3.17) can be used to make an assertion and that an assertion made with (3.17) is satisfied on the condition that the sky is blue.

\[(3.17) \text{ The sky is blue.}\]

\(^{48}\) Recanati (1993), p. 16
It would be a mistake to assume that on this view the linguistic meaning of an utterance is just a
description of the truth conditions of that utterance, however, for in cases where the utterance in
question contains, for example, a context-sensitive expression, the linguistic meaning of the
utterance will indicate the truth conditions for any context of utterance. In a sense, then, this
view of the linguistic meaning of an utterance is extremely similar to that underlying the two-
dimensional theory.

Recanati views the linguistic meaning of an utterance as doing more than just indicating
truth conditions. He tells us that:

The hearer who understands the utterance must know its truth-condition under a certain
mode of presentation…
The mode of presentation under which the truth-condition has to be identified for the
utterance to be understood depends on the meaning of the utterance. The utterance, by
virtue of its linguistic meaning, gives indications concerning its truth-condition; it
presents the truth condition in a certain way… The meaning of the utterance provides a
‘route’ to the truth-condition. In order to understand the utterance, what the hearer must
grasp is not its truth-condition simpliciter, but its truth-condition as it is presented by the
utterance itself. This – the truth condition as it is present by the utterance itself – is what I
call the proposition expressed.49

The linguistic meaning, then, not only indicates the truth conditions of the utterance, but it does
so in a particular way. And it is the truth conditions of the utterance under this mode of
presentation that constitute the proposition expressed for Recanati here.

In addition to relying on different understandings of the proposition expressed and the linguistic
meaning of an utterance than that underlying the two-dimensional theory, the synecdoche theory
also assumes that utterances are used to convey thoughts. Thus, unlike the two-dimensional
theory, the synecdoche theory attempts to sketch a view of the psychological content associated

with utterances containing definite descriptions. Recanti appears to view utterances as conveying thoughts by expressing propositions. For the most part, then, the propositions expressed by utterances (particularly utterances containing directly referential expressions) are bare shadows of the thoughts those utterances are meant to express.

With a sketch of the basic framework Recanati is relying on, we are now in a position to layout the synecdoche theory. As we noted earlier, Recanati abandoned the two-dimensional theory because it was incapable of accounting for functionally used descriptions. He attributes this failure of the two-dimensional theory to the fact that it analyzes the linguistic meaning of definite descriptions in much the same way as Russell did. This quantificational account of descriptions, Recanati claims, seems to preclude an account of functional uses. Instead of relying on Russell’s account, then, Recanati suggests that a description ‘the F’ expresses a *descriptive concept*, that is a concept of a unique F. Recanati views this descriptive concept as being part of both the thought expressed by an utterance containing a description and the linguistic meaning of the such an utterance.

So, a description ‘the F’ expresses a descriptive concept. If the description is being used attributively, this is the end of the story. In the case of a referentially used description, however, Recanati claims that a transfer from the descriptive concept to a *de re* concept is triggered. This transfer is a move from part to whole – a synecdochic transfer. As Recanati sees it, *de re* concepts are essentially dossiers containing whatever information is possessed about the object of the dossier (that is, the ultimate causal source of the information). My *de re* concept of George Carlin, for example, contains, among others, concepts and mental representations of properties

---

50 Recanati (1993), p. 293
such as MALE, AMERICAN, FAMOUS STAND-UP COMEDIAN, PROponent of FREE-SPEECH RIGHTS, and HUSBAND of SALLY WADE. It is a collection of the large number of concepts and properties I associate with George Carlin – descriptions I take to identify him, properties I attribute to him, etc.

To get a better picture of what the Recanati is suggesting here, suppose that, in conversation with me, Dan utters (3.18) with the intention of expressing his de re thought about George Carlin And suppose further that the description in (3.18) is clearly being used referentially.

(3.18) The husband of Sally Wade was funny.

If I am to understand Dan’s utterance of (3.18), Recanati claims I must satisfy two conditions. First, I must recognize that Dan is expressing a de re thought – a thought which represents a singular state of affairs. And second, I must identify the singular state of affairs in question. As these two conditions should make clear, Dan utterance of (3.18) expresses a singular proposition – a singular proposition which represents the same state of affairs represented by the de re thought which Dan’s utterance of (3.18) was intended to express.

The satisfaction of the first condition, on this view, triggers a transfer from the descriptive concept expressed by the description to the de re concept. According to Recanati, once I see Dan’s utterance of (3.18) as intended to express a de re thought, I will view the descriptive concept expressed by the expression ‘the husband of Sally Wade’ as standing for a

51 We will use small caps to indicate concepts. An entire phrase in small caps should be understood as indicating a structured strings of concepts.

52 This is Recanati’s characterization of de re thought. See, e.g., Recanati (1993), p. 296. While Recanati does not indicate just what he takes to be a singular state of affairs, it seems likely that he means the type of state of affairs indicated by a singular proposition.
broader *de re* concept of which it is a part. And it would seem that this transfer from the
descriptive concept to the *de re* concept aids in the satisfaction of the second condition. For in
order to identify the singular state of affairs in question, I must identify the object which is a
constituent of that state of affairs. In order to do that, Recanati claims, I must form a *de re*
thought about it. It is not necessary that my *de re* thought employ the same *de re* concept which
occurred in the thought Dan originally intended to express. It is necessary, however, that my *de re*
concept contain the descriptive concept expressed by description occurring in the utterance.
Recanati contends, remember, that the proposition expressed by an utterance consists in the truth
conditions of that utterance *presented in a certain way*. The truth conditions of Dan’s utterance
of (3.18) are *George Carlin was funny*. But the description used in (3.18) presents those
conditions under a certain mode of presentation. Thus, in order to grasp the proposition
expressed by Dan’s utterance, I must not only identify the singular state of affairs it represents, I
must identify that state of affairs under the mode of presentation indicated by the linguistic
meaning of the utterance. In other words, I must form a *de re* thought using a *de re* concept
which contains the descriptive concept expressed by the description ‘the husband of Sally
Wade’.

3.4 **Critique of the Synecdoche Theory**

Recanati’s synecdoche theory contains a number of positive elements. For example, in
attempting to give an account of definite descriptions in terms of both their semantics at the
interpersonal level and their cognitive import at the personal level, the synecdoche theory takes
strides toward covering ground that the other theories of descriptions we have seen so far have
completely ignored. Additionally, unlike the other theories we have thus far examined, the synecdoche theory provides us with an explanation of the role the descriptive content associated with referentially used descriptions plays in the propositions expressed by utterances containing such descriptions. While there are those who would balk at the notion that the descriptive content associated with referentially used descriptions has any role to play beyond that of identifying the referent of the description, our intuitions in cases of referentially used misdescription, along with the sorts of concerns raised by Frege’s problem, would seem to provide us with some justification for, at the very least, making room for the possibility that the descriptive content associated with referentially used descriptions does more than just fix their referent.

Despite these beneficial aspects of the account, however, the synecdoche theory has a number of substantial weaknesses. First, despite the fact that Recanati abandoned the two-dimensional theory for the synecdoche theory because the two-dimensional theory could not account for the functional use of definite descriptions, the synecdoche theory as currently stated provides no account of such uses of definite descriptions. Not only does it leave the functional use unexplained, the theory also fails to account for both the generic use and the predicative use of definite descriptions. Nor does an account of such uses of descriptions appear to be readily available on the basis of the claims of the theory as it stands.

The synecdoche theory’s account of the referential/attributive distinction is also incomplete. Recanati gives us an in depth explanation of how the referential use arises out of the descriptive concept expressed by a description ‘the F’. In certain contexts, he claims, we will see the descriptive concept expressed by the description as standing for a de re concept, thus identifying a singular state of affairs under a particular mode of presentation. In other contexts,
the transfer from the descriptive concept to the \textit{de re} concept will not be triggered, and the utterance containing the description will express a proposition of which the descriptive concept is a part. But just as we saw with the two-dimensional theory, the claim that certain contexts will result in a change in how the linguistic meaning of the utterance is interpreted is nothing more than a stipulation unless we have some account of how the context determines the change in interpretation.

In response to this charge, the proponent of the synecdoche theory might remind us that one of the conditions which must be satisfied for a hearer to understand an utterance containing a referentially used definite description is recognition of the fact that the speaker intends for the utterance to express a \textit{de re} thought. Those contexts which allow the hearer to make this recognition, it might be argued, are the contexts in which a transfer from the descriptive concept to the \textit{de re} concept are triggered. And it is those features of the context which cause the hearer to recognize the speaker’s intention to express a \textit{de re} thought that trigger the transfer.

This would seem to be a promising account of the relationship between the conversational context and interpretation of the linguistic meaning of a description. It does leave one to wonder, however, which features of the context cause the hearer to recognize the speaker’s intentions? We will need to know what the relevant features of context are if we are to determine whether or not the transfer from descriptive concept to \textit{de re} concept is triggered in any particular case, and thereby test the theory against concrete cases.

The most likely candidate of the various features of conversational context would seem to be something like mutual knowledge between conversational participants that the description
occurring in the utterance applies to a particular object. After all, referentially used descriptions appear to arise almost exclusively in those situations in which it is mutually known by speaker and hearer that the description being used is satisfied by the referent of that description. There are problems with claiming that it is this feature of context which triggers the transfer from the descriptive concept to the *de re* concept, however. For example, suppose that Jake and Julie are going through their mail when Jake opens a letter notifying him that he has won a contest which he had completely forgotten about entering. He immediately hands the letter to Julie, who begins reading it. After a few moments, Jake asks what prize he’s won. Julie responds by uttering (3.19).

(3.19) The winner gets a set of encyclopedias.

In this context, it is mutual knowledge between Jake and Julie that Jake is the winner. Nonetheless, the intuitive interpretation of the description in (3.19) is attributive, not referential. Clearly, then, even with the rather solid basis of an account of how context triggers a transfer to the *de re* concept from the descriptive one suggested above, there are still a number of kinks that must be worked out.

Cases of mutual knowledge are not the only source of problems for the synecdoche theory. Because it provides such closely related accounts of both the cognitive import and the semantics of definite descriptions, it is opened up to worries about cases in which the speaker and the hearer are not privy to the same information. To see this, imagine (3.20) uttered by a

---

53 Mutual knowledge, here, should be understood in the standard sense. That is, if *p* is mutually known to A and B, then A knows that *p* and B knows that *p* and they each know that the other knows. In the place of mutual knowledge, one might instead require that *p* be part of the shared cognitive environment. See Sperber and Wilson (1986). The important point here is that the conversational participants share some awareness about the fact in question. The details of how the condition is cashed out are unimportant for our present purposes.
speaker who has a particular individual in mind, but whose audience has no knowledge that the individual in question satisfies the description. 54

(3.20) The man living next door to Paul has some very strange habits.

Now, if the speaker does not expect his audience to associate the description in (3.20), we have a case of what Ludlow and Neale call the specific use of definite descriptions. 55 Recanati’s synecdoche theory seems well equipped to handle such cases, for if the speaker does not expect the hearer to associate a particular individual with the description used, then it is hard to see why he would intend for the description he uses to trigger a transfer from descriptive concept to de re concept. Instead, he must intend for the description to be used attributively, and this is how it will be interpreted by the hearer.

But now imagine a slightly different case. Suppose the speaker utters (3.20) with a particular individual in mind, and mistakenly believes that his audience associates this same individual with the description used. Relying on this false belief, the speaker intends the description in (3.20) to be used referentially. If we suppose that the hearer does not associate any particular individual with the description ‘the man living next door to Paul’, however, then it seems clear that the hearer will be unable to give a referential reading to the description. Let us suppose further that the hearer has no idea that the speaker had a specific person in mind. In terms of the synecdoche theory, the speaker will have uttered (3.20) in an attempt to express a de re thought, but the hearer will interpret it as expressing a thought containing a descriptive concept – we might call it a descriptive thought. At this point, it seems that an utterance of (3.20) in the circumstances described would be nothing more than a minor miscommunication, one that

54 This example is inspired by one offered by George Powell. See Powell (2001), p. 109.
55 Ludlow & Neale (1991)
would go unnoticed by the conversational participants unless something later in the conversation hinged on the particular identity of the man living next to Paul. And this seems to be an accurate assessment of the situation. But, from the perspective of the synecdoche theory, this cannot be all there is to the story. For according to the theory, there is more in such cases of communication than just the thought the speaker attempts to express and the thought the speaker arrives at via interpretation of the utterance. There is also the proposition expressed. But just what would be the proposition expressed in this case?

Recall that, according to Recanati, the proposition expressed by an utterance is the truth conditions of that utterance under a certain mode of presentation. But which truth conditions? The thought the speaker intended to communicate by uttering (3.20) is associated with different truth conditions from the thought the hearer interpreted the utterance of (3.20) as expressing, and Recanati gives us no explicit basis on which to make a determination between the two sets of truth conditions. Some of Recanati’s claims would seem to indicate that it is the speaker’s intention that dictates the content of proposition expressed. With respect to cases of misdescription, for example, Recanati tells us

…we have no clear intuition of what is said when there is a divergence between the two factors which normally concur into determining what is said, namely the semantic properties of the sentence on the one hand and the intention of the speaker on the other hand. What is being said when I point to Jones and utter ‘Smith is raking the leaves’? There is no clear and definite answer to this question, except one based on stipulation.\(^56\)

Since Recanati describes the two factors that determine what is said as the semantic properties of the sentence and the intention of the speaker, we might be tempted to conclude that in cases in which the speaker and hearer interpret the utterance differently, such as our problematic

utterance of (3.20), it is the speaker’s intentions, and not the hearer’s interpretation, which indicate the proposition expressed.\textsuperscript{57} But a close look at Recanati’s claim about our intuitions in cases of misdescription reveals this conclusion to be hasty. For Recanati specifically says that in cases where the speaker’s intentions diverge from the semantics of an uttered sentence, there is no definite answer to the question of what was said. In other words, when the speaker’s intentions do not line up with the linguistic meaning, there is no clear determinant of the proposition expressed except stipulation.

But what is our problematic utterance of (3.20) except a case where the semantics diverge from the speaker’s intentions? Remember, according to the synecdoche theory, a description ‘the F’ expresses a descriptive concept – both in cognitive and semantic terms. The linguistic meaning of ‘the F’ is a descriptive concept. But it is, \textit{ex hypothesi}, the intention of the utterer of (3.20) that a \textit{de re} concept replace that descriptive concept. He does not wish to express a descriptive thought and a general proposition, but rather a \textit{de re} thought and a singular proposition. If Recanati is correct in claiming that in cases where speaker’s intentions and linguistic meaning part company there is no proposition expressed, then, it would seem, there is no proposition expressed in the case of our problematic utterance of (3.20).

Indeed, the problem may extend far beyond cases in which the utterer of a description sentence has a referent in mind while the audience does not. If the synecdoche theory is correct, then all instances of referentially used descriptions are cases where the speaker’s meaning diverges from the linguistic meaning of the utterance. The linguistic meaning of a description

\begin{footnote}
\textsuperscript{57} The linguistic meaning of an utterance (i.e., the linguistic meaning of the uttered sentence) and the semantic properties of uttered sentences are the same for Recanati. Hence the terms are used interchangeably here. Similarly, for Recanati, what is said is the same as the proposition expressed. Hence we will treat them as identical for the remainder of the section.
\end{footnote}
‘the F’, recall, is a descriptive concept. But, according to the synecdoche theory, the utterer of a referentially used description intends that descriptive concept to stand for a *de re* concept. In such cases, the speaker intends a meaning that differs from that linguistically expressed by the sentence he utters. If this is an accurate portrayal of the synecdoche theory, then it would seem to be inconsistent with Recanati’s claim that there is no proposition expressed when the speaker’s intentions differ from the semantic properties of the sentence he has uttered. According to the synecdoche theory, utterances containing referentially used descriptions express singular propositions. But, when taken together, the theory and Recanati’s claim imply that an utterance containing a referentially used definite description fails to express any proposition at all.

If correct, the foregoing objection poses a substantial threat to the synecdoche theory. There are a number of potential responses that the synecdoche theorist might make, however. First, it might be argued that the objection rests on a misinterpretation of Recanati’s claim. The objection assumes that *any* sort of divergence between linguistic meaning of an utterance and the speaker’s communicative intentions prevents the utterance from expressing a proposition. But we might understand Recanati’s claim as holding that any *significant* divergence between linguistic meaning and speaker’s intentions prevents the expression of a proposition. The difference between a descriptive concept and a *de re* concept, it might be argued, is not significant. After all, according to Recanati, the descriptive concept is part of the *de re* concept. The concept that is linguistically expressed is a component of the concept that the speaker intended to be part of what was said. Hence, Recanati’s claim would not commit the synecdoche theorist to the conclusion that utterances containing referentially used descriptions fail to express propositions. And since the speaker’s intentions are a factor in determining what is said, the synecdoche
theorist may claim that our problematic utterance of (3.20) expresses the singular proposition which the speaker intended.

This response, however, rests on a rather tenuous claim – that it is only a significant divergence between linguistic meaning and speaker’s intentions that prevents the expression of a proposition. Just how significant does this divergence have to be? Recanati uses the example of an inapplicable name. Assuming (as Recanati does) that names are directly referential, his example is one in which the gap between linguistic meaning and speaker’s intentions is quite wide. But where does one draw the line? The expressions ‘professor’ and ‘instructor’ have closely related meanings. If a speaker intends to communicate that Bob will be an excellent swimming instructor, will his utterance of ‘Bob will be an excellent swimming professor’ express a proposition to that effect? If the synecdoche theorist plans on handling the objection in this fashion, he owes us an explanation of how divergent linguistic meaning and speaker’s intentions can be before they prevent the expression of a proposition and, moreover, a justification for drawing the line where he does.

Even if we grant that such explanation and justification could be provided, there may still be a problem, for even the narrowest interpretation of Recanati’s claim that would still suit the synecdoche theorist’s purposes may be too broad. In order to ensure that utterances containing referentially used descriptions are still viewed by his theory as expressing propositions, the synecdoche theorist must hold, at the very least, that in cases in which the linguistic meaning of an expression comprises part of the concept the speaker intended to express the divergence is not significant enough to prevent the utterance containing the expression from expressing a proposition. But now, consider (3.21).
(3.21) Molly believes that Tom is in Spain.

It is generally accepted that part of what it means to know something is to believe it, that part of the concept KNOWS is the concept BELIEVES. But suppose you ask Lisa where Tom is, and she replies by uttering (3.21). And suppose further that in uttering (3.21), Lisa intended to convey that Molly knows that Tom is in Spain. Would her utterance express the proposition that Molly knows Tom is in Spain? That is, do the words that she uttered mean, on that occasion of utterance, that Molly knows that Tom is in Spain? Intuitively, the answer to these questions is no. I suspect that most others will share reaction, though perhaps with a little additional uncertainty. And this makes even the synecdoche theorist’s narrowest modification of Recanati’s claim suspect. The case of ‘believes’ and KNOWS with regard to (3.21) precisely parallels the synecdoche theorist’s depiction of the relationship between the descriptive concept and the de re concept in the case of a referentially used description. So it should be just as feasible for Lisa’s utterance of (3.21) to express the proposition she intended as it is for the utterer of a description sentence to express a singular proposition when he intends it. But it is not.

Of course, there are a number of situations in which Lisa’s intended meaning might reach her audience. On the basis of extra linguistic clues or previous knowledge, you might reach the conclusion that a knowledge claim was Lisa’s intended meaning, and in such a case, her utterance might serve to help you determine that Lisa was trying to communicate the proposition that Molly knows that Tom is in Spain. That is, you might infer her intended meaning with some help from the utterance itself. But the same is true of any case in which a proposition is conversationally implicated. It is also true in Recanati’s example, discussed below, of the
speaker who points to Jones and says ‘Smith is raking the leaves’. This, then, is no indication that Lisa’s utterance of (3.21) expresses the proposition that Molly knows that Tom is in Spain.

It would seem, then, that the synecdoche theorist cannot avoid the inconsistency between his theory and Recanati’s claim by drawing a distinction between different extents of divergence. In order to get around the objection, then, it might be necessary to simply reject Recanati’s claim, and instead maintain that, all else equal, a proposition is expressed by an utterance even when the speaker’s intended meaning and the linguistic mean differ. Rejecting Recanati’s claim very neatly eliminates the synecdoche theorist’s problems with inconsistency, but it may do so at a substantial cost. For there seems to be some good evidence in support of Recanati’s claim. Recanati’s own example of the use of an inapplicable name is, on its own, rather compelling. To see this, let us expand upon the case a bit. Suppose you, along with Smith, Jones, and Brown, have decided to act charitably and clean up the front yard of an infirm elderly neighbor. Arriving a bit late, you find Brown trimming a tree and Smith pulling weeds. You do not immediately see Jones. You approach Brown and ask if you should rake the leaves. He points toward Jones, and utters (3.22).

(3.22) Smith is raking the leaves.

Does Brown’s utterance of (3.22) express a proposition? And, if so, which one? Does he express the false proposition that Smith is raking the leaves, or the true proposition that Jones is? To many, it is not at all clear. You are certainly in a position to figure out just what thought Brown intended to convey – that that man (Jones) is raking the leaves. And on the basis of that, you can also glean the implicature (or intended implicature) that you should take up a task other than

74
raking leaves. But what was actually expressed by the utterance, if anything? For many of us, our intuitions supply us with no answer.

Misapplications of names are not the only cases which seem to support Recanati’s claim. Consider, for example, an utterance of (3.23) made by someone who thinks ‘entomologist’ refers to etymologists and ‘etymologist’ refers to entomologists, or an utterance of (3.24) made in a term paper by a Philosophy 101 student in reference to an obviously fallacious argument.

(3.23) An etymologist would know what kind of bug this is.
(3.24) This argument is false. The conclusion doesn’t follow from the premise.

Does (3.23) express the true proposition that an entomologist would know the kind of bug, or the false proposition that an etymologist would know? Does (3.24) express the false (or possibly incoherent) claim the argument in question is false, or the true claim that the argument is no good? In both cases, my intuitions pull in opposite directions, leaving things in a state of uncertainty. And perhaps the best explanation for this would be that neither (3.23) nor (3.24) expresses a proposition, as the two factors which determine what proposition is expressed by an utterance are at odds with one another in each case.

But some might find the foregoing examples unpersuasive, and there may very well be justification for thinking that Recanati’s claim is false. Even if there reasons to reject Recanati’s claim, the synecdoche theorist will face trouble if he chooses to deny it, for it will leave him without an explanation for the intuitive tension we encounter in cases of misdescription. Our intuitions seem to pull us in two directions when it comes to interpreting utterances containing inaccurate descriptions. Suppose, for example, that Charles utters (3.25) while indicating a man who is most attentive to Megan and has been very kind and caring to her all evening.
(3.25) Her husband is very kind to her.

Suppose that the man Charles has indicated – call him Eric – is in fact very kind to Megan. But suppose further that the man Charles indicated is not Megan’s husband, and that her actual husband is, in fact, most cruel to her. Is Charles’ utterance true or false? For many, the answer to that question is not intuitively easy to come by. And even for those who have an intuitive inclination one way or the other, there is a tension. But if the synecdoche theorist rejects Recanati’s claim, then he must hold that (3.25) expresses the singular proposition that Eric is kind to Megan. And it is not clear how he can explain our intuition, or rather, our pull toward the intuition that there is something inappropriate about Charles’ utterance.

The synecdoche theorist might attempt to explain the tension by appealing to the fact that, for him, the proposition expressed by Charles utterance is more than just simple truth conditions. It is truth conditions presented a certain way – in the way indicated by the description ‘her husband’. Since Eric is not Megan’s husband, the synecdoche theorist might argue, it is this mode of presentation that we find inappropriate about Charles’ utterance of (3.25).

But this strategy will not provide a full explanation of the tension in our intuitions with respect to (3.25). The mode of presentation in question is, according to the synecdoche theorist, a way of thinking about the truth conditions associated with the utterance. Since it is an inapt portrayal of the truth conditions, the presence of this mode of presentation in the proposition expressed would certainly explain why we might feel that, despite expressing a true proposition, Charles has done something incorrect in uttering (3.25). But that is not all there is to the intuitive tension arising out of cases of misdescription. For while some have the intuition that, in uttering (3.25), Charles has expressed a truth, others have the opposing intuition that Charles has, strictly
speaking, expressed a falsehood. The presence of an inappropriate mode of presentation in a true proposition is not sufficient to explain those intuitions. There is no reason, on the synecdoche account, for us to have the intuition that Charles’ utterance of (3.25) is false. At best, then, the synecdoche theorist has only a partial account of the tension in intuitions that arises out of cases of misdescription.

One final worry plagues the synecdoche account. The core of the synecdoche theory’s treatments of the referential/attributive distinction lies in two claims: 1) the referential reading is an instance of synecdoche and 2) synecdochic transfer in the case of a referentially used description impacts the proposition expressed by the utterance in which the description occurs. It is the second claim that generates concern. Let us grant the first claim. After all, given Recanati’s characterization of the relationship between descriptive concepts and de re concepts and his belief that the attributive use of a description expresses the latter while the referential use expresses the former, it seems highly plausible to claim that moving from the descriptive concept originally expressed by the description to a de re concept is a case of synecdoche, as it is a move from part to whole.  

The second claim – that the result of the synecdochic transfer in question impacts the content of the proposition expressed – seems to be largely motivated by Recanati’s prior acceptance of the notion that the referential/attributive distinction is truth conditionally relevant. It is an open question, however, whether or not cases of synecdoche in general, or this case of it  

---

58 Of course, synecdoche also includes instances in which the whole is used to stand for the part, as in, for example ‘England beat Brazil in the semi-final’, where the countries stand for the teams which represent them. This means that it would also be plausible to hold that the referential reading is linguistically expressed and, in certain contexts (say, those in which the referent of description is unknown to speaker and hearer), it triggers the attributive reading. It’s not clear what justification, if any, there is for holding that the relationship goes one way rather than the other.
in particular, are properly treated as bearing on the truth conditions of the utterance they are associated with. Synecdoche is generally classified as a figurative use of language, and is often associated with antonomasia and metonymy.\textsuperscript{59} Synecdoche bears such a similarity to metonymy that it has been suggested both that synecdoche is a subclass of metonymy and that metonymy is a subclass of synecdoche.\textsuperscript{60} All three figurative uses of language – synecdoche, metonymy and antonomasia – have been viewed as types of metaphor.\textsuperscript{61} Synecdoche belongs to a family of communicative phenomena, and the philosophical analyses that are typically given of the members of that family cast some doubt on the notion that instances of synecdoche are relevant to the content of the propositions expressed by the utterances in which they are found.

Take, for instance, the following cases of antonomasia:

(3.26) Pollyanna is finally marrying Prince Charming.
(3.27) She refuses to let man’s best friend into the house.
(3.28) A bitter little spinster here at work won’t let me leave early.

Suppose (3.26) is uttered in reference to a Melanie, who is always sweet natured and upbeat, and her dashing fiancé, Brian, while (3.27) is uttered by a man whose wife refuses to allow him to get a dog and (3.28) is uttered in reference to the speaker’s middle-aged, unmarried boss who is known for her foul temper. There are, of course, many ways in which such utterances could be analyzed. But for a sizable number of thinkers, these three utterances would be classified as communicating something different from what was said via implicature. (3.26), for example, would be taken as, say, failing to express a proposition because the names used lack referents, but nonetheless generating an implicature that Melanie is finally marrying Brian. The two

\textsuperscript{59}McArthur (1992), pp. 402-404, 656, 1014; Lanham (1991)
\textsuperscript{60}McArthur (1992), p. 656
\textsuperscript{61}Ibid. p. 653
instances of antonomasia, on this sort of analysis, do not impact the truth conditions of the proposition expressed by an utterance of (3.26), but rather affect the content that is extra-linguistically communicated by it. The same sort of analysis would be given by many to both (3.27) and (3.28).

Now examine the following examples of metonymy:

(3.29) His only solace is the bottle.
(3.30) The pen is mightier than the sword.
(3.31) The suits on Wall Street ruined everything.

Again, (3.29)-(3.31) could easily be analyzed as cases in which the result of the metonymical transfer is found in the proposition communicated by an utterance, and not the proposition expressed. Indeed, this would seem to be the best analysis for handling (3.30). Surely we wouldn’t want to say that (3.30) explicitly expresses the proposition that literary power exceeds that wielded by the military (or that words are more powerful than physical violence). Intuitively, this is something that is communicated by the utterance, but not expressed by it.

Since synecdoche is of the same class as metonymy and antonomasia, it seems likely that cases of synecdoche would function in a similar fashion. Hence synecdochic transfer might very well impact the content communicated by an utterance, but not the proposition expressed. In Grice’s terms, it would bear on what was meant rather than what was said. Of course, this is not to say that the analysis we have suggested for (3.26)-(3.31) is undoubtedly the correct view, or that synecdoche must be treated as bearing on communicated content rather than linguistically expressed content. Rather, it is merely meant to demonstrate that there is a plausible view on which synecdoche is not truth conditionally relevant. The proponent of the synecdoche theory of descriptions must justify his assumption that synecdoche is truth conditionally relevant by
showing that an account which treats it as such is more plausible than one which treats it as impacting only what is communicated.

Recanati does briefly argue that cases of transfer (including synecdochic transfer) are best understood as impacting the truth conditions of an utterance, but his argument is rather unsatisfying. He rejects a Gricean treatment of these phenomena as counterintuitive, claiming that we do not hear, say, (3.29) as expressing the absurd proposition that the individual in question is only solaced by a unique bottle. Intuitively, he claims, we hear the proposition that the man is solaced only by strong liquor. I do not share his intuitions, however. Moreover, the fact that one might jokingly respond to (3.29) by asking “Oh, which bottle?” and the fact that the joke would be understood would seem to provide at least some indication that the absurd proposition is conveyed. Since it lines up with the linguistic meaning of the sentence uttered, it seems reasonable to assume that it is the proposition expressed. At the very least, the notion that the absurd proposition is expressed, and the intended meaning extra-linguistically communicated, should not be so hastily dismissed.

In fact, this notion seems even more plausible when we turn away from dead metonyms like those in (3.29) – (3.31) and focus instead on a live metonym, such as that in (3.32).

(3.32)  His only solace is his 8 inch French knife.\(^\text{62}\)

Suppose that (3.32) is uttered in reference to a chef. According to Recanati, such an utterance would not express the absurd proposition that the chef only finds solace in the knife. Of course, that proposition is most likely not the content that the speaker primarily intended to communicate in uttering (3.32); in this case, it seems plausible that the speaker intended to communicate the

\^\text{62} \text{Many thanks to Wayne Davis, both for pointing that a live metonym would make for a more convincing example and for the example itself.}
proposition that the chef’s only solace is cooking. Nonetheless, what we saw in the case of the
dead metonyms also holds here, as a joke stemming from the absurd proposition would be just as
likely in this case as it was previously.

Recanati’s synecdoche theory, then, seems to face a number of serious objections. But it
is not entirely without merit, for it hints at two interrelated avenues that might be pursued by
those wishing to account for definite descriptions. First, it relies, like the two-dimensional
theory, on the notion that a single linguistic meaning can result in the expression of different
propositions in different contexts. Thus, we might capture the intuition that the
referential/attributive distinction is relevant to the proposition explicitly expressed by an
utterance without resorting to semantic ambiguity. Second, the synecdoche theory hints at the
idea that there is something cognitively significant about the different uses of definite
descriptions and that revealing the cognitive significance of the various uses could help explain
their overall semantics and pragmatics. These two avenues are pursued more fully by the type of
theory we shall examine next – relevance theoretic accounts of definite descriptions.
Chapter 4
Relevance Theoretic Accounts

Relevance theoretic accounts of definite descriptions utilize much the same strategy as contextualist accounts. Like their contextualist cousins, relevance theoretic accounts assume a univocal semantics for definite descriptions which results in the expression of different propositions in different contexts. What makes the relevance theoretic accounts different is their reliance on the tenets of relevance theory to characterize the linguistic meaning of definite descriptions and explain how that meaning yields different propositions in different contexts of utterance.

A number of relevance theoretic accounts of definite descriptions have been recently been offered. These accounts differ from one another in several respects. Not only do the theoretical details differ significantly from account to account, but in some cases the explanatory scope or the theoretical goal also differs. Nonetheless, the core of these views, based on relevance theory, is the same across the different accounts. From the relevance theoretic viewpoint, a definite description, ‘the F’ has a single linguistic meaning, but this linguistic meaning is incomplete in the sense that a sentence containing a definite description fails to determine a proposition or complete thought independent of a context of utterance. The linguistic meaning of any sentence containing a definite description must be pragmatically supplemented in order for an utterance of that sentence to express a proposition. In some cases, this contextual supplementation leads to a referentially used description, in others it leads to an attributive use, or a generic use, etc.
In this chapter we will examine the two most developed accounts of definite descriptions which are based around this core relevance theoretic picture. Each view has its own problems, but as we shall see, there are some difficulties that seem to apply across the board. The observations we make about the theories evaluated in this chapter will then be used to aid in constructing a new relevance theoretic account of definite descriptions in chapter 6. Before discussing the various relevance theoretic accounts of definite descriptions, however, it is important to understand relevance theory itself. Let us begin, then, with a sketch of the relevance theoretic understanding of communication.

4.1 Relevance Theory

Sperber & Wilson’s relevance theory is an account of linguistic communication. Admittedly, this theory is not without its problems, and it may require expansion or minor modifications if it is to truly provide an adequate basis for explaining particular communicative phenomena. Chapter 5 is dedicated to laying out relevance theory in greater depth and providing a more detailed characterization of certain aspects of the theory in an attempt to eliminate certain weaknesses. The relevance theoretic accounts of definite descriptions currently on offer, however, are based on the standard, unmodified characterization of relevance theory. Hence it is that characterization that we will sketch briefly here.

---

63 This is a simplification. In truth, relevance theory is an account of a particular kind of communication. Linguistic communication falls within the class of communicative phenomena that relevance theory attempts to explain, but this class also includes non-linguistic communicative acts. We will address the precise scope of the theory in greater detail in chapter 5. For our present purposes, it is enough to know that definite descriptions, as linguistic phenomena, fall within the scope of the theory.
The relevance theoretic account of communication is based in the relevance theorist’s view of human cognition as a whole. According to relevance theory, human cognition is characterized by the Cognitive Principle of Relevance.

**Cognitive Principle of Relevance**

Human cognition tends toward the maximization of relevance.

The term ‘relevance’ as it occurs in this principle is a technical term that has little to do with the everyday sense of the expression. Relevance, here, is a ratio of benefits to costs; specifically, it is the ratio of cognitive effects to cognitive processing cost. In order to process an input to a cognitive process – like a visual experience, an auditory stimulus, or an assumption – the individual will have to expend cognitive effort. He or she will have to utilize cognitive resources, applying a portion of his or her limited energy and time to the task. The amount of effort that the individual must exert in order to process an input is the cognitive processing cost.

But why should any individual exert cognitive effort processing inputs? According to the relevance theorist, it is because those stimuli yield cognitive effects. A cognitive effect is a change in one’s belief set, a change in one’s understanding of the world. A new belief is a cognitive effect, as is the elimination of a belief from the set. The strengthening of the justification for an extant belief is a cognitive effect, as is the weakening of justification for a belief.

So, there are two factors determining the relevance of any given input to a cognitive processing system: the effects arising out of processing that stimulus and the cost of processing it. All else equal, the greater the number of effects that result from processing, the greater the relevance of the input. On the other hand, however, all else being equal, the greater the cost of
processing, the lower the relevance of the input. What the Cognitive Principle of Relevance essentially says, then, is that human cognition tends to be geared toward maximizing cognitive effects while minimizing processing costs.

It is from this claim about human cognition that the relevance theorist’s account of communication derives. Utterances are, after all, potential inputs to cognitive processing. If the Cognitive Principle of Relevance is correct, then, hearers will approach the interpretation of utterances in the same way that they approach the processing of other cognitive inputs – with an eye toward maximizing cognitive effects while minimizing processing cost. Given this fact, if speakers are to achieve their goal of communicating content, they will need to make their utterances relevant. The production of an utterance is in itself an indication of the speaker’s intention to convey content. Thus, when a speaker produces a communicative stimulus – like an utterance – the hearer is justified in assuming that the stimulus is relevant. This, according to the relevance theorist, is the principle which guides utterance interpretation – the Communicative Principle of Relevance.

**Communicative Principle of Relevance**
Every ostensive stimulus conveys a presumption of its own optimal relevance.

An ostensive stimulus is essentially a stimulus produced in order to communicate. An utterance is one form of ostensive stimulus. We will discuss ostensive-stimuli in greater detail shortly. What the Principle says, then, is that upon encountering an utterance or other ostensive stimulus, a hearer is justified in assuming that that stimulus is optimally relevant. But what exactly is the presumption of optimal relevance described in the principle? Sperber & Wilson’s characterization of the presumption of optimal relevance has undergone some revision since
relevance theory first appeared on the scene. The latest revision, and the one which is relied on by the two relevance theoretic accounts of descriptions we will examine, however, is detailed in (POR).

(POR) Presumption of Optimal Relevance
(a) The ostensive stimulus is relevant enough to be worth the audience’s processing effort, and
(b) It is the most relevant one compatible with the communicator’s abilities and preferences.\(^6^4\)

Relevance theorists take (a) to mean that for an utterance to be optimally relevant, processing it must yield enough cognitive effects to be worth the cognitive resources required to process it, and (b) to mean that the stimulus given must be the most relevant one that the speaker could produce, given his abilities and preferences. And, if the Communicative Principle of Relevance is correct, all ostensive stimuli indicate to the addressee that they meet these criteria.

We said earlier that an ostensive stimulus was a stimulus produced to communicate. There is actually quite a bit more to the notion than that. To understand it, however, one must understand the relevance theoretic picture of the communicator’s involvement in an instance of communication. According to relevance theory, a communicator engaging in linguistic communication does so in an attempt to fulfill two intentions – the informative intention and the communicative intention. The informative intention is an intention to make manifest (or more manifest) an assumption or set of assumptions to the audience. An assumption is manifest to an individual if it is possible for him or her to mentally represent the assumption and accept it as true, or at least probably true. The communicative intention is an intention to make the informative intention manifest to the audience. In essence, then, the relevance theorist views the

speaker engaged in communication as intending his audience to not only form a particular thought, but also recognize the speaker as intending that the hearer form that thought. This view of linguistic communication is very similar to that put forward by Grice.\textsuperscript{65} The main difference is that Sperber & Wilson put the informative intention in terms of making manifest rather than informing.

With the definitions of the informative and communicative intentions in hand, we can now give a more precise definition of an ostensive stimulus. An ostensive stimulus is a stimulus which is intended to make manifest an intention to make something manifest. That is, an ostensive stimulus is a one which is meant to make an informative intention manifest. In the simplest of terms, such a stimulus is one which is clearly communicative. Since ostensive stimuli are, by their very nature, clearly acts of communication, it is clear how they would convey their own relevance. Communication only succeeds when the hearer processes the communicative stimulus. As we noted above, a speaker who intended to communicate would have no reason to produce such a stimulus were it not relevant enough for the hearer to bother processing it. To do so would be to simply waste effort. Thus, when an ostensive stimulus is produced, the hearer is licensed in assuming that that stimulus is optimally relevant.

Together with the two principles of relevance, the relevance theorist’s attribution of the informative and communicative intentions to a speaker forms the core of the relevance theoretic account of linguistic communication. But there is one aspect of the theory yet remaining for us to address – the relevance theoretic view of the cognitive processing of an utterance by a hearer. According to relevance theory, there are two stages to utterance interpretation which occur in

\textsuperscript{65} Grice (1989)
two distinct modules of the mind. The first stage is the processing of linguistic meaning, which occurs in a linguistics processing module. The input to this module is the data given as output by the auditory or visual processing centers (depending on whether the utterance is spoken or written or signed). The linguistics module takes this data and constructs a mental representation of the linguistic meaning of the utterance, which relevance theorists call the logical form or LF. The LF, which includes both syntactic and lexical content, is the output of the module.

The linguistics processing module is limited with respect to the information it relies on in performing its task, however. In constructing the LF of an utterance, the linguistics processing module has access only to the hearer’s store of linguistic knowledge – syntactic and lexical information. As a result of these limits on the linguistic processing module, the LF of an utterance is often insufficient to determine a complete proposition, and it will not amount to a fully truth conditional thought. For example, the LF of an utterance containing an indexical expression will include the linguistic meaning of the indexical, but this alone will not be sufficient to determine a proposition. The referent of the indexical must be identified, and that requires contextual information which the linguistics processing module does not have access to. Similarly, an utterance containing other context sensitive expressions or syntactic or semantic ambiguities will have an LF that fails to determine a proposition.

Those features of the meaning of an utterance which cannot be determined by the linguistics processing module are filled in during the second stage of utterance interpretation – pragmatic processing. At this stage, those gaps in the content which are found in the input to

---

66 It may very well be possible for the relevance theorist to construct an account of utterance interpretation without assuming that the mind is modular. Sperber & Wilson’s account of utterance processing assumes the modularity of mind, however, and so we, too, will assume the mind is modular and proceed on that basis. This issue will be discussed in greater detail in chapter 5.
pragmatic processing – the LF – are filled by reference to information about the context of utterance and standard conversational practice, in Sperber & Wilson’s terms, the LF is enriched into a propositional form, that is, a mental representation of the explicit content of the utterance which does determine a proposition. According to relevance theory, this pragmatic processing is guided by the presumption of optimal relevance. Ambiguities are resolved, contextually sensitive meanings are determined, and referents are identified on the basis of the assumption that the complete interpretation of the utterance will be optimally relevant. Enrichment of the LF is only a part of what occurs during pragmatic processing, however. Guided by the presumption of relevance, pragmatic processing also determines what content, if any, is implicitly communicated by an utterance. The result of pragmatic processing, then, is a complete interpretation of the content conveyed, both explicitly and implicitly, by an utterance.

On the relevance theoretic picture, the interpretation of an utterance is a set of assumptions which can be categorized into two types roughly corresponding to the distinction between explicit and implicit content. Some of the assumptions comprising an interpretation are explicatures; others are implicatures. Relevance theory defines an explicature as an assumption which derives from the enrichment of LF. An implicature, on the other hand, is an assumption which can be inferred from the explicature(s) of an utterance together with what Sperber & Wilson call the contextual assumptions – that is, a set of background assumptions drawn from

---

67 In some cases, specific expectations of relevance – that is, expectations as to how the utterance will achieve relevance – may also guide the interpretation process. We will discuss specific expectations of relevance in chapter 5.

68 This is the ideal scenario. A complete interpretation is not always achieved when pragmatic processing is complete. In cases in which the utterance is not fully comprehensible to the hearer, the output of pragmatic processing may be a partial interpretation.
those assumptions which are mutually manifest to speaker and hearer – and could not be inferred from the contextual assumptions alone.

The foregoing is only a brief outline of relevance theory which glosses over many of the details. We will return to a more in depth discussion of relevance theory in chapter 5. For now, however, this standard sketch of relevance theory will be sufficient for examining previous attempts to use the theory to account for the referential/attributive distinction and the semantics of definite descriptions. It is to this task we will now turn.

4.2 Bezuidenhout’s Account

Anne Bezuidenhout’s account of definite descriptions differs in focus from the other accounts on offer. While the accounts we have looked at so far focus solely on descriptions, Bezuidenhout’s account is meant to encompass both definite descriptions and indexicals. The reason Bezuidenhout extends her focus to indexicals is that the main target of her theory is not the semantics of definite descriptions, but rather the referential/attributive distinction. And according to Bezuidenhout, indexicals exhibit referential and attributive uses in the same way that descriptions do.

Before moving into the details of Bezuidenhout’s account, it is important to understand the theoretical framework she is operating with. First, she is working under the assumption that referentially used descriptions are referring expressions, whereas attributively used descriptions are denoting expressions. In other words, she takes utterances containing referentially used descriptions to express object-dependent propositions and utterances containing attributively used descriptions to express general propositions. An object-dependent proposition is not, on
Bezuidenhout’s view, the same as a singular proposition. Thus, she is not claiming that the
descriptive content of the description is completely absent from a proposition expressed by an
utterance containing a referentially used description. Rather, her claim is that such an utterance
expresses a proposition containing what she calls an *object-dependent mode of presentation*.

It is less clear what Bezuidenhout takes to be the nature of propositions expressed by
utterances containing attributively used descriptions. She clearly holds that such propositions are
general propositions, but it is not entirely clear what she takes their precise nature to be. Her
classification of attributively used descriptions as denoting rather than referring would seem to
indicate that such descriptions express the existentially quantified complexes envisioned by
Russell. She also claims, however, that the existential presupposition associated with a
description is pragmatically determined and is not always present in the meaning recovered by a
hearer.  

Since we do not normally think of presuppositions as part of the proposition expressed
by an utterance, this would seem to indicate that the existential claim is not part of the
proposition expressed by an utterance containing an attributively used description, but is rather
something communicated in addition to that proposition. In all likelihood, then, Bezuidenhout is
assuming that the proposition expressed by an utterance containing an attributively used
description contains something like a descriptive concept – a mode of presentation that is not
essentially tied to any particular object.

In addition to these basic assumptions about how to best characterize the general
phenomena of uses of definite descriptions that we are attempting to explain, Bezuidenhout’s
account also takes for granted that a particular approach to incomplete descriptions is more or

---

69 Bezuidenhout (1997), p. 394
less correct. Since incomplete descriptions can be used either attributively or referentially, she argues, all incomplete descriptions require a contextual delineation of the domain from which their referent is to be drawn. And Bezuidenhout seems to treat this particular type of interaction between the context of utterance and the proposition ultimately expressed as completely distinct from her own account of the referential/attributional distinction.

Perhaps the most important aspect of the framework Bezuidenhout is operating with is the constraints she places on an adequate account of the referential/attributional distinction. First, an adequate account must distinguish referential and attributional uses of an expression on the basis of something other than whether or not that expression requires contextual completion. Second, any adequate theory must account for attributional uses of indexicals.70

The justification for the first adequacy constraint lies in Bezuidenhout’s acceptance of the notion that all incomplete descriptions require contextual completion, whether they are referentially or attributionally used. An incomplete definite description, remember, is a description that fails to uniquely identify a referent. As we have seen, all the accounts of how such descriptions succeed hold that they pick out their referents by drawing on information in the conversational context. Since both referentially and attributionally used descriptions might require supplementation from the context in order to express a proposition, we cannot use the need for such supplementation to mark the difference between the types of uses.

The second constraint arises out of Bezuidenhout’s acceptance of the notion that indexicals can be used both referentially and attributionally. Bezuidenhout offers a number of

---

70 There is another adequacy constraint set forth by Bezuidenhout which states that an adequate account must deal with the distinction on the basis of pragmatics. Since all the accounts at issue in this chapter are pragmatic accounts, this constraint is not relevant to our present discussion. See Bezuidenhout (1997), p.385
examples of this phenomenon that were originally offered by Nunberg and Schiffer. One of Schiffer’s examples asks that we imagine a context in which we have just stumbled upon a huge footprint in the sand and I exclaim (4.1).

(4.1) He must have been a giant!

The interpretation of ‘he’ in (4.1) does not seem to fit the standard model of indexical reference. Rather, what I seem to mean in uttering (4.1) is that the man who made the footprint, whoever he is, must be a giant. My use of ‘he’ in this case, then, would seem to be attributive. If indexicals can have both referential and attributive uses, Bezuidenhout argues, then any account of the referential/attributive distinction should include an explanation of its application to indexicals.

Bezuidenhout’s account is geared toward meeting the second of these constraints by providing a unitary analysis of both definite descriptions and indexicals. The core of her account is, of course, the standard relevance theoretic line: the LF of utterances containing definite descriptions or indexicals fails to determine a proposition. This LF serves as the input to pragmatic processing, where, on the basis of the hearer’s beliefs and assumptions in the context of utterance, the LF is altered and supplemented in various ways until the completed process yields a proposition. Bezuidenhout characterizes this process in terms of levels of processing and interpretation. The LF of the utterance is the first level of interpretation and is the result of the first level of processing (syntactic and semantic processing). The explication yielded at the end of pragmatic processing – which she identifies with the proposition expressed – is the second level of interpretation, and the pragmatic processing that yields it is the second level of processing.

---

71 Nunberg (1993); Schiffer (1995)
According to Bezuidenhout, both descriptions and indexicals are analyzed in terms of descriptive concepts at the first level of interpretation. In the case of indexicals, this means that the linguistically encoded meaning of ‘I’ is something like “the utterer of this token”, while the meaning of ‘she’ is “the female in the context”. The descriptive analysis of the expression at this level is the same for both referentially and attributively used expressions. In a significant way, then, Bezuidenhout’s account gives one semantic analysis for both referentially and attributively used descriptions. It is the second level of processing – the pragmatic processing – that determines whether an utterance containing a definite description or indexical expresses an object-dependent proposition or a general proposition.

While Bezuidenhout does not provide us with a detailed characterization of the LF of utterances containing indexicals, she does provide us with a sketch of the LF of utterances containing definite descriptions. On her account, the LF of the form *The F is G* looks something like (4.2):

\[
(4.2) \quad \text{[[Feature F is instantiated <uniquely/accessibly> by an x which is G]]}
\]

With the exception of the features of uniqueness and accessibility, all of the parts of the LF are concepts or conceptual elements structured in a specific way. Uniqueness and accessibility, on the other hand, are not conceptual elements, but rather indicate procedural constraints on the pragmatic derivation of the referent of the description. According to Bezuidenhout, ‘the’ marks two important aspects of the referent it aids in identifying – first, that the referent is unique in its satisfaction of the description used. Of course, this uniqueness is relative to a contextually set domain. The second thing that the definite article indicates about the referent of a description is that the referent is easily accessed in the context. The way that the definite article succeeds in so
marking the referent of a description is via the procedural instructions it encodes in the LF of any utterance containing a definite description. These instructions indicate to the pragmatic processing center that the procedure of identifying the referent of the description must identify a referent with the properties of uniqueness and accessibility.

According to Bezuidenhout’s account, this LF is an input to pragmatic processing, and the end result of that processing is either an object-dependent proposition (in the case of a referentially used description) or a general proposition (in the case of an attributively used description). Bezuidenhout’s characterization of this level of interpretation leaves a good many questions unanswered, but the basic picture is that drawn by relevance theory. Pragmatic processing proceeds under the guidance of the presumption of optimal relevance. As a result of this assumption, the processing of the utterance is directed toward uncovering an interpretation that is optimally relevant. According to Bezuidenhout, “An interpretation of an utterance is optimally relevant if it has adequate cognitive effects for no gratuitous processing effort.”\(^{73}\)

There will, on this account, be several important tasks performed during the pragmatic processing of an utterance of the form *The F is G*. Obviously, the pragmatic processing system must determine how the description is being used. In addition, the domain of objects that the referent of the description is to be drawn from must be demarcated. If the description is being used attributively, these two tasks will be sufficient to yield a proposition. If the description is being used referentially, however, an object-dependent mode of presentation must be selected to replace the descriptive concepts and the variable in the grammatical subject of the utterance, and that replacement must be performed, yielding an object-dependent proposition. The result of the

performance of each of these tasks is constrained and guided by the presumption that the final product of processing should be optimally relevant.

As Bezuidenhout sees it, the task of determining how a description is used requires determining the speaker’s intentions with respect to the use of the descriptive information in the LF. The general idea is that speakers using descriptions have what Bezuidenhout calls a *directive intention*, that is, an intention as to how the hearer should use the descriptive information in the LF during pragmatic processing. If the speaker intends the information to be used to identify the referent, but not to occur in the proposition expressed, his directive intention is that the information be used in an *identifying way*. If he intends the information to be used as part of the proposition expressed – as the mode of presentation under which the hearer thinks of the description’s referent – his directive intention is that the information be used in a *criterial way*. As should be clear, a directive intention for identifying use is an intention that the hearer take the description to be used referentially and thus interpret the utterance as expressing an object dependent proposition, while a directive intention for criterial use is aimed at an attributively used description and a general proposition. On this account, then, the distinction between referential and attributive uses of descriptions is marked by speakers’ intentions. Which type of intention the speaker has determines which proposition he means to express; and which type of intention the hearer takes the speaker to have determines which proposition he takes the speaker to have expressed.

According to Bezuidenhout, the hearer’s identification of the speaker’s intention is done via inference on the basis of various contextual assumptions. Of particular import will be assumptions about which facts are accessible in the context, which propositions the speaker
knows, and which propositions the speaker takes the hearer to know. The identification of the speaker’s directive intention is done on the basis of assumptions present in the context that are themselves about the context. In other words, in order for the hearer to identify the speaker’s directive intention, he must have mental representations of the conversational context and also be able to draw conclusions about the speaker’s intentions from those representations. And this process is, like all other pragmatic processes, guided toward obtaining the optimally relevant utterance interpretation.

Bezuidenhout views the pragmatic processing of utterances containing indexicals as occurring in much the same way the processing of descriptions does, though some differences arise in cases in which indexicals can be put to uses that descriptions cannot. Since our main focus at this point is descriptions, we will leave the discussion of this aspect of Bezuidenhout’s account for another time. Nonetheless, it is important to note that the ability of the theory to account for both definite descriptions and indexicals is a significant point in its favor.

4.3 Problems with Bezuidenhout’s Account

Like any theory, however, Bezuidenhout’s account has flaws. The first and perhaps most troublesome deficiency in the theory is the lack of detail and precision in the picture it offers of the pragmatic processing of descriptions. This is a deficiency which Bezuidenhout’s theory likely inherits from relevance theory itself. Recall that Bezuidenhout’s understanding of optimal relevance is as the property of yielding sufficient cognitive effects to be worth the processing effort. But this definition is hopelessly vague and, as a result, of virtually no use in applying the account to specific cases. What is a cognitive effect worth? How is the worth of cognitive
effects determined? How is the amount of processing effort determined? And how do we know whether a certain amount of cognitive effects is worth a certain amount of processing effort? These are questions about the meaning of concepts central to the definition, and in the absence of answers to them, the concept of optimal relevance on offer would seem to be largely devoid of content.

Remember, though, that an interpretation with the property of optimal relevance is supposed to be the goal of the pragmatic processing of an utterance, and that the optimally relevant interpretation is the interpretation that relevance theory predicts will be the natural interpretation. Indeed, optimal relevance is a core concept of the overall theory, and in the absence of a precise definition, two monumental gaps appear in the account of definite descriptions. The first gap is in the theory’s explanation of pragmatic processing level of utterance interpretation. Such processing is supposed to be guided by an expectation that the final interpretation will be optimally relevant. What these processes look like and which intermediate interpretive results they reach in a specific case of utterance interpretation, then, would seem to be shaped by the expectation of optimal relevance. If we do not have a clear picture of what optimal relevance really is, though, our explanation of the pragmatic processes in the case of particular utterances (like utterances containing definite descriptions) will be incomplete and shallow.

The second gap in the theory of descriptions which results from a vague definition of optimal relevance opens up in the area of providing evidence for the theory. Theories are disconfirmed or backed up by confirming evidence on the basis of testable predictions. If a theory predicts that a certain situation will obtain and that situation does obtain, that lends some
evidential support to the account. If the theory predicts a certain situation and it does not obtain, then the theory, or some part of it, is disconfirmed. Thus, proper support for a theory requires that the theory be capable of generating testable predictions. Since, according to relevance theory, which interpretation of an utterance is the natural interpretation is a matter of which interpretation is optimally relevant, the ability of the theory to generate testable predictions hinges, at least in part, on the possibility of identifying the optimally relevant interpretation. Of course, this is not possible if we do not have a clear understanding of what is required for optimal relevance.

It might be argued that, although we cannot construct an argument whose conclusion is a prediction and whose premises are the core claims of the theory and the facts about a particular case, we can still use the theory to generate testable predictions about specific cases on the basis of our intuitive judgments as to whether or not an interpretation meets the criteria for optimal relevance. In fact, Bezuidenhout attempts to perform this task herself, and she indicates that it should be possible to perform the same task for other specific examples of referentially or attributively used expressions. Here is the first part of her argument:

In all the examples I gave above, it should be possible to explain why a certain interpretation is the correct one by appeal to considerations of optimal relevance. Take Schiffer’s example “He is a giant” said in the presence of a huge footprint in the sand. This is presumably highly visually salient, and so will be a part of the listener’s mental context. Moreover, if it is categorized as a human male footprint, its size will suggest that the person who left the footprint is larger than average. All this information will be a part of the listener’s cognitive environment. In this cognitive environment, the speaker can assume that the listener will enrich the character associated with the pronoun “he” so as to yield the representation [[male who left this huge footprint in the sand]]. This will be an interpretation which can be accessed with little processing effort.\footnote{Bezuidenhout (1997) p. 398}
All of the facts that she lists as part of the context are intuitively linked with the ultimate interpretation of the utterance. And the steps of inference that she gestures at are also intuitively part of the process linking the context and the utterance to the interpretation. That much of her explanation is backed by appeal to intuition. But this set of facts alone provides us with no reason, intuitive or otherwise, to conclude that the interpretation requires little processing effort. Bezuidenhout describes the tools that would be used by a hearer in this situation, but she gives us no indication of how those tools are to be used. How can we possibly draw a conclusion as to the amount of effort required to perform a task when we have no idea what performance of the task involves? The fact that certain assumptions are salient in the context tells us little unless we know how just how salient those facts are, how salience affects processing effort, and what has to be done with the assumptions in question. Making an intuitive judgment may not require the possession of precise data, but it does require the possession of relevant data. If you ask me to judge how hard it will be to push a particular stone uphill, I may not need to know the precise mass of the stone, but my intuition will be of no use to me if I do not have some sense of the stone’s dimensions and density.

So the claim about the processing cost of the interpretation in the example is not, as the objector would have it, backed by intuition. This is half of what is involved in optimal relevance. What about cognitive effects? We might be able to count the effects of a particular interpretation. But even on the assumption that we can, what is important in the case of determining optimal relevance is that the number of cognitive effects be sufficient to be worth the processing effort required. Even if you have been given a list of the cognitive effects and you know their precise number, however, this would give you no reason to make a judgment about whether those effects
were worth the cost of processing. After all, we have been given no clear idea of what processing
is worth relative to cognitive effects and we have no real idea of how much processing is actually
required. It would seem, then, that intuition is useless with respect to applying this half of the
notion of optimal relevance as well.

The problem of demonstrating that the natural interpretation of an utterance is the
optimally relevant interpretation is exacerbated by the fact that the account does not provide us
with a clear definition of the notion of an interpretation. Generally, relevance theory views the
interpretation of an utterance as including all of its various explicatures and implicatures, and so
we would seem to be justified in using that understanding of an interpretation with respect to
Bezuidenhout’s account. But her attempt to apply her account to a specific example focuses
solely on the relevance of the proposition expressed – the explicature (presumably the primary
explicature). Of course, this limited focus makes sense in light of the fact that she is primarily
interested in the propositional form of utterances containing definite descriptions or indexicals,
but it is entirely possible that different uses of descriptions or indexicals yield interpretations that
differ with respect to more than just the content of the propositional form. To demonstrate that
the natural interpretation of an utterance is in fact the optimally relevant interpretation, then, we
need to show that the cognitive effects generated by the entire interpretation are worth the
processing cost for the entire interpretation. This means that we must identify the cognitive
effects and processing cost for each explicature and implicature that makes up the interpretation,
and the difficulties we face in making these determinations in the case of the propositional form
will arise again for each of the other aspects of the interpretation. It would seem, then, that the
prospects of determining which interpretation is optimally relevant on the basis of Bezuidenhout’s account, and thereby deriving testable predictions from the theory, are bleak.

In fact, it is not clear that Bezuidenhout’s account would predict a single natural interpretation of an utterance containing a definite description. She defines optimal relevance, recall, as the property of producing adequate cognitive effects without demanding gratuitous processing effort. But surely it is possible for there to be multiple interpretations of an utterance which do not require gratuitous processing effort, and yet yield sufficient cognitive effects to be worth the effort. Bezuidenhout’s characterization of optimal relevance does not guarantee that a unique interpretation of an utterance will be optimally relevant. And yet, in general, we intuitively select only one interpretation as the meaning of an utterance.

Even if the theory could generate testable predictions as to the unique natural interpretation of an utterance, there is good reason to think it would make the wrong predictions. The reason for this lies in Bezuidenhout’s analysis of the LF of utterances containing descriptions. On this analysis, the LF of any utterance of the form The F is G is composed of descriptive information which can be used in one of two ways during the pragmatic processing phase of interpretation: it can become part of the proposition expressed or it can be used to identify the speaker’s intended referent and thereby lead the hearer to select an object-dependent mode of presentation of that referent as a component of the proposition expressed. These two different uses lead to the expression of different propositions, and they would also seem to require a different number of cognitive steps. Both attributively used descriptions and referentially used descriptions require determination of the speaker’s directive intention and delineation of the relevant contextual domain. In the case of an attributively used description,
where the descriptive information in the LF is put to use as a constituent of the proposition expressed, these two cognitive processes are all that is required. Once it is determined that the descriptive content present in the LF will be part of the proposition expressed, the LF needs only be enriched to indicate the relevant contextual domain before it determines a propositional form which is the output of pragmatic processing. When the descriptive information in the LF is being used in a criterial way, no processing need be done on the descriptive aspect of the LF as it is enriched into a proposition. Essentially, the criterial use of the descriptive information is tantamount to leaving that information alone.

In the case of referentially used descriptions, however, the two steps that are required in the enrichment of utterances containing attributive descriptions must be performed, but they are not enough to determine a proposition. The reason is that in this case, the determination of the speaker’s directive intention triggers a pragmatic process via which the speaker’s intended referent is identified, an object-dependent mode of presentation of that referent is selected, and the descriptive information in the LF is replaced with the selected mode of presentation. We could easily claim that three extra processes are required to enrich the LF of an utterance containing a description into an object-dependent proposition, but the problem will still arise even on the assumption that those three seemingly distinct steps are in fact one process, since this still leaves us with the conclusion that the referential interpretation of an utterance requires the same processes that the attributive interpretation does plus an additional process. What this means is that the referential interpretation will always require more processing cost than the attributive interpretation.
The disparity in the processing cost of the different interpretations becomes an issue when we think about the number of cognitive effects likely to be produced by either interpretation. Neither interpretation seems to convey more information than the other, since both convey the claim that a certain object has a certain property. Nor does there appear to be any reason why one interpretation would serve to influence the previously held assumptions of a hearer any more than the other. Of course, there will certainly be particular contexts where the different interpretations will have different effects, but there does not appear to be any reason to suspect that, in general, one interpretation would have more cognitive effects than the other. What this means, however, is on Bezuidenhout’s account the attributive interpretation should generally be more relevant than the referential interpretation, and hence should generally be the natural interpretation of utterances containing descriptions. But this is not the case. If anything, the reverse is true – descriptions are more often used referentially than they are attributively. Bezuidenhout’s theory, then, would seem vulnerable to incorrectly predicting the preferred interpretations of specific utterances.

This vulnerability of Bezuidenhout’s theory appears to stem from the fact that her account seems to favor the attributive reading of descriptions. Indeed, her characterization of the LF of utterances containing descriptions depicts it as essentially an incomplete general proposition that includes instructions as to how it should be filled into a full general proposition. The descriptive component of the general proposition expressed by an attributively used description is, on this view, already present in the LF of any descriptive utterance, regardless of how the description in the utterance is ultimately used. So, while very little need be done to the LF to enrich it into a general proposition, quite a bit must be done to enrich it into an object-
dependent proposition. Of course, the same problem would arise in reverse if we were to propose an account that favored the referential interpretation of descriptions. What this would seem to suggest is that an account of definite descriptions that will make accurate predictions as to the preferred interpretations of utterances in context should avoid favoring one interpretation over the other.

There might be a way for Bezuidenhout to alter her account such that it avoids favoring the attributive interpretation of definite descriptions over the referential. But this is not the only problem facing the account. Bezuidenhout’s view gives us an account of the LF of utterances containing descriptions which does not appear consonant with the content expressed by attributively used descriptions. In other words, it is not clear that the LF as Bezuidenhout describes it would actually be enriched into a proposition matching that which we would intuitively ascribe to an utterance containing an attributively used description.

Bezuidenhout’s picture of the LF of an utterance of the form *The F is G* is given in (4.2).

(4.2) \[\text{[Feature F is instantiated <uniquely/accessibly> by an x which is G]}\]

Given this understanding of the LF, an utterance of (4.3) would have the LF indicated in (4.4)

(4.3) The cat is black.
(4.4) \[\text{[[Being a cat is instantiated <uniquely/accessibly> by an x which is black]]}\]

But the LF given in (4.4) represents a different grammatical structure than that exemplified by (4.3). In (4.3) *being black* is predicated of *the cat*. In (4.4) in contrast *being instantiated by an x which is black* is being predicated of *being a cat*. We might be comfortable with claiming that *the cat* and *being a cat* amount to the same thing, but the LF given by (4.4) involves an entirely different predicate than that occurring in (4.3). (4.4), then, cannot be the LF of (4.3).
In addition to the concerns about testability and disconfirmation, one might also worry that Bezuidenhout’s account is unable to account for the uses of definite descriptions other than referential and attributive uses. Remember, not only can definite descriptions be used referentially or attributively, but they can also be used functionally, generically or specifically. It’s not clear exactly how Bezuidenhout’s account would allow her to deal with such uses of descriptions.

Bezuidenhout’s theory also faces a potential objection on the basis of theoretical parsimony. According to Bezuidenhout, we can only determine which type of mode of presentation will occur in the proposition expressed on the basis of conclusions about the speaker’s intentions, and those conclusions about the speaker’s intentions are drawn on the basis of assumptions about the context which are themselves available in the context. But one might wonder why it wouldn’t be possible for the hearer to select the correct mode of presentation on the basis of contextual assumptions — perhaps even assumptions about the context. That is, it seems as though the hearer might very well be able to use contextual cues alone to arrive at a representation of the proposition the speaker intended to express without first determining what the speaker’s intentions are. The easy access of certain assumptions or concepts could guide the hearer to the speaker’s intended meaning, and speakers who wish to succeed in communicating are certainly likely to take their hearer’s contextual assumptions into account when crafting their utterances. If this is possible, then Bezuidenhout’s characterization of the pragmatic processing of descriptions and indexicals is more complex than necessary, and the hearer’s determination of the speaker’s directive intention becomes superfluous. In accordance with the notion that theoretical parsimony is preferable, then, this aspect of the theory should be discarded.
As with any theory, Bezuidenhout’s account of the referential/attributive distinction has its problems. Its main problem is the problem of testability – which seems to be the result from the account’s reliance on vague definitions of key concepts, and would plague any relevance theoretic account that did not first sharpen the characterization of those concepts. There are several problems that are peculiar to Bezuidenhout’s account, however. First, it looks as though the account would face disconfirmation even if it were capable of generating testable predictions and second, it seems to include notions that are not explanatorily necessary.

4.4 Powell’s Account

George Powell has offered a relevance theoretic account of definite descriptions and the referential/attributive distinction that appears to avoid some of the problems facing the views of Recanati, Ramachandran, and Bezuidenhout. Its unique view of the LF of definite descriptions allows the account to avoid a variety of difficulties and makes for an elegant and simple theory. On the other hand, the account also leaves many questions unanswered, and it appears to fall prey to the same testability worry that has so far plagued the relevance theoretic approach in general.

As with any relevance theoretic account of definite descriptions, Powell’s theory takes the semantics of definite descriptions to be univocal but insufficient to determine a proposition. Where his view diverges from Bezuidenhout’s is in its depiction of the LF of definite descriptions. In contrast to the other relevance theoretic account on offer, Powell’s theory views the contribution of a definite description to the LF of the utterance in which it occurs to be purely procedural in nature. On this view, definite descriptions encode an instruction to retrieve an
individual concept which is compatible with the content of the description, and this instruction exhausts the semantic contribution of the description. The selection of the individual concept occurs at the level of pragmatic processing. An individual concept is simply a concept that is taken by its possessor to represent an individual or set of individuals. According to Powell, such concepts contain information that uniquely identifies their referents and can be either descriptive or de re. Whether a definite description is interpreted referentially or attributively depends on which type of concept is selected – the selection of a descriptive concept yields an attributive interpretation while the selection of a de re concept yields a referential interpretation.

As individual concepts, descriptive concepts and de re concepts both contain content that uniquely identifies their referents and both refer to individuals or sets of individuals. The difference between a descriptive concept and a de re concept lies in the relationship the concept bears to its referent. In the case of descriptive concepts, the relationship is satisfactional; the referent of a descriptive concept is the individual or set of individuals that satisfies the concept’s descriptive content, and such satisfaction is all that is necessary to link the concept to its referent. In the case of a de re concept, however, the relationship between concept and referent is generally thought to involve a causal connection. A variety of ways of cashing out this causal relationship between de re concepts and their referents have been suggested, and there is considerable debate as to the best way to characterize the relationship. Powell holds that de re concepts must be linked to their referents via a perception-based causal chain.

Because descriptive concepts and de re concepts are related to their referents in different ways, they will differ in their effects on truth conditions. The truth conditions of a proposition involving a descriptive concept will be dependent on the individual that satisfies the descriptive
content of the concept. The truth conditions of a proposition involving a *de re* concept will be
dependent on a particular individual, namely the one that is anchored to the concept via a causal
chain. And this difference between propositions involving descriptive concepts and propositions
involving *de re* concepts appears to be just the difference we see between attributively used
descriptions and referentially used descriptions. The truth conditions of the propositional form of
an utterance containing an attributively used description depend on whatever object satisfies the
description, while the truth conditions of the propositional form of an utterance containing a
referentially used description are dependent on a particular object largely independently of
whether it satisfies the descriptive content of the concept or not. Thus, on Powell’s view,
descriptions are interpreted attributively when a descriptive concept is selected and interpreted
referentially when a *de re* concept is selected.

Whether a descriptive or *de re* concept is selected during pragmatic processing is,
according to Powell, guided by considerations of relevance. Which type of concept is selected
depends on which interpretation is optimally relevant – that is, yields the most cognitive effects
for the least processing effort. Powell does not give us a more precise characterization of optimal
relevance, nor does he provide us with any standard by which to evaluate the processing costs of
various utterances. This would seem to leave us in the same position we faced with
Bezuidenhout’s account – without a justifiable means of using the theory to predict which
interpretation would be preferred in a given case. While we might claim that certain
interpretations are easily processed or yield a good number of effects, these claims would be
largely unsubstantiated.
In attempting to apply his theory to different cases, however, Powell may very well have pointed the way toward constructing a relevance theoretic account that is testable. He attempts to apply his theory to (4.5) uttered in a context in which speaker and hearer are in a pit lane, watching Jones scheme with the organizer of the race, and it is mutually manifest that Jones is the Ferrari driver.

(4.5) The Ferrari driver has an unfair advantage.

Assume that the description “the Ferrari driver” is being used referentially here to pick out Jones. With respect to how the hearer would select the relevant individual concept in this case, Powell tells us:

There are two…obvious candidates, both of which are likely to be reasonably salient: one is a descriptive concept something like ‘the Ferrari driver in this race’ and the other is a de re concept of Jones, standing in front of him. The question is, which one does he accept? And this answer to this, on a relevance theoretic picture, is going to depend on which concept offers him more contextual effects. It’s not hard to see, then, which this will be in the proposed context: since the hearer believes Jones to be the Ferrari driver, any contextual effects carried by the descriptive interpretation come along for free with the de re interpretation. However, the de re interpretation also has additional contextual effects: it might, for instance, contain the information that Jones is wearing a red helmet, in which case the hearer could infer that the driver with a red helmet has an unfair advantage and so on. In this context, then, the de re interpretation is going to be optimally relevance since it offers the hearer more contextual effects for no extra processing effort.\(^\text{75}\)

Now, in order to justifiably draw the conclusion he reaches, Powell would have to do a good deal more work. His application of the theory assumes the salience of the two concepts in question without giving us any account of what makes for salience and it also assumes, without providing any significant justification, that the de re interpretation has cognitive effects that the descriptive interpretation does not and that the two interpretations have the same processing cost.

\(^{75}\) Powell (2001), p.120
What is promising here is the method that Powell uses to determine which interpretation is optimally relevant. More often than not, relevance theorists claim that the processing cost of one interpretation would be low while the cost of another interpretation would be higher and that one interpretation has more cognitive effects than another. But without a numerical assessment of the cost and the effects, such claims are not enough to conclude that one interpretation is more relevant than another. Powell’s application of the theory, however, holds one aspect of the determination of relevance fixed across the interpretations being evaluated – he assumes that the different interpretations have the same processing cost and then seeks to show that one interpretation has more cognitive effects than the other, thus making it more relevant. If this method is truly available in the case of definite descriptions, that is, if we can reasonably assume that a set of potential interpretations are identical on one dimension of the relevance ratio, then it may be possible to compare the interpretations along the other side of the ratio and thereby determine which interpretation is optimally relevant.

As it stands, however, Powell gives us no reason to think that it is actually possible to use this method of evaluating the relevance of utterance interpretations, and his claims about the salience of interpretations and their cognitive effects are unjustified. Since Powell does not provide us with a detailed account of how individual concepts are selected or the basis on which they would be chosen, any claims one would be inclined to make with respect to which interpretations the theory predicts as the natural interpretation would be unsubstantiated. Some significant additions to the theory and a good deal of argumentative work would have to be done, then, before his relevance theoretic account of definite descriptions could truly be considered a testable theory.
This problem with testability is not limited to just the pragmatic aspect of the account. Powell also leaves us without a clear understanding of one of the key concepts of the semantic component of the theory. The account tells us that definite descriptions encode an instruction to retrieve an individual concept that is compatible with the content of the description used. If we are to use the theory to predict an interpretation of a particular definite description, however, we need to know what counts as being compatible with the content of the description and what does not. Unfortunately, we are given no clarification of the notion of compatibility by Powell. Of course, the notion that Powell is trying to get at is one of similarity of content. We want the individual concept selected to have a mode of presentation consistent with the description used. A concise definition of compatibility is necessary to ensure that the theory properly predicts which concept will be selected for a particular use of a definite description. Without such a definition, we cannot use the account to generate predictions as to the interpretation of an utterance containing a definite description, even if we have a complete and well defined account of the pragmatic side of the equation.

Despite its problems with testability, there appear to be a number of ways in which Powell’s account is superior to its competitors. After all, while it may not be capable of generating testable predictions as it stands, it does avoid the problems with theoretical parsimony faced by Recanati and Bezuidenhout’s theories. Unlike Recanati’s theory, it does not require us to add new linguistic items to our theory. Instead, it relies on theoretical constructs that are already well established and in use in other theories. And unlike Bezuidenhout’s account, Powell’s account does not require us to drag in questions about the speaker’s intentions. This allows Powell’s account greater theoretical parsimony. The way Bezuidenhout sees it, we must
discern speaker’s intentions and then use that to determine how the description is being used. But discerning speaker’s intentions requires drawing inferences based on the context of utterance. On Powell’s view, on the other hand, we can use those same contextual clues to determine how the description is being used without first having to infer what the speaker’s intentions are. Bezuidenhout’s account requires an extra step in processing interpretations which Powell’s does not.

Powell claims an additional benefit for his account, arguing that it can handle more than just the referential and attributive uses of definite descriptions. Functional uses of definite descriptions, he claims, can also be covered by the account. A functional use of a definite description, remember, is one that refers to an individual fulfilling a role or function, such as the use of ‘the President of the United States’ in Powell’s example of (4.6).

(4.6) The President of the United States changes every five years.

According to Powell, uses of definite descriptions like that in (4.6) should be interpreted metarepresentationally. On his view, an utterance of (4.6) “… will express a proposition to the effect that there is a denotation of the descriptive individual concept corresponding to ‘the President of the United States’ and it changes every five years.”

While Powell’s proposal does appear, at first glance, to adequately explain what is occurring when descriptions are used functionally, it is not clear why this should be taken to be a benefit of his theory, since the metarepresentational interpretation of functionally used definite descriptions does not appear to be a consequence of the theory, but rather an awkward addition to it. Remember that, according to Powell, definite descriptions encode an instruction to retrieve

---

76 Powell (2001), p. 122
an individual concept. In both the referential and attributive cases, the concept is retrieved for
*use* in the propositional form being constructed. In the functional case, if Powell is right, the
class is being mentioned and something stated about it. It is not clear, however, that one and
the same instruction for concept retrieval would lead to a use of the concept in one case and a
mention of it in another.

Even if Powell’s account can handle functional uses of definite descriptions, it is still
unable to cope with generic uses (such as ‘the grey wolf’ in ‘The grey wolf ranges throughout
the American Southwest’). Nor does he give us an account of specific uses – those uses in which
the speaker has a particular individual in mind but does not expect the hearer to identify that
individual. Since Powell appears to be mainly targeting the referential/attributive distinction and
only indirectly addressing the semantics of definite descriptions, the other uses of definite
descriptions would appear to be tangential to his main project. If we are to give an accurate
account of the semantics of definite descriptions, however, we must take all the various uses of
such descriptions into account. This is something that Powell’s account, like its fellow
contextualist views, is unable to do.

In sum, then, Powell’s account faces the same problems with testability and vagueness
that Bezuidenhout’s account does. And, like the other relevance theoretic accounts, it cannot
handle uses of definite descriptions beyond the referential/attributive distinction. On the other
hand, the account does seem to point the way to a strategy for achieving testability, and it makes
some headway in attempting to deal with functional uses of definite descriptions. Additionally, it
avoids the problems with theoretical parsimony faced by its competitors. So while Powell’s
account may have its shortcomings, it would also seem to point us in the right direction.
As we have seen, the relevance theoretic accounts on offer seem to face similar difficulties. Perhaps the most troubling of these is a lack of testability, which seems to be due, largely, to vagueness in the definitions of key terms and descriptions of key aspects of the theory. Moreover, none of the accounts can handle all the uses of definite descriptions. In addition to these rather devastating issues, each of the accounts faces its own particular problems. But each account also has its own benefits, and we may be able to use the beneficial aspects of the theories while being wary of the pitfalls they face in our analysis of definite descriptions. From Powell’s account, in particular, we may be able to draw suggestions as to how to construct a testable relevance theoretic account of definite descriptions. Before we can do any work on the account of definite descriptions, however, we must attempt to deal with the problematic aspects of relevance theory itself. It is to this task that we now turn.
Chapter 5
Relevance Theory

As we saw in the previous chapter, relevance theoretic accounts of definite descriptions face a significant problem which arises out of relevance theory itself. If we are to use relevance theory to build a viable account of definite descriptions and the referential/attributive distinction, then, we must first address the issues with the theory. The next few chapters are dedicated to doing just that. Since the overall aim here is a theory of the semantics and pragmatics of definite descriptions, however, the focus of the discussion will be limited to aspects of relevance theory which might prove problematic for constructing such an account. Our goal is not to defend relevance theory as a whole or offer a significantly modified version of the theory. Such a task is well beyond the scope of our current project. Rather, our aim here is merely to smooth the way for the construction of an account of definite descriptions. Thus, there will likely be potential objections to relevance theory, or facets of the theory which stand in need of clarification, which will go untreated here.

Most of our treatment of relevance theory will consist in further developing basic aspects of the theory or providing more detailed characterizations of core concepts. We have already seen a brief exposition of relevance theory in the previous chapter. As a result, some of the discussion of the theory given here will seem repetitive. In many places, however, the description of relevance theory that follows will be more detailed than that already given. We sketched relevance theory in chapter 4 for the purpose of elucidating the basis for the relevance theoretic accounts of descriptions found in the literature. But we will need a more complete and
in-depth view of the theory if we are to determine which aspects of it must be clarified or modified to pave the way for a defensible relevance theoretic account of definite descriptions.

5.1 Ostensive Inferential Communication

Relevance theory is an account of ostensive inferential communication. According to Sperber & Wilson, the central feature of ostensive inferential communication is the communicator’s possession of two intentions – the informative intention and the communicative intention. An informative intention is an intention to make some assumption or set of assumptions manifest or more manifest. An assumption is manifest to an individual to the extent that he or she is capable of mentally representing the assumption and accepting it as true or probably true.\(^{77}\) The communicative intention is an intention to make mutually manifest to the audience and communicator that the communicator has the intention to make some set of assumptions manifest.\(^{78}\) An assumption is mutually manifest to individuals \(A\) and \(B\) if the fact that the assumption is manifest to \(A\) and \(B\) is itself manifest to \(A\) and \(B\). Sperber & Wilson treat this in terms of cognitive environments. The cognitive environment of an individual at a time is the set of assumptions which are manifest to that individual at that time. When the cognitive environments of two individuals intersect, the intersection is a shared cognitive environment. A mutual cognitive environment is a shared cognitive environment for which it is manifest which individuals shared that environment. That is, for every assumption in the shared environment, the fact that that assumption is manifest to the people sharing the environment is itself manifest. In

\(^{78}\) Ibid., p. 61.
other words, the mutual cognitive environment of two individuals consists in the set of assumptions which are mutually manifest to those individuals.\textsuperscript{79}

In order to convey his or her meaning, however, it is not enough for a communicator to simply have both the informative and communicative intentions. The communicator must also produce some stimulus whereby his or her meaning can be conveyed. In cases of linguistic communication, we will call this stimulus an \textit{utterance}.\textsuperscript{80} When the stimulus is intended to make manifest an intention to make manifest – when it is intended to make manifest an informative intention – it is called an \textit{ostensive stimulus}. Ostensive stimuli are produced as means for realizing communicative intentions, and thereby also fulfilling informative intentions. An ostensive stimulus, then, is essentially a communicative stimulus – it is a stimulus which is meant to serve as a means of communication.

Whether or not the communicator’s informative intention is fulfilled depends on the audience. For the communicator’s informative intention to be fulfilled, the audience must interpret the ostensive stimulus produced as making manifest the same assumptions which the communicator intended to make manifest. But for any given stimulus, there are a number of different ways that stimulus might be interpreted. According to relevance theory, however, the audience’s choice of interpretation is not random. On the contrary, it is directed by the search for relevance.

\textsuperscript{79} Ibid., pp. 41-42
\textsuperscript{80} This is, perhaps, not the colloquial use of ‘utterance’. As it is meant here, an utterance is clearly not limited to verbalized language. For our purposes, an utterance can be written, spoken or signed.
5.2 The Principles of Relevance

One of the core claims of relevance theory is that human cognition, thanks, presumably, to various selection pressures, is characterized by the drive toward efficiency. This claim is embodied in the relevance theorist’s Cognitive Principle of Relevance.

(CPP1) Cognitive Principle of Relevance
Human cognition tends toward the maximization of relevance.

Relevance is a property of inputs to cognitive processes, a measure of the cognitive benefit of an input relative to the cost required to process it. The cognitive processing of any input will require the expenditure of cognitive resources, and different inputs will require different levels of expenditure. The processing cost of an input is the amount of cognitive effort that is required to process that input.

To understand the cognitive benefits side of the relevance ratio, we first need to understand contexts and contextual effects. On the relevance theoretic view, a context is a set of assumptions. When an individual interprets an utterance in a context, the context is a subset of that individual’s overall set of assumptions about the world which is drawn from the mutual cognitive environment of the communicator and the audience. A contextual effect is a particular sort of modification of a context, one which results from the interaction of old and new information.\(^{81}\) Strictly speaking, contextual effects need not necessarily be effects resulting from the interaction of a context and an utterance interpretation. They may arise out of any contextualization of new information in the context of old information. Since utterance interpretations are our concern here, however, we will focus on contextual effects arising out of

---

\(^{81}\) Sperber & Wilson (1986), p. 109
the interaction between the set of assumptions comprising an utterance interpretation (the “new information”) and a context (the “old information”).

Now, according to Sperber & Wilson, there are four types of contextual effect: contextual implications, strengthenings, weakenings, and the elimination of assumptions.\(^8^2\) They provide the following definition of a contextual implication:

\begin{quote}
\textbf{Contextual Implication}
A set of assumptions $P$ contextually implies an assumption $Q$ in the context $C$ if and only if

(i) the union of $P$ and $C$ non-trivially implies $Q$,

(ii) $P$ does not non-trivially imply $Q$, and

(iii) $C$ does not non-trivially imply $Q$.\(^8^3\)
\end{quote}

In the case we are interested in here – that of utterance interpretation – $P$ is the set of assumptions comprising the interpretation of the utterance and $C$ is the context that interpretation is processed in. The definition of non-trivial implication is given in (5.1).

\begin{quote}
(5.1) A set of assumptions $P$ \textit{logically and non-trivially implies} an assumption $Q$ if and only if, when $P$ is the set of initial theses in a derivation involving only elimination rules, $Q$ belongs to the set of final theses.\(^8^4\)
\end{quote}

Elimination rules, according to Sperber & Wilson, are interpretive; their outputs explicate or analyze the content of their inputs.\(^8^5\) This contrasts with introduction rules, which introduce new content (or reiterate content already present). (5.2) and (5.3) exemplify introduction rules, while (5.4) and (5.5) exemplify elimination rules.

\(^8^2\) In point of fact, Sperber & Wilson list only three types of contextual effects: contextual implications, contextual strengthenings, and removal of contextual assumptions. The weakening of confirmation values they treat as an indirect effect of this last type. See Sperber & Wilson (1986) p.194 n2d. I have listed all four here so that we can clearly see all the ways that contextual effects may arise.


\(^8^4\) Ibid. p. 97

\(^8^5\) Ibid.
As Sperber & Wilson see it, the human deductive device is limited to the use of elimination rules. The second and third kinds of contextual effect – strengthenings and weakenings – are changes in the confirmation values of old assumptions. These arise, essentially, when assumptions in the interpretation of the utterance serve as evidence for or against an assumption in the context. Finally, the last kind of contextual effect – removal of old assumptions – comes about when an assumption present in the context is contradicted by an assumption with a higher confirmation value found either in the utterance interpretation or in the contextual implications of that interpretation.

We are now in a position to discuss the cognitive benefits side of the relevance ratio. Cognitive benefits come in the form of positive cognitive effects. Cognitive effects are the psychological counterparts of contextual effects – they are contextual effects occurring in an individual. A positive cognitive effect is a cognitive effect that makes a positive contribution to the fulfillment of cognitive functions or goals.86 Sperber & Wilson assume that at least one of the

86 Sperber & Wilson (1995), p. 265. Sperber & Wilson seem to assume that all inputs to cognitive processing – or, at least, all those capable of possessing the property of relevance – are fit to serve as premises in inferences. This assumption, when taken by itself, seems to me to be implausible. It fits well, however, with the picture of human cognition that appears in the background of Sperber & Wilson’s characterization of relevance theory. Their focus on assumptions, inference and even manifestness (which includes the ability to accept an assumption as true) paints the
cognitive goals for humans is the improvement of their knowledge of the world. So, a newly acquired justified, de-gettierized, true belief would be a positive cognitive effect, while a newly acquired unjustified belief would not be.

So, the first principle of relevance theory – the cognitive principle of relevance – says that human cognition tends toward the maximization of relevance, where relevance is a ratio of positive cognitive effects to cognitive processing effort. The principle which is key to the relevance theoretic account of ostensive inferential communication, however, is the second principle – the communicative principle of relevance.

(CPP2) Communicative Principle of Relevance
Every ostensive stimulus conveys a presumption of its own optimal relevance.

The idea behind CPP2 is that by producing an utterance, a communicator is licensing the hearer to expect that utterance will be optimally relevant. It is this presumption of optimal relevance which guides the interpretation of an ostensive stimulus. But what does it mean for a stimulus to be optimally relevant? As it turns out, answering this question is not as simple as it might at first appear. Various problems have led Sperber & Wilson to alter the definition over time, and, as we saw in chapter 4, different relevance theorists have relied on different understandings of optimal relevance in their own work.

Sperber & Wilson do not actually directly define optimal relevance. Instead, they give us a characterization of the presumption of optimal relevance. According to the communicative

---

principle of relevance, remember, every ostensive stimulus conveys a presumption of its own optimal relevance. Sperber & Wilson’s most recent version of the presumption of optimal relevance is given in (POR).

(POR) Presumption of Optimal Relevance
(a) The ostensive stimulus is relevant enough to be worth the audience’s processing effort, and
(b) It is the most relevant one compatible with the communicator’s abilities and preferences.\(^{88}\)

Clause (a) of (POR) tells us that an optimally relevance stimulus is relevant enough to be worth the audience’s effort. According to Sperber & Wilson, an ostensive stimulus is relevant enough to be worth processing if it is more relevant than any alternative input at the time.\(^{89}\) Optimal relevance, then, is a comparative notion: an input is optimally relevant relative to other available inputs.

Clause (b) of (POR) requires that an optimally relevant stimulus be the most relevant of the available stimuli that are compatible with the communicator’s abilities and communicative preferences. The restriction in clause (b) is necessary for making sense of, for example, the interpretation of the utterances of small children and non-native speakers who lack facility with the language. To see this, imagine two different occasions of an utterance of (5.6).

(5.6) Me want juice now.

---

\(^{89}\) Wilson & Sperber (2004). Sperber & Wilson have changed on this over time. They originally maintained that “what counts as relevant enough… varies with the way information is accessible… to the addressee over time.” Sperber & Wilson (1986), p. 161). The more recent version, it seems to me, will ultimately make for a better theory, though as we shall see, it may force us to modify Sperber & Wilson’s overall definition of optimal relevance. Additionally, it seems to me that it might make more sense to treat the new version as a revision of clause (a), rather than as a clarification of “relevant enough to be worth processing”, since “relevant enough to be worth processing” and “more relevant than alternatives” do not appear to mean the same thing.
In the first case, imagine a very small child tugging on her mother’s pant leg and uttering (5.6). Now imagine an adult, native speaker of English dressed as a caveman for Halloween and uttering (5.6) to the host at a Halloween party. The differences in the linguistic abilities of the two speakers lead to differences in interpretation of their utterances. The small child’s utterance will be interpreted as a completely serious request for juice from the mother. The party guest’s utterance could be interpreted as a request for juice, though not a completely serious one, but it could also be interpreted as self-derision, mocking of others, or even mere play acting, depending on the full context. Because the child’s linguistic capabilities are limited, we do not take the ungrammaticality of her utterance to indicate a joke – her linguistic capabilities limit the set of relevant interpretations. On the other hand, we need to get a better understanding of the preferences of the party guest in order to get a better understanding of what he was trying to get across.

There is a small problem with clause (b) as it is currently worded, however, for it limits the interpreter to a set of potential interpretations which are consistent with the communicator’s actual abilities and preferences. Clearly, limiting potential interpretations to those consistent with the communicator’s actual abilities and preferences would make it very difficult to explain the hearer’s interpretation process, as hearers can be mistaken or ignorant, and hence will not necessarily have access to the facts about the speaker. Interpretation, as a cognitive process, cannot occur using assumptions that the interpreter lacks cognitive access to. Instead of limiting potential interpretations to those consistent with the communicator’s actual abilities and preferences, clause (b) should rather limit potential interpretations to those consistent with what the hearer assumes about the communicator’s abilities and preferences. We can solve this
difficulty with a minor change to the clause. The new presumption of relevance can be found in (POR’).

(POR’) Presumption of Optimal Relevance
(a’) The ostensive stimulus is relevant enough to be worth the audience’s processing effort, and
(b’) It is the most relevant one compatible with the audience’s assumptions about the communicator’s abilities and preferences.

Let us make one last clarification regarding (b’). The current wording of the clause might make it seem as though what is required is that the ostensive stimulus be the most relevant of all the ostensive stimuli compatible with the interpreter’s assumptions about the communicator’s abilities and general attitudes on speech and expression. This is not how Sperber & Wilson understand it, however. Instead, they take the clause to mean that the ostensive stimulus is the most relevant of those stimuli which would be likely to communicate what the communicator wants to communicate.\(^90\) It is the most relevant one the communicator could have used to satisfy his informative intention, given his or her abilities and preferences.

There is still one problem facing (POR’). According to relevance theory, remember, the pragmatic interpretation process is guided by the presumption of optimal relevance. As it stands, however, (POR’) would be of very little use in aiding the interpreter in selecting an interpretation. Clause (a’) would appear to be of virtually no use whatsoever to the interpreter in selecting an interpretation of the stimulus. The requirement that the ostensive stimulus be the most relevant of the available stimuli tells the interpreter that the stimulus needs to achieve a certain degree of relevance, but multiple different interpretations of the same stimulus might

\(^90\) See Sperber & Wilson (1995), p. 270. This is also in line with clause (b) as it occurs in Sperber & Wilson’s original characterization of the presumption of optimal relevance. See Sperber & Wilson (1986), p. 158.
serve to satisfy that requirement, particularly if other available stimuli have a low degree of relevance.

At first glance, the only aid that (b’) would give in the interpretation selection process would be to limit the set of interpretive hypotheses to those consistent with the communicator’s abilities and preferences. So, in our example above, the interpreter of the small child who utters (5.6) will not consider interpretations which would be considered by the interpreter of the costumed adult at the Halloween party. While certainly useful, this would not be enough to select a complete interpretation in many cases. In the case of either utterance of (5.6), for example, the requirement that stimulus be the most relevant of those the communicator could have used to fulfill his communicative intention is of no use in determining the extent of the time-span referred to by the indexical ‘now’.

It might be argued that the presumption of optimal relevance serves to guide interpretation by serving as the basis for context specific expectations of relevance. *Specific expectations of relevance* are expectations about how an utterance will achieve relevance in the context. To better understand this, consider the following example from Sperber & Wilson:

(5.7) Peter: Did John pay back the money he owed you?  
     Mary: No. He forgot to go to the bank.

According to Sperber & Wilson, Mary’s utterance raises the specific expectation of relevance in (5.8) for Peter.

(5.8) Mary’s utterance will achieve relevance by explaining why John has not repaid the money he owed her.

---

91 Wilson & Sperber (2004), p. 615-617
Sperber & Wilson claim that this expectation, like other specific expectations of relevance, is raised by the presumption of optimal relevance together with facts about what would be most relevant to the hearer – in this case, an explanation of the fact that John did not pay the money back. According to Sperber & Wilson, this expectation aids in the disambiguation of ‘bank’. Since forgetting to go to a financial institution explains why John did not pay back the money he owed, while forgetting to go to a river bank does not, selecting the financial institution sense of ‘bank’ will result in an interpretation which satisfies expectations of relevance. It might be, then, that it is specific expectations of relevance, arising out of the presumption of optimal relevance, that guide the interpretation process.

There are two problems with this line of argument. The first is that it is unclear that specific expectations of relevance arise out of the presumption of relevance. As we just saw, Sperber & Wilson claim that the specific expectation in (5.8) arises out of the presumption that Mary’s utterance will be optimally relevant and the fact that an explanation of John’s failure to pay would be most relevant to Peter. It is by no means certain, however, that this explanation would be most relevant to Peter at the time of Mary’s utterance. A statement about how John’s failure will affect Mary might be equally relevant, as might additional information about when John was supposed to pay the money back.

Of course it might be claimed that what Sperber & Wilson mean when they say that an explanation of John’s failure would be most relevant to Peter is that an interpretation which included such an explanation would be the optimally relevant interpretation of Mary’s utterance. This does not fit with what Sperber & Wilson themselves say, however, for they tell us that

92 Ibid., p. 616
“Since what [Peter] wants to know at this point is why John did not repay the money he owed, he assumes… that Mary’s utterance will achieve relevance by answering this question.”

This would seem to indicate that they take the specific expectation of relevance in question to arise out of facts about what is most relevant to Peter at the time, not facts about the optimally relevant interpretation of Mary’s utterance. In fact, if specific expectations of relevance are meant to aid the audience in identifying the optimally relevant interpretation of the utterance, then Sperber & Wilson cannot be claiming that such expectations are formed on the basis of what would be the optimally relevant interpretation of the utterance, for this would mean that the specific expectations of relevance are formed on the basis of the very thing they are supposed to help identify.

This is not to say that hearers do not have expectations as to how utterances will achieve relevance, however. There are assumptions which are generally held by adults participating in conversation with one another about how conversations typically proceed and the kinds of speech acts or conversational tasks which are typically performed at various points in a conversation. In other words, we have general expectations regarding the sorts of goals speakers intend to achieve in different sorts of conversational circumstances. When an utterance is a response to a question, for example, it is generally assumed that it will answer the question. We also generally expect a statement which follows directly on the heels of an independent assertion to explain, justify, or expand upon that assertion. These sorts of expectations are defeasible, of course, and can be abandoned during the interpretation process – just as Sperber & Wilson take specific expectations of relevance to be. If there are specific expectations of relevance which

93 Ibid., p.
guide the interpretation process, it seems more plausible to suppose that they are of this sort, rather than expectations which arise out of clause (b’) of (POR’).

The second problem with claiming that (POR’) is the source of interpretation guidance via the generation of specific expectations of relevance is that even if (POR’) is in fact the basis for specific expectations of relevance, such expectations would only serve to guide the interpretation process in those cases in which they are actually present. But there are undoubtedly cases in which an utterance does not carry with it any specific expectations of relevance. It is unlikely, for example, that we would have any specific expectations as to how a discourse initial utterance would achieve relevance.\(^9\) Nonetheless, we are perfectly capable of interpreting such utterances. If (POR’) is responsible for guiding the interpretation process, then, it cannot be via specific expectations of relevance alone.

It would likely be argued at this point that (POR’) guides the interpretation process by giving rise to the interpretation strategy which Sperber & Wilson claim is used in the interpretation of utterances. According to Sperber & Wilson, the interpretation strategy used in pragmatic processing involves testing interpretive hypotheses in order of accessibility, and stopping when expectations of relevance are satisfied. They tell us that “Given clause (b) of the presumption of optimal relevance, it is reasonable for the hearer to follow a path of least effort because the speaker is expected (within the limits of her abilities and preferences) to make her

\(^9\) Sperber & Wilson themselves admit that this is the case. They tell us that they use the example of Peter and Mary for ease of exposition, since it is a case in which the preceding discourse creates a specific expectation of relevance, but that in cases of discourse initial utterances considerations of effort will likely determine the selection of the interpretation. See Wilson & Sperber (2004), pp. 629-630fn14
As they see it, the hearer is entitled to this expectation by their original clause (b) (and, hence, our (b')). But clause (b') does not say that the ostensive stimulus is the easiest to understand of those compatible with the communicator’s abilities and preferences. It says that the ostensive stimulus is the most relevant one. Processing cost is only one side of the relevance ratio. Take two ostensive stimuli – A and B, and suppose that A requires less processing effort than B. If B generates sufficient contextual effects to overcome its higher processing cost, then B will be the more relevant stimulus of the two. Given this fact, it would appear that (b’) does not license the hearer to expect that the communicator will make the stimulus as easy as possible to understand and hence does not support the interpretation strategy as Sperber & Wilson claim it does.

It would seem, then, that as it is currently characterized the interpretation strategy which Sperber & Wilson claim is the means by which interpreters select interpretations is actually incompatible with the assumption that (POR’) governs the pragmatic interpretation process. Now, one possible move at this point would, of course, be to abandon (POR’) and maintain that the interpretation strategy alone is what governs interpretation. Doing this, however, would essentially mean abandoning a relevance theoretic account of utterance interpretation in favor of what we might call an accessibility theory account. If we wish to maintain a relevance theoretic account, then, we will need to make sense of how (POR’) guides the process of interpretation selection.

I believe that the cognitive principle of relevance provides the answer here. According to the cognitive principle of relevance, remember, human cognition tends toward the maximization

---

95 Wilson & Sperber (2004), pp 613-614
of relevance. If this is correct, then the tendency of human cognition would be to assign the most relevant interpretation available to any stimulus which is being interpreted, whether it be an ostensive stimulus or otherwise. Given this fact, it is in the interest of a communicator who wishes to be understood to make sure, within the limits of his abilities and preferences, that his intended interpretation is the most relevant available interpretation of the stimulus he uses.

This actually lines up with comments made by Sperber & Wilson about what the presumption of optimal relevance licenses an audience to expect. According to them, “[The presumption of optimal relevance] says that the addressee is entitled to expect a level of relevance high enough to warrant his attending to the stimulus, and which is, moreover, the highest level of relevance that the communicator was capable of achieving given her means and goals.” For the stimulus to yield the highest level of relevance that the communicator is capable of achieving given her abilities and preferences, the stimulus must be assigned the most relevant available interpretation consistent with the communicator’s abilities and preferences. Call this the optimally relevant interpretation. Given the foregoing analysis, as well as Sperber & Wilson’s comments, then, it would seem that the presumption of optimal relevance licenses the addressee to expect that optimally relevant interpretation of the stimulus will be more relevant than other available stimuli and the most relevant compatible with the communicator’s abilities and preferences.

Once we understand the presumption of optimal relevance as licensing the communicator to expect the highest level of relevance compatible with the communicator’s abilities and preferences, it is clear how the presumption of relevance could serve to guide the pragmatic

---

interpretation process. The interpreter will assign the optimally relevant interpretation. But how might the interpreter go about doing this? We will examine this question in the next section.

5.3 Cognitive Processing

CPP2 tells us that every ostensive stimulus conveys the presumption of its own optimal relevance. According to relevance theory, utterance interpretation is guided by this presumption of optimal relevance. But how does the process of interpretation proceed? The complete interpretation process will differ depending on whether the ostensive stimulus is linguistic or not, as linguistic stimuli require processing that non-linguistic stimuli do not. Additionally, the way in which the set of potential interpretations for a linguistic stimulus is delimited will likely differ from the way that set is delimited for a non-linguistic stimulus, since, in the former case, there is linguistic information which constrains potential interpretations, while only contextual information is available in the latter. Since our main focus is definite descriptions – a linguistic phenomenon – we will limit our discussion to the process of interpreting linguistic ostensive stimuli – that is, utterances.

According to relevance theory, the interpretation of an ostensive stimulus occurs in different stages, taking place in different modules of the mind. While it would likely be possible to construct a relevance theoretic picture of utterance interpretation which did not assume that the mind is modular, this is the picture of the mind that is embraced by Sperber & Wilson and other relevance theorists. In particular, Sperber & Wilson assume that the mind is massively modular.
5.3.1 Massive Modularity

The massive modularity view of mind encompasses a wide number of claims. We will limit our focus here to those aspects of thesis which will be important to bear in mind with respect to the relevance theoretic view of the cognitive processing of utterances. The massive modularity thesis about the mind holds that the human mind is largely (if not entirely) composed of modules – that is, distinct processing systems. In line with Carruthers, relevance theorists view the mind as consisting in “…a great many different processing systems, which exist and operate to some degree independently of one another. Each of these systems will have a distinctive function or set of functions; and each will have a distinct neural realization.” The massive modularity theorist also holds that cognitive modules are, to a greater or lesser degree, *informationally encapsulated* and *inaccessible* to other modules.

When we say that a module is inaccessible to other modules, we mean that the internal operations of that module cannot be accessed by other modules. That is, the internal operations of a module are such that other modules can neither utilize nor acquire information about them. The original picture of cognitive modules provided by Fodor held that the internal processing of a module was virtually always completely inaccessible. On the massive modularity picture of the mind, however, the property of inaccessibility is a matter of degree, and is subject to other

---

97 Carruthers (2006), p. 62. The neural realization of a module need not be the same across individuals. Some modules, particularly those which are part of basic human cognitive architecture (such as visual processing systems, etc) may be realized in the same neural type across individuals. Massive modularity theory, however, leaves open the possibility that some cognitive modules develop in order to perform functions which are learned by the individual (such as reading). While we would expect those modules to have a distinct neural realization in each individual, we wouldn’t expect it to be the same neural realization.

98 Ibid.

99 Fodor (1983). While Fodor states that the characteristic properties of cognitive modules are matters of degree, this is not reflected in his overall treatment of modules. Discussion of these so-called “Fodor-modules” seems to ignore any potential gradation in the properties of inaccessibility or informational encapsulation. Instead, Fodor-modules are usually depicted as always being completely inaccessible and informationally encapsulated.
aspects of the function and composition of the module in question. According to the massive modularity thesis, not only is the mind largely comprised of distinct modules, but modules themselves may be composed of sub-modules, and in some cases modules may very well share parts. Such circumstances would serve to limit the degree to which a module’s inner workings are inaccessible to other modules. A sub-module’s processing would likely be accessible to the module of which it is a part, and modules sharing parts would also clearly have to share access to operation of those parts. For the massive modularity theorist, then, cognitive modules will generally be inaccessible to other modules, but that inaccessibility need not be complete.

The massive modularity theorist has a similarly qualified view of the informational encapsulation of cognitive modules. The property of informational encapsulation was a principal feature – perhaps the principal feature – of the cognitive modules envisioned by Fodor. A module is informationally encapsulated, for Fodor, if it is limited only to information contained within the system itself (with the exception, of course, of the input to processing). Cognitive modules as Fodor understands them – so-called “Fodor-modules” – cannot draw upon any information external to the system. Since these modules are also inaccessible to other modules, a Fodor-module relies on a dedicated database of information which only it can access. Since general reasoning processes, such as utterance interpretation, belief formation, or the development of a plan of action, must draw upon all sorts of different information – indeed, potentially upon the entire belief set of the individual – such processes could not be performed by Fodor-modules. As a result, Fodor holds that the modularity of the mind is limited to input systems, such as visual or auditory processing, and a central processing system is responsible for

100 Carruthers, (2006), p. 60
101 Fodor (1983)
all those cognitive processes that require information that would have to be widely shared, such as the subject’s set of beliefs. \(^{102}\)

In contrast, the massive modularity theorist takes a different view of encapsulation. Carruthers has argued that there are two types of informational encapsulation – narrow-scope encapsulation and wide-scope encapsulation. \(^{103}\) A processing system has the property of narrow-scope encapsulation if it is unable to access information external to itself. Narrow-scope encapsulation is the sort of encapsulation that Fodor assumes cognitive modules possess. A processing system has the property of wide-scope encapsulation, on the other hand, if there is certain information that it cannot be affected by. A cognitive module which is widely encapsulated, then, could very well draw upon information shared with other modules, but would be still be limited in the amount of information it would have to consult in performing its function. According to Carruthers, the limits on the information affecting a module would likely result from heuristic-guided searches over databases of information, perhaps even over the complete set of information available to the individual. \(^{104}\)

Relevance theory holds that multiple cognitive modules are involved in the cognitive processing of linguistic stimuli.

### 5.3.2 Processing Modules

Linguistic stimuli, according to Sperber & Wilson, convey content in two different ways – via linguistic encoding and via contextually aided inference on the part of the hearer. \(^{105}\) Since linguistic stimuli come in the form of auditory or visual stimuli, before an utterance can be

\(^{102}\) Ibid.

\(^{103}\) Carruthers (2006)

\(^{104}\) Carruthers (2006), p. 59

processed for encoded or inferentially derivable content, it must first undergo processing in one of the input systems of the mind. So, for example, a verbal utterance, as an auditory stimulus, is first processed by the auditory processing system. This initial processing, which may occur in multiple stages, takes basic auditory data and yields a set of phonetic data. This data then undergoes two more stages of processing, corresponding to the two different kinds of content that the relevance theorist believes a linguistic stimulus to have.

Since it is a linguistic stimulus, the output of the auditory processing of an utterance is sent to a *linguistics processing module*. The function of this module is to take phonetic (or graphetic) data and construct mental representations of the linguistic meaning of the utterance. The operation of the linguistics processing module is guided by the subject’s linguistic competence – his knowledge of the syntactic and semantic rules of the language. In the case of a verbal utterance, the module draws upon a database of rules linking various phonetic sequences with the relevant syntactic and lexical information. Guided by the rules of the relevant language, the module constructs a mental representation of the syntactic and lexical content of the utterance. This representation is called a *logical form* or *LF* of the utterance, and it represents the *linguistic meaning* of the utterance.

Since many utterances are syntactically or lexically ambiguous, there will be multiple LFs generated by the linguistics processing module for a single utterance. Thus the production of LFs by the linguistics processing module is only the first step in the interpretation of an utterance. The second step – pragmatic processing – is performed by a pragmatic processing module. In working to complete the interpretation of an utterance, the pragmatic processing module must select from among the LFs generated during linguistic processing. The selection of
an LF is not necessarily the end of the story, however. In many cases the linguistic meaning of an utterance alone will not be sufficient to determine a proposition. The LF of an utterance expressing a complete proposition but containing an indexical expression, for example, will not be truth-evaluable; it will not determine a proposition. Further processing of the utterance in light of the conversational context will be necessary if the interpretation of the utterance is to be complete. The situation is the same with respect to utterances containing gradable adjectives or other context sensitive expressions. According to relevance theory, this processing is performed by a metacommunicative or *pragmatics processing module* (PPM). Because pragmatic processing involves the use of a good deal of information about what is manifest to both conversational participants, Sperber & Wilson have argued that the PPM is a sub module of the metapsychological (theory of mind) module of the mind.\textsuperscript{106} On the relevance theoretic picture, this pragmatics processing module will use contextual information to fill in the selected LF in order to fully interpret the utterance in question.

So, after processing in the linguistics module is completed, the LF or LFs are sent to the pragmatic processing module (PPM). And this is where considerations of relevance finally come into play. Up until this point, processing of the utterance has been guided and constrained by factors other than relevance. The uptake of the utterance is constrained by the inputs received by the auditory or visual processing systems. The LFs that are generated by the linguistic processing module are processed in accordance with morphology, syntax, and semantics of the language of the utterance’s origin. Given Sperber & Wilson’s claim that human cognition tends to maximize relevance, one might think that all aspects of utterance processing must be directed by

\textsuperscript{106} Wilson & Sperber (2002)
considerations of relevance. But this is not so. Sperber and Wilson explicitly tell us that the communicative principle of relevance does not apply to “straightforward coded communication.” Since linguistic processing is a matter of straightforward decoding, it seems reasonable to suppose that the communicative principle of relevance would not apply to that stage of processing. Thus, despite their general claim about human cognition, it would seem that Sperber & Wilson see the decoding of individual linguistic stimuli as occurring independently of considerations of relevance.

5.3.3 Accessibility and Availability

According to relevance theory, the interpretation assigned to an utterance is the optimally relevant one, and as we have just seen, Sperber & Wilson hold that the selection of an interpretation is made by a pragmatics processing module (PPM). The PPM’s selection of the optimally relevant interpretation is claimed to be a clear cut process. The module evaluates potential interpretations of an utterance – interpretive hypotheses – in order of accessibility, stopping when it reaches an interpretation which satisfies expectations of relevance. This is the interpretation which is assigned to the utterance.

According to Sperber & Wilson, “a more accessible assumption is one that is easier to recall.” What they seem to mean by this is that a more accessible assumption is one that is more readily retrieved from encyclopedic memory. While conceding that there are still many questions about the nature of encyclopedic memory and memory retrieval left to answer, and that

---

107 Ibid, p. 158
110 Sperber & Wilson (1986), p. 77
cognitive psychology will have to provide those answers, Sperber & Wilson suggest that the most accessible assumptions will be those which remain in the memory of what they call the ‘deductive device’ - the system which they claim is used in spontaneous inference, and, in particular, in utterance comprehension. According to Sperber & Wilson, at the end of any inference process, the deductive device will retain in its memory all newly derived conclusions and all premises which were involved in the inference process or were affected by it. Thus, the most accessible assumptions, on this view, are those involved in the most recent inference process performed. In the case of the interpretation of an utterance following immediately on the heels of a previous utterance, this would be the interpretation of the previous utterance, the assumptions involved in generating that interpretation, and its cognitive effects.

Sperber & Wilson further claim that there is likely a short term memory store other than that of the deductive device available, as our ability to switch back and forth between two different tasks like watching television and discussing family affairs would seem to require a short-term memory store distinct from that of the deductive device. Assumptions removed from the memory of the deductive device, they claim, are transferred to this general short-term memory store. This store will contain the interpretations of earlier utterances along the assumptions involved in generating those interpretations as well as other information or thoughts the individual has recently attended to. Assumptions in this store are less accessible than those in the memory of the deductive device, but more accessible than those stored in long term memory.

111 Ibid., p. 170
112 Ibid., pp. 93-108
113 Ibid., p. 139
114 Ibid., pp. 139-140
Accessibility of assumptions is one thing, but what we are interested in is the accessibility of interpretive hypotheses. According to Sperber & Wilson the accessibility of an interpretive hypothesis is linked with its processing cost. They claim that “… given the cognitive environment, given the initial context, and given the stimulus, some hypotheses are more accessible than others, and this means that they require less processing effort.”  

Presumably, this is because the accessibility of an interpretive hypothesis is determined by the accessibility of the assumptions involved in processing that hypothesis – in other words, the context in which the hypothesis is processed. But the accessibility of the context is only one of the factors in the processing cost associated with an interpretive hypothesis. The number of inferences required to process the hypothesis in the context is also a factor. According to Sperber & Wilson, accessing a context requires effort. The more accessible the context is, the less effort is required, and vice versa.

The accessibility of a context is much the same as the accessibility of an assumption (indeed, the accessibility of assumptions seems to be derivative on the accessibility of contexts). The assumptions remaining in the memory of the deductive device form an initial context which can then be expanded in various ways. At this point, according to Sperber & Wilson, …

…the individual has at his disposal a particular set of accessible contexts. This set is partly ordered: each context (apart from the initial context) contains one or more smaller contexts, and each context (apart from the maximal contexts) is contained in one or more larger contexts. The set of accessible contexts is thus partly ordered by the inclusion relation. This formal relation has a psychological counterpart: order of inclusion corresponds to order of accessibility.

---

115 Ibid., p. 167. Given that there is more to processing cost than just the accessibility of the processing context, this claim should probably read as holding only when all else is equal.
116 Ibid., p. 142
117 Ibid., p 140
118 Ibid., p 142

140
So, the accessibility of an interpretive hypothesis is determined by the accessibility of the context in which the hypothesis is processed, and the effort involved in processing that hypothesis is at least partially determined by the effort required to access the context. Since the less accessible a context is, the more effort is required to access it, this means that, at else equal, the less accessible the context for processing an interpretive hypothesis, the more effort is required to process that hypothesis.

According to relevance theory, the interpretation of an utterance (and hence an interpretative hypothesis) will frequently be composed of multiple mental representations. Each of these mental representations will either be an explicature or an implicature. An explicature is a mental representation resulting from the enrichment of LF. Explicatures, then, are mental representations of the content expressed by an utterance which derive from the linguistic meaning of that utterance and have been enriched on the basis of contextual information. Explicatures can be divided into two types. The first is the propositional form, which is a non-higher-level semantically complete enrichment of the LF of the utterance being interpreted. By saying that the propositional form is semantically complete we mean that it is capable of being true or false. By saying the propositional form is non-higher-level we are contrasting it with what, following Carston, we will call higher-level explicatures. A higher-level explicature is derived from the propositional form together with mutually manifest assumptions about the speaker’s propositional attitudes with respect to the propositional form and cannot be derived from the propositional form or the mutually manifest assumptions about the speaker’s propositional attitudes alone. The sort of assumptions Sperber & Wilson have in mind here are

\[119\] Carston (2004)
essentially assumptions describing the speech acts being performed. So, for example, if Allen utters the sentence “The sky is blue” the higher-level explicatures associated with that utterance would likely include assumptions like *Allen said that the sky is blue*, *Allen asserted that the sky is blue*, *Allen believes that the sky is blue*, and so on.

In contrast with an explicature, an *implicature* does not result from the enrichment of logical form. Sperber & Wilson define an implicature as a “…contextual assumption or implication which a speaker, intending her utterance to be manifestly relevant, manifestly intended to make manifest to the hearer.” An assumption $A$ is an implicature, then, if it is the case that it is mutually manifest that the speaker intended to make $A$ manifest. Implicatures will frequently arise when the context gives rise to expectations about how an utterance will achieve relevance. If the speaker is responding to a question, for example, the expectation is that his utterance will respond to the question. If the propositional form of his utterance does not itself answer the question, but does make accessible an assumption $B$ which does answer the question, that may, depending on the rest of the context, make it manifest that the speaker intended to make $B$ manifest.

We saw earlier that the optimally relevant interpretation is the most relevant of the available interpretations. We are now in a position to say a bit about when an interpretation is available. Now, it seems fairly clear that on the relevance theoretic view, for a mental representation to be available as an interpretation of an utterance for an individual, two things must be the case. First, since relevance theory takes interpretations to consist solely in mental representations of one of two types, an available interpretation will also have to consist in mental

---

representations of these types. That is, a mental representation in an available interpretation must either be a propositional form, a higher-level explicature, or an implicature. Second, the mental representation should be cognitively accessible for the individual in question. Given this, we can say that an interpretation is available when it meets the conditions listed in (IA).

(IA) Interpretation Availability
An interpretation of an utterance is available for an individual if and only if:
(1) It is cognitively accessible to that individual, and
(2) Each mental representation in the interpretation satisfies the definition of a propositional form, a higher-level explicature, or an implicature.

Now, the definition of a propositional form tells us that the propositional form of an utterance is an enrichment of the LF of that utterance. In order for a mental representation to be an enrichment of LF, it must conform to linguistically encoded content of the LF. For a mental representation to satisfy the definition of a propositional form, then, all of the concepts in that mental representation which serve as enrichments of encoded content must conform to that encoded content. And, per clause (1) of (IA), for the interpretation containing that propositional form to be available, the concepts it contains must be cognitively accessible. Let us say that a concept which is cognitively accessible and conforms to encoded content is available as an enrichment of that encoded content. In order for an interpretation to satisfy both clauses of (IA), then, the concepts present in the propositional form in that interpretation which serve as enrichments of components of the LF must be available as enrichments of whatever encoded content they serve to enrich. We will discuss the availability of concepts as enrichments in much greater detail in chapter 7.
5.3.4 The Interpretation Selection Process

So, an interpretation of an utterance can be composed of a propositional form, higher-level explicatures, and implicatures. Now, as we saw earlier, Sperber & Wilson claim that the interpretation selection process is one in which interpretive hypotheses are tested in order of accessibility, and testing stops when the system encounters an interpretive hypothesis that satisfies expectations of relevance. On the face of things, this would seem to indicate that the process used by the PPM is something like the following: the PPM begins by testing the most accessible of the available interpretations for relevance. If the system determines that this interpretation meets expectations of relevance – including the expectation of optimal relevance – then testing stops and that interpretation is selected. If the system finds that the interpretation does not meet the expectation of optimal relevance, then it turns to testing the next most accessible interpretation for relevance, and so on.\textsuperscript{121} Call this the \textit{Individual Testing Process} (ITP).

Now, the optimally relevant interpretation, remember, is the interpretation which yields the most cognitive effects per unit of processing cost of the available interpretations. If this is correct, however, then the ITP cannot be the process used to select the optimally relevant interpretation of an utterance or other stimulus in cases in which there are no specific expectations of relevance. According to ITP, the processing system stops testing interpretive hypotheses when it discovers an interpretation that meets the expectation of optimal relevance. ITP presumes that the processing system tests only a subset of the available interpretations. If meeting the expectation of optimal relevance – that is, being optimally relevant – requires being

the most relevant of the available interpretations, however, then the system could not make a
determination without having tested all the available interpretive hypotheses. One cannot
rationally pick an object out of a set and make a claim about its status relative to the rest of the
set without having examined the rest of the objects in the set or having knowledge about them.
This would be much like a professor reading the very first paper on the top of the stack of papers
to be graded and declaring it to be the best paper of the class without having read the rest of the
papers in the stack; the probability of accuracy with such a method is very low.

Furthermore, ITP is not, on its face, geared toward the maximization of relevance, since it
is not clear how it would make the selection of interpretations with the most relevance more
likely. While it is true that more accessible interpretations have lower processing costs than their
less accessible competitors, ITP seems to allow for the possibility that a system might select the
most accessible interpretation even when a slightly less accessible (yet still relatively accessible)
interpretation yielded a great deal more cognitive effects for very little added processing effort.
Given an understanding of optimal relevance in comparative terms, however, such a situation
should not be possible. It would seem, then, that ITP and the definition of optimal relevance
cannot both be consistently held by the relevance theorist.

At this point, then, the relevance theorist looks to be impaled on the horns of a dilemma.
He must either abandon or alter the definition of optimal relevance or he must reject the picture
painted by ITP and replace it with something different. Each of these options constitutes a
different interpretation of relevance theory, and, as is generally the case with dilemmas, neither
horn appears particularly appealing. If the definition of optimal relevance is abandoned, it is not
clear what might replace it and whether any replacement would be precise enough to allow the
theory to generate testable predictions. There is no clear scale on which we might quantify relevance, and no reasonable justification for drawing the line of “relevant enough” in one place rather than another.\textsuperscript{122}

If, on the other hand, the definition of optimal relevance is preserved, and ITP is rejected, it looks as though ITP would have to be replaced with a process that involved testing \textit{all of the} available interpretations for an utterance, comparing them, and then selecting the most relevant of the set. The problem with ITP, after all, was that it did not test all the available interpretations, and it was not clear how the PPM could find the \textit{most} relevant interpretation in a set without testing all the interpretations in that set. The problem with claiming that the PPM tests all the available interpretations of an utterance or other stimulus is that there is some reason to think that that the number of available of interpretations might exceed the limited capabilities of the PPM to generate an interpretation in the time normally taken to interpret an utterance. Take, for example, the task of reference assignment – only one of many tasks involved in the generation of utterance interpretations. For any token of ‘he’, for example, there are thousands of potential referents. Add to this all the potential narrowings or broadenings of concepts, the semantic and syntactic ambiguities, and all of the potentially implicitly communicated contents, and it is unclear whether the human PPM would be capable of completing the processing of any particular utterance in a reasonable period of time, if at all. Of course, there are factors that limit the set of available interpretations – the LF and the speaker’s capabilities and preferences, for example, but while the existence of such factors might give us some reason to suspect that the

\textsuperscript{122} Even if we were to claim that the standard of “relevant enough” is contextually variant, we would still need some account of what makes an interpretation relevant enough in a context if relevance theoretic treatments of communicative phenomena are to be testable, and it is unclear what such an account would look like.
typical set of available interpretations associated with an utterance is small enough for the PPM to handle, it is by no means sufficient to rule out the possibility.

So, neither horn of the dilemma currently facing the relevance theorist is particularly appealing. It might not be necessary to select between the horns, however. Instead, the relevance theorist might attempt to pass through them by arguing for an understanding of pragmatic processing that fits with the currently held picture but differs enough from the ITP to make escape from the dilemma feasible. Remember that the reason the ITP does not fit with the definition of optimal relevance is that it appears to determine whether an interpretation is optimally relevant simply by looking at that individual interpretation, rather than comparing the relevance of interpretations. To fit with the definition of optimal relevance, then, a proposed process of interpretation testing must be a comparative process. As we saw above, however, we don’t want a process that involves testing every last available interpretation. We want a process that does what the relevance theorist says it does – that is, a process that stops when the expectation of optimal relevance (and specific expectations of relevance, if there are any) is likely to have been met.

In light of this, I suggest that we view the interpretation process as one on which the PPM assesses the relevance of interpretations in order of their accessibility, but determines whether an interpretation is the optimally relevant one by comparing it with other interpretations. Call this the Comparative Interpretation Process (CIP). Of course, given what has been said so far, it is by no means obvious that CIP is a process that stops when the expectation of optimal relevance is satisfied. After all, as we noted earlier, it does not seem rational to make a determination about the status of a member of a set relative to the other members if one has not evaluated all the
members of the set. Call a judgment that a member of a set is the most $P$ of that set a superlative judgment. Intuitively, superlative judgments about objects in a particular set cannot be made without the relevant information about each member of the set.

If we look beyond intuition, however, it becomes clear that while I cannot make a superlative judgment on the basis of examination of one member of a set, it is possible to make a superlative judgment about a member of a set without examining all the members of that set. Using inductive reasoning, one may be able to rationally make a superlative judgment about an object in a set after having evaluated only a subset of the entire set. Imagine walking into a ballroom filled with people and wandering around. After having seen, say, 80% of those in the ballroom, you would seem to be justified in claiming that the tallest man you have seen so far - call him Chris - is the tallest man in the room. You are probably justified in making the claim even if you have only seen 70% of the people in the ballroom. At 60%, it’s less clear whether you are justified or not, and even less clear at 50%. Your justification here is inductive, not deductive, and, as such, the rationality of drawing your conclusion will obviously decrease as the percentage of the set that you have evaluated decreases. The point is that, after having observed a certain percentage of the individuals in the room, you are licensed in drawing a superlative conclusion about one person in the room without having observed everyone in the room.

We can easily view the CIP as utilizing these sorts of inductive inferences in the process of testing utterance interpretations. Moving in order of accessibility, the PPM tests interpretive hypotheses. As testing continues, inductive evidence accumulates, eventually making it possible for the PPM to rationally conclude that the most relevant of the interpretations it has tested is the most relevant interpretation of the set – i.e., the optimally relevant interpretation. At the point at
which this inductive inference is licensed, the expectation of optimal relevance has likely been met, and the PPM can stop testing interpretations.

There may be a bit of a problem here, however. Inductive inferences of the sort we have envisioned operating here are only valid if the sample being used is representative. But if the PPM tests hypotheses in order of accessibility, then the sample set of interpretive hypotheses will be significantly biased. Of course, the simplest way of dealing with this issue would be to abandon the idea that the PPM tests in order of accessibility and claim instead that the order of testing is random. This approach faces a serious difficulty, however, for the size of the set of potential interpretive hypotheses for many utterances could turn out to be extremely large, potentially too large for the PPM to test the required percentage of the set in the time normally given for utterance interpretation in everyday conversation. We can circumvent this difficulty by assuming that there is some limitation on testing time or the number of hypotheses tested. In that case, however, it would then become difficult to explain why communication succeeds so frequently. For why should it be the case that the optimally relevant interpretation of an utterance is nearly always found in a limited set of randomly selected interpretive hypotheses?

We may be able to make things work without abandoning the assumption that the PPM tests interpretive hypotheses in order of accessibility, however, for biasing the set of tested interpretive hypotheses towards those hypotheses with greater accessibility should make it more likely that the optimally relevant interpretation is among the first interpretive hypotheses to be tested. Assuming that audiences test in order of accessibility and that there is a time limit on testing, communicators who wish to be understood have good reason to make their intended interpretation one of those which are most accessible to the audience. If relevance theory is
correct, then communicators are also attempting to make their intended interpretation optimally relevant. When the communicator is successful in his task, then, the optimally relevant interpretation is likely to be among those most accessible to the audience. When the communicator fails to make his intended interpretation optimally relevant to his or her audience, communication is likely to fail.

Given the assumption that, in cases in which the communicator has succeeded in his tasks, the optimally relevant interpretation is more likely to be among those interpretations which are most accessible to the audience, we have two options if we wish to salvage the basic picture of the CIP testing process given above, though it seems to me that they ultimately boil down to the same thing. The first option would be to simply bite the bullet and concede that the induction being used during the testing process is invalid. After all, humans are notoriously poor at certain sorts of reasoning and frequently fall prey to fallacies of various kinds. And given that the bias toward accessibility in the set of tested hypotheses actually increases the likelihood that the optimally relevant interpretation will be in that set, the fact that the inductive inference involved is an invalid one becomes of significantly less import. When it comes to the CIP testing process, what is important is explaining how the audience manages to successfully find the optimally relevant interpretation for a communicator’s utterance most of the time, not the logical validity of the process used.

The second option would be to abandon the label ‘induction’ and instead view the CIP testing process as a search process similar to those which have been found effective in work on
artificial intelligence.\textsuperscript{123} The actual testing process would be the same – interpretive hypotheses would be tested in order of accessibility until the time limit for testing has been reached, and then the most relevant hypothesis in the tested set would be selected. Instead of thinking of the testing process as one in which a conclusion is drawn on the basis of a sample set, however, we could instead think of the PPM as performing a search for the optimally relevant interpretation. Since testing occurs in order of accessibility, the search process is an \textit{informed} one, as it uses a heuristic function to order the search. Indeed, the search for the optimally relevant interpretation would seem to be largely analogous to what is known as a greedy best-first search.\textsuperscript{124} A greedy best-first search is one which proceeds by first evaluating those potential solutions which are closest to the goal. Basically, the idea is to evaluate first those potential solutions which are most likely to be the optimal solution available. The main difference between the standard greedy best-first search and the search for the optimally relevant interpretation as we have envisioned it is the heuristic function. The standard greedy best-first search is ordered on the basis of the estimated cost to reach the goal from the current state (since potential solution paths with the least cost are most likely to be the optimal solution), while the search for the optimally relevant interpretation would be ordered on the basis of accessibility of the interpretive hypotheses to the individual at the time the search is being performed. The basic principle is the same, however, since, as we saw above, the optimally relevant interpretation is likely to be among the interpretive hypotheses which are most accessible to the audience.

\textsuperscript{123} We do not have space here for a detailed discussion of the various AI approaches to problem solving via searches. For an overview see Russell & Norvig (2010), chapters 3-6.
\textsuperscript{124} See Russell & Norvig (2010), pp. 92-93
As should be clear, treating the CIP testing process as a greedy best-first search occurring within a set time limit is not all that different from our original characterization of the process as a case of inductive inference. In both cases, the PPM will evaluate interpretive hypotheses in order of their accessibility for whatever set time is allotted to pragmatic processing of utterances. And, in both cases, when the time limit has been reached, the most relevant interpretive hypothesis in the set will be selected as optimally relevant and used as the interpretation for the utterance in question. The main difference is that in treating the CIP testing process as a search process rather than an inductive inference, we are avoiding the need to claim that a valid conclusion is drawn on the basis of the tested set of interpretive hypotheses. Notice that this CIP account fits nicely with the standard understanding of the interpretation selection process as a process of testing in order of accessibility and stopping when the expectation of optimal relevance is satisfied. Notice also that the CIP account incorporates the comparative understanding of optimal relevance and that it is does so without generating unresolved inconsistencies or puzzling incoherence in the account as a whole, something the ITP account is incapable of doing. At this point, then, the CIP account seems to be the best available way to fill in the relevance theoretic account of how interpretive hypotheses are evaluated and the optimally relevant interpretation selected.

Before moving on, there is one last facet of the pragmatic processing step of utterance interpretation that we should take note of. The foregoing discussion may leave the reader with the impression that the interpretation of an utterance is selected in its entirety after the utterance is completed. This is not quite correct, however. According to relevance theory, an individual involved in interpreting an utterance does not wait until the utterance is complete to begin
interpretation; his or her interpretation is ongoing as the utterance is being produced. This is why we are capable (to a point) of interpreting half-finished or partially heard utterances. It also explains our initial incorrect interpretations in garden path utterances such as those in (5.9) and (5.10).

(5.9) The man who hunts ducks out on weekends.
(5.10) She told me a little white lie will come back to haunt me.

Sperber & Wilson suggest that such on-line interpretation is done, at least in part, on the basis of anticipatory hypotheses about the logical or syntactic structure of the utterance. On this view the PPM provisionally selects the interpretation for the part of the utterance already heard which would, given the hypothesized structure of the utterance, yield the optimally relevant overall interpretation.

At first blush, one might worry that this view of how interpretation proceeds is incompatible with the CIP interpretation process. After all, as we have characterized it, CIP is an interpretation process in which available interpretations of the entire utterance are assessed for relevance and compared with one another, resulting in the selection of the most relevant interpretation of the entire utterance. But the on-line interpretation process we are currently looking at depicts the interpretation process as occurring in a piecemeal fashion, where parts of the utterance are assigned interpretations before the utterance itself is complete. It is important to keep in mind, however, that while the on-line interpretation process does in fact provisionally

---

125 Sperber & Wilson (1986) pp. 205-208. It is unclear where Sperber & Wilson take these hypotheses to be generated. It seems likely that it is the PPM, however. Given that syntactic processing is the domain of the linguistics processing module, one might assume that these anticipatory hypotheses are generated there and then delivered to the PPM along with the bits of LF that give rise to them. However, the linguistics processing module may not have access to the information necessary to make anticipatory hypotheses about logical structure, since that information is stored in the logical entry for a concept. The linguistics processing module has access to conceptual addresses, but Sperber & Wilson give us no reason to think it has access to the information stored at those addresses.
assign interpretations to only parts of the utterance, it does so on the basis of an assessment of the relevance of potential interpretations of the whole utterance, or, at least, the whole utterance as depicted by anticipatory hypotheses about its structure. Given this, the notion that interpretation of utterances is an on-line process is consistent with the claim that relevance assessment occurs via the CIP.

With a detailed characterization of relevance theory in hand, we are now in a position to put forward our relevance theoretic account of definite descriptions. We will turn to this task in the following chapter.
Chapter 6
An Account of Definite Descriptions

With a more detailed picture of relevance theory, we are now in a position to construct our relevance theoretic account of definite descriptions. While the account developed in this chapter will bear many similarities to previous contextualist and relevance theoretic accounts of descriptions, it will also differ in a number of significant ways, hopefully avoiding the pitfalls and problems of these previous accounts. I will begin by giving a general sketch of the overall view in §6.1. §6.2 will be dedicated to a discussion of descriptive, de re and type concepts. §6.3 and §6.4 attempt to expand the account to still further uses of descriptions. Finally, in §6.5, we will look at a rather serious problem which threatened previous relevance theoretic accounts of definite descriptions and which threatens the account sketched here as well – the problem of testability. Much of the work in the remaining chapters will be dedicated to addressing this problem.

6.1 The Relevance Theoretic Account

As we have seen, definite descriptions have both referential and attributive readings. According to Russellians, the linguistic meaning of a definite description is its attributive reading – utterances containing definite descriptions always explicitly express general propositions. Referentially used descriptions, on this view, arise when a singular proposition is implicated. In contrast, referentialists hold that definite descriptions are ambiguous and hence utterances containing them explicitly express different propositions in different contexts. Essentially, referentialists hold that the different uses of definite descriptions arise via a semantic ambiguity
to impact the level of what is said, while Russellians hold that the different uses arise via Gricean pragmatic factors to impact the level of what is meant, or what is implicated. Both Russellians and referentialists seem to limit their focus to only two levels of meaning – what is said and what is meant (or, in relevance theoretic terms, what is explicitly communicated and what is explicitly communicated).

The referentialist view of definite descriptions is largely motivated by the intuition that in a context in which a description is used referentially, the singular proposition is what is said – that is, the singular proposition is explicitly expressed. The Russelian contends, however, that positing an extra linguistic sense for definite descriptions is theoretically unnecessary, and hence should be avoided. This perspective becomes even more plausible when we recognize that there are multiple uses of definite descriptions beyond the referential/attributive distinction, each of which would have to be a different linguistic sense on the referentialist’s view. The contextualist view is motivated by the concerns of both the referentialist and the Russelian. And, as we saw in earlier chapters, by expanding the focus beyond the what is said and what is meant to include a level of linguistic meaning, the contextualist can claim that definite descriptions have a single linguistic meaning which will be supplemented differently in different contexts, and hence will give rise to different explicitly communicated contents on different occasions.

Relevance theoretic accounts of definite descriptions rely on the same contextualist strategy. According to relevance theory, remember, an utterance encodes a linguistic meaning which then undergoes pragmatic processing, ultimately resulting in explicatures (including a propositional form) and also, potentially, implicatures. The relevance theoretic accounts of descriptions offered by Bezuidenhout and Powell both maintain that definite descriptions are
univocal. On these views, a sentence containing a definite description encodes a linguistic meaning which fails to determine a proposition, and will give rise to different propositional forms in different contexts. In some contexts, pragmatic processing of the LF of an utterance containing a description will result in an attributive reading of the description, and in others, it will lead to a referential reading. In this section, we will attempt to construct a relevance theoretic account of definite descriptions by building on the accounts offered by Bezuidenhout and Powell. Since both Bezuidenhout and Powell claim that definite descriptions encode procedural information, we will begin by taking a closer look at procedural meaning.

6.1.1 Procedural Meaning

Most linguistic expressions encode concepts or conceptual information. The notion that some expressions might linguistically encode procedural information instead of (or in addition to) conceptual information was first suggested by Diane Blakemore in an attempt to give a relevance theoretic account of discourse connectives.\textsuperscript{126} According to Blakemore, the notion of linguistically encoded procedural information is made plausible by the relevance theoretic framework. If relevance theory is correct, then the output of linguistic processing serves as the input for pragmatic processing. It is entirely possible, then, that some of the information decoded during linguistic processing would be procedural rather than conceptual. The idea is that since there is no reason why the entire content of an LF must be conceptual, it is feasible to suppose that sometimes an LF contains content which serves to constrain the processing of the conceptual representation rather than being part of the conceptual representation itself. Indeed, if constraints on the pragmatic processing of a conceptual representation were necessary for one reason or

\textsuperscript{126} Blakemore (1987)
another, it would make sense for those constraints to come along with that conceptual representation.

Of course, the simple fact that procedural meaning is possible, or even plausible, given the relevance theoretic account of communication is not sufficient reason for thinking that there is such a thing. But Blakemore also contends that the notion of procedural meaning is theoretically valuable. In particular, she argues that the relevance theorist can account for discourse connectives of all sorts by relying on the notion that such connectives encode procedural constraints – that is, constraints on the pragmatic processing of the LF of an utterance containing such expressions.\(^\text{127}\) Robyn Carston has similarly relied on the notion of procedural meaning in accounting for discourse connectives.\(^\text{128}\) Procedural meaning has also been used to account for mood indicators, illocutionary particles and attitudinal particles.\(^\text{129}\) And Sperber & Wilson have relied on the notion of procedural meaning to account for pronouns.\(^\text{130}\) Since the notion of procedural meaning has such broad ranging theoretical utility for relevance theory, it would seem that the relevance theorist has reason to accept it. Of course, theoretical utility only gives us reason to accept the notion of procedural meaning if the theories using it are successful ones, and it is by no means certain that the relevance theoretic theories that rely on procedural meaning are in fact successful. But procedural meaning isn’t theoretically useful to the relevance theorist alone.

In addition to its utility in explaining various linguistic phenomena from the relevance theoretic perspective, the idea that certain expressions encode procedural information helps to

\(^{127}\) Blakemore (1987), (2002)
\(^{128}\) Carston (2002)
\(^{130}\) Wilson & Sperber (1993)
explain why we are unable to give definitions for some expressions in our language. While it is certainly true that native speakers might disagree on the definitions of certain terms like ‘freedom’ or ‘justice’ or ‘marriage’, they are at least capable of providing paraphrases of these terms which share elements in common. Discourse connectives like ‘but’, ‘however’ or utterance initial ‘well’, on the other hand, are notoriously difficult to paraphrase. Moreover, unlike expressions like ‘hungry’ and ‘thirsty’ or ‘dog’ and ‘puppy’, it is not possible to determine whether, say, ‘but’ and ‘however’ are synonymous without looking at whether or not they are intersubstitutable in a context. According to Blake, the reason that we cannot provide paraphrases for these terms or determine their synonymy relations without a substitutability test is that they encode procedural rather than conceptual information. Conceptual information is available to consciousness, but since encoded procedural information is meant to constrain pragmatic processing – a subconscious processes – procedural meaning is not available to consciousness. Hence we cannot consciously access the meanings of expressions which encode procedural information in the same way that we can access the meanings of terms which encode conceptual information.

It would seem that the notion of procedural meaning has earned a place in the relevance theoretic framework. By using it in a relevance theoretic account of definite descriptions, then, it looks like we would merely be relying on a piece of the relevance theoretic machinery which is already much in use. Moreover, given what has been said about procedural meaning, there is some reason to think that the definite article is among those expressions which encode procedural information rather than conceptual information. Ask any native speaker of English to give you a definition of the word ‘the’, and you’ll likely get silence. And the years of debate
between Russellians and referentialists should be enough to tell us that determining synonymy relations between ‘the’ and other expressions (such as ‘there exists exactly one’) is not something that can be consciously done by native speakers. Both of these features are, according to Blakemore, markers of expressions that encode procedural rather than conceptual information.

It may be worth noting that all those expressions which relevance theorists have thus far accounted for in terms of procedural meaning belong to the class of expressions commonly known as function words. Function words, as opposed to content words, are notoriously difficult to define, and they are members of a closed-class. They are frequently described as functioning to provide grammatical structure and indicate grammatical relations between the content words in the sentence in which they occur. The fact that those expressions which have so far been pegged as encoding procedural information are all function words is suggestive. While function words are often described as merely providing grammatical structure, it cannot be the case that this is necessarily all that they do, or even that this is necessarily what they primarily do. In the vast majority of sentences in which it occurs, the discourse connective ‘however’ is syntactically superfluous; it can be left in or removed without any impact on the syntactic relations between the other expressions in the sentence. Of course, as a discourse connective, one might argue that it functions to indicate grammatical relations between sentences. But this is true only to the extent that semantics is a part of grammar, for the expression ‘however’ is used to indicate a certain relation between the contents of sentences in a discourse (that of some sort of contrast). It does not serve to indicate some sort of structural relationship. Perhaps, then, what function words truly have in common is not that they serve only to provide syntactic structure and mark off
syntactic relationships, but rather that they all encode procedural rather than conceptual information.

I cannot argue this point completely here. Suffice it to say that it might be worth looking into the possibility that all function words encode procedural rather than conceptual information. What is important for our purposes here is the fact that, just like the expressions which relevance theorists have so far accounted for in terms of procedural meaning, the definite article is a function word. Since the definite article belongs to the same class as these expressions which have been accounted for in terms of procedural meaning and which exhibit similar properties, there is good reason to think that we should be able to account for the definite article in the same way.

6.1.2 The LF

Both Bezuidenhout and Powell treat definite descriptions in terms of procedural meaning, but they provide us with different pictures of the LF of utterances containing such descriptions. Bezuidenhout takes the LF of an utterance of the form *the F is G* to be something like (6.1), where what occurs inside the arrow brackets should be understood as encoded procedural constraints on pragmatic processing.

\[(6.1) \quad \llbracket \text{Feature } F \text{ is instantiated } \langle \text{uniquely/accessibly} \rangle \text{ by an } x \text{ which is } G \rrbracket\]

The accessibility constraint here would seem to be superfluous. According to relevance theory, the audience of an utterance utilizes an interpretation strategy which involves testing interpretive hypotheses in order of accessibility. Thus, referents with high levels of accessibility are sure to be considered, with or without an accessibility constraint built into the procedural information encoded by the definite article. Of course, the accessibility constraint here might be motivated by
the idea that the referent of a description must be something highly accessible. It is not clear to me that this assumption is correct. Even if it is, unless we discover that considerations of relevance alone are insufficient to ensure that the referent of the description is sufficiently accessible, there will be no reason to hold that the LF of descriptions contains an accessibility constraint.

Bezuidenhout’s account of the LF of descriptions faces a much more significant problem, however. As we saw in our discussion of her view in chapter 4, her characterization of the LF is problematic insofar as it privileges the attributive use of descriptions over the referential use. There is no reason to think that the attributive use is generally more likely to occur, or that it is more prevalent than the referential use. There is also no reason, at least on the face of things, to think that attributively used descriptions requires less processing than referentially used ones. As we saw in the previous chapter, this would seem to suggest that Bezuidenhout’s characterization of the LF of descriptions would yield incorrect predictions as to the proper interpretation of utterances containing definite descriptions. We will want to avoid making the same mistake in constructing our account of the LF of descriptions. Thus, our account of the LF should be one on which the same type of processing can yield the different interpretations of definite descriptions without privileging one interpretation over the others, rather than one like Bezuidenhout’s which holds that different types of processing arise out of a determination as to how the description is being used.

Powell’s understanding of the LF of definite descriptions would seem to meet this requirement. According to Powell, “definite descriptions are univocal and encode procedures for the retrieval of propositional constituents; they instruct a hearer to access an individual concept,
whether descriptive or *de re*, which includes the descriptive content of the description."\(^{131}\) He also characterizes the content encoded by a definite description as an “instruction to access an individual concept.”\(^{132}\) An individual concept, remember, is one which is taken by its possessor to represent an individual and contains information which uniquely identifies its referent. On Powell’s view whether a description is interpreted as being used referentially or attributively depends on what type of individual concept is selected. In the case of an attributive use, a descriptive concept is selected. In the case of a referential use, a *de re* concept is selected. We will discuss these different types of concepts in detail in the following section. What is important for our purposes here is that, in contrast to Bezuidenhout’s view, Powell’s characterization of the LF of descriptions does not privilege one interpretation over another.

In light of this, Powell’s view would seem to make for a better starting point in constructing our own view of the LF of descriptions, though there are a few details that merit a closer look. Unlike Bezuidenhout, Powell does not provide us with a definite representation of what he takes to be linguistically encoded by definite descriptions. As a result, we are left with a somewhat unclear picture of just what Powell takes the encoded procedural content to be. As we just saw, Powell characterizes the encoded procedural content as an “instruction” – presumably to the pragmatics processing module – to access an individual concept. But what exactly does this instruction look like? And what reason do we have for thinking that in addition to (or instead of) the sorts of constraints that relevance theorists have previously understood as characteristic of procedural meaning, there are also instructions? Powell does not provide us with answers to these questions.

\(^{131}\) Powell (2001) p. 118
\(^{132}\) Ibid p. 123
There is no reason for us to take on these difficulties, however. Blakemore’s original understanding of procedural content views it as encoded information which limits the output of pragmatic processing of the LF in which it occurs. This is the view of procedural content which Sperber & Wilson and Bezuidenhout rely on. Taken in this light, we could understand the content encoded by the definite article as essentially a label attached to the NP, marking it as an individual concept. Thus, we can represent the LF of an utterance of the form *the F is G* using the bracketed diagram given in (6.2) where the arrow brackets indicate a procedural constraint applicable to the entire NP.

\[(6.2) \quad [[[\text{<individual concept>}}][F]][[\text{is}][G]]\]

The idea is, essentially, that the definite article encodes a procedural constraint indicating that the NP of which it is a part should be interpreted as an individual concept. Since this view relies on Blakemore’s original view of procedural content, we do not have to worry about making sense of encoded instruction, as would have been necessary had we taken on Powell’s characterization of the encoded procedural content. And since the class of individual concepts includes both descriptive and *de re* concepts, this view of the LF does not privilege one reading of descriptions over another, as Bezuidenhout’s view does.

At this point, then, we have an account which treats the English definite article as encoding a procedural constraint which limits the interpretation of the NP to an individual concept. The different readings of definite descriptions result from the selection of different types of individual concept as enrichments of the element of LF encoded by the description. The attributive reading results from the selection of a descriptive concept, while the referential reading results from the selection of a *de re* concept. This is essentially the same as the account
given by Powell, except that instead of treating the encoded procedural content as an instruction
to the pragmatics processing module, it views this content as a constraint on processing. We will
see in the next section, however, that fully accounting for definite descriptions will require
making additions to this account.

6.2 Individual Concepts

A key element of the account we have just given of the linguistic meaning of definite
descriptions is Powell’s notion of an individual concept. Since on our account the different
readings of descriptions are the result of the selection of different types of individual concepts
during pragmatic processing, we will need to have a clear understanding of the ways in which
these types of concepts differ if we are to be able to test our account using the method proffered
in the previous chapter. We will attempt to give a more detailed picture of these types of
concepts in this section.

First, however, we should say a few things about the relevance theoretic view of
concepts. Sperber & Wilson view concepts as addresses in memory at which three different types
of information are stored: lexical, logical, and encyclopedic. The lexical entry of a concept
contains information about which natural language expressions encode the concept, while the
encyclopedic entry of a concept contains information about the extension or denotation of the
concept. The logical entry contains the set of deductive elimination rules associated with the
concept. This will include both rules of logic as well as rules associated with the meaning of the
concept. So, for example, at the conceptual address for the concept DOG would be a rule allowing

---

133 Sperber & Wilson (1986) p. 86
the deductive device to infer \textit{x is an animal} from \textit{x is a dog}.

With this in mind, let us turn to a discussion of the different types of individual concept.

### 6.2.1 Descriptive and \textit{De Re} Concepts

An individual concept, remember, is a concept which is taken by its possessor to represent an individual or set of individuals and which contains information that uniquely identifies its referent. As both descriptive and \textit{de re} concepts have individuals or sets of individuals as their referents, they are both types of individual concepts.

The most basic way of understanding the difference between descriptive concepts and \textit{de re} concepts is in terms of the types of thoughts in which they occur. \textit{De re} concepts occur in singular thoughts. Descriptive concepts occur in non-singular thoughts about individuals. The difference between these types of concepts, then, will ultimately boil down to the difference between singular and non-singular – or descriptive – thoughts. The different types of concept can also be distinguished in terms of the relationship between concept and referent which is characteristic of them. A descriptive concept bears a satisfactional relationship to its referent. With descriptive concepts the referent of the concept is whatever object satisfies the descriptive content of the concept. With \textit{de re} concepts, on the other hand, the referent of the concept is an individual which stands in a non-satisfactional relationship to the content of the concept. In particular, the content of the concept is causally connected to the referent of the concept.

This understanding of the difference between descriptive and \textit{de re} concepts parallels the common

\footnotesize
\begin{itemize}
  \item \textsuperscript{134} Cf. Sperber & Wilson (1986) pp. 90-93
  \item \textsuperscript{135} For support of the claim that acquaintance is a condition on singular thought, see Bach (1987), (2010); Donnellan (1979); Evans (1982); Kaplan (1989a); Recanati (1993), (2010); Reimer (2004); Salmon (1987).
\end{itemize}
understanding of the difference between singular and non-singular and descriptive thought. \textsuperscript{136}

We will see later than descriptive and \textit{de re} concepts may not be the only sorts of individual concept, but first, let’s take a closer look at what these types of concepts would look like in light of the relevance theoretic view of concepts in general.

Relevance theory treats concepts as addresses in memory at which different types of information is stored. Given this view of concepts, one might wonder just what it means for the referent of a descriptive concept to bear a satisfacional relationship to its content. While we generally assume that the information stored in the encyclopedic entry of a concept accurately describes the concept’s referent, what is important in the case of a descriptive concept is that the referent of the concept uniquely satisfies a certain descriptive content, or would uniquely satisfy it if there were one. Let us use dashes to indicate a descriptive concept with a particular content. So, for example, \textsc{author-of-beowulf} is a descriptive concept whose referent satisfies the content \textsc{author of beowulf}. On the relevance theoretic view of concepts, the logical entry of the concept contains rules associated with the meaning of the concept. It seems to me, then, that the logical entry of a descriptive concept \textit{X-Y-Z} will contain a rule indicating that the referent of the concept, if there is one, satisfies the content \textit{X Y Z}. So, for example, in the logical entry of the descriptive concept \textsc{author-of-beowulf} we would find a rule indicating that the referent of the concept, if there is one, satisfies the content \textsc{author of beowulf}. Having this sort of rule in its logical entry is what distinguishes a descriptive concept from other sorts of individual concepts.

As we saw above, the general consensus is that \textit{de re} concepts are distinguished from other concepts by the fact that their contents are causally connected with their referents. This is

\textsuperscript{136} For example, see Jeshion’s characterization of the different types of thought in Jeshion (2010) pp. 1-2
sometimes known as the acquaintance condition on singular thought. There is some
disagreement as to the precise nature of the causal relationship in question, however. The notion
of acquaintance as a condition on singular thought had its genesis with Russell. On Russell’s
view, we can only have singular thoughts about those things which we are directly acquainted
with, where direct acquaintance with an object requires complete, unmediated awareness which
allows us to identify the object without risk of error. 137 This acquaintance condition would
appear to be entirely too strong. If we must have the sort of direct acquaintance envisioned by
Russell in order to have singular thoughts, then the only possible objects of singular thoughts are
sense data. But, intuitively, we can have de re thoughts about people, places or objects. They are
the sorts of thoughts we take ourselves to be expressing when we use demonstratives, names or
referentially used descriptions. A weaker acquaintance condition than that envisioned by Russell,
then, is necessary.

The main three alternative accounts of singular thought in terms of acquaintance come
from Evans, Bach, and Recanati. Each of the accounts is closely related. Evans for example,
accounts for singular thought in terms of what he calls “information based” thoughts – thoughts
whose content is perceptually derived from their object. 138 Recanati accounts for singular
thought in terms of mental files whose content derives from being linked to the object by a
special sort of causal chain – one which allows for the flow of information. 139 Bach maintains
that singular thoughts arise out of a perception-based causal links to their object. 140 This is the
acquaintance view which Powell endorses. We do not have space here to evaluate all these

137 Russell (1905); (1911); (1912)
138 Evans (1982)
139 Recanati (1993), (2010)
different views, nor do I think that the debate over which is the best view can be quickly or easily
resolved. Fortunately, we do not need to resolve the issue here. These main views of the
acquaintance relation are all based around the same core ideas – that the acquaintance
relationship is causal and that the content of a *de re* concept is acquired via that relationship.
Where the views differ is in their understanding of what sort of causal relationship is required.
For Evans, the causal relationship must be perceptual. For Bach, the causal relationship need not
be wholly perceptual, but it must have a perceptual base. For Recanati, the causal relationship
need not be perceptual at all. But on all of these views, there must be some sort of causal link
between the content of a *de re* concept and the referent of that concept. To put it another way,
each of these views holds that *de re* concepts must be object-dependent, in the sense that the
content of the concept is causally dependent on its object.

It seems to me that we may leave open the question of which of these accounts of the
acquaintance condition on *de re* thought is correct. Since the three main views share a core
understanding of the acquaintance relationship, they will largely agree on whether or not a
particular thought is a singular one. Where they will differ is with respect to those thoughts about
which our intuitions are already uncertain. Suppose I am visually attending to a book on the
desk, and I have the thought *that book is blue*. This is a standard example of a singular thought,
and all three views would categorize it as such. The same would be true if the thought was had as
a result of my memory of a perception of the book on the desk. Now, take a different example. I
have never personally perceived Aristotle, but I have acquired my concept of Aristotle as a result
of a causal-historical chain based in someone’s actual perception of Aristotle. Suppose I have the
thought *Aristotle was a great philosopher*. Is my thought singular? Intuitions in this case are mixed. So, too, are the pronouncements of the different theories of acquaintance.

Since the referential and attributive uses of definite descriptions are tied to *de re* and descriptive thoughts, a lack of clarity in our intuitions about the classification of a thought as descriptive or *de re* will make for a lack of clarity in the identifying the proper interpretation of a use of a description as either referential or attributive. If I cannot determine whether the thought *Aristotle was a great philosopher* is singular thought, then I cannot determine whether or not an utterance of the sentence “Aristotle was a great philosopher” expresses a singular thought. The situation would be the same in the case of a definite description. Given this, there is no need for a theory of descriptions to include commitments as to the precise nature of the causal relationship involved in *de re* concepts. We want our account of descriptions to make sense of our intuitions as to the natural interpretations of definite descriptions. Cases in which it is unclear what would be the natural interpretation of an utterance involving a particular linguistic construction are not suitable for testing an account of that construction. In order to test such an account, one needs cases in which the intuitively preferred interpretation is clear, so that it is possible to determine whether the interpretation predicted by the account is, in fact, the natural interpretation. Thus, since our interest will ultimately be in testing our relevance theoretic account of descriptions, those cases in which it is unclear whether a thought is *de re* or not will be of no interest to us. If we are only concerned about those cases in which we can clearly identify a thought as descriptive or *de re*, we need not worry about those differences in the accounts of acquaintance which result in different categorizations in the more problematic cases. The core understanding of *de re* thought shared by all the theories should be sufficient for our purposes.
There is one last point we should note about *de re* concepts. Such concepts are generally thought of as the cognitive counterparts of directly referential terms. A directly referential term is one which contributes solely its referent (as opposed to something like an intension or a Fregean sense) to the proposition expressed. Indexical expressions like ‘I’ or ‘you’ are typical examples. Proper names are also frequently considered directly referential expressions. It is generally assumed that the cognitive representation of the content of an utterance containing a directly referential term will contain a *de re* concept which serves as the interpretation of the directly referential term. As a result of this link between directly referential terms and *de re* concepts, it is tempting to assume that since directly referential terms contribute solely their referents to the proposition expressed, the case is much the same with *de re* concepts – that *de re* concepts involve a referent and nothing more.

There is a significant problem with making this assumption about the content of *de re* concepts, however, one which we can make clear via a case borrowed from Mark Richard.\(^{141}\) In Richard’s case, there are two women, let us call them Anna and Betty, who are speaking on the phone. Anna has no reason to think that the woman she is speaking on the phone with is in danger, and hence forms the belief *she is not in danger* about the woman she is speaking with. However, Anna then takes a look out of her window and sees a woman in a telephone booth across the way that is clearly in danger. As a result, she forms the belief *she is in danger* about the woman in the booth. Unbeknownst to her, this woman is Betty.

If the content of *de re* concepts is limited to solely their referents, then Anna has inconsistent beliefs, as each of the two beliefs in question contains a *de re* concept of Betty.

\(^{141}\) Richard (1983)
Anna believes both that Betty is in danger and that Betty is not in danger. Most of us, however, have the intuition that Anna is not guilty of irrationally holding two inconsistent beliefs. Rather, it would seem that Anna’s two beliefs differ from one another in such a way as to make it possible for Anna to hold them without being irrationally inconsistent. The *de re* concepts employed in each belief must have content beyond simply their referents. Following Bach, let us call this content the *mode of presentation* of the concept.

One might be worried that by introducing modes of presentation, we risk losing any real distinction between descriptive concepts and *de re* concepts. After all, what could the mode of presentation associated with a particular *de re* concept be besides some sort of descriptive content? As it turns out, there is a way to construe the modes of presentation associated with *de re* concepts without appeal to descriptive content. According to Recanati, for example,

> Modes of presentation are now construed as *ways the object is given to the subject*, and an object may be given either directly, in experience, or indirectly, via descriptions. Non-descriptive modes of presentation are ways the object is (directly) given to the subject in experience, while descriptive modes of presentation are ways the object is (indirectly) given via properties which it uniquely instantiates. When, facing Monte Blanc, the subject thinks ‘That mountain is less than 4,000 metres high’, she thinks of Mont Blanc under a non-descriptive mode of presentation based on her perceptual relation to Mont Blanc.\(^{142}\)

The mode of presentation associated with a *de re* concept, then, is a way the object is presented to the subject via experience. Thus, the mode of presentation in such cases is bound up with the causal link between concept and referent via which the concept acquires its content – that is, via the acquaintance relation. To put it another way, we might think of *de re* modes of presentation as *ways of being acquainted*. In the case of Anna and Betty, then, Anna is not guilty of holding the clearly inconsistent pair of beliefs *Betty is in danger* and *Betty is not in danger*. Betty is

---

presented to Anna in two different ways, and hence her two beliefs contain two different *de re* concepts with different modes of presentation.

In summary, then, descriptive concepts and *de re* concepts are types of individual concepts which differ from one another in how they relate to their referents. A descriptive concept stands in a satisfactional relationship with its referent, whereas *a de re* concept bears a causal relationship to its referent. As a result of this difference, the two types of concepts stand in different relationships to the linguistic constructions which express them. The content of an attributively used description matches the content of the associated descriptive concept, while the content of a referentially used description serves as an aide in the identification of the referent of the associated *de re* concept and potentially in the identification of the concept’s mode of presentation as well.

### 6.2.2 Type Concepts

If we were solely concerned with the referential/attributive distinction, descriptive and *de re* concepts would be all we needed. But as we have seen definite descriptions have uses beyond the referential and the attributive; they can also be used generically, as in an utterance of (6.3), predicatively, as in an utterance of (6.4) (where the ‘is’ is the ‘is’ of predication), or functionally, as in an utterance of (6.5).

(6.3) The lobo is endangered.
(6.4) Shakespeare is the greatest playwright of all time.
(6.5) The President of the U.S. changes every 4 years.

Definite descriptions may have another use as well. Imagine an utterance of (6.6) made by a speaker who has a particular individual in mind, but has no intention that his or her audience think of that particular person.
(6.6) The man I live next door to has some strange habits. According to Ludlow and Neale, such uses, which they call *specific*, are neither referential nor attributive.143 We’ll discuss the specific, functional, and predicative uses of definite descriptions in §6.3. In this section, our concern will be the generic use.

In the literature on linguistic generics, referential expressions like the description in (6.3) are called *generic NPs* or a *kind-referring NPs*.144 These are NPs which pick out well established kinds (as opposed to individual objects). Because there are a number of verbal predicates – ‘be extinct’, ‘evolve from’, or ‘invent’, for example – that require this sort of NP in at least one argument place, and appear to be applicable only to kinds and not to members of those kinds, it seems clear that these NPs are referring to kinds.145 In order to make sense of kind-referring descriptions – and possibly other kind referring NPs – I would like to suggest that in addition to descriptive and *de re* concepts of particular concrete objects, we also have individual concepts of kinds. These concepts are not merely descriptive or *de re* concepts which have kinds as their referents, but rather are a distinct class of concept falling under the rubric of individual concepts, in Powell’s sense of the term. Let us call these *type concepts*. Appealing to this sort of concept should allow the relevance theorist to make sense of the cognitive representations of utterances like (6.3). On our account of descriptions, remember, the definite article encodes a procedural constraint which limits interpretation of the NP to individual concepts. My suggestion is that kind-referring descriptions like that in (6.3) refer to kinds as individuals because they are

143 Ludlow & Neale (1991)
144 See, e.g., Krifka et. al. (1995) p. 2
145 Krifka et. al. (1995) pp. 63-64
mentally represented by type concepts, that is, the kind-referring reading of a description results from the selection of a type concept.

Of course, the mere fact that kind-referring NPs refer to kinds does not necessarily mean that the mental representations of utterances containing these NPs will involve type concepts. We might think that interpretations of utterances containing kind-referring NPs are mentally represented using descriptive or *de re* concepts whose referents are types rather than objects. As it turns out, however, neither descriptive nor *de re* concepts are suitable for making sense of these mental representations. The reason descriptive concepts are not suitable is the simple fact that the relationship between a kind-referring definite description and its referent does not appear to be a satisfactional one. The description ‘the lobo’ does not pick out the species *canis lupus baileyi* because that species satisfies the descriptive content ‘lobo’. Indeed, it seems strange to even call this descriptive content, as it does not describe so much as name. In fact, one popular theory of kind-referring NPs in linguistics treats them as proper names.\(^{146}\)

Given the apparent similarity with names here, one might be led to expect that *de re* concepts would be suited for making sense of the mental representations of generic descriptions. After all, names are usually thought to be directly referential terms, and it is common to account for the mental representations of directly referential terms using *de re* concepts. This would certainly simplify things for us here. If we made sense of the mental representations associated with utterances containing generic descriptions in terms of *de re* concepts, however, we would be unable to account for all the various situations in which generic descriptions can be used. To see this, we must remember that it is an essential quality of a *de re* concepts that they are causally

\(^{146}\text{Carlson (1977); Heyer (1985); Krifka et. al. (1995)}\)
connected to their referents. Types are not the sorts of things that we are causally connected to. That said, it may be the case that causal contact with tokens is sufficient to count as causal contact with the type, at least for the purposes of connecting a *de re* concept with its referent. But it is not uncommon for us to use generic descriptions to refer to types even though we have never had contact with any tokens of that type. Take, for example, the generic description in (6.7).

(6.7) The Italian city-state allowed for the rise of a merchant class.

As the vast majority of Italian city states vanished before the 17th century, and were certainly gone by the end of Italian unification in 1871, it is highly unlikely that anyone today has had causal contact with any token of the type *Italian city-state*. Since *de re* concepts are causally linked with their referents, we cannot be said to have *de re* concepts of the type Italian city-state. Yet we seem perfectly capable of interpreting and understanding (6.7).

It would seem, then, that in addition to descriptive and *de re* concepts there must be a third kind of concept which allows us to refer to kinds – a type concept. Such concepts can be viewed as individual concepts in Powell’s sense, for they take as their referents individual kinds. Additionally, there is reason to suppose that type concepts involve modes of presentation of their referents. First, the other the other classes of individual concept – *de re* and descriptive concepts – both involve modes of presentation of their referents. We should not be surprised, then, to find that type concepts do as well. Moreover, our conceptualization of types would seem to be subject to Frege’s puzzle in much the same way that our conception of concrete objects is. For example, as a child, I frequently saw insects known to many in the southwestern United States as ‘children

---

147 Nor would it be sufficient to have had causal contact with something that *once* tokened the type. We would not say that you have been in causal contact with sand because you have touched a glass window, even if some of the material the window is made of once tokened the type *sand.*
of the earth’. Given their frightening aspect and the many (ultimately mistaken) things I was told, I came to believe that children of the earth are dangerously venomous creatures. Later on, I read about Jerusalem crickets, and I learned that they are sometimes called ‘potato bugs’, they are largely harmless, and they are not venomous. So, I had the following two beliefs:

(6.8) The child of the earth is a very dangerous, venomous creature.
(6.9) The Jerusalem cricket is a largely harmless, non-venomous creature.

Now, as it turns out, children of the earth and Jerusalem crickets are just different names for the same genus of insect – *stenopelmatus*. Thus, the propositions expressed by (6.8) and (6.9) would seem to be inconsistent, and yet we would not wish to say that I was irrational in holding both beliefs before learning that children of the earth are in fact just Jerusalem crickets. The standard way around this puzzle is to assume that the different beliefs involve concepts with different modes of presentation of the subject. This would seem to indicate that type concepts, like their descriptive and de re counterparts, involve modes of presentation of some sort.

Most of the time, however, the mode of presentation of a type concept will be irrelevant to successful communication. Generally, identification of the type is all that is needed. Take (6.10) for example.

(6.10) I’ve got an infestation of potato bugs in my garden.

Now, generally, I could think of *stenopelmatus* under a child-of-the-earth mode of presentation or under a Jerusalem-cricket mode of presentation and interpret (6.10) perfectly well either way. What is important for understanding (6.10), in most contexts at least, is identifying the type.

So, if what we have said so far is right, hearers interpret generically used descriptions by selecting a type concept rather than a descriptive or de re one. There are some cases of
generically used descriptions where this might not seem to work, however. Consider, for example, the description ‘the lobo’ as it occurs in (6.11).

(6.11) The lobo hunts at night.

An utterance of (6.11) would seem to be telling us something about the type, and not just tokens of that type. Yet it seems strange to say that ‘the lobo’ as it occurs in an utterance of (6.11) is referring to the type, as types are not the sorts of things that hunt; tokens are. Given this fact, it would seem that if we assume that ‘the lobo’ as it occurs here is referring to the type *canis lupus baileyi*, then (6.11) should be false. Yet, intuitively, it is true.

We’ll return shortly to the problem just raised. First, however, we must address a different issue with sentences like (6.11). In addition to predicking something of a type which would seem to be satisfiable only by tokens, (6.11) and sentences like it also differ in an important way from sentences like (6.3) and (6.7).

(6.3) The lobo is endangered.
(6.7) The Italian city-state allowed for the rise of a merchant class.

(6.3) reports a particular fact. (6.7) describes a specific (if temporally extended) event. But (6.11) describes a regularity which summarizes specific events. Something similar can be seen in (6.12) which reports a regularity which summarizes particular facts.\(^{148}\)

(6.12) The potato contains vitamin C.

In the work in linguistics on generics, sentences like (6.11) and (6.12) are called *characterizing sentences*. They exhibit a kind of genericity which differs from what we might call the referential genericity of kind-referring uses definite descriptions and which applies to the whole sentence,

\(^{148}\) The example is taken from Krifka et al. (1995) p. 78
not just the subject NP or the predicate.\textsuperscript{149} Because characterizing sentences like (6.11) and (6.12) exhibit genericity at the sentential level, their syntactic structure differs significantly from particular sentences like (6.3) and (6.7). We saw the basic syntactic structure of sentences like (6.3) and (6.7) in our description of the LF of definite descriptions overall. The syntax of a characterizing sentence, in contrast, involves the GEN operator. GEN functions as a dyadic adverbial quantifier, relating one set of conditions which contains a free variable to another set of conditions.\textsuperscript{150} The two sets of conditions are known as the \textit{restrictor} and the \textit{matrix}. The syntax of a characterizing sentence thus involves the tripartite structure shown in (G).

\[(G) \quad \text{GEN[Restrictor; Matrix]}\]

In the case of a characterizing sentence of the form \textit{The F is G} where ‘the F’ is kind-referring the syntax will be like that given in (KC).

\[(KC) \quad \text{GEN}(x;)[R(x,F); x \text{ is } G]\]

R is a realization relation relating types to their tokens.\textsuperscript{151} Note, however, that the presence of R does not eliminate the need for a type concept in making sense of the interpretation of an utterance of the form \textit{The F is G} which is characterizing and involves a kind-referring NP in the subject slot. A concept of the type $F$ will be necessary for filling in the restrictor in such cases.

Thus, we should still see the instruction to select an individual concept as encoded by the definite article and part of the LF of an utterance of a characterizing sentence containing a kind-referring definite description. The LF, then, would look something like (KC\textsuperscript{*}).

\[(KC^*) \quad \text{GEN}(x;)[R(x,<\text{individual concept}>F); x \text{ is } G]\]

\textsuperscript{149} Krifka et. al. (1995), pp. 8 and 19
\textsuperscript{150} Ibid., pp. 25-26; Leslie (2012), pp 357-358
\textsuperscript{151} Ibid., p. 66
As (KC) should make clear, utterances of sentences like (6.11) should no longer pose any problem. The syntax for (6.11) is given in (6.13).

(6.11) The lobo hunts at night.
(6.13) GEN(x;)[R(x, lobo); x hunts at night]

Indeed, it seems likely that the apparent problem with utterances of sentences like (6.11) derives simply from the fact that these are characterizing sentences. Remember that the problem raised with utterances of (6.14) was that while such an utterance would seem to be about the type lobo, the type lobo does not hunt. Only individual lobos do. But this would seem to simply be another way of saying that sentences like (6.11) report regularities summarizing particular facts or events. In the case of (6.11), this regularity of token lobos hunting at night is attributed to the type lobo via the GEN operator.

If the foregoing is correct, then it looks like our relevance theoretic account can make sense of the referential, attributive, and generic uses of definite descriptions in terms of the different types of individual concepts. There are several uses of definite descriptions that we have not yet given a treatment of, however. We will attempt to make sense of these uses of descriptions in the following section.

6.3 The Specific, Functional and Predicative Uses

In the previous two sections, we showed how a relevance theoretic account of definite descriptions can give a unified treatment of the referential, attributive, and generic kind-referring uses of definite descriptions. On this account, remember, the linguistic meaning encoded by utterances containing definite descriptions fails to determine a proposition. This is because the
definite article encodes not conceptual, but procedural information. In particular, it encodes a constraint on the interpretation process limiting interpretation of the NP to an individual concept. During pragmatic processing, an individual concept is selected on the basis of considerations of relevance. The selection of a descriptive concept yields an attributive reading of the description, the selection of a *de re* concept yields a referential reading of the description, and the selection of a type concept yields a generic kind-referring reading of the description. These are not the only possible readings of a definite description, however. There are also specific, functional, and predicative uses which we have yet to account for. This section is dedicated to giving a treatment of these uses.

6.3.1 The Specific Use

Suppose that Dee and Meier are having a conversation and Dee utters (6.14) with a particular individual – say, Marcus – in mind.

(6.14) The man I live next door to has some strange habits.

Now, let us suppose that Dee is perfectly aware of the fact that Meier has never met Marcus. As a result, he has no intention that Meier interpret him as referring specifically to Marcus. Nonetheless, we may imagine that the thought that Dee had in mind when uttering (6.14) was one involving a *de re* concept of Marcus. Is the description in (6.14) being used referentially or attributively in this case? Ludlow and Neale treat cases like this as distinct from referential and attributively uses, indicating that they take such uses to form a category of their own, which they
call the “specific use”. And Powell seems to agree, claiming that “Such a use seems neither truly referential nor truly attributive”.

For accounts of descriptions which do not rely on relevance theory, the specific use would seem to pose a problem, as it is not clear just what the semantics of such a use would be. The speaker has a singular proposition in mind, and hence the description would seem to be a referential one. But the audience can only interpret the description attributively. Of course, one might get out of this predicament by noting that that whatever thought the speaker actually has in mind, he intends the audience to interpret the description attributively. Since this is in fact how the audience will interpret the description, calling this an attributively used description would not seem out of line. But this line of reasoning only works if we assume that it is the speaker’s intentions that determine which proposition is expressed by an utterance (or which implicatures or presuppositions are associated with the utterance). This assumption has not been explicitly endorsed by any of the views of descriptions we have looked at so far, nor is it one that would be easily adopted in every case. Recanati’s views, for example, seem to place heavy emphasis on the hearer’s interpretation, and the Russelian would seem to be tied to holding that whatever Gricean or Neo-Gricean mechanism he endorses is what determines which implicatures are associated with an utterance.

Ultimately, however, it is not clear that tying the proposition expressed or implicatures generated to the speaker’s intentions will solve the problem, for the only reason this move serves to deal with an utterance of (6.14) in the context we have described is because the speaker’s

---

152 Ludlow & Neale (1991). Though they are concerned with indefinites there, the issue arises for definite descriptions as well.
intentions and the audience’s interpretation line up in that case. But what the problem is
supposed to be in the case of a specific use, it would seem, is that things on the speaker’s side do
not line up with things on the hearer’s side. So, let us change the example a bit. Suppose that Dee
utters (6.14) with Marcus in mind and that he assumes that Meier is not familiar with Marcus.
Thus, Dee does not intend for Marcus to figure into Meier’s interpretation of his utterance. But
now imagine that Meier saw Marcus outside of Marcus’ house earlier in the day mowing his
lawn in a top hat. So, when Dee utters (6.14), Meier interprets the description as being used
referentially to pick out Marcus. \(^{154}\) Now, is the description in this case attributive or referential?
The speaker intended the description to be understood attributively. But since the audience
interpreted the description referentially, and thereby understood the utterance as expressing the
very singular proposition that the speaker had in mind, it seems odd to claim that a general
proposition was in fact expressed by the utterance of (6.14). At the same time, however, since the
speaker intended his utterance of (6.14) to express a general proposition, there is something
intuitively inappropriate about claiming that it in fact expresses a singular proposition. Because
of the discrepancy between the speaker’s intentions and the hearer’s interpretation, the
description in this utterance of (6.14) does not appear to be truly referential or truly attributive. It
is unclear how the accounts of descriptions we have examined which are not based on relevance
theory would be able to handle cases like this.

The relevance theoretic account of descriptions can easily handle the specific use of
definite descriptions, however. After all, the problem posed by the specific use of definite

\(^{154}\) Some might argue that this interpretation would be out of place, since Meier presumably knows that Dee thinks
he is unacquainted with Marcus. This can be circumvented by imagining that Meier mistakenly believes that Dee is
aware of Meier’s familiarity with Marcus.
descriptions only arises if we assume, as the accounts not based on relevance theory all do, that there is a single correct way of characterizing the content associated with an utterance. To put it another way, specific uses of definite descriptions are only problematic to those who hold that there is a particular what is said and a particular what is meant for every utterance. The conflict between the speaker’s intentions and the hearer’s interpretation found in a specifically used description prevents a clean determination as to what content is associated with the utterance in question. We might ultimately claim that the speaker’s intent is the relevant determining factor, or we might claim that it is the hearer’s interpretation, but either way we will be left with some intuitive unease with respect to the correctness of the content we claim to be the content associated with the utterance.

For the relevance theorist, however, there is no specific what is said or what is meant by an utterance. The views we have looked at which are not based on relevance theory all assume that there is some content associated with an utterance which is independent of the cognitive systems of the parties to the conversation in question. Since relevance theory gives a cognitive account of communication, the relevance theorist will view the content(s) associated with a particular utterance as tied to the cognitive systems of the conversational participants. Thus, on the relevance theoretic view, there may be multiple contents associated with a particular utterance. There is what the speaker intended to convey with his utterance, and there is what the audience interpreted the utterance as conveying. When these two are more or less lined up, communication is successful, but they can differ from one another. And this is precisely what has occurred in the case of a specifically used description. The speaker intended his utterance to convey certain content, while the hearer interpreted the utterance as conveying something
slightly different. As for the question of whether the description is referential or attributive, the relevance theorist need not pick one or the other, since he is not committed to the notion that there is a single specific content associated with the utterance which is independent of the speaker’s intended meaning and the audience’s interpretation.

6.3.2 The Functional Use

As we just saw, the specific use of definite descriptions posed virtually no problem at all for the relevance theorist. The functional use, exemplified in (6.15)-(6.18), is not likely to be as easily handled, however.

(6.15) The President of the U.S. changes every four years.
(6.16) The instructor of Logic 101 differs from semester to semester.
(6.17) The temperature varies widely in the Spring here.
(6.18) The price of gas has increased considerably in the last few years.

Clearly, the descriptions in utterances of (6.15)-(6.18) would be neither referentially nor attributively used. (6.15) does not tell us that Barack Obama (the current U.S. president) changes every four years any more than (6.17) tells us that 72 degrees Fahrenheit (the temperature outside as I write this) varies widely in the Spring. The first thought when faced with examples like (6.15)-(6.18) might be to attempt to handle them in much the same way as the generic descriptions we looked at earlier. Thus, ‘the president of the U.S.’ in (6.15) refers to the type *president of the U.S.* and similarly for the other examples. But this won’t work either. A typical utterance of (6.15) does not mean that the type *president of the U.S.* changes every four years. What it means is that which individual tokens the type *president of the U.S.* changes every four years.
So how might a relevance theoretic account of definite descriptions make sense of descriptions like those in (6.15)-(6.18)? On Powell’s relevance theoretic account, remember, functionally used descriptions are given a metarepresentational treatment. According to Powell, an utterance of, say, (6.15) would express the proposition that there is a denotation of the descriptive concept associated with ‘the President of the United States’ and it changes every four years. As we saw in chapter 4, however, rather than following from Powell’s account as a whole this treatment of functionally used descriptions seems like something of an awkward add on. Powell gives an analysis of functionally used descriptions which is largely distinct from his analysis of the attributive and referential uses, and he fails to give us any explanation of how the metarepresentational content he associates with functionally used descriptions might arise out the selection of an individual concept. As a result, it is unclear whether his account provides us with a single unified analysis of definite descriptions or two different analyses.

There is something right in Powell’s treatment of functionally used descriptions, however, for the functional use does appear to be metarepresentational in nature. We earlier characterized this in terms of types and tokens, claiming that (6.15) understood functionally expressed the proposition that which individual tokens the type President of the U.S. changes every four years. Powell’s analysis not only captures the same basic sense, but does so in a way which still treats functionally used definite descriptions as referring expressions. Our treatment of the functional use in terms of types and tokens, on the other hand, would not appear to allow this.

There may be a way to supplement Powell’s treatment of the functional use so that it avoids the problems we saw earlier, however. Sperber & Wilson, as well as Wilson & Carston,
have argued that lexical interpretation frequently involves the construction of *ad hoc* concepts – occasion specific senses which result from the narrowing and/or broadening of linguistically encoded concepts.  

So, consider the sense of ‘drinking’ in an utterance of (6.19) in response to a question about the safety of the speaker’s driving home from a party the evening of the utterance.  

(6.19) I’m not drinking tonight.

The speaker of (6.19) could be understood as saying either that he will not drink any alcohol or that he will not drink significant amounts of alcohol. To those like Wilson & Carston who see this as an instance of lexical adjustment, both of these interpretations involve a narrowed, occasion specific sense of ‘drinking’. According to Wilson & Carston, we can understand the interpretation of (6.19) with its occasion specific sense of ‘drinking’ in terms of *ad hoc* concepts constructed via adjustment of the lexical sense of the term – in this case, by restricting the denotation. On their view, instead of involving the concept DRINKING, that is, the concept associated with the encoded meaning of ‘drinking’, the interpretation selected by the audience of (6.19) will instead be either that the speaker is not DRINKING*\(^1\) – not drinking alcohol – tonight or that the speaker is not DRINKING*\(^2\) – not drinking a significant amount of alcohol – tonight.  

This process of adjusting the encoded sense is sometimes called “conceptual narrowing and broadening” or “conceptual modulation”.

---


156The example is taken from Wilson & Carston (2007)

157 Following Wilson & Carston (2007), I use ‘*’ to indicate *ad hoc* concepts.
Following this view of lexical interpretation, Papafragou has argued that *ad hoc* concepts are at work in the interpretation of utterances involving metonymy, which she treats as a species of interpretive use. A representation with a propositional form is used interpretively, according to relevance theory, when it represents another representation with a propositional form by virtue of the resemblance between the propositional forms. Individual concepts and conceptual constructs can be used interpretively as well. A characteristic example of this can be seen in the introduction of names, as in (6.20) and (6.21).

(6.20) This is a ‘trombone’.
(6.21) Hi, ya, Beautiful!

According to Papafragou, the phrase “a ‘trombone’” in an utterance of (6.20) is “… used interpretively to mean something like ‘the thing that can appropriately be called trombone’; it introduces the representation of a representation, a ‘self-referring’ linguistic expression.” On Papafragou’s view, this ‘self-referring’ linguistic expression is interpreted via the construction of an *ad hoc* concept of the referent whose content is limited to that expressed by the self-referring expression (in the case of (6.20), the value CAN APPROPRIATELY BE CALLED ‘TROMBONE’) In cases where the interpretively used expression is also metonymical and referring, as in (6.21),

---

158 Papafragou (1995), (1996). Both Sperber & Wilson and Papafragou originally labeled this phenomenon the ‘echoic use’. Papafragou abandons this label in favor of ‘interpretive use’, as some instances of the phenomenon do not appear to involve echoing any particular thought, but instead “echo” ideas which have not yet been expressed or, perhaps, entertained.

159 Papafragou (1995), (1996); Sperber & Wilson (1986); Wilson & Sperber (1988b)

160 The examples are taken from Papafragou (1996), p. 180 and 181. (6.23) was originally Kaplan’s. See Kaplan (1989a) pp. 560-561

161 Papafragou (1996), p. 180. Her use of the definite article in her characterization of the meaning of ‘a ‘trombone’’ in this instance seems rather odd given that the original phrase is indefinite. It would perhaps be better to characterize the meaning of ‘a ‘trombone’’ here as “a thing that can be appropriately called ‘trombone’”
Papafragou holds that the constructed *ad hoc* concept helps the interpreter gain access to the relevant individual concept of the referent.\(^\text{162}\)

The reader may have noticed a similarity between interpretively used expressions and Powell’s suggested treatment of functionally used descriptions – both are metarepresentational in character. Just as Papafragou treats the interpretation of interpretively used language in terms of *ad hoc* concepts, then, we might, relying on Powell’s analysis, give a similar treatment of the interpretation of functionally used descriptions. The suggestion, then, is that in the interpretation of utterances containing functionally used definite descriptions an *ad hoc* concept with the metarepresentational content suggested by Powell (or something very similar) is constructed and selected by the PPM during pragmatic processing. In the case of an utterance of (6.15), for example, the functionally used description ‘the President of the U.S.’ would be interpreted via an *ad hoc* type concept with the content "DENOTATION OF ‘THE PRESIDENT OF THE UNITED STATES’". Now, here we diverge a bit from Powell’s characterization of the proposition expressed by an utterance containing a functionally used definite description. For according to Powell, (6.15) expresses the proposition that there is a denotation of ‘the President of the U.S.’ and it changes every four years. Because we have characterized the interpretation of the functionally used description in terms of an *ad hoc* type concept with the metarepresentational content suggested by Powell, there is no need to cash out the full proposition expressed by an utterance of (6.15) existentially. Powell does this, presumably, to prevent the metarepresentational content from taking on a particular reading – that is, to block a reading on which “the denotation of ‘President of the U.S.’ picks out the current denotation of the phrase (i.e. the current President of the U.S.).

\(^{162}\) Papafragou (1995), p. 156
In taking this metarepresentational content to belong to a type concept, this reading is already blocked.

One might wonder what benefit there is to avoiding an existential treatment of the proposition expressed by utterances containing functionally used descriptions. There are two reasons we might wish to do so, however. First, Powell’s existential interpretation of such utterances does not fit well with the overall relevance theoretic view of the definite article as encoding an instruction to select an individual concept. If utterances containing functionally used descriptions express existentially quantified propositions as Powell suggests, what role, if any, is this instruction playing? Powell must either provide us with an answer to this question – a task which does not seem promising – or abandon the hope of giving a unified account of the definite article.

Additionally, a proper treatment of those utterances containing functionally used descriptions, in the vast majority of cases, cannot be existential, as they appear to exhibit characterizing genericity.\(^{163}\) Thus, the syntax for an utterance of (6.15) would be something like that in (6.22)\(^{164}\).

\[(6.22) \text{GEN}(x; )[x \text{ is denotation of ‘the President of the U.S’}; x \text{ changes every 4 years}]\]

It is unclear how Powell’s existential reading could be made to fit here, as the relevant variable is already bound by the GEN operator. It would seem that a characterization of the interpretation of utterances containing functionally used descriptions in terms of *ad hoc* type concepts with the

---

\(^{163}\) A possible exception here would be (6.18), which appears to have both a characterizing and a particular reading. The characterizing reading reports a regularity of price-increasing events. The particular reading treats the increase in price as a single (though, perhaps, extended) event. An existential interpretation of the proposition expressed can capture the particular reading, but is inconsistent with the standard treatment of the characterizing reading. The view I have suggested here, however, would seem to be capable of capturing both readings.

\(^{164}\) On the account we are suggesting, the restrictor would actually be something like “\(x = \text{denotation of ‘the President of the U.S’}\)” where the right hand side of the identity picks out the type.
metarepresentational content suggested by Powell, then, is more plausible than the existential interpretation he puts forward.

Given our account so far, however, it is unclear just why such *ad hoc* concepts might be constructed during the interpretation process. According to Wilson & Carston, the process of *ad hoc* concept construction is triggered by the search for relevance. During the interpretation of a description, then, *ad hoc* concept construction would be triggered when the relevant descriptive, *de re*, and type concepts were either unavailable or resulted in an interpretation of the utterance which was not relevant. So, for example, in the case of (6.17), we do not normally have ready to hand a *de re* concept of the temperature. In addition, ‘varies widely’ cannot be appropriately predicated of the number, whichever it is, that satisfies the content of the descriptive concept of the temperature, and there is no well-established kind which would be picked out by ‘the temperature around here’. The idea, then, is that in such a case, *ad hoc* concept construction would be triggered. Given the fact that functionally used descriptions are not uncommon in discourse, it seems likely that a metarepresentational move would be among the first made, along with, perhaps, the sort of metonymical constructions envisioned by Papafragou. If relevance theory is correct, then an interpretation of an utterance containing a definite description which involves an *ad hoc* metarepresentational concept of the sort described will be selected when the ratio of the cost of processing the interpretation to the cognitive effects derived is better than the ratios of processing cost to cognitive effects for other available interpretations. There are a few reasons for thinking this would be the case in contexts in which the functional reading of the description is the intuitively correct reading. First, as we noted earlier, in many such cases, the interpretations involving the descriptive, *de re* and type concepts will be unavailable either
because they are not possessed or easily acquired by the interpreter or because they are semantically unacceptable. Second, in those contexts in which the functional reading of the description is intuitively preferred, the interpretation involving the *ad hoc* metarepresentational concept will likely generate more cognitive effects, and the processing cost involved in generating those cognitive effects will be lower. So, for example, (6.17) could be given a referential or attributive reading. But the contexts in which those readings are preferred are generally ones in which the accessible contextual assumptions pertain to the individual who is currently President of the United States. In these contexts, cognitive effects will be more readily generated by an interpretation on which the particular individual who is President is being referred to or denoted. On the other hand, contexts in which the functional reading is preferred are those in which assumptions about the office of the Presidency or the role of being President are more accessible. In these contexts, an interpretation in which it is the role of being president being picked out will more readily generate cognitive effects.

The foregoing is meant merely to make plausible the idea that an interpretation of an utterance containing a definite description which involves an *ad hoc* metarepresentational concept will be generated and selected as optimally relevant in those contexts in which the functional reading of the description is the intuitively preferred reading. Unfortunately, we will not be able to go beyond these rather vague gestures at the idea that such a reading would be optimally relevant. As we shall see in the §6.4, this is a worry that applies to our relevance theoretic treatment of the definite article as a whole, and not just the account we offer of the functional use of descriptions. Before we get to that, however, there are a few more uses of definite descriptions that we should look at.
6.3.3 The Predicative Use

As we saw in chapter 2, Ramachandran has identified a use of definite descriptions which differs from the referential and attributive uses. This use is exemplified by the description in a typical utterance of (6.4).

(6.4) Shakespeare is the greatest English playwright.

Even if we are assuming that it accurately describes him, the description ‘the greatest English playwright’ as it occurs in such an utterance does not denote Shakespeare. It would seem, then, that the description is neither referentially nor attributively used. Nor is it a kind-referring reading, as it does not refer to the kind greatest English playwright (if there is such a kind) as an individual. Rather, an utterance of (6.4) in which the copula is interpreted as the ‘is’ of predication would seem to communicate a proposition which predicates of Shakespeare the property of being the greatest English playwright. Ramachandran thus labels this use of descriptions the predicative use.

As we have seen, Ramachandran contends that the predicative use of descriptions is a non-referring use. While it is unclear what the content of a non-referring use of a description would look like, it seems unlikely that the relevance theoretic account of descriptions on offer here would be able to account for non-referring descriptions. According to our account, remember, definite descriptions are interpreted via individual concepts. And individual concepts are concepts which are taken by their possessors to represent individuals. They are essentially the cognitive counterparts of definite referring expressions. Whatever the content of a non-referring use of a description might end up being, then, it seems unlikely that we would be able to account for it in terms of individual concepts.
I do not believe we should be too hasty to accept the notion that predicatively used descriptions are non-referring, however. The fact that a predicative use of ‘the greatest English playwright’ does not refer to *Shakespeare* does not mean that the description does not refer at all. Since an utterance of (6.4) in which the description is used predicatively predicates the property of being the greatest English playwright of Shakespeare, it seems plausible to suppose that the description refers to the property. If this is the case, then it may be possible to treat the predicative use of descriptions as a special case of the attributive use. Of course, the property of being the greatest English playwright does not satisfy the conceptual content $\text{GREATEST \text{- ENGLISH - PLAYWRIGHT}$. Since the referent of a descriptive concept must satisfy the content of that concept, this means that the predicative use of the description in (6.4) cannot be interpreted via a descriptive concept whose descriptive content is comprised of just the conceptual content linguistically encoded by the description. However, if we approach the predictive use as we did the functional use – in terms of *ad hoc* concept construction – it should be possible to overcome this difficulty. The referent of a predicative use of ‘the greatest English playwright’ would uniquely satisfy a conceptually modulated version of $\text{GREATEST-ENGLISH-PLAYWRIGHT}$ which narrowed the extension of the concept to properties. We can account for a predicative use of the description, then, in terms of an *ad hoc* descriptive concept with this content. The predicative reading of a description, on this view, results from the selection of an *ad hoc* descriptive concept which picks out a property or state.

So, what reason might we have for thinking that such a concept would be selected in contexts in which a predicative reading of the description is the natural reading? There are two things to note here. The first thing is that a speaker who intends to identify $A$ with *the F* will
typically use a sentence of the form *The F is A*. A speaker who wished to communicate that Shakespeare and the greatest English playwright are the same person would typically use the sentence in (6.23) rather than that in (6.4).

(6.23) The greatest English playwright is Shakespeare.

The second thing to note, and perhaps the reason that a speaker who wished to make an identity statement would typically use an utterance of the form *The F is A*, is that in the absence of any contextual information, the predictive reading appears to be the default reading of descriptions in utterances of the form *A is the F*. In the case of (6.4) above, for example, we are given no contextual information, and our intuitive reading of the description is the predicative reading. This may perhaps be the result of a general tendency in the generation of interpretive hypotheses by hearers. As we saw in the previous chapter, relevance theory holds that interpretation is an online process. According to the theory, a hearer (or interpreter) engaged in the interpretation of an utterance will generate interpretive hypotheses as the process is ongoing and then use those interpretative hypotheses as guides in constructing an overall interpretation. If, upon encountering a referring expression in the subject slot, hearers tend to generate the interpretive hypothesis that something will be predicated of the subject, this would explain why the default reading of a description in an utterance of the form *A is the F* is the predicative reading. The hypothesis that the utterance will predicate something of the subject would lead to a provisional interpretation on which ‘is’ is treated as the ‘is’ of predication and the description is interpreted as picking out something that can be predicated of a subject. In other words, the description would be interpreted via an individual concept whose referent was a property or state. In cases in which the conceptual content encoded by the description alone is not satisfied by a property or
state, an *ad hoc* concept would have to be selected. And if we are correct in thinking that hearers typically form the interpretive hypothesis in question, then this sort of concept will be selected in any case in which the context does not provide the hearer with reason for thinking that the interpretive hypothesis is incorrect. If the context *does* provide a reason for thinking that the interpretive hypothesis is incorrect – say, because an interpretation which predicated something of the subject failed to satisfy specific expectations of relevance – then the description will be given a non-predicative reading. If the foregoing is correct, then relevance theory gives us reason to think that an individual concept referring to a property or state will be the default interpretation of the description in an utterance of the form *A is the F*, and the description will only be given a different type of interpretation when the context gives the hearer reason to believe that the utterance does not communicate a proposition which predicates something of the subject.

In this section, we have looked at three uses of descriptions which do not fall within the standard referential/attributive distinction, and we have attempted to show how our relevance theoretic account of descriptions might handle them. There are a few ways in which it is possible that descriptions are used which we have not yet looked at, and these possible uses of descriptions may indicate the existence of individual concepts other than those we have already examined. This is the topic of the next section.

### 6.4 Other Individual Concepts

So far we have identified three different types of individual concept – descriptive concepts, *de re* concepts and type concepts. Since these three types of individual concept are
central to our account of the interpretation of the different uses of definite descriptions, they will be our main focus. But they may not be the only types of individual concepts. Since we take proper names like ‘George Washington’ or ‘Santa Claus’ to pick out single individuals, it seems plausible to suppose that these expressions are interpreted via individual concepts. Indeed, it is frequently assumed that names are interpreted via \textit{de re} concepts. While this would seem to work in cases in which we can identify a causal connection between the hearer’s \textit{de re} concept and the referent of the name, it is problematic in the case of names of fictional characters. Since these names lack actual referents, we cannot be causally connected to those referents. As a result, they cannot be interpreted via \textit{de re} concepts. And if an utterance of a sentence like (6.23) can be used to mean \textit{Santa Claus likes gingerbread cookies}, the description in that utterance cannot be interpreted via a \textit{de re} concept either.

(6.23) The jolly old man who brings gifts on Christmas likes gingerbread cookies.

It is not entirely clear to me that a description like the one in (6.23) can in fact be used to convey an object dependent proposition about Santa Claus. My intuitions here are unclear. Perhaps the description in (6.23) and others like it are attributively used, in which case we would treat them as being interpreted via descriptive concepts.

Names of fictional characters might also be given a similar treatment. There are well known problems with descriptive theories of names, however. While it is not necessarily the case that all of these problems will also trouble the view that names with fictional referents are interpreted via descriptive concepts, it seems likely that at least some of the same problems would arise. For example, according to the account of descriptive concepts we suggested earlier, a descriptive concept has a rule in stored in its logical entry which indicates that the referent of
the concept, if it has one, satisfies a certain descriptive content. So, what would this descriptive
content be in the case of ‘Santa Claus’? Any choice of some particular descriptive content over
another would seem to be completely arbitrary.

If we cannot account for names which pick out non-existents in terms of descriptive
content, then the relevance theorist who wishes to treat referring expressions in terms of
individual concepts will have to account for them in terms of something other than descriptive
concepts. Since neither de re concepts nor type concepts are suitable for this task, there would
have to be some type of individual concept in addition to those we have already identified. This
new type of individual concept could be completely different from those we have already
examined, or it might turn out that it is merely a special species of one of the types of individual
concepts we have already looked at. For example, here is one suggestion: De re concepts
actually come in two varieties. Let us call these two varieties connected de re concepts and
disconnected de re concepts. Connected de re concepts are the sort of de re concept we have
been concerned with so far – concepts of individuals whose referents bear a causal relationship
to their contents. Disconnected de re concepts are structurally just like connected de re concepts,
except that their contents are not causally related to their referents. These proposed types of de re
concepts share in common that they are individual concepts with empty logical entries (or, if
individual concepts contain a rule indicating that the referent is an individual, this is the only rule
in the logical entry). The idea would be that names and referentially used descriptions are
interpreted via de re concepts. Most are interpreted via connected de re concepts, but those
which refer to non-existents are interpreted via disconnected de re concepts.
This is only a suggestion, of course. The problem of reference to non-existents and the question of how to best account for proper names are both difficult issues which we do not have space to attempt to address in any significant way here. The point is that, depending on how we ultimately choose to answer the question of how reference to non-existents occurs, relevance theoretic treatments of referring expressions in terms of individual concepts may be required to postulate a kind of individual concept other than the descriptive, type or standard \textit{de \textit{re}} concepts that we discuss above.

A similar situation arises when we consider reference to numbers. Some expressions referring to numbers may be treated as interpreted via descriptive concepts. For example, an attributive use of the description ‘the smallest prime greater than 140’ is presumably interpreted via a descriptive concept. But it is possible that in a conversation in which it was mutually manifest to the participants 149 is the smallest prime greater than 140, the description might be used in conveying an object dependent proposition. Even if such a description cannot be used to convey an object dependent proposition, it would still nonetheless be necessary to make sense of how number expressions like ‘149’ are interpreted. Since we cannot be causally connected to numbers, however, the description cannot be interpreted as a standard \textit{de \textit{re}} concept. Unless numbers are kinds, then, and can be the referents of type concepts, the class of individual concepts will have to include some whose referents are numbers, and possibly other mathematical objects as well. Since the question of how we conceptualize numbers and other mathematical objects, like the question of how to account for reference to non-existents, is a large and complex one, we will not be able to address it here. But it is clear that in addition to
descriptive, de re and type concepts, the class of individual concepts likely includes number concepts as well.

In order to account for reference to numbers and non-existents, a relevance theoretic account which treats reference in terms of individual concepts may have to hold that there are distinct types of individual concepts in addition to descriptive, type, and standard de re concepts, but until we have a better picture of the best way to address these questions, it will not be possible to give a more detailed account of the nature of these additional types of individual concepts. The inability to provide a full account of descriptions which are used in utterances expressing object dependent propositions about numbers and non-existents is, without question, a weakness of the account of descriptions proposed here. It is, however, a weakness that is shared across the available accounts of definite descriptions. Since philosophical work on descriptions has been focused on the referential/attributional distinction, the major approaches to definite descriptions have largely neglected to consider these issues.

6.5 The Problem of Testability

In this chapter, I have attempted to construct a relevance theoretic account of definite descriptions which, while drawing on the work of other relevance theorists, avoids the problems faced by other relevance theoretic accounts. The picture that has emerged is one on which the LF of an utterance containing a definite description fails to determine a proposition on its own, at least in part because the definite article encodes procedural rather than conceptual content. In particular, the definite article encodes a procedural constraint which limits interpretation of the NP to an individual concept. In the case of a referentially used description, this individual

200
concept will be a \textit{de re} concept, which stands in a causal relationship to its referent. In the case of an attributively used description, it will be a descriptive concept, which bears a satisfactional relationship with its referent. And in the case of a generically used description, it will be a type concept, which picks out a kind.

The relevance theoretic account constructed here avoids many of the problems we saw with previous attempts to use relevance theory to account for definite descriptions. In contrast to Bezuidenhout’s view, our account characterizes the LF of an utterance containing a description in a way which does not privilege one reading of descriptions over another and which, when enriched, would result in an assumption with the same structure as that possessed by our intuitive interpretation of such an utterance. Moreover, in contrast with both Powell’s and Bezuidenhout’s theories, our account provides explanations of uses of descriptions other than just the referential and attributive. Unfortunately, while our account may have avoided some of the difficulties faced by previous relevance theoretic views, it has not avoided the most significant problem facing these accounts – testability.

The standard way to test accounts of communicative phenomena like the relevance theoretic account on offer here is to apply them in philosophical example cases. We construct an example utterance in a stipulated context and then determine whether the interpretation of the utterance predicted by the account is the same as the interpretation we find to be intuitively correct or natural. If it is, this is evidence in favor of the theory. If it is not, this is evidence against the theory. Now, just as the accounts offered by Bezuidenhout and Powell do, our account maintains that the audience of an utterance containing a definite description will assign to the description whichever available interpretation is optimally relevant. To show that our
account predicts the correct interpretation of an utterance containing a description, then, we would need to show that the interpretation which is intuitively correct in the case is the optimally relevant one – that is, the one with the most cognitive effects for the least amount of processing effort. Without a means of quantifying either cognitive effects or processing effort, however, it is unclear how we might perform such a task. The best we can provide at this point, it seems, would be the same sort of vague gestures in the direction of optimal relevance that we saw offered by Bezuidenhout and Powell.

This is a serious problem. An untestable theory is of little value to the theorist. And since this testability problem is likely one facing any attempt to use relevance theory to account for particular communicative phenomena, it is one that serves to weaken relevance theory overall. While there may be other ways to test relevance theory itself – by testing the individual principles of relevance, for example – an inability to test applications of the theory to particular phenomenon lessens the value of relevance theory as an explanatory engine in general.

Much of the remainder of this work is dedicated to attempting to solve the testability problem for the relevance theoretic account of definite descriptions offered here. The goal of the next chapter is to sketch a strategy for showing that a particular interpretation of a definite description is optimally relevant in a given case. While the strategy in question is specifically offered for testing our account of definite descriptions, it may be possible to modify it for use in testing other relevance theoretic accounts of particular communicative phenomena. In chapter 8 we will attempt to use the strategy constructed in chapter 7 to put the relevance theoretic account of descriptions to the test.
Chapter 7
Comparative Assessment of Relevance

According to relevance theory, a hearer’s natural interpretation of an utterance is the optimally relevant interpretation of that utterance. In order for a relevance theorist to show that his or her theory predicts that a particular interpretation of an utterance is the natural interpretation, then, he or she must demonstrate that the interpretation in question is the optimally relevant one – that it is more relevant than the other available interpretations. The relevance of an interpretation, remember, is a function of its potential cognitive effects and the cognitive effort required to process it. Unfortunately, we do not have a way to precisely quantify the cognitive effects and cognitive effort associated with an interpretation. In order to assess the comparative relevance of the available interpretations of an utterance, then, we will have to find a way to compare them that does not rely on a precise quantification of relevance. In this chapter, we will outline a strategy for doing just that. In the next chapter, we will attempt to use this strategy to compare the relevance of the available interpretations of several utterances involving different uses of definite descriptions.

7.1 A Basic Strategy

To some, the possibility of uncovering a method of comparing the relevance of different interpretive hypotheses which does not rely on a quantification of relevance may seem remote. But it is important to keep in mind that what the relevance theorist is attempting to supply is an account of the cognitive processing of utterances. That cognitive processing itself is not likely to
operate on the basis of a precise quantification of relevance. After all, the vast majority of our
cognitive processing of stimuli operates on the basis of intuitive assessments of values, and not
on quantifications of those values. This is largely because we are incapable, in many cases, of
obtaining a precise quantification of the values we are working with. For example, imagine a
scenario in which you are playing catch with someone, and your partner throws the ball wide.
Now, you have two potential courses of action – you can move to attempt to catch the ball, or
you can let it pass by. Setting aside questions of whether you feel too tired to try, etc, which
decision you make will be based on several factors, in particular, how wide the throw was, how
fast the ball is moving, how far the ball is from your current position, and how fast you can
move. The decision you come to is one that would appear to be made at a subconscious level, or,
at the very least, the calculations the decision is based on are subconscious. Most of us, at any
rate, are unaware of performing any calculations as to whether or not we could reach the ball in
time, and yet this consideration would be cited by most of us when explaining our choice in such
a situation. Those who would attempt to catch the ball would likely justify this by claiming that
they thought they would be able to catch it, and those who would forgo the attempt would likely
justify their decision by claiming that they couldn’t have reached the ball in time.

It would seem, then, that in such a situation, we subconsciously evaluate the various
pertinent factors (the speed of the ball, its distance and trajectory), and on the basis of these
determine whether or not we would be able to move quickly enough to catch the ball. But these
determinations are certainly not made on the basis of precise quantifications of the values
involved. We are not capable of obtaining the numbers in time (or really, at all, without the aid
of tools of measurement). Such scenarios of cognitive processing using various values in the
absence of precise quantifications of those values are not at all uncommon. And, hence, we should not be surprised if our cognitive assessments of relevance are done without precise quantifications of the number of cognitive effects or the processing cost involved.

What is problematic, then, is not the idea that the relevance of an interpretive hypothesis could be assessed in the absence of assignments of numerical values to its expected cognitive effects and its processing cost, but rather that without quantification of the values in question, the relevance theorist does not appear to have any solid basis for the claim that one potential interpretation of an utterance is more relevant than another. If we are to make it possible for our relevance theoretic account of definite descriptions to generate testable predictions, then, we must find a way to draw justifiable conclusions as to the optimally relevant interpretation of an utterance containing a definite description without reliance on quantifications of either the expected cognitive effects or the processing cost of the available interpretations of that utterance.

It is likely true that for some sorts of utterances, the theorist will be unable to give a comparative assessment of the relevance of the various available interpretations in the absence of a means of precisely quantifying processing costs and cognitive effects. There are certain situations, however, in which it might be possible. If one side of the relevance ratio is the same across the interpretive hypotheses under consideration, then we need only compare interpretations with respect to the remaining side of the ratio. For example, if we know that all of the available interpretations require the same amount of cognitive effort, we should be able to compare the cognitive effects of each interpretation and thereby determine which of the set is most relevant. Conversely, if we know that all of the available interpretations have the same number of cognitive effects, then we might be able to compare the cognitive effort required for
each interpretation and thereby discover which of the set is most relevant. In the case of example utterances containing definite descriptions, I believe we will be able to show that the available interpretations have the same cognitive effects and then compare their processing costs. In order to do this we will need to first have the means of accomplishing three tasks. First, we will need to have a means of delineating the set of available interpretations. Second, we will need to have some way of showing that the different available interpretations of the utterance are all the same along the cognitive effects side of the relevance ratio. Finally, we will need to provide some means of comparing the available interpretations along the processing cost side of the relevance ratio. We will spend most of the remainder of this chapter attempting to accomplish these three tasks. First, however, we must take a closer look at the nature of the example utterances whose natural interpretations we will be attempting to predict.

### 7.2 Contrived Cases and Interpreters

It would be extremely difficult, if not impossible, for the relevance-theorist to give a comparative assessment of the relevance of interpretations of an actual utterance made in real conversational conditions. Just delineating the set of available interpretations for such an utterance would require not only that the theorist be able to identify all the assumptions which were mutually manifest to the conversational participants, but also that he or she have knowledge of the extent of the audience’s mental lexicon. Thankfully, however, the philosophical method of testing semantic and pragmatic theories does not require any real world testing. Rather, such theories are most often tested against our intuitions about the natural interpretations of example
utterances in contrived philosophical examples. The conversational context is a matter of the theorist’s stipulation in these cases.

When we test our semantic or pragmatic theories using example utterances, we test the predictions made by the theory about the example in question against our natural interpretation of the utterance in the example. Take, for instance, Donnellan’s famous example of an attributively used description. A and B happen upon Smith’s corpse and A utters (7.1).

(7.1) Smith’s murderer is insane.

When we wish to test a theory of definite descriptions against this example, we first determine what the theory would say about the content of ‘Smith’s murderer’ in this case, and then test it against our natural interpretation of the utterance, which, in this case, is something like the murderer of Smith, whoever he is, is insane.

Now, it may be tempting to think that in generating our interpretation of (7.1) we are merely standing in B’s shoes, as it were, and interpreting the utterance just as he would. In a certain sense, this is true, since, as readers of the contrived case, we are attempting to do the same thing that a hearer in a conversation attempts to do – we are attempting to uncover the speaker’s meaning. But when we interpret an example utterance in one of these contrived cases, there are two key differences between our position and that of a hearer in an actual conversation. The first lies in the fact that the extent of the context in a contrived case is considerably limited when compared to that of an actual conversation. Actual conversational contexts are extremely rich things, filled with assumptions about the immediate physical environment and the conversational participants themselves. Actual conversational contexts also include whatever assumptions have been made mutually manifest by previous utterances in the conversation, in
addition to those assumptions which we attribute to adults generally. The context in a contrived example case, in contrast, is limited to those assumptions which we attribute to adults in general and whatever assumptions the stipulation of the case makes clear are part of the context.

The second key difference between the position of a reader attempting to interpret an utterance in a contrived example case and that of a hearer in an actual conversation is that, unlike the hearer, the readers are not participants in the conversation. According to relevance theory, speakers attempt to make their utterances optimally relevant to their hearer(s). Speakers’ utterances are crafted for their hearers. Thus, when it comes to contrived example cases, we are much like eavesdroppers, interpreting utterances which were not meant for us.

Since the relevance theoretic machinery that has been developed thus far attempts to explain how hearers interpret utterances, it would seem that relevance theory is not currently equipped to properly handle cases like those we are currently considering – that is, cases in which the individual interpreting the utterance is not a participant in the conversation. As we saw in chapter 5, relevance theory claims that hearers select the optimally relevant interpretation after testing available interpretations in order of their accessibility. The set of available interpretations in question is the set of interpretations that are available to the hearer. Similarly, the accessibility ordering at issue is the accessibility ordering for the hearer. And, according to relevance theory, speakers construct their utterances on the basis of their understanding of what is available to their hearers and what is most accessible to their hearers. But the set of interpretations available to an individual attempting to interpret an example utterance in a

---

165 I believe this applies to cases of eavesdropping just as much as it does to cases of interpretation of example utterances in contrived cases. The extra machinery I will sketch here should, I believe, allow relevance theory to handle both kinds of cases. Since our worry here is generating testable predictions in example cases, however, it is this kind of case that I will focus on in what follows.
contrived case will not necessarily be the same as the set of interpretations available to the hearer in that contrived case. Nor is it guaranteed that the accessibility ordering would be the same for both the hearer and the interpreter. If the interpreter of an example utterance is to successfully uncover what the speaker meant, he cannot be selecting an interpretation on the basis of what is optimally relevant to him, as we have defined optimally relevant. It would seem, then, that if relevance theory is to be able to account for cases of utterance interpretation undertaken by individuals who are not themselves participants in the conversation, some additions will have to be made to the existing relevance theoretic machinery.

Let us call an individual who is attempting to interpret an utterance which has occurred in a conversation in which he is not a participant an interpreter. The reader of a philosophical example case, then, is an interpreter. Interpreters are attempting to do the same thing that hearers are – uncover the speaker’s meaning. According to relevance theory, remember, the assumptions which make up the context used by a hearer to interpret an utterance are drawn from those assumptions which are mutually manifest to the conversational participants. Speakers who wish their utterances to be successfully interpreted will attempt to craft their utterances such that the intended interpretation relies solely on such assumptions. To successfully discern what a speaker intended to communicate via an utterance, then, the interpreter should rely on the assumptions which he or she has good reason to believe are mutually manifest to the participants in the conversation in which the utterance occurred. Let us call a context from which the interpreter draws the assumptions relied on in the interpretation of an utterance an interpretive context. We may say, then, that an assumption A is part of an interpreter I’s interpretive context for an
utterance produced in a conversation \( C \) if and only if \( I \) has good reason on balance to believe that \( A \) is mutually manifest to the conversational participants in \( C \).

Now, interpreters are attempting to uncover the speaker’s meaning. As we have seen, relevance theory holds that hearers select the optimally relevant interpretation of an utterance and a speaker who wishes to be understood attempts to make his or her intended interpretation optimally relevant to the hearer(s). The optimally relevant interpretation to the hearer, remember, is the most relevant of those available to the hearer. A speaker who wishes to be understood, then, will attempt to make his intended interpretation the most relevant of those which he or she judges to be available to the hearer. In light of this, we can define optimal relevance in an interpretive context as follows:

**Optimal Relevance in an Interpretive Context (ORIC)**

An interpretation of an utterance \( U \) in a conversation \( C \) is optimally relevant in an interpretive context if and only if it is the most relevant of the set of interpretations of \( U \) that the interpreter has good reason on balance to believe the speaker would judge to be available to the hearer(s) in \( C \).

An interpreter’s natural interpretation of an utterance will be the interpretation that is optimally relevant in the interpretive context. When we turn to attempts to use our theory of definite descriptions to predict the natural interpretations of example utterances in contrived cases, then, we will rely on the following predictive principle:

**Predictive Principle (PP)**

In any interpretive context \( C \) in which an utterance \( U \) has one or more available, relevant English interpretations, \( P \) is the most natural English interpretation of \( U \) if and only if \( P \) is the optimally relevant interpretation in \( C \).

When we turn to the task of using the relevance theoretic account of definite descriptions to predict the natural interpretations of standard examples of the different uses of definite
descriptions, then, what we will be doing in each case is attempting to demonstrate that it follows from the claims of the theory that a particular interpretation is the most relevant of those the interpreter has good reason to believe the speaker would judge to be available to the hearer.

### 7.3 Available Interpretations

If we are to show that a particular interpretation of an utterance containing a definite description is optimally relevant in a particular interpretive context, then, the first thing we will need to do is delineate the set of available interpretations. As we briefly noted above, delineating the set of available interpretations for a naturally produced utterance in a natural conversational context – with or without a definite description – could easily prove to be an extremely arduous task for the theorist to undertake. Not only might the number of interpretations turn out to be extremely high, but even if the number of available interpretations were small, it is unclear that the theorist would truly have access to the information necessary to determine the available interpretations. In order to do so, after all, he would need to know the cognitive capabilities of the audience (the concepts he or she possesses and the inferences he or she is capable of making) in addition to knowing which assumptions were mutually manifest to the conversational participants. This is the sort of knowledge which could be difficult to acquire without at least being acquainted with the audience.

These will not be problems for the theorist who is attempting to delineate the set of available interpretations for a philosophical example case, however, for, as we noted in the previous section, readers interpret such example utterances in an interpretive context which contains only those assumptions that the interpreter has reason to believe are mutually manifest.
to the conversational participants. In contrived philosophical examples, the case is stipulated by
the theorist, and hence he or she has access, as an interpreter, to the necessary information for
constructing the available interpretations. Similarly, the reader of a stipulated case also has
access to all of the necessary information. That said, the task of laying out the entire set of
available interpretations might still turn out to be problematic for a number of different types of
utterances. After all, there are utterances involving both syntactic ambiguity and multiple cases
of semantic ambiguity, polysemy, or vagueness, resulting in a high number of possible
permutations for just the basic enrichment of LF (the propositional form), and that is before we
even get to the multiple explicatures and implicatures that make up an interpretation. There is
some reason to think that this will not prove to be a problem for us here, however. To see this,
we will have to look at the different aspects of potential utterance interpretations.

7.3.1 Provisional Interpretation and Independence

According to relevance theory, remember, the interpretation of an utterance may contain
a propositional form, higher-level explicatures, and implicatures. For an interpretation to be
available, then, each of the assumptions contained in it must satisfy the description of either a
propositional form, a higher-level explicature, or an implicature. A propositional form is a
mental representation arising out of the enrichment of LF. Higher-level explicatures arise out of
the embedding of the propositional form in a propositional attitude description, and implicatures
are contextual assumptions or implications which the speaker manifestly intended to make
manifest. Moreover, as we saw in chapter 5, for an interpretation to be available any concepts
used as enrichments of components of the LF of an utterance must be available as enrichments of
those components. For any given utterance, then, it might seem as though there will be a
potentially daunting number of available interpretations. And to compare the relevance of all of
the available interpretations of an utterance containing a definite description, one might assume
that we would have to be able to assess the processing cost associated with the interpretation of
the predicate, implicatures, and explicatures as well as that associated with the interpretation of
the description itself. Thankfully, there is good reason to think this will not be a problem facing
us here.

The ultimate goal here, remember, is to use the relevance theoretic account of definite
descriptions offered in chapter 6 to predict the natural interpretations of standard examples of
attributive, referential, and kind-referring readings of definite descriptions. These standard
examples occur in utterances of the form *The F is G*. The description occurs at the very
beginning of the utterance. Now, as we saw in chapter 5, relevance theory views interpretation as
an on-line process, one which begins before the utterance is completed. Provisional
interpretations are assigned as the utterance is ongoing. As the interpretation of the utterance
continues, these provisional interpretations can be revised in light of other elements of the
utterance interpretation. If the other elements of the utterance interpretation provide no reason to
alter a provisionally assigned interpretation, then that interpretation will stand.

In the case of our standard examples of utterances containing definite descriptions, this
means that the interpreter will assign a provisional interpretation to ‘The F’ which can be revised
later in light of the interpretation of the other elements of the utterance. If we can show that the
other elements of the interpretation of the example utterance would not provide a reason to alter
whatever provisional interpretation is assigned to the ‘The F’, this would allow us to focus solely
on the interpretation of the description when comparing different available interpretations for
relevance. Perhaps the easiest way to do this would be to show that the elements of the interpretation of the utterance other than the interpretation of the description are irrelevant to the question of which of the available enrichments of the content encoded by the description is selected. If the predicate would be given the same interpretation no matter which of the available enrichments of the description were to be selected, then the interpretation of the predicate could not yield any reason to revise whatever provisional interpretation is assigned to the description. Similarly, if the available interpretations were to have the same implicatures and higher-level explicatures no matter which of the available enrichments of description was selected, then these elements of the interpretation of the utterance would not provide a basis for revising the provisional interpretation of the description during the interpretation process.

Let us say that when the elements of an utterance interpretation other than the description are irrelevant to the question of which interpretation of the description is selected, the interpretation of the description is *independent* of those other elements of the interpretation. If we can show in our test cases that the interpretation of the description is independent of the other elements of the utterance interpretation, then we can focus solely on the interpretation of the description when comparing available interpretations for relevance. In order to be able to show this, however, we must be able to determine which enrichments of the content encoded by the description are available.

### 7.3.2 Concept Availability

According to our account of definite descriptions, the definite article encodes a procedural constraint on the interpretation process which indicates that the NP should be interpreted as an individual concept. The individual concept selected is an enrichment of the LF
of the utterance in which the description occurs. But we don’t want to say that the content encoded by a description can potentially be enriched by just any old individual concept. ARISTOTLE is not a potential enrichment of the content encoded by the description “the first woman to land a triple axel in competition”. Pragmatic enrichments enrich encoded content; they must be tied to that content in some way. But conformity to encoded content is not enough. Japanese figure skater Midori Ito was the first woman to land a triple axel in competition, but we wouldn’t want to say that the de re concept MIDORI ITO is available as an enrichment of the description to a hearer who has never heard of Midori Ito, and hence has no concept of her. Hearers must have cognitive access to concepts in order to use them as enrichments. In order to account for this, we will need to outline conditions on when a concept is available as an enrichment of encoded content.

Let us begin with the issue of cognitive accessibility. For a hearer to use a concept as an enrichment of encoded content, he or she must be able to cognitively access that concept. More than that, it must be mutually manifest to the conversational participants that the hearer can cognitively access the concept. This is so because a speaker who wishes to be understood will attempt to limit the concepts in their intended interpretation to those for which it is manifest to the speaker that the hearer is capable of using them in mental processes. In light of this, the hearer who wishes to discover the speaker’s meaning will only rely on concepts for which it is mutually manifest to both speaker and hearer that they are available to the hearer. In order for a concept to be available as an enrichment of encoded content, then, it must be mutually manifest to the conversational participants that the concept can be cognitively accessed by the hearer.
There are two ways in which a concept might be cognitively accessible to an individual at a particular time. First, the concept could already be stored in the individual’s memory. When this is the case, we will say that the concept is *retrievable* for the individual. The definition of retrievability is given in (R).

**Retrievability (R):**
A concept $C$ is retrievable for an individual $I$ at $t$ if and only if $C$ is part of $I$’s mental lexicon at $t$.

Concepts not present in an individual’s mental lexicon may still be cognitively accessible to that individual, however. A concept will also be cognitively accessible to an individual at a time if it is possible for the individual to construct or otherwise acquire the concept at the time in question. When this is the case, we will say that the concept is *derivable* for the individual. We can acquire concepts in a number of different ways. The definition of derivability given in (D) attempts to capture all of the different methods of deriving a concept.

**Derivability (D):**
A concept $C$ is derivable for an individual $I$ at a time $t$ if and only if one of the following conditions holds:

1. $I$ could acquire $C$ by composing concepts present in $I$’s mental lexicon at $t$,
2. $I$ could acquire $C$ via narrowing, broadening or otherwise modifying a concept in $I$’s mental lexicon at $t$.
3. $I$ could acquire $C$ via some perception based interaction with $I$’s physical environment at $t$,
4. $I$ could acquire $C$ via some combination of (1)-(3)

When we acquire new concepts by having them defined or explained for us in terms of other concepts, it is presumably the method identified in clause (1) – the composition of concepts – which is in use. Since descriptive concepts are constructed as interpretations of NPs which encode multiple concepts, we will assume that such concepts are constructed via this process. Some type concepts may also be acquired in this way. We will see later on that the relevance
theoretic picture of concept composition differs in some important respects from more traditional understandings of this process of concept acquisition. In fact, I will suggest that the process of deriving a concept via composition of other concepts is extremely similar to that indicated in clause (2) – narrowing and/or broadening of concepts already possessed. We will discuss this method of concept acquisition in greater detail shortly. The method of concept derivation listed in clause (3) of (D) is the one which yields many of our de re concepts.166 In combination with the methods listed in (1) and (2), it may also be the source of some of our type concepts. Unfortunately, in order to say anything substantial about what sort of cognitive process or processes are involved in the method of concept acquisition indicated in (3), we would need to have a much better understanding of the relationship between perception and concepts than we currently possess. Work on the relevance theoretic understanding of lexical interpretation, however, should help to provide us with a basic picture of what is involved in the processes of concept acquisition mentioned in clauses (1) and (2).

The method of concept derivation listed in clause (2) of (D) – narrowing and/or broadening – is a key component of the relevance theoretic account of lexical interpretation. Along with other relevance theorists, Sperber & Wilson claim that lexical interpretation frequently occurs via the construction of ad hoc concepts.167 Arising out of work focused almost exclusively on loose or non-literal uses of language, the relevance theoretic picture of ad hoc concepts views them as occasion specific senses, and treats the process of ad hoc concept

166 On a highly restrictive view of the sort of causal connection required between the referent and the content of a de re concept, the process indicated in (3), or some combination of (3) and either (1) or (2), might be the only way that such concepts could be acquired. On more permissive views, de re concepts might be acquired via (1).
construction as one of conceptual narrowing and/or broadening of a linguistically encoded
concept. Since relevance theory treats concepts as addresses in memory at which different types
of information is stored, this narrowing or broadening of a concept – also sometimes referred to
as “conceptual modulation” – occurs via the elimination or addition of rules and/or assumptions
in the concept’s logical and encyclopedic entries. To make this idea a little more concrete,
consider the following example from Wilson & Carston (2006)¹⁶⁸:

(7.2) Peter: Will Sally look after the children if we get ill?
     Mary: Sally is an angel.

The LF of Mary’s utterance contains the encoded concept ANGEL. The logical entry of this
concept contains rules which would allow various inferences to be drawn from a literal
interpretation of Mary utterance - for example, the inference to the conclusion that Sally is a
supernatural being. The encyclopaedic entry of ANGEL, according to Wilson & Carston, attributes
to angels properties like those listed in (7.3), among others.

(7.3) EXCEPTIONALLY GOOD AND KIND
     WATCHES OVER HUMANS AND HELPS THEM WHEN NEEDED
     VIRTUOUS IN THOUGHT AND DEED
     MESSENGER OF GOD

Now, Mary’s utterance is a response to a question. Typical, adult conversational
participants generally expect that the interpretation or contextual implications of an utterance
produced in reply to a question will respond to that question. Thus, given the assumption that
Mary and Peter are adults, Mary’s utterance carries with it the specific expectation of relevance
that its interpretation will respond to Peter’s question. Contextual implications such as Sally is
exceptionally good and kind or Sally helps humans when needed would allow the derivation of

Sally will look after the children if we get ill, and hence an interpretation which contained these would satisfy Peter’s expectations of relevance. Since it is manifest that Sally is not an angel, however, a literal interpretation would fail to properly warrant such contextual implications. According to the relevance theoretic view of lexical interpretation, Peter thus constructs an *ad hoc* concept ANGEL* by broadening the extension of the concept ANGEL.\(^\text{169,170}\) This broadening involves selecting only those rules and assumptions in the logical and encyclopaedic entries of ANGEL which apply to Sally and would be of use in deriving the contextual implications necessary to satisfy Peter’s expectations of relevance. So, for example, the rule in the logical entry of ANGEL which would allow the inference from Sally is an angel to Sally is a supernatural being would be left out of the *ad hoc* concept ANGEL*, as would the encyclopaedic property MESSENGERS OF GOD. The properties EXCEPTIONALLY GOOD AND KIND and WATCHES OVER HUMANS AND HELPS THEM WHEN NEEDED, on the other hand, would be included, since these can be used to derive the contextual implications which would satisfy Peter’s specific expectation of relevance in this case.

I would like to suggest that the process of deriving a concept via composition of other concepts occurs in much the same way that conceptual narrowing and broadening does – via the selection of rules and assumptions from the concepts which are the inputs to the derivation process. This may seem to be a rather significant departure from what most would normally think of as conceptual composition, which is likely something similar to the picture of concept acquisition painted by the classic view of conceptual structure. According to the classical view, a

\(^{169}\) Following Wilson & Carston, and others, we will use the asterisk (*) to indicate a concept derived via conceptual modulation.

\(^{170}\) According to Wilson & Carston, the modulation of the encoded concept ANGEL in this case actually involves both narrowing and broadening – adding some rules, and omitting others. The process involved is nonetheless the same.
complex concept is constructed out of other concepts in much the same way a structure might be constructed out of building blocks. The concepts used in the construction of the new complex concept are definitional constituents of it. So, in the case of the standard example BACHELOR, the classical theory holds that UNMARRIED and MALE are constituents of the concept, and an individual falls under the concept BACHELOR only if it falls under all of the constituents of the concept as well. Relevance theory does not treat concepts as definitions or potential definitional constituents, however, but rather views them as addresses in memory at which lexical, logical, and encyclopedic information is stored. As a result, the theory cannot treat conceptual composition as a matter of stacking up constituent concepts to form a new concept in the way that the classical theory does. Stacking up addresses in memory does not give you a new address in memory. In a certain sense, then, it is impossible for concepts of the sort envisioned by relevance theory to be combined into a new concept. Even if addresses in memory cannot be combined to form a new address, however, the information stored at those addresses can be combined and stored under a new address. This, I think, is the only way the relevance theorist can make sense of a process of conceptual composition.

Let us call concepts which serve as inputs to a process of concept derivation *contributing concepts*. On the relevance theoretic view I am suggesting, acquiring a new concept via conceptual composition involves creating a new address in memory, copying the rules and assumptions from the contributing concepts and modifying them so that they apply to the new concept, and then writing the rules and assumptions to the new address. So, take the well-worn example of BACHELOR. Let us suppose for the moment that this concept can be acquired by
composing the concepts UNMARRIED and MALE.\(^{171}\) Acquisition of the concept would begin by creating a new address in memory – BACHELOR. The different entries at this address would then be filled with rules or assumptions from the contributing concepts. If the individual acquiring the concept learned it because someone told him or her “a bachelor is an unmarried male”, then the information that the word ‘bachelor’ encodes the concept would be added to the lexical entry of the new address. Now, let us assume that the logical entry for MALE contains a rule like that in (7.4). During the acquisition of BACHELOR, this rule would be copied, modified if necessary, and written to the logical entry of BACHELOR so that the logical entry of the new concept contained a rule like that in (7.5).\(^{172}\)

\[
\begin{array}{c|c|c}
\text{(7.4)} & \text{MALE}(x) & \text{~FEMALE}(x) \\
\text{(7.5)} & \text{BACHELOR}(x) & \text{~FEMALE}(x)
\end{array}
\]

A similar process would happen with assumptions in the encyclopedic entries.

The example of BACHELOR may make it appear as though the relevance theoretic view treats the concept composition process as simply a matter of copying and modifying rules from the entries of the contributing concepts. It may turn out that way in some cases, where the concept being derived is like BACHELOR in having what looks almost like a definitional structure.

\(^{171}\) In all likelihood the concept BACHELOR is acquired via a combination of composition of concepts and conceptual modulation over an extended period of time. An account of the true content of BACHELOR would either have to hold that there are contributing concepts other than UNMARRIED and MALE or would have to hold that BACHELOR is the result of composing the concepts UNMARRIED and MALE\(^*\), where MALE\(^*\) has a narrowed extension picking out humans of the male gender who are eligible for marriage. We should be able to get a clear picture of how the relevance theoretic account of concept composition views the composition process without concerning ourselves with these sorts of complications, however, so we will stick to the simplified understanding of BACHELOR.

\(^{172}\) In a case like that of BACHELOR where the concept would seem to apply to an individual only if the constituents apply to the individual, it might be the case that the rules in the constituent concepts are not moved to the new concept, but rather new rules are created permitting the move from \textit{x is a bachelor} to \textit{x is a male} and \textit{x is a bachelor} to \textit{x is unmarried}. Even if this is so in the case of bachelor, however, there are many concepts that would appear to be acquired by composing other concepts that do not have this seemingly definitional structure. In the case of these concepts, the process would have to be like that which we have described for BACHELOR here.
But in most cases, the derived concept will not have such a structure. In these cases, the composition process will not only involve copying and modifying rules and assumptions, but also selecting which rules and assumptions found in the logical and encyclopedic entries of the contributing concepts should be written to the entries of the new concept. This may sound a great deal like the process of conceptual modulation discussed earlier, and it is. The main difference is that according to the relevance theoretic view of conceptual modulation, which rules and assumptions are eliminated from or added to a contributing concept depends on what contextual implications are required for an utterance to satisfy the hearer’s expectations of relevance. In the case of conceptual composition, on the other hand, which rules and assumptions are included in the entries of the derived concept – and hence the content of the derived concept – depends on how the contributing concepts are taken to relate to one another.

For example, take the concept VOCALOID. A vocaloid is a singing voice synthesizer. Someone who had never heard of a vocaloid before and, upon hearing what it is, attempted to construct a concept by composing the concepts SINGING, VOICE and SYNTHESIZER would end up with different rules and assumptions stored in the entry of the new concept depending on how he or she took these concepts to relate to one another. Someone who understood VOICE as modifying SYNTHESIZER and SINGING as modifying VOICE-SYNTHESIZER would end up with something like the rule in (7.6) in the logical entry of the new concept. Someone who took SINGING to modify VOICE and understood SINGING-VOICE as modifying SYNTHESIZER would end up with a rule like that in (7.7).

\[
\begin{array}{c|c}
\text{VOCALOID}(x) & \text{PRODUCES VOICE} \\
\text{VIA SYNTHESIS}(x) & \\
\end{array}
\]

\[
\begin{array}{c|c}
\text{VOCALOID}(x) & \text{PRODUCES SINGING-VOICE} \\
\text{VIA SYNTHESIS}(x) & \\
\end{array}
\]
Since a singing voice is a type of voice, someone who understood a vocaloid to be a synthesizer of singing voices would probably include both of the rules in (7.6) and (7.7) in the logical entry of the concept. Someone who took a vocaloid to be a voice synthesizer which has been made to sing, however, would not include the rule in (7.7). Since vocaloids are, in fact, synthesizers of singing voices, the exclusion of this rule would mean a misconceptualization of what a vocaloid is. A difference in how the contributing concepts are taken to relate to one another, then, results in a difference in the content of the resulting concept. And in this case the resulting difference has significant consequences.

If the foregoing is correct, then the composition of concepts on the relevance theoretic view is very much like the modulation of concepts via broadening and narrowing. The only difference between the two processes is the basis used for determining which rules should be included in the new derived concept. Just as the modulation of concepts is a form of ad hoc concept construction when performed in order to interpret a lexical item, so too, I believe, is the composition of concepts when performed to interpret a component of LF. Ad hoc concepts, remember, are occasion specific senses. And this is essentially what the relevance theoretic account of definite descriptions depicts descriptive concepts as being when they are constructed in order to interpret an attributively used description – a sense constructed specifically for the purpose of interpreting the content encoded by a description in a context.

It seems likely that these two methods of ad hoc concept construction would both involve greater processing cost than would simple concept retrieval. The retrieval of a concept involves locating the conceptual address in memory. On the other hand, the two methods of ad hoc
concept construction require, at the very least, accessing the rules and assumptions stored in the entries of the contributing concept, determining which rules and assumptions should be in the new concept, and writing those new rules and assumptions to the entries of the new concept. *Ad hoc concept* construction, then, not only requires a greater number of cognitive processes than a simple concept retrieval, but the main process – that of selecting the rules and assumptions which should be written to the entries of the new concept – would appear to be rather more complex and involved than the search for a conceptual address required to retrieve a concept.

The idea that *ad hoc* concept construction requires greater processing cost than concept retrieval fits with Wilson & Carston’s claim that *ad hoc* concept construction is triggered by the search for relevance, as the hearer would only consider interpretations requiring the extra processing effort of *ad hoc* concept construction if interpretations that did not require it failed to satisfy his or her expectations of relevance.\(^{173}\) Let us follow Wilson & Carston, then, in holding that *ad hoc* concept construction is triggered during the interpretation process by the search for relevance. Now, since derivation of concepts occurs via *ad hoc* concept construction, concepts will only be derived if *ad hoc* concept construction is triggered. In other words, in order for a derivable concept to be cognitively accessed – in order for it to be available as an enrichment of encoded content – *ad hoc* concept construction must be triggered.

As we saw earlier, the first condition on the availability of a concept as an enrichment of a component of encoded content was that it be mutually manifest to the conversational participants that the concept is cognitively accessible to the hearer. A concept will be cognitively accessible to an individual at a time only when it is either retrievable or derivable for that

individual at that time. As we just saw, however, a derivable concept will only be cognitively accessible on the relevance theoretic view if *ad hoc* concept construction has been triggered. In light of this, we can say that in order for a concept $C$ to be available to an individual $I$ as an enrichment of a component of the content encoded by an utterance $U$ at a time $t$, $C$ must satisfy the condition laid out in (CA1)

(\text{CA1}) \text{ It is mutually manifest to all conversational parties to } U \text{ that one of the following holds:}

a. $C$ is retrievable for $I$ at $t$, or
b. *Ad hoc* concept construction has been triggered and $C$ is derivable for $I$ at $t$.

Since derivable concepts will only be available when the *ad hoc* concept construction process has been triggered, it will be important for us to have a clear picture of what it takes to trigger that process. Let us call an interpretation which does not require *ad hoc* concept construction an *unconstructed interpretation*. If Wilson & Carston are correct, then *ad hoc* concept construction will only be triggered if the available unconstructed interpretations fail to satisfy the hearer’s expectations of relevance. The hearer’s expectations of relevance, remember, are the expectation of optimal relevance and any specific expectations as to how the utterance might achieve relevance. An interpretation is optimally relevant when it is the most relevant of the available interpretations. Since only unconstructed interpretations will be available prior to *ad hoc* concept construction, as long as there at least one unconstructed interpretation available prior to *ad hoc* concept construction which has relevance, the hearer’s expectation of optimal relevance will be satisfied. Given this, we can say that *ad hoc* concept construction will be triggered only when at least one of the conditions in (ACC) holds.
**Ad Hoc Concept Construction Triggering Conditions (ACC)**

1. No unconstructed interpretation of the utterance is available.
2. All available unconstructed interpretations of the utterance fail to have relevance.
3. There are specific expectations of relevance which the available unconstructed interpretations fail to satisfy.

We now have a more or less clear picture of when a concept would satisfy the cognitive accessibility requirement on availability as an enrichment of encoded content. But being cognitively accessible is only half of what is necessary. As we saw earlier, in order for a concept to be available as an enrichment of a component of LF it must actually be an *enrichment* – that is, the interpretation of the component of LF as that concept must be compatible with the linguistically encoded content being enriched. The enriched interpretation must conform to the encoded meaning it enriches. A concept which satisfied both this condition and the one laid out in (CA1) would, I believe, be available as an enrichment of encoded content. (CA) gives the full set of conditions on the availability of a concept as an enrichment of encoded content.

**Concept Availability (CA):**

When a component of LF $E$ is undergoing enrichment at time $t$ as part of an individual $I$'s interpretation of an utterance $U$, a concept $C$ is an available enrichment of $E$ at $t$ if and only if:

1. It is mutually manifest to all conversational parties to $U$ that one of the following holds:
   a. $C$ is retrievable for $I$ at $t$, or
   b. *Ad hoc* concept construction has been triggered and $C$ is derivable for $I$ at $t$.

2. It is mutually manifest to all conversational parties to $U$ that the interpretation of $E$ as $C$ is compatible with any syntactic, lexical or procedural information encoded at $E$.

What counts as being compatible with encoded content, of course, will depend on the encoded content in question. According to our account of definite descriptions, remember, the definite article encodes a constraint on the interpretation process limiting the interpretation of the
NP to individual concepts. And as we saw in chapter 6, the different types of individual concepts relate to the content of a description in different ways. What this means is that the different types of individual concept will satisfy the compatibility requirement in clause (2) of (CA) in different ways. The three clauses of (COM) below attempt to specify the conditions which must be met for different individual concept interpretations of a description to be compatible with the content encoded by that description. Since the relevance theoretic account of descriptions on offer here views a description as encoding both conceptual and procedural information, an interpretation of the description as a concept will have to conform to both the conceptual and procedural content if it is to be compatible with the description’s encoded content. The (a) clauses in (COM) indicate what is required for compatibility with encoded conceptual content. The (b) clauses, on the other hand, indicate what is required for compatibility with the encoded procedural content – in essence, they indicate what is necessary for a concept to be an individual concept of the type in question.

**Compatibility (COM):**

(1) The interpretation of a description as a descriptive concept is compatible with the conceptual and procedural content encoded by that description only when:
   a. The encoded conceptual content, or content derived from it via conceptual broadening or narrowing, constitutes the descriptive content of the concept, and
   b. The referent of the concept uniquely satisfies the content of the concept, or would uniquely satisfy that content if there were a referent.

(2) The interpretation of a description as a de re concept is compatible with the conceptual and procedural content encoded by that description only when:
   a. The encoded conceptual content, or content derived from it via conceptual broadening or narrowing, serves to uniquely identify the referent of the concept.
   b. The referent of the concept bears a causal relationship to its content.

(3) The interpretation of a description as a type concept is compatible with the conceptual and procedural content encoded by a description only when:
a. The extension of the encoded conceptual content or of content derived from it via conceptual broadening or narrowing contains all and only tokens of the type which is the referent of the type concept.

b. The referent is a kind.

According to our account of definite descriptions, the attributive use of a description results from the selection of a descriptive concept, the referential use results from the selection of a de re concept, and the kind-referring use results from the selection of a type concept. Now, the generally held view of the attributive use of descriptions holds that the content of the description is part of the proposition expressed and the referent of the description satisfies that content. Clause (1) of (COM) attempts to capture this by requiring the conceptual content encoded by the description, or content derived from it via conceptual narrowing or broadening, to constitute the descriptive content of the descriptive concept. Thus, the content of the description will be part of what is explicitly communicated. And since the referent of a descriptive concept uniquely satisfies the descriptive content of that concept, this means that for a descriptive concept interpretation to be compatible with encoded content, the referent of the concept must satisfy the content of the description.

In the case of referential uses of definite descriptions, the general view is that the content of the description identifies the referent of the description, but is not part of the proposition expressed. Clause (2) of (COM) attempts to capture this by requiring the referent of a de re concept to be uniquely identified by the conceptual content encoded by the description, or content derived from it via conceptual narrowing or broadening, in order for an interpretation of the encoded content as that de re concept to be compatible with that encoded content. Since the
content of the *de re* concept is not constituted by the content of the description, however, the content of the description will not be part of what is explicitly communicated.

Finally, a generic kind-referring use of a description refers to whatever kind is identified in the content of the description as an individual. Clause (3) of (COM) attempts to capture this by requiring that the extension of the content encoded by the description be equal to the set of members of the kind referred to by a type concept.

The conditions on concept availability that we have laid out in this section should allow us to demonstrate which of the potential interpretations of a definite description are available as enrichments of the content encoded by the description in our test cases and which are not. And if the test cases turn out to be as we suggested in §7.3.1, then the interpretation of the description will be the only aspect of the interpretation we will need to look at when comparing available interpretations for relevance. As indicated earlier, the plan is ultimately to show that the different available interpretations of the definite description in the test cases yield the same cognitive effects and then compare them along the processing cost side of the relevance ratio. The next section is dedicated to examining the possibility of showing that the different available interpretations have the same cognitive effects.

### 7.4 Justification and Cognitive Effects

There is one significant obstacle facing any relevance theorist who wishes to show that a set of available interpretations yield the same cognitive effects. Cognitive effects come either in the form of contextual implications or in the form of changes in the confirmation values of assumptions which are already part of the context. As we saw in chapter 5, the contextual
implications of an interpretation are conclusions which can be derived via elimination rules from the interpretation together with the contextual assumptions, and cannot be derived from the interpretation or the contextual assumptions alone. Counting up assumptions inferable from the interpretation together with the contextual assumptions or, better yet, showing that same assumptions could be inferred from different available interpretations should not pose any significant problems, at least in theory. The problem lies in the other sort of cognitive effect – changes in the confirmation values of assumptions already present in the context.

According to Sperber & Wilson, the confirmation values of assumptions in the context are non-quantitative. As such, they cannot be measured; nor could any changes to those confirmation values. If we cannot measure changes in the confirmation values of assumptions, however, then it is unclear how we could go about showing that two different interpretations have the same cognitive effects. This is only a problem, however, if changes in the confirmation values of contextual assumptions are actually a factor that we must consider. Sperber & Wilson’s claim that confirmation value changes are cognitive effects is a claim about how relevance is assessed in natural conversational circumstances. But philosophical test cases are far from natural. They are contrived cases, ones in which we, as interpreters, are generally given no indication whatsoever about the confirmation values of any of the assumptions in the context or the assumptions comprising the interpretation of the utterance. Without information about the confirmation values of the assumptions making up the interpretation of the utterance or the contextual assumptions themselves, we as interpreters cannot determine what changes, if any, might be made in the confirmation values of the contextual assumptions. Thus, when we give our

---

174 Sperber & Wilson (1986), pp. 110&111
intuitive interpretations of the example utterances in these contrived cases, it would appear that we are doing so without any reference to the confirmation values of the assumptions involved. If this is the case, however, then the theorist who is attempting to show that relevance theory predicts that a particular interpretation will be optimally relevant in the interpretive context of one of these contrived cases need not worry about confirmation values when comparing the cognitive effects of the various available interpretations. After all, if the interpreter whose interpretation the theorist is attempting to predict does not use confirmation values when giving his or her interpretation, then there are no confirmation values for the theorist to be concerned with when attempting to predict that interpretation.

One might worry that in the face of a lack of information about the confirmation values of assumptions in one of these contrived cases, the interpreter simply assigns confirmation values to the assumptions himself, and then assesses the relevance of available interpretations on the basis of those confirmation values. There are a number of reasons to suspect that this is not the case, however. To see this, let us once again consider Donnellan’s famous example of a referentially used description. Suppose that Jones has been placed on trial for the murder of Smith. Suppose further that there has been some discussion of Jones’ strange behavior at the trial, in light of which A, in conversation with B, utters (7.1).

(7.1) Smith’s murderer is insane.

If the arguments in §7.1 are on the right track, then your intuitive interpretation here is the result of interpreting the utterance in an interpretive context. It is likely that your intuitive interpretation of (7.1) – as an interpreter of an utterance in a contrived case – is something like Jones is insane. Now, assuming that all other conditions for knowledge are met, as yourself: does
B’s justification for his belief that Jones is the murderer of Smith rise to the level of knowledge? Can we attribute to B the knowledge that Jones behaved strangely at this trial? I suspect that you do not have answers to these questions. If you, as an interpreter, had assigned confirmation values to these contextual assumptions, however, then you should be able to make an intuitive determination as to whether those confirmation values are sufficient for knowledge.

But it might be the case that confirmation values are assigned to the contextual assumptions here at the subpersonal level (i.e., by your pragmatics processing submodule) and then only for the provisional purpose of interpretation. As such, perhaps we should not expect you to be able to determine whether those confirmation values rise to the level of knowledge. There is reason to think it is not how things work, however. Confirmation values might sometimes be assigned on the basis of general information about how the world works (for example, assigning a high confirmation value to the assumptions that are stipulated to be perceptually available to the conversational participants), but at least some of the assumptions in the context would have to be assigned confirmation values almost completely arbitrarily.

For example, in the case above, one of the contextual assumptions is clearly that Jones is the murderer of Smith. While this assumption is not explicitly stated (a fact which makes the case murkier than is ideal for the theorist), it is necessary to get the referential interpretation of the description that we intuitively give in the case. Given that the only evidence we have for this assumption is that Jones is on trial for Smith’s murder, different confirmation values might be assigned by different interpreters. Those who tend to trust the efficacy of criminal justice systems might assign a higher confirmation value to this assumption, while more skeptical interpreters might assign a lower confirmation value. Now, if relevance theory is correct – as we
are assuming it is for our purposes here – then we would expect to see interpreters who assigned different confirmation values to the assumptions in the context stipulated for (7.1) reaching different intuitive conclusions as to the interpretation of (7.1). For example, those who assigned a higher value to the assumption that Jones is the murderer of Smith would find the referential interpretation of the description intuitive, while those who assigned a lower value to this assumption might have unclear or uncertain intuitions as to the proper interpretation. Now, intuition pumps in philosophy are frequently contentious, but that is not the case here. The near universal intuitive interpretation of the description is definitely a referential one.

Given this, one might be inclined to claim that the pragmatics processing submodule in such cases simply assigns a confirmation value of certain – that is, the highest confirmation value possible – to all the assumptions given in these cases – both the contextual assumptions and the assumptions comprising the interpretation. After all, these assumptions are treated as given. If this is the case, however, then the theorist attempting to evaluate cognitive effects in one of these contrived cases need not be concerned about changes in confirmation values, as there are likely to be none.

According to Sperber & Wilson, remember, there are two types of cognitive effect: contextual implications and confirmation value changes. Each of these types of cognitive effect comes in two varieties. A contextual implication can be either the addition of an assumption to the context or the removal of an assumption from the context. And a confirmation value change can either be a strengthening of the confirmation value of an assumption or a weakening of its
confirmation value.\textsuperscript{175} Now, to make our discussion here a bit clearer, let us use $I$ to refer to the assumptions forming a given utterance interpretation in a context $C$, and let us use $K$ to refer the contextual implications of $I$ in $C$.

Now, the addition of any assumption $A$ to the context occurs via contextual implication. The removal of an assumption $A$ from the context occurs when $A$ is contradicted by an assumption of a higher confirmation value found either in $I$ or $K$.\textsuperscript{176} Strengthening of a contextual assumption $A$ occurs when the following two conditions are met: (1) $A$ is implied (non-trivially) by premises drawn from \textit{both} $I$ and $C$, but not by premises drawn just from $I$ or just from $C$, and (2) the confirmation values of those premises are sufficient to raise the confirmation value of $A$. Finally, the weakening of a contextual assumption $A$, according to Sperber & Wilson, occurs when an assumption serving as support for $A$ is removed from the context as a result of a contradiction.\textsuperscript{177}

If we assume that all assumptions in both $I$ and $C$ are given confirmation values of \textit{certain}, then no strengthening of the confirmation values of contextual assumptions can take place, as all contextual assumptions already have the highest confirmation value possible. Assuming that Sperber & Wilson are correct in their claim that contextual implications arising out of premises whose confirmation values are certain will themselves have confirmation values of certain, this means that the introduction of contextual implications cannot serve to strengthen

\textsuperscript{175} As noted in chapter 5, Sperber & Wilson actually list only three types of contextual effects: contextual implications, contextual strengthenings, and removal of contextual assumptions. The weakening of confirmation values they treat as an indirect effect of this last type. See Sperber & Wilson (1986) p.194 n2d. I have again listed all four so that we can clearly see \textit{all} of the ways that contextual effects may arise.
\textsuperscript{176} When there is a contradiction between $A$ and an assumption from $I$ or $K$, but $A$ has a higher confirmation value, there is no contextual effect, as the rejected assumption(s) will be those with the lower confirmation value. See Sperber & Wilson (1986) p. 115
\textsuperscript{177} Sperber & Wilson (1986) pp. 108-117
any the confirmation values of other assumptions. Weakening of confirmation values is not possible either, though, for, \textit{ex hypothesi}, the confirmation values of the assumptions in $I$ and $C$ in these contrived cases are provisionally assigned for the purpose of interpretation, not estimated on the basis of whatever justification is actually possessed. Thus, the loss of a supporting assumption should have no effect on the confirmation values of any of the assumptions in $I$ or $C$, since those confirmation values were not dependent on support from other assumptions in the first place.

It would seem, then, that even were we to assume that, \textit{qua} given assumptions, the assumptions in $I$ and $C$ in a contrived case are assigned confirmation values of \textit{certain}, we would not need to worry about changes in justification as potential contextual effects of the interpretations to be compared when it comes time to attempt to show that one interpretation of an utterance containing a definite description is more relevant than other available interpretations. As we saw above, it is implausible to assume that an interpreter's pragmatics processing submodule is arbitrarily making provisional assignments of different confirmation values to the different assumptions present in a contrived case. Given this, it would appear that our only reasonable options when it comes to confirmation values in a contrived case are either to treat confirmation values as irrelevant, or to treat them as though they are given provisional values of \textit{certain}. Either way, changes in confirmation value will not be a factor in the cognitive effects associated with the various interpretations of an utterance.

If the foregoing is correct, then the only cognitive effects we will need to consider when attempting to compare interpretations of an utterance containing a definite description will be contextual implications and removal of assumptions from the context. Theoretically speaking,
there shouldn’t be any serious obstacles to comparing the number of contextual implications, but what about the removal of assumptions from the context? There may be a bit of a problem there, as removal of an assumption $A$ from $C$ comes about only when $A$ is contradicted by an assumption with a higher confirmation value in either $I$ or $K$. But if what we have said is right, then if we encounter a contradiction between an assumption $A$ in $C$ and one in either $I$ or $K$ we will be unable to determine whether $A$ should be rejected or whether the assumption in $I$ or $K$ should be rejected, as we will be treating both assumptions as either certain or lacking confirmation values entirely. In other words, we will not be able to determine whether there is a contextual effect or not. I see no way around this problem. All this means, however, is that we will be unable to compare the relevance of interpretations of an utterance in a context when one of the interpretations or its contextual implications contradicts an assumption already in the context. While we will not be able to say for certain that this will not hinder our project until we look at particular cases, I do not think it is likely that we will encounter many cases involving such contradictions. After all, the contexts of contrived cases are stipulated simply for the purposes of generating an intuitive interpretation and are smaller than natural conversational contexts, presenting fewer opportunities for contradiction. Additionally, the standard test cases involving definite descriptions, like those involving utterances of (7.1), do not appear on the face of things to involve inconsistencies between the contextual assumptions and the interpretations or their contextual implications. Even so, I do not expect this to convince the reader that our project is safe from cases involving such contradictions. Whether or not this will prove fatal for our project cannot be determined until we attempt our analysis of individual contrived cases.
If the foregoing is correct, then it will not be necessary to consider confirmation value changes when assessing the cognitive effects associated with the different available interpretations of utterances containing definite descriptions in contrived cases. This should make it possible for us to compare the cognitive effects associated with available interpretations in our test cases. If the different available interpretations of descriptions in our test cases are, as I suspect, the same along the cognitive effects side of the relevance ratio, then it will be possible to show that one interpretation is more relevant than another by showing that it requires less processing cost. A method for comparing the processing costs of different interpretations of definite descriptions is presented in the next section.

### 7.5 Comparing Processing Costs

The relevance theoretic account of definite descriptions on offer here holds that definite descriptions are interpreted as individual concepts. The processing cost associated with a particular interpretation of a description, then, is the processing cost involved in accessing the relevant individual concept. If we are to compare the processing cost of different interpretations of descriptions, then, we will have to have some idea of the processing effort involved in accessing individual concepts. The goal of this section is not to give a complete account of how to measure the processing cost associated with accessing different individual concepts. It will not be possible to give such an account until we have a far more detailed picture of the cognitive processes involved in memory and perception. Rather than giving a full account, then, the goal of this section is to determine what sort of assumptions we might reasonably make about the
processing costs associated with individual concepts which will aid us in comparing the processing costs of different interpretations of descriptions in example utterances.

Sperber & Wilson treat concepts as addresses in memory under which three types of information are stored: logical, encyclopedic, and lexical. The encyclopedic information stored at a conceptual address consists in information about the extension or denotation of the concept, while the lexical information consists in information about the natural language items(s) that express the concept. The logical information consists in the set of deductive elimination rules associated with the concept. According to Sperber & Wilson, this will include rules associated with the meaning of the concept. So, for example, at the conceptual address for DOG would be a rule allowing the deductive device to infer \( x \) is an animal from \( x \) is a dog.

Now, as we have seen, there are two ways that a concept might be cognitively accessed: either the concept could be retrieved from memory, or it could be constructed. Let’s start by looking at the processing cost associated with retrieving a concept. According to Sperber & Wilson, when a concept appears in an LF, it provides access to the information stored at that conceptual address. Retrieval of a concept on this view, then, amounts to a recovery of the conceptual address. Thus, how much processing effort is required to retrieve a concept is directly related to how accessible that conceptual address is. Sperber & Wilson’s discussion of reference assignment indicates that they view the accessibility of concepts in the same light as the accessibility of assumptions and interpretive hypotheses. In response to the question of how a hearer should construct and evaluate referential hypotheses about an instance of ‘it’, they claim “…he should first consider the immediate context, see if any of the concepts of a non-human...

---

178 Sperber & Wilson (1986) p. 86
entity represented in this context…yields a propositional form consistent with the principle of relevance; if not, he should extend the context and repeat the procedure.” What Sperber & Wilson are depicting here is a hearer using the ITP testing process to evaluate referential hypotheses. In other words, they are claiming that the hearer engaged in reference assignment tests referential hypotheses – concepts – in order of accessibility, stopping when the expectation of relevance is met.

As we saw in the chapter 5, ITP cannot be the testing process used by hearers in their interpretation of utterances. But what is important for our purposes here is not the testing process itself, but the order of testing depicted. Just as they do with assumptions, Sperber & Wilson locate the most accessible concepts in the immediate given context – the set of assumptions remaining in the deductive device after the last deductive process. And they assume that the next most accessible concepts are located in the next most accessible context. It would seem, then, that they take the accessibility ordering they give for assumptions to apply to concepts as well. If we treat concept accessibility for hearers in the same manner as assumption accessibility, then the most accessible concepts are found in the memory of the deductive device, the next most accessible are in the short term memory store, and the next most are in long term memory. The accessibility of concepts within these memory stores is further ordered by reference to their frequency of use. Just as more frequently processed assumptions are, according to Sperber & Wilson, more accessible, more frequently processed concepts are also more accessible.

180 Ibid., p. 187
This accessibility ordering is the accessibility ordering for a hearer in an actual conversation. What we will be attempting to predict in the next chapter, however, is the natural interpretation of the reader of a contrived example in an interpretive context. So, what accessibility ordering is utilized in an interpretive context? As we saw in §7.2, the interpreter is attempting to uncover the speaker’s meaning. The speaker, however, constructs his or her utterance for the hearer(s). Hence, a speaker who wished to be understood would tailor his or her utterance with the hearer’s accessibility ordering in mind. To uncover the speaker’s meaning, then, the interpreter should rely on the accessibility ordering he or she has good reason on balance to believe the speaker would judge to be the hearer’s accessibility ordering.

So, the processing cost associated with retrieving a concept is directly related to how accessible it is. The less accessible the concept, the greater the processing cost required to retrieve it. But what if a concept is not retrievable, but is instead derivable? (D), which was introduced in §7.3.1, attempts to capture all the ways in which we might expect a concept to be derived.

**Derivability (D):**
A concept $C$ is derivable for an individual $I$ at a time $t$ if and only if one of the following conditions holds:

1. $I$ could acquire $C$ by composing concepts present in $I$’s mental lexicon at $t$,
2. $I$ could acquire $C$ via narrowing, broadening or otherwise modifying a concept in $I$’s mental lexicon at $t$,
3. $I$ could acquire $C$ via some perception based interaction with $I$’s physical environment at $t$,
4. $I$ could acquire $C$ via some combination of (1)-(3)

Unfortunately, without a much better understanding of the relationship between concepts and perception than we currently possess, it will not be possible to make any claims about what sort of processing effort is involved in constructing concepts via perception based interactions with
the physical environment. We will thus be unable to say anything about the processing costs associated with concepts derived via the methods described in clauses (3) and (4) of (D).\textsuperscript{181}

There are a few reasonable assumptions we can make about concepts derived via the methods described in clauses (1) and (2), however.

Let us begin by considering concepts derived via the method described in (1) – that is, concepts derived via composition of other retrievable concepts. As we saw in §7.3.2, it seems reasonable to suppose that deriving a concept by composing other concepts requires greater processing effort than retrieving a concept because there are multiple processes involved in composing a concept, one of which is, on the face of things, considerably more involved than the basic search process required to retrieve a concept from memory. According to the relevance theoretic view of concept composition we offered in §7.3.2, the derivation of a new concept via the composition of other concepts involves selecting rules and assumptions from the logical and encyclopedic entries of the contributing concepts, or constructing new rules and assumptions, on the basis of how the contributing concepts relate to one another. This would require at least three steps: retrieving the information stored at the logical entries of each concept used in the composition process, determining which deductive rules should be stored in the logical entry of

\textsuperscript{181} Taken within the relevance theoretic framework, the claim that the process of \textit{ad hoc} concept construction is not triggered unless the unconstructed interpretations fail to satisfy the hearer’s expectations of relevance would seem to rely on the assumption that any process of \textit{ad hoc} concept construction will require greater processing cost than a process of concept retrieval. If we assume that all of the methods of deriving new concepts are types of \textit{ad hoc} concept construction, this would mean that the methods of concept derivation listed in clauses (2) and (3) of (D) require greater processing cost than concept retrieval. It seems plausible to me to think that any method of \textit{ad hoc} concept construction will involve more processing effort than a simple concept retrieval would, since any method of deriving a concept will involve a process of selecting or constructing the rules and assumptions which will be stored in the concept’s logical and encyclopedic entries. We will not be explicitly relying on this claim here, however. Those who are uncomfortable with the assumption that all of the various methods of \textit{ad hoc} concept construction involve more processing effort than concept retrieval should treat Wilson & Carston’s claim that \textit{ad hoc} concept construction is triggered by the search for relevance as a stipulated assumption of the relevance theoretic theory of \textit{ad hoc} concept construction.
the new concept (and possibly constructing new rules), and then writing the new rules to the logical and encyclopedic entries of the concept being derived. Given this, it will be the case that the greater the number of concepts being used in the composition process, the higher the overall processing cost of deriving the new concept will be. This is so for two reasons. First, each additional contributing concept adds an additional step of cognitive processing. If two concepts are being composed, the operation of accessing information stored in memory will have to be performed twice. If three concepts are being composed, the operation will have to be performed three times, and so on. Second, each additional contributing concept will make the process of determining which deductive rules should be stored in the logical entry of the new concept more involved.

So, the greater the number of constituent concepts, the greater the processing cost associated with the derivation of the complex concept. Additionally, since the processing effort required to access the logical entry of a constituent concept will vary depending on how accessible that concept is, the overall processing cost required to construct a concept via composition will depend on how accessible the constituent concepts are. The less accessible the information stored in memory is, the more processing effort is required to access it. Thus, the less accessible the constituent concepts, the greater the overall processing cost involved in constructing the complex concept will be.

If what we have said so far is correct, then, there will be two factors impacting the construction of a descriptive concept via composition of retrievable concepts: the number of concepts being composed, and the accessibility of those concepts. But what about concepts derived via conceptual narrowing and broadening? As we saw in §7.3.2, a relevance theoretic
approach to *ad hoc* concept construction views the process of deriving new concepts via conceptual modulation as similar in important respects to that of deriving a concept via composition. But while the process of selecting which rules and assumptions will be stored in the entries of a new composed concept proceeds on the basis of how the concepts being composed are taken to relate to one another, the selection of rules and assumptions during conceptual modulation is based on what is necessary to get whatever contextual implications are required to satisfy the hearer’s expectations of relevance. This process involves a sort of backward inference, beginning with the expected conclusion and working out the premises necessary to reach that conclusion. Unfortunately, we do not currently possess a clear enough picture of how the human cognitive system engages in these sorts of reasoning tasks to be able to make any general claims about what would increase or decrease the processing effort required to derive a concept via conceptual narrowing and broadening. Nor are we in a position to attempt to construct an account of such reasoning processes here, as this is undoubtedly a task best left to cognitive scientists working on reasoning.

We may not currently be in a position to say anything about what would increase or decrease the processing effort required to derive an *ad hoc* concept via narrowing and broadening, but there is one assumption that it would seem reasonable to make about concepts whose derivation involves, in part, conceptual narrowing and broadening. All else equal, constructing a concept via both composition and conceptual narrowing or broadening will require more processing effort than constructing a concept via only composition. The reason for thinking this is simple – when all else is equal, the derivation of a concept via both composition and conceptual modulation involves an extra cognitive process that derivation via composition
alone does not. So, for example, take the concepts MURDERER-OF-SMITH and MURDERER*-OF-SMITH. Suppose that there are no narrowed or broadened versions of MURDERER which are retrievable. Constructing MURDERER-OF-SMITH in such a case would require composing three concepts. The construction of MURDERER*-OF-SMITH would also require composing three concepts, but it would also require the derivation of MURDERER* via the narrowing and/or broadening of MURDERER. This additional process involves additional processing effort, meaning that MURDERER*-OF-SMITH will have a higher processing cost than MURDERER-OF-SMITH.

Let us turn now to type concepts. It seems plausible to me to think that a type concept is constructed via some sort of conceptual modulation of the concept of the kind which ends up being the referent of the type concept. It is also entirely possible, though, that type concepts are constructed via the composition of concepts, or even via some combination of perceptual interaction with a member of the kind the concept refers to and conceptual combination or modulation. It may even be that some type concepts are constructed in one way and some in another. Sorting out this issue, it seems to me, would ultimately require a deeper understanding of the relationship between the content of a type concept and its referent. And this would require understanding the referents of type concepts themselves – kinds. I do not believe it will be necessary for us to commit ourselves to any particular view of which methods are used to derive type concepts in order to accomplish our main goal of using the relevance theoretic account of descriptions to predict the natural interpretation of a description in standard examples, however. The standard examples of generic kind-referring uses of definite descriptions always involve types that we expect adults to already possess concepts of, so construction of type concepts would not be an issue in these cases. And the standard examples of referential and attributively
used descriptions are generally such that their content could not be taken as referring to anything which adults generally assume to be a kind. In these cases, then, there will be no type concept – constructed or otherwise – which satisfies the compatibility requirement on concept availability laid out in clause (2) of (CA).

There is one reasonable assumption we may be able to make about the processing cost associated with type concepts that could prove useful when we turn to the task of using our relevance theoretic account to predict the natural interpretations of standard examples of uses of descriptions. It turns out that certain kinds can more readily serve as the referents of generic kind-referring descriptions than others. Work on generics has revealed that the kind-referring interpretation will only be naturally assigned to a description if the descriptive content is associated with a well-established kind.\textsuperscript{182} To see this, consider the following example from Krifka, et. al.\textsuperscript{183}:

\begin{enumerate}
\item[a.] The Coke bottle has a narrow neck
\item[b.] ??The green bottle has a narrow neck
\end{enumerate}

There is a marked contrast between (a) and (b). The kind-referring reading is unproblematic in (a). Indeed, it is the natural reading, though we could also very easily assign an attributive or referential reading to the description. The same cannot be said for (b). The natural reading of the description in (b) is an attributive reading, though it would also be unproblematic to assign the description a referential reading. While may be able to force a kind-referring reading onto the description in (b), however, it is considerably more difficult than hearing the description as being used referentially or attributively. The reason for this contrast between (a) and (b) is that the kind

\textsuperscript{182} Nunberg & Pan (1975); Carlson (1977); Dahl (1975), (1985)
\textsuperscript{183} Krifka, et. al. (1995), p. 11
referred to by the description in (a) is well-established – it is conventionally conceived of and talked about as a kind – whereas the kind referred to by the description in (b) is not.

Now, since the general idea behind cognitive processing cost is the amount of cognitive effort required to perform a task, if some cognitive task is more difficult than another, it seems reasonable to assume that the former involves greater cognitive processing cost than the latter. In light of this, we can say that assigning a kind-referring reading to a description when the kind referred to is not well-established requires greater processing cost than assigning an attributive or referential reading. In the terms of our account of definite descriptions, this means that retrieving or constructing the type concept of a kind that is not well-established requires greater processing cost than retrieving or constructing a descriptive or *de re* concept.

The foregoing discussion has pinpointed several assumptions that we might reasonably make about the processing cost associated with selecting different individual concepts in different situations. As indicated at the beginning of this section, the goal here was not to give a complete account of the processing cost associated with the selection of concepts, or even the processing cost associated with the selection of individual concepts. Such a task will have to wait until our understanding of various aspects of human cognition is considerably more advanced than it currently is. Until that time, we will have to rely on those general assumptions about comparative processing cost that are justified by what we do currently know. (PIC) details the assumptions we have argued for in this section.
Processing Cost of Concepts in an Interpretive Context (PIC):
(1) If an interpreter has good reason to believe that a concept is retrievable, the processing cost of that concept will be directly related to its accessibility in the interpretive context.
(2) Retrieving a complex concept requires less processing cost than constructing it.
(3) The cost of composing a descriptive concept is directly related to (a) the number of its components and (b) the accessibility of its components.
(4) Other things equal, constructing a concept via both composition and conceptual narrowing or broadening requires greater processing cost than constructing a concept via only composition.
(5) Retrieving or constructing a type concept T-K requires greater processing cost than retrieving or constructing a descriptive concept K, unless the type is well-established.

Hopefully, these assumptions – along with Sperber & Wilson’s claims about the accessibility ordering of assumptions and concepts in memory – will help us to compare the processing costs that would be associated with the selection of different individual concepts in our examples of utterances containing definite descriptions.

We have seen in this chapter that there is a potential method available for comparing the relevance of different interpretations of definite descriptions without the need to quantify either cognitive effects or processing costs. In the next chapter, we will employ this method in an attempt to use our relevance theoretic account of definite descriptions to predict the reader’s natural interpretations of standard examples of the attributive, referential, and generic kind-referring uses of definite descriptions.
Chapter 8
The Test Cases

In the previous chapter, I suggested a strategy for comparing the relevance of interpretations and constructed the theoretical machinery that would allow us to employ that strategy in using relevance theory to make predictions about the natural interpretations of utterances in philosophical example cases. In this chapter, I will employ this machinery in an attempt to use the relevance theoretic account of definite descriptions offered in chapter 6 to make predictions about the natural interpretations of utterances containing definite descriptions. §8.2 will look at a standard example of an attributive use of a definite description, while §8.3 will examine a standard referential case. §8.4 deals with a generic kind-referring use of a definite description. If we can use the relevance theoretic account of definite descriptions to generate predictions as to the natural interpretations in these cases, we will have overcome what is perhaps the most significant problem facing that account. Before diving into our task, however, it might be helpful to briefly review the key theoretical claims we will be relying on in the attempt show that the relevance theoretic account of definite descriptions accurately predicts the natural interpretations of definite descriptions in standard examples. That is the objective of the next section.

8.1 Key Theoretical Assumptions

According to the account offered in chapter 6, remember, the definite article encodes a procedural constraint on the interpretation process which indicates that the NP in which the
article occurs should be interpreted as an individual concept. Since the NP contains descriptive content, the individual concept selected must be compatible with that descriptive content. We classify individual concepts into, in total, three types: descriptive concepts, *de re* concepts, and type concepts.\(^{184}\) Descriptive concepts bear a satisfactional relationship with their referents, when they have them. *De re* concepts stand in a causally dependent relationship with their referents. The content of a *de re* concept is acquired via a causal relationship with its referent. Finally, type concepts refer to kinds as individuals. As individual concepts, all three of these different types of concepts include content which enables unique identification of the referent. According to the account of descriptions on offer, the attributive reading of a description results from the selection of a descriptive concept during pragmatic processing; the referential reading results from the selection of a *de re* concept, and the generic kind-referring reading results from the selection of a type concept.

On the relevance theoretic view, the individual concept that will be selected by a hearer as the interpretation of a description is the one which would yield the optimally relevant interpretation in the context. The optimally relevant interpretation, remember, is the most relevant of the interpretations available to the hearer, where relevance is a ratio of cognitive

\(^{184}\) As noted when we introduced these classifications, it seems plausible to think that there are at least two other types of individual concept – number concepts and non-descriptive concepts of fictional entities, which are encoded by numerals and names of fictional entities. If this is the case, then there may be non-attributive readings of descriptions which refer to such entities. My account of descriptions would be able to handle this by treating non-attributive readings of descriptions referring to numbers as arising out of the selection of number concepts and non-attributive readings of descriptions referring to fictional entities as arising out of the selection of fictional entity concepts. At this point in time, however, I do not have in depth characterizations of the nature of these types of individual concepts. As a result, it will not be possible to use the theory to predict the natural interpretations of utterances containing descriptions for which these number concepts or fictional entity concepts are potential interpretations. The main task of this chapter is to predict the natural interpretations of standard examples of utterances containing referential, attributive and generic kind-referring readings of descriptions. Since the linguistic meanings of these standard examples of descriptions are such that neither number concepts nor fictional entity concepts are plausible potential interpretations, however, omitting these types of individual concepts from our discussion should not impact our ability to derive predictions.
effects to processing costs. An interpretation of an utterance, according to relevance theory, consists in the propositional form, any higher-level explicatures, and any implicatures. As we saw in chapter 7, for an interpretation of an utterance to be available to a hearer the conditions laid out in (IA) must be met:

(IA) Interpretation Availability
An interpretation of an utterance is available for an individual if and only if:
1. It is cognitively accessible.
2. Each mental representation in the interpretation satisfies the definition of a propositional form, a higher-level explicature, or an implicature.

The definitions of propositional form, higher-level explicature and implicature are as follows:

**Propositional Form:**
A semantically complete enrichment of the LF of an utterance.

**Higher-level explicature:**
An assumption derived from the propositional form and mutually manifest assumptions about the speaker’s propositional attitudes which could not be derived from the propositional form or the mutually manifest assumptions about the speaker’s propositional attitudes alone.

**Implicature:**
A contextual assumption or implication for which it is manifest that the speaker intended to make it manifest.

As we saw in chapter 5, it follows from the conditions in (IA) that for an interpretation to be available, any concept used in that interpretation as an enrichment of linguistically encoded content must be available as an enrichment of that encoded content. ‘Available as an enrichment’ here is meant in a technical sense which we defined in the previous chapter. That definition is given in (CA).
(CA) Concept Availability
When a component of LF $E$ is undergoing enrichment at time $t$ as part of an individual $I$'s interpretation of an utterance $U$, a concept $C$ is an available enrichment of $E$ at $t$ if and only if:

1. It is mutually manifest to all conversational parties to $U$ that one of the following holds:
   a. $C$ is retrievable for $I$ at $t$, or
   b. Ad hoc concept construction has been triggered and $C$ is derivable for $I$ at $t$.

2. It is mutually manifest to all conversational parties to $U$ that the interpretation of $E$ as $C$ is compatible with any syntactic, lexical or procedural information encoded at $E$.

In order to understand clause (1) of (CA), we must understand retrievability and derivability. The definitions of these are given in (R) and (D).

(R) Retrievability
A concept $C$ is retrievable for an individual $I$ at a time $t$ if and only if $C$ is part of $I$'s mental lexicon at $t$.

(D) Derivability
A concept $C$ is derivable for an individual $I$ at a time $t$ if and only if one of the following conditions holds:

1. $I$ could acquire $C$ by composing concepts present in $I$'s mental lexicon at $t$,
2. $I$ could acquire $C$ via narrowing, broadening or otherwise modifying a concept in $I$'s mental lexicon at $t$.
3. $I$ could acquire $C$ via some perception based interaction with $I$'s physical environment at $t$,
4. $I$ could acquire $C$ via some combination of (1)-(3)

Derivation of concepts, remember, occurs via ad hoc concept construction. According to clause (1) of (CA), a derivable concept will be available as an enrichment of a component of linguistically encoded content only when ad hoc concept construction has been triggered during the interpretation process. We argued in the previous chapter that the conditions for triggering ad hoc concept construction during the interpretation process are those given in (ACC).
Ad hoc concept construction will only be triggered during the utterance interpretation process if one of the following conditions holds:

1. No unconstructed interpretation of the utterance is available.
2. All available unconstructed interpretations of the utterance fail to have relevance.
3. There are specific expectations of relevance which the available unconstructed interpretations fail to satisfy.

Clause (2) of (CA) tells us that to for a concept to be available as an enrichment of a component of LF, the interpretation of that component as that concept must be compatible with the linguistically encoded content. Basically, what this means is that any enrichment of encoded content must conform to that encoded content. Because the different types of individual concept work in different ways, there are differences in how they conform to linguistically encoded content. (COM) gives the specific requirements for compatibility of interpretations involving individual concepts with the linguistic content our account assigns to descriptions. The (a) clauses give the requirements for compatibility with encoded conceptual content while the (b) clauses give the requirements for compatibility with encoded procedural content.

(COM) Compatibility

1. The interpretation of a description as a descriptive concept is compatible with the conceptual and procedural content encoded by that description only when:
   a. The encoded conceptual content, or content derived from it via conceptual broadening or narrowing, constitutes the descriptive content of the concept, and
   b. The referent of the concept, when there is one, uniquely satisfies the content of the concept.

2. The interpretation of a description as a de re concept is compatible with the conceptual and procedural content encoded by that description only when:
   a. The encoded conceptual content, or content derived from it via conceptual broadening or narrowing, serves to uniquely identify the referent of the concept.
   b. The referent of the concept bears a causal relationship to its content.

3. The interpretation of a description as a type concept is compatible with the conceptual and procedural content encoded by a description only when:
a. The extension of the encoded conceptual content or of content derived from it via conceptual broadening or narrowing contains all and only members of the kind which is the referent of the type concept.

b. The referent is a kind.

Conceptual broadening and narrowing, remember, are one type of ad hoc concept construction.

What we will be attempting to predict in each of our cases is the reader’s natural interpretation of the description in an example utterance in a stipulated context. Since the reader will not be a party to the conversations in our examples, his or her interpretation will occur in what we have called an interpretive context. As we saw in the previous chapter, an assumption is part of an interpreter’s interpretive context if and only if the interpreter has good reason on balance to believe that the assumption is mutually manifest to the conversational participants. (PP) details the relevance theoretic view of the natural interpretation of an utterance in an interpretive context.

(PP) Prediction Principle
In any interpretive context $C$ in which an utterance $U$ has one or more available, relevant English interpretations, $P$ is the most natural English interpretation of $U$ if and only if $P$ is the optimally relevant interpretation in $C$.

Given (PP), our relevance theoretic account of descriptions will predict that a particular interpretation would be the reader’s natural interpretation of an utterance when that interpretation is the optimally relevant one in the interpretive context in question. The definition of optimal relevance in an interpretive context is given in (ORIC).

(ORIC) Optimal Relevance in an Interpretive Context
An interpretation of an utterance $U$ in a conversation $C$ is optimally relevant in an interpretive context if and only if it is the most relevant of the set of interpretations of $U$ that the interpreter has good reason on balance to believe the speaker would judge to be available to the hearer in $C$. 

253
Part of relevance is processing cost, remember, and if our test cases are as we suggested in chapter 7, the aspect of processing cost that will matter for our purposes will be that surrounding the retrieval or derivation of the individual concept. (PIC) details our assumptions about processing cost in an interpretive context.

**(PIC) Processing Cost of Concepts in an Interpretive Context**

(1) If an interpreter has good reason to believe that a concept is retrievable, the processing cost of that concept will be directly related to its accessibility in the interpretive context.

(2) Retrieving a complex concept requires less processing cost than constructing it.

(3) The cost of composing a descriptive concept is directly related to (a) the number of its components and (b) the accessibility of its components.

(4) Other things equal, constructing a concept via both composition and conceptual narrowing or broadening requires greater processing cost than constructing a concept via only composition.

(5) Retrieving or constructing a type concept \( T \)-\( K \) requires greater processing cost than retrieving or constructing a descriptive concept \( K \), unless the type is well-established.

Clause (1) of (PIC) indicates that the processing cost of a retrievable concept is directly related to its accessibility in an interpretive context. (AIC) lays out the accessibility ordering in an interpretive context.

**(AIC) Accessibility Ordering in an Interpretive Context**

The accessibility ordering used by an interpreter in interpreting an utterance \( U \) in an interpretive context \( IC \) will be that which the interpreter has reason to believe the speaker would judge to be the accessibility ordering for the hearer. When \( U \) is immediately preceded by another utterance, \( T \), the most accessible set of assumptions and concepts will consist in:

(1) Assumptions constituting the interpretation of \( T \) which the interpreter has good reason on balance to believe the speaker would judge to be the hearer’s interpretation of \( T \).

(2) Contextual implications derived from (1).

(3) The premises used in deriving (2).

(4) Concepts located in the assumptions or rules found in (1), (2) and (3).

The next most accessible assumptions will be those which the interpreter has reason to believe the speaker would judge to be in the short term memory of the hearer. The last
most accessible will be those the interpreter has reason to believe the speaker would judge to be in the long term memory of the hearer.

Together, (PIC) and (AIC) will aid us in comparing the processing costs of the different available interpretations of the descriptions in the test cases we will look at here, and will hopefully make it possible to compare the processing costs of interpretations in other cases as well. But processing cost is only one side of the relevance ratio. The other side is cognitive effects. We argued in chapter 7 that the only cognitive effects that are significant in a stipulated interpretive context like those we will be examining are contextual implications. As we saw in chapter 5, a set of assumptions P contextually implies an assumption Q in a context C if and only if: (i) the union of P and C non-trivially implies Q, (ii) P does not non-trivially imply Q, and (iii) C does not non-trivially imply Q. Non-trivially implication, remember, is implication that proceeds solely via the use of elimination rules.

With this theoretical machinery in hand, we should now be able to use our relevance theoretic account of definite descriptions to generate predictions as to the natural interpretations of utterances in contrived cases. We will begin with a standard example of an attributively used description in §8.2. We will look at a standard example of a referentially used description in §8.3, and in §8.4 we will look at a generic kind-referring use of a description.

8.2 The Attributive Case

Consider the following case: Let us suppose that it is mutually manifest to both Lucy and Elsa that they are typical adults. Let us further suppose that it is mutually manifest that both Lucy and Elsa are acquainted with Smith. Now, Lucy and Elsa happen upon Smith’s horribly
mutilated corpse. Let us further suppose that it is mutually manifest that Smith has been murdered by a single individual. Upon discovering this scene, Elsa utters (8.1). Let us assume that Elsa’s intonation is that typically used for assertions. Finally, let us suppose that there is no individual known to Lucy and Elsa such that it is mutually manifest to Lucy and Elsa that this individual is responsible for Smith’s death.

(8.1) Smith’s murderer is insane.\(^{185}\)

We will make the following assumptions about the case:

A. The natural interpretation of the definite description in Elsa’s utterance here is the attributive reading.

B. A reader’s intuition about the natural interpretation of the description in this case is the result of interpreting the utterance in an interpretive context.

C. Given our definition of an interpretive context, assumption B about the case, and the stipulated context for (8.1), the interpretive context for Elsa’s utterance of (8.1) includes:
   * The assumption that Lucy and Elsa are adults.
   * The assumption that Elsa is making an assertion.
   * The assumption that Smith has been murdered by a single individual.
   * All assumptions which we attribute to adults generally, unless the interpreter has good reason on balance to believe that they are not mutually manifest to Lucy and Elsa.

D. The stipulated context for Elsa’s utterance of (8.1) is the only evidence available to the interpreter in interpreting the utterance.

E. The set of assumptions that the interpreter has good reason to believe Elsa would judge to be available to Lucy in is the same as the set of assumptions the interpreter has good reason to believe are available to Lucy in this case.

F. The accessibility ordering the interpreter has reason to believe Elsa would judge to be the Lucy’s ordering is the same as the accessibility ordering that the interpreter has reason to believe is Lucy’s ordering.

G. Per (ORIC), the optimally relevant interpretation in an interpretive context is the most relevant of those which the interpreter has reason to believe the speaker would judge to be available to the hearer. Given assumption E about the case, this means that the optimally relevant interpretation of (8.1) will be the most relevant of those the interpreter has reason to believe are available to Lucy.

\(^{185}\) Following others in the literature on descriptions, I am assuming that ‘Smith’s murderer’ linguistically encodes the same content linguistically encoded by ‘the murderer of Smith’. Those inclined to reject this assumption should change the example utterance to ‘The murderer of Smith is insane’.

256
As noted in assumption A, the natural interpretation of the definite description in Elsa’s utterance is the attributive reading. To show that the attributive reading is optimally relevant in the interpretive context – that is, to show that it is the most relevant of those the interpreter has reason to believe are available to Lucy – we will use the strategy outlined in chapter 7. This will require completing steps (I)-(IV).

(I) Determine which interpretations of the description the interpreter has reason to believe are available to Lucy. Show that the attributive interpretation is one of them.

(II) Show that interpretation of the description is independent of other elements of the interpretation of Elsa’s utterance.

(III) Show that the attributive interpretation has the same or greater cognitive effects than the other available interpretations from (I).

(IV) Show that the attributive interpretation requires the least amount of processing cost of the interpretations from (I).

We will attempt to complete these steps in turn.

8.2.1 Step I – Available Enrichments

The first step is to determine which interpretations of the description the interpreter has good reason on balance to believe are available to the hearer and show that the attributive interpretation is one of them. For an interpretation of an utterance to be available, remember, all of the concepts serving as enrichments of components of the LF of the utterance must be available as enrichments of whatever encoded content they serve to enrich. To be available as an enrichment, a concept must satisfy both clauses of (CA). Thus, an interpreter has a good reason on balance to believe that a concept is available to the hearer as an enrichment of encoded content if and only if he has good reason to believe that the conditions in (CA) are satisfied for the hearer. According to our account of descriptions, there are three types of individual concept
that may be available: *de re* concepts, descriptive concepts, and type concepts. The attributive interpretation arises out of the selection of a descriptive concept.

Clause (1) of (CA) tells us that for a concept to be available as an enrichment of the content encoded by the description, it must be mutually manifest to Lucy and Elsa either that the concept is retrievable for Lucy, or that *ad hoc* concept construction has been triggered and the concept is derivable for Lucy. Given the conditions for triggering *ad hoc* concept construction laid out in (ACC), we will not be able to determine whether *ad hoc* concept construction has been triggered until have first determined which individual concepts, if any, are retrievable without it and whether or not the interpretations involving those concepts have relevance. We will begin, then, by attempting to ascertain if there are any retrievable *de re*, descriptive, or type concepts which satisfy the second condition on concept availability laid out in clause (2) of (CA) prior to any *ad hoc* concept construction.

Let us begin with *de re* concepts. The interpreter lacks a good reason on balance to believe that a *de re* concept satisfying clause (2) of (CA) is retrievable for Lucy prior to any *ad hoc* concept construction. Clause (2) of (CA) tells us that for a concept to be available, it must be mutually manifest that the interpretation of the description as that concept is compatible with the content of the description. As indicated by clause (2) of (COM), in the case of a *de re* concept this means that it must be mutually manifest that the referent of the concept is uniquely identified by the encoded content or content derived from it via narrowing or broadening. Since we are currently looking at the availability of concepts in the absence of *ad hoc* concept construction, if a *de re* concept is to be available, the referent of the *de re* concept must be uniquely identified by the conceptual content encoded by the description. The encoded conceptual content in this case
is MURDERER OF SMITH. Since there is no individual known to Lucy and Elsa such that it is mutually manifest to Lucy and Elsa that this individual is the murderer, however, there are no de re concepts such that the interpreter has reason to believe it is mutually manifest to Lucy and Elsa that their referents are uniquely identified by MURDERER OF SMITH. Thus, the interpreter lacks good reason to believe that a de re concept is retrievable for Lucy in the absence of ad hoc concept construction in this case.

Let us next consider whether the interpreter has good reason on balance to believe that a type concept which satisfies clause (2) of (CA) is retrievable for Lucy prior to any ad hoc concept construction. For a concept to satisfy clause (2) of (CA) prior to ad hoc concept construction in this case, the interpretation of the description as that concept would have to be compatible with the conceptual content encoded by the description – MURDERER OF SMITH.

According to clause (3) of (COM), for the interpretation of a description as a type concept to be compatible with the encoded content, it must be the case that (a) the extension of the encoded conceptual content or of content derived from it via conceptual broadening or narrowing contains all and only members of the kind which is the referent of the type concept, and (b) the referent of the concept is a kind. Thus, for a type concept to be compatible with the content encoded by the description in Elsa’s utterance, the referent of the concept would have to be a kind, and the extension of MURDERER OF SMITH would have to contain all and only members of that kind. But

---

186 We will continue in this chapter to rely on the conventions established in previous chapters for identifying different concepts and conceptual structures. Small caps are used for concepts. Phrases in small caps are used for strings of syntactically structured concepts. Dashes are used for descriptive concepts, and ‘\(\Phi\)’ indicates a type concept referring to the kind \(\Phi\). So, ‘MURDERER OF SMITH’ picks out a string of concepts structured syntactically, ‘MURDERER-OF-SMITH’ picks out the descriptive concept which refers to whoever is the murderer of Smith, and ‘WHALE’ picks out the type concept whose referent is the kind whale.
it is not clear that there is a kind murderer of Smith or that there would be in the circumstances described.

Unlike panda or doctor or tax collector, for example, murderer of Smith is not intuitively a kind. Unfortunately, the intuitive difference between murderer of Smith and classes like tax collector which are intuitively kinds may be all we have to go on when it comes to the question of whether or not there is a kind murderer of Smith. While there has been a good deal of work done on the nature of natural kinds, far less attention has been given to the nature of kinds broadly speaking, and what makes a set of objects a kind in general still remains something of a mystery. If murderer of Smith is not a kind, then there will be no type concept compatible with the content encoded by the description in Elsa’s utterance, as clause (3) of (COM) requires that the referent of such a concept be a kind.

Of course, since what makes something a kind is currently unclear, it may turn out that murderer of Smith does in fact constitute a kind. But this is not enough. Given clause (2) of (CA), if a type concept is to be available to Lucy as an enrichment of the content encoded by the description, it must be mutually manifest to Elsa and Lucy that there is a kind murderer of Smith which is the referent of that type concept. As we noted above, however, murderer of Smith is not intuitively a kind. Given this, it seems highly unlikely that we would attribute to adults generally the assumption that there is a kind murderer of Smith. Since there is nothing else in the stipulated context which would give the interpreter reason to believe that it is mutually manifest to Lucy and Elsa that murderer of Smith is a kind, the interpreter lacks good reason to believe that there

---

187 It is important to note here that the question is not whether we can conceive of a world in which murderer of Smith would plausibly be treated as a kind, but rather whether it is plausible to consider it a kind in the scenario surrounding the example utterance.
is a type concept which satisfies clause (2) of (CA). The interpreter thus lacks good reason to believe that there is a type concept available to Lucy as an enrichment of the content encoded by the description in Elsa’s utterance of (8.1) prior to any *ad hoc* concept construction.

Let us now turn to the question of whether a descriptive concept is available prior to *ad hoc* concept construction. In order for a descriptive concept to satisfy clause (2) of (CA), an interpretation of the description as that concept would have to be compatible with the encoded content MURDERER OF SMITH. Clause (1) of (COM) tells us that for an interpretation of a description as a descriptive concept to be compatible with the content encoded by that description, the encoded conceptual content or content derived from it via conceptual broadening or narrowing must constitute the descriptive content of the concept. For a descriptive concept to satisfy clause (1) of (COM) prior to *ad hoc* concept construction in this case, then, the encoded content MURDERER OF SMITH must constitute the descriptive content of the concept. It will be manifest to anyone who possesses or is capable of acquiring the concept MURDERER-OF-SMITH that it satisfies this condition and is the only descriptive concept to satisfy it. Since we attributive to adults generally the assumption that typical adults are capable of acquiring concepts like MURDERER-OF-SMITH, the interpreter has reason to believe it is mutually manifest to Lucy and Elsa that the descriptive concept MURDERER-OF-SMITH satisfies clause (2) of (CA) prior to any *ad hoc* concept construction, and is the only descriptive concept to do so. The question, then, is whether or not the interpreter has good reason to believe that it is mutually manifest that this concept is retrievable for Lucy. It seems to me that he or she does not.

Before diving into the question of whether or not the interpreter has good reason to believe it is mutually manifest to Elsa and Lucy that the descriptive concept MURDERER-OF-
SMITH is retrievable for Lucy at the time of Elsa’s utterance, there are two things we should note. The first is that if the descriptive concept is in fact retrievable, then it would be the selected interpretation of the description. According to (ACC), ad hoc concept construction will only be triggered if one of the following holds: (a) no unconstructed interpretation of the utterance is available, (b) all available unconstructed interpretations of the utterance fail to have relevance, or (c) there are specific expectations of relevance which the available unconstructed interpretations fail to satisfy. The available interpretation of the description in Elsa’s utterance is unconstructed, and there would be an unconstructed interpretation of the predicate available as well.

Furthermore, the utterance interpretation involving the selection of MURDERER-OF-SMITH in this case would yield cognitive effects (Smith’s murderer may not be competent to stand trial, for example). As a result, the interpretation would have relevance. Since there do not appear to be any specific expectations of relevance in the case, none of the three conditions for triggering ad hoc concept construction given in (ACC) would be satisfied. The descriptive concept would thus be the only available interpretation of the description, and by default the optimally relevant one.

Since the attributive reading of the description results from the selection of a descriptive concept, this means that if the descriptive concept were retrievable for Lucy in this case, our relevance theoretic account would accurately predict an attributive reading of the description in Elsa’s utterance.

The second thing to keep in mind here is that, as we have seen in previous chapters, the content encoded by the description – MURDERER OF SMITH – is a distinct cognitive entity from the descriptive concept MURDERER-OF-SMITH. An individual who has mentally represented the former has not necessarily constructed the latter. The encoded conceptual content MURDERER OF
SMITH is a syntactically structured string of concepts. This same structured string of concepts would be the conceptual content encoded by the indefinite description ‘a murderer of Smith’. The descriptive concept MURDERER-OF-SMITH, on the other hand, is a single individual concept, taken by its possessor to refer to an individual.

With this in mind, let us consider whether the interpreter has reason to believe it is mutually manifest to Elsa and Lucy that the descriptive concept MURDERER-OF-SMITH is retrievable for Lucy. For this concept to be retrievable for Lucy at the time of Elsa’s utterance, it must be in her mental lexicon at that time. In other words, it must be the case that Lucy had constructed the concept at some point prior to Elsa’s utterance. Now, one might think that the mere fact that the stipulated context tells us that it is mutually manifest to Lucy and Elsa that Smith has been murdered by a single individual would give the interpreter reason to believe that it is mutually manifest that Lucy has already constructed the descriptive concept MURDERER-OF-SMITH. After all, it may seem reasonable on the face of things to assume that as soon as someone learns that there is some individual they do not already possess a concept of who can be uniquely identified with a certain descriptive content, he or she automatically constructs a descriptive concept of that individual. There are two problems with this line of thought, however. First, the stipulated context tells us merely that the assumption that Smith has been murdered by a single individual is manifest. This is not the same as saying that it is believed or even mentally represented. An assumption is manifest to an individual, remember, when that individual is capable of mentally representing it an accepting it to be true or probably true. An assumption can be manifest to an individual at a time, then, even if that individual has not yet mentally represented the assumption at that time. If at the time of Elsa’s utterance Lucy had not yet
mentally represented the assumption that Smith has been murdered by a single individual or some similar assumption which represents the same fact, she would not have constructed the descriptive concept MURDERER-OF-SMITH either.

Of course, one might be inclined to argue that while it may be possible that at the time of Elsa’s utterance Lucy had not yet mentally represented the assumption that Smith has been murdered by a single individual or constructed a concept of Smith’s murderer, it would nonetheless be typical for the cognitive system of adult like Lucy to mentally represent the assumption and construct the descriptive concept in the circumstances described. If it would indeed be typical for an adult’s cognitive system to construct the assumption and concept in such circumstances, the argument goes, then the interpreter has good reason to believe it is mutually manifest to Lucy and Elsa that Lucy has done so.

This argument ultimately relies on the claim that we attribute to adults generally the assumption that an adult in the circumstances described in the case would form a concept of Smith’s murderer (which, because the murderer is unknown, would have to be a descriptive concept). It is this claim that justifies the conclusion that it would be mutually manifest to Lucy and Elsa that Lucy would have constructed a descriptive concept of Smith’s murderer in this case. It is not clear that we do attribute to adults generally the assumption that an adult in the described circumstances would form a concept of Smith’s murderer, however. It is, perhaps, plausible to claim that we would generally assume that an adult in the circumstances would form a concept of Smith’s murder, or perhaps Smith’s death, but it is less certain that we assume that he or she would form a concept of the murderer simply in virtue of being at the scene of the crime. Many of the new assumptions which we would expect to be acquired in such
circumstances could be stored in the encyclopedic entry of the concept MURDER-OF-SMITH or even the encyclopedic entry of the de re concept SMITH without requiring the additional outlay of cognitive processing effort involved in constructing MURDERER-OF-SMITH. Ultimately, it seems to me that whether or not an individual in Lucy’s position would form a concept of Smith’s murderer simply in virtue of being at the scene of the crime depends on the direction the individual’s thoughts take in those circumstances. Some individuals’ thoughts would immediately turn to the murderer. Some individuals, however, would think solely of Smith and the consequences of his death, and never think of the murderer. Still others might be too psychologically traumatized by the scene to have much in the way of rational thought at all. Different adults will have different cognitive reactions differently to such circumstances. And while we do generally assume that certain sorts of adults – detectives, for example – would very quickly form a concept of the murderer in such circumstances, it does not seem to me that we assume that a typical adult would. If this is correct, then the stipulated context does not provide the interpreter would good reason to believe it is mutually manifest that the descriptive concept MURDERER-OF-SMITH is retrievable for Lucy.

If the foregoing is correct, then the interpreter lacks good reason to believe that any de re, type, or descriptive concept is available to Lucy as an enrichment of the content encoded by the description in Elsa’s utterance prior to any ad hoc concept construction. (ACC) tells us that if no unconstructed interpretation of an utterance is available, ad hoc concept construction is triggered. The task facing us now is to discover which, if any, de re, type, or descriptive concepts are available with the aid of ad hoc concept construction.
As it turns out, there are no *de re* concepts which the interpreter has reason to believe are available to Lucy as enrichments of the description’s encoded content, even with the aid of *ad hoc* concept construction. According to clause (2) of (COM), remember, for a *de re* concept interpretation of encoded content to be compatible with that content, that encoded content or content derived from it via conceptual narrowing or broadening must uniquely identify the referent of the concept. But the stipulated context does not give the interpreter reason to believe that any narrowing or broadening of *MURDERER OF SMITH* would uniquely identify the referent of any *de re* concept which is mutually manifestly retrievable or derivable for Lucy. Conceptual narrowing would serve to make the content more specific – for example, narrowing *MURDERER* to an *ad hoc* concept picking out only axe murderers. But since there is no individual known to Lucy and Elsa such that it is mutually manifest that this individual is responsible for Smith’s death, greater specificity will not be of any use in yielding the unique identification of a referent of a *de re* concept which manifestly retrievable for Lucy. And since the stipulated context tells us nothing about what is mutually manifest to Lucy and Elsa about Smith’s life, a broadening of the encoded content will not help either. *MURDERER* might be broadened to, say, include those who have dealt their victim some crushing defeat. But the stipulated context gives the interpreter no reason to believe that there are any *de re* concepts retrievable or derivable for Lucy for which it is mutually manifest that their referents would uniquely satisfy any such broadened content. It would seem, then, that the interpreter lacks good reason to believe that there are any *de re* concepts available as enrichments of the content encoded by the description in this case.

So, what about type concepts? Per clause (2) of (CA), remember, for a concept to be available as an enrichment of the content encoded by a description, it must be mutually manifest
that the interpretation of the description as that concept is compatible with the encoded content. Given clause (3) of (COM), this means that for a type concept to be available as an enrichment in this case, it must be mutually manifest to Lucy and Elsa that the referent of the concept is a kind and that the extension of MURDERER OF SMITH or content derived from it via conceptual narrowing or broadening contains all and only members of that kind. As we saw earlier, since murderer of Smith is not intuitively a kind, the interpreter does not have reason to believe that there is a type concept for which it is mutually manifest to Lucy and Elsa that the referent of the concept is the kind murderer of Smith. It seems likely that a similar problem will arise for any content derived from MURDERER OF SMITH via conceptual narrowing or broadening. Axe murderer of Smith or brutal defeater of Smith are no more intuitively kinds than murderer of Smith is. In light of this, it is highly unlikely that we would attribute to adults generally the assumption that these are kinds. In the absence of any such assumptions attributed to adults generally, there is nothing in the stipulated context that would give the interpreter reason to believe that it is mutually manifest to Lucy and Elsa that there are kinds like axe murderer of Smith. Since it must be mutually manifest to the conversational participants that the referent of the concept is a kind if a type concept is to be available as an enrichment, this means that the interpreter lacks good reason to believe any type concepts are available as enrichments of the content encoded by the description in this case, with or without ad hoc concept construction.¹⁸⁸

Let us now turn to the issue of which, if any, descriptive concepts are available as enrichments of the content encoded by the description in this case. To be available, a concept

---

¹⁸⁸ If type concepts were available here, per clause (5) of (PIC), they would nonetheless require greater processing cost than the available descriptive concepts. As a result, a descriptive concept, and not a type concept, would be selected provisionally as interpretation of the utterance was ongoing. And it seems unlikely that there would be any reason for the interpreter to revise this selection after the predicate was interpreted.
must satisfy both conditions laid out in (CA). Clause (2) of (CA) tells us that in order for a concept to be available as an enrichment of encoded content, it must be mutually manifest to the conversational participants that interpretation of the description as that concept is compatible with that encoded content. For a descriptive concept to be available, then, it must be mutually manifest to the conversational participants that the concept satisfies clause (1) of (COM). Clause (1a) of (COM) tells us that interpretation of a description as a descriptive concept is compatible with encoded content when the descriptive content of the concept is constituted by the encoded conceptual content or content derived from it via conceptual broadening or narrowing. This condition is very clearly satisfied by the descriptive concept MURDERER-OF-SMITH. It is also clearly satisfied by any descriptive concept whose descriptive content is constituted by content derived from MURDERER OF SMITH via conceptual narrowing or broadening. The interpreter thus has good reason to believe that it is mutually manifest to Lucy and Elsa that clause (1a) of (COM) is satisfied by MURDERER-OF-SMITH and any descriptive concept whose descriptive content is a narrowing or broadening of MURDERER OF SMITH. Clause (1b) of (COM) tells us that the referent of the concept, if there is one, must uniquely satisfy the descriptive content of the concept. For MURDERER-OF-SMITH and descriptive concepts resulting from broadening or narrowing of the encoded concept MURDERER OF SMITH to be available, then, it must be mutually manifest to Lucy and Elsa that the referent of the concept uniquely satisfies the concept’s descriptive content. Since the stipulated context tells us that it is mutually manifest to Lucy and Elsa that Smith has been murdered by a single individual, the interpreter has good reason to believe that MURDERER-OF-SMITH satisfies clause (1b) of (COM). Our next task, then, will be to determine if any descriptive concepts resulting from narrowing or broadening of the encoded
content MURDERER-OF-SMITH are such that the interpreter has good reason to believe it is mutually manifest to Lucy and Elsa that they satisfy clause (1b) of (COM).

For the sake of concision, let us use ‘altered descriptive concepts’ to refer to descriptive concepts whose contents are constituted by the result of narrowing or broadening the encoded content MURDERER OF SMITH. In order to satisfy clause (1b) of (COM), an altered descriptive concept would have to be such that the referent of the concept uniquely satisfies the content of the concept. Now, no matter how the encoded content MURDERER OF SMITH is narrowed or broadened, the resulting content will still pick out an individual who has brought about some harm, defeat, or overwhelming of Smith or his interests. Any altered concept which satisfies clause (1b) of (COM), then, will have a referent which can be uniquely identified as having brought about some harm, defeat or overwhelming of Smith or his interests. Per clause (2) of (CA), if a concept is to be available as an enrichment, it must be mutually manifest to the conversational participants that (COM) is satisfied. If an altered descriptive concept is to be available as an enrichment here, then, it must be mutually manifest to Lucy and Elsa that the referent of the concept can be uniquely identified as having brought about some harm, defeat or overwhelming of Smith or his interests. There is only one individual who is such that the stipulated context could potentially give us reason to believe it is mutually manifest to Lucy and Elsa that this individual can be uniquely identified as having brought about some harm, defeat or overwhelming of Smith or his interests – the individual who murdered Smith. The only altered descriptive concepts which the interpreter has reason to believe are available as enrichments of the description, then, will be such that it is mutually manifest that Smith’s murderer is their referent.
If the foregoing is correct, then the interpreter has good reason to believe that condition (2) of (CA) is satisfied in this case by the descriptive concept MURDERER-OF-SMITH as well as any altered descriptive concepts such that it is mutually manifest that their descriptive contents are uniquely satisfied by whoever murdered Smith. These descriptive concepts will be available as enrichments of the content encoded by the description in Elsa’s utterance, then, if the stipulated context also gives the interpreter reason to believe that they satisfy clause (1) of (CA). We argued earlier that *ad hoc* concept construction has been triggered in this case. Given this, one of these descriptive concepts will satisfy clause (1) of (CA) as long as it is mutually manifest to Lucy and Elsa that the concept is either retrievable or derivable.

We saw earlier that the stipulated context does not provide the interpreter with good reason to believe that it is mutually manifest to Lucy and Elsa that the descriptive concept MURDERER-OF-SMITH is retrievable for Lucy. In fact, we concluded in that earlier discussion that we do not attribute to adults generally the assumption that someone in Lucy’s circumstances would form a concept of Smith’s murderer. Such an assumption would be the sole basis for concluding that it is mutually manifest to Lucy and Elsa that a typical adult in Lucy’s circumstances would construct any concept of Smith’s murderer – including an altered descriptive concept of Smith’s murderer. In the absence of this assumption, the stipulated context fails to give the interpreter good reason for believing it is mutually manifest to Lucy and Elsa that the altered descriptive concepts referring to Smith’s murderer are retrievable for Lucy. It may not be mutually manifest to Lucy and Elsa that MURDERER-OF-SMITH or the altered descriptive concepts referring to Smith’s murderer are retrievable for Lucy, but the interpreter does have good reason to believe they are derivable. The stipulated context tells us that it is
mutually manifest to Lucy and Elsa that they are acquainted with Smith, and we attribute to adults generally the assumption that typical adults are in possession of concepts like MURDERER. Since we also attribute to adults generally the assumption that typical adults are capable of the cognitive processes involved in utterance interpretation – which, on our account, includes the derivation of concepts via composition or narrowing and broadening – the interpreter has good reason to believe that it is mutually manifest to Lucy and Elsa that Lucy can derive the descriptive concept MURDERER-OF-SMITH and its associated altered descriptive concepts. These concepts thus satisfy clause (1) of (CA).

If the foregoing is correct, then the interpreter has reason to believe that the following descriptive concepts are available to Lucy as enrichments of the content encoded by the description in Elsa’s utterance: the descriptive concept MURDERER-OF-SMITH and any altered descriptive concepts which refer to Smith’s murderer and for which it is mutually manifest that the Smith’s murderer satisfies the content of the concept. Intuitively, the natural interpretation of the description in Elsa’s utterance is the basic attributive interpretation, which would result from the selection of the descriptive concept MURDERER-OF-SMITH. It may be worth noting at this point, however, that all of the concepts which are available as enrichments of the content encoded by the description are descriptive concepts. According to our relevance theoretic account of definite descriptions, the selection of a descriptive concept yields an attributive reading of the description. No matter which of the available concepts yields the optimally relevant interpretation, then, our account will predict that the natural interpretation of the description is an attributive one. Before we can tackle task of comparing the relevance of the different available descriptive concepts, however, we will first need to show that the
interpretation of the description is the only aspect of the interpretation of Elsa’s utterance we need to focus on when comparing the relevance of available interpretations. This is the goal of the next section.

8.2.2 Step II – Independence of the Description

According to relevance theory, remember, utterance interpretation is an online process. Provisional interpretations are assigned as the utterance is ongoing, and can be revised later in light of considerations of relevance. Since the description occurs at the very beginning of Elsa’s utterance, then, an individual concept would be selected as a provisional interpretation of the description as the utterance was ongoing. This provisional interpretation could be revised in light of considerations of relevance later on. One might think that this would force any theorist attempting to predict the natural interpretation of a description to give an assessment of the relevance of all the available full interpretations of the utterance in which the description occurs. Doing this would require, at the very least, the machinery to compare the processing cost involved in interpreting simple predicates like ‘is insane’ or ‘is bald’ – machinery which we do not currently possess. As we saw in chapter 7, however, as long as the interpretation of the description is independent of the other elements of the interpretation of the utterance in the case it question, it should be possible to predict the natural interpretation of the description without assessing the relevance of other aspects of the interpretation of the utterance. The task of this section, then, is to show that the interpretation of the description ‘Smith’s murderer’ is independent of the other elements of the interpretation of Elsa’s utterance of (8.1).

According to (IA), an available interpretation of an utterance will contain only assumptions which fit the definition of a propositional form, a higher-level explicature, or an
implicature. An implicature remember, is a contextual assumption or implication for which it is manifest that the speaker intended to make them manifest. As far as I can tell, there is nothing in the stipulated context which indicates that it is manifest that Elsa intended to make manifest any contextual assumption or implication. If this is correct, then the interpreter lacks good reason on balance to believe that there are available interpretations containing implicatures which might impact the interpretation of the description.

One might contend, however, that by uttering (8.1) Elsa is in fact implicating that Smith’s murderer should be apprehended and put away. This is certainly a contextual implication of Elsa’s utterance, given the assumptions we attribute to adults generally about insane murderers who are on the loose. It is not clear to me what in the stipulated context would give us reason to believe that that it is manifest that Elsa intended to make this contextual assumption manifest. Nonetheless, if this implicature is in fact generated by Elsa’s utterance of (8.1), it would be generated independently of which available individual concept is selected to enrich the content encoded by the description. All of the available individual concepts, remember, refer to Smith’s murderer. So, no matter how the description is interpreted, the propositional form of the utterance will be a representation of the assumption that Smith’s murderer is insane, and together with assumptions we attribute to adults generally about how insane murderers should be handled, it will yield the contextual implication that Smith’s murderer should be apprehended and put away.

So, the interpretation of the description is independent of any implicatures which might be part of the interpretation of Elsa’s utterance of (8.1). What about higher-level explicatures? A higher-level explicature, remember, is derived from the propositional form of the utterance
together with mutually manifest assumptions about the speaker’s propositional attitudes with respect to the utterance and its content. The context in this case provides no basis for attributing to Elsa propositional attitudes toward the propositional form of her utterance other than those associated with the fact that Elsa’s utterance of (8.1) is an assertion. Her utterance will be an assertion regardless of whether ‘Smith’s murderer’ is given one reading rather than another, however, and the attitudes which might be ascribed to Elsa with respect to the propositional form of her utterance will be the same no matter which descriptive concept of Smith’s murderer is selected. So which higher-level explicatures end up included in the interpretation of the utterance is independent of which interpretation the description is given.\footnote{There is a certain sense in which the higher-level explicatures which end up in the interpretation of Elsa’s utterance will differ depending on which interpretation of the description is selected. For example, suppose that one of the higher-level explicatures which ends up in the full interpretation of the utterance is one which attributes to Elsa a belief in the propositional form of the utterance. If MURDERER-OF-SMITH is selected as the interpretation of the description, this higher-level explication would thus be ELSA BELIEVES THAT MURDERER-OF-SMITH IS INSANE. If, on the other hand, an altered descriptive concept involving a narrowing of MURDERER is selected, then the higher-level explication in question will be ELSA BELIEVES THAT MURDERER*-OF-SMITH IS INSANE. This difference need not concern us, however. The only way in which the differences between these higher-level explicatures might impact which interpretation of the description is selected would be by making available different cognitive effects. We will argue in the next section that any cognitive effect derivable from an interpretation involving an altered descriptive concept will also be derivable from the interpretation involving the basic descriptive concept MURDERER-OF-SMITH.} 

Finally, let us consider whether the description is independent of the other aspects of the propositional form of Elsa’s utterance of (8.1). The propositional form of the utterance in this case consists in the interpretation of the description in the subject slot and the interpretation of the predicate. The interpretation of the predicate ‘is insane’ would appear to be the same no matter how the description is interpreted. The predicate ‘is insane’ does admit of a number of different interpretations arising from a broadening or narrowing of extension. It could mean, for example, is legally insane or is mentally unstable or is highly irrational. But the selection of one
of these interpretations rather than another would not appear to have any potential impact on which of the available interpretations of the description is selected.

It would seem, then, that interpretation of the description in Elsa’s utterance of (8.1) is independent of the interpretation of the other elements of the propositional form of the utterance. Since it is also independent of both the selection of higher-level explicatures and the selection of implicatures, it is independent of the interpretation of all of the other elements of the interpretation of Elsa’s utterance. In light of this, we can focus solely on the interpretation of the description when assessing the relevance of different available interpretations. To show that the basic attributive interpretation arising out of the selection of MURDERER-OF-SMITH is the most relevant of the available interpretations using the strategy outlined in chapter 7, we still need to do two things. First, we must show that the basic attributive interpretation yields the same or greater cognitive effects than the other available interpretations. Then we must show that the basic attributive interpretation requires the least processing cost. The issue of cognitive effects is the topic of §8.2.3, while the processing costs of the available interpretations are compared in §8.2.4.

**8.2.3 Step III – Cognitive Effects**

We concluded in §8.2.1 that the following concepts are available as enrichments of the content encoded by the description in Elsa’s utterance: the descriptive concept MURDERER-OF-SMITH and any altered descriptive concepts which refer to Smith’s murderer and for which it is mutually manifest that the Smith’s murderer satisfies the content of the concept. The goal of this section is to show that the cognitive effects resulting from the interpretation of Elsa’s utterance involving selection of the descriptive concept MURDERER-OF-SMITH would be greater than or
equal to the cognitive effects resulting from the other available interpretations. We argued in chapter 7 that the only cognitive effects that are significant in a stipulated interpretive context like that in which the reader is interpreting Elsa’s utterance of (8.1) are contextual implications. As we saw in chapter 5, a set of assumptions P contextually implies an assumption Q in a context C if and only if: (i) the union of P and C non-trivially implies Q, (ii) P does not non-trivially imply Q, and (iii) C does not non-trivially imply Q.

Now, as we saw in §8.2.1 all of the altered descriptive concepts available as enrichments in this case are such that it is mutually manifest to Lucy and Elsa that their referent is Smith’s murderer. For a descriptive concept to be available as an enrichment, remember, it must be the case that it is mutually manifest that the referent uniquely satisfies the concept’s descriptive content. So, for any altered descriptive concept with descriptive content A which is available as an enrichment of the content encoded by the description it will be mutually manifest to Lucy and Elsa that Smith’s murderer satisfies A. Thus, any contextual implication which can be derived from an interpretation involving an altered descriptive concept will also be derivable from the interpretation involving the basic descriptive concept MURDERER-OF-SMITH. Since contextual implications are the only cognitive effects which are significant, this means that the basic attributive interpretation resulting from the selection of the descriptive concept MURDERER-OF-SMITH will have at least as many cognitive effects as the other available interpretations. Given this, if we can show that the selection of the descriptive concept MURDERER-OF-SMITH requires less processing cost than the other concepts which are available as enrichments, we will have shown that the basic attributive interpretation of the description in Elsa’s utterance is the
optimally relevant interpretation. The next section is dedicated to comparing the processing costs of the various available enrichments of the content encoded by the description.

8.2.4 Step IV—Processing Costs

The concepts available as enrichments of the content encoded by the description in Elsa’s utterance of (8.1), remember, are the descriptive concept MURDERER-OF-SMITH and any altered descriptive concepts for which it is mutually manifest that Smith’s murderer uniquely satisfies their descriptive content. We argued in §8.2.1 that none of these concepts are retrievable. The processing cost associated with these concepts, then, will be whatever cost is required to derive them.

The cost of deriving a concept depends, of course, on what is involved in the derivation. All that would be involved in deriving the basic descriptive concept MURDERER-OF-SMITH in this case would be composing the concept. The same cannot be said for the altered descriptive concepts, however. Those concepts, remember, are ad hoc descriptive concepts whose descriptive content is constituted by narrowed or broadened versions of the encoded content. As descriptive concepts, their derivation will involve composition, but it will also involve conceptual narrowing or broadening. (PIC) details our assumptions about the processing costs associated with retrieving and deriving individual concepts, and according to clause (4) of (PIC), other things equal, constructing a concept via both composition and conceptual narrowing or broadening requires greater processing cost than constructing a concept via only composition. Since derivation of the altered descriptive concepts involves both narrowing or broadening and composition, while the derivation of MURDERER-OF-SMITH involves only composition, the altered descriptive concepts require greater processing cost than does MURDERER-OF-SMITH. Since the
altered descriptive concepts and MURDERER-OF-SMITH are the only concepts available as enrichments, this means that MURDERER-OF-SMITH requires the least amount of processing cost of the available enrichments of the content encoded by the description. And since the basic attributive interpretation results from the selection of MURDERER-OF-SMITH, this interpretation requires the least amount of processing cost of those the interpreter has reason to believe are available to the hearer.

8.2.5 Conclusion

The natural interpretation of the description in Elsa’s utterance of (8.1) is a basic attributive one. The goal of this section has been to show that this is the natural interpretation predicted by our relevance theoretic account of definite descriptions. In §8.2.1 we showed that the interpretations of the description in Elsa’s utterance of (8.1) which the interpreter has reason to believe are available to the hearer are a basic attributive interpretation resulting from the selection of the descriptive concept MURDERER-OF-SMITH and attributive interpretations which result from the selection of a limited set of altered descriptive concepts. In §8.2.2 we showed that interpretation of the description is independent of the rest of the elements in the interpretation of Elsa’s utterance. This made it possible for us to focus solely on the interpretation of the description when comparing relevance. In §8.2.3 we showed that the basic attributive interpretation yields at least the same cognitive effects as the other available interpretations. Since relevance is a ratio of cognitive effects to processing cost, this means that if the basic attributive interpretation has lower processing cost than the other interpretations the interpreter has reason to believe are available, it will be the most relevant of the interpretations the interpreter has reason to believe are available. In §8.2.4 we showed that the basic attributive
interpretation resulting from the selection of the descriptive concept MURDERER-OF-SMITH requires less processing cost than the other available interpretations. According to assumption G about the case, the optimally relevant interpretation of (8.1) in the interpretive context will be the most relevant of those the interpreter has reason to believe are available to Lucy. The basic attributive interpretation it is the optimally relevant interpretation of Elsa’s utterance of (8.1) in the interpretive context. The prediction principle laid out in (PP) tells us that in an interpretive context, the natural interpretation of an utterance is the optimally relevant one. Since the reader is interpreting Elsa’s utterance in an interpretive context, our account thus predicts that the reader’s natural interpretation of the description in Elsa’s utterance of (8.1) is the basic attributive interpretation. This prediction lines up with our intuitive reading of (8.1).

8.3 The Referential Case

Let us assume that it is mutually manifest to Gray and Leo that they are both typical adults. Further assume that the following assumptions are mutually manifest: Both Gray and Leo know who Jones is, Smith was murdered by a single individual, and Jones murdered Smith. A brief news spot comes on and the news anchor states that Jones has engaged in some behavior which is very bizarre. In reaction to this, Leo utters (8.2). Suppose that the tone of Leo’s utterance makes it manifest that he is making an assertion.

(8.2) Smith’s murderer is insane.

We will make the following assumptions about the case:

A. The natural reading of the definite description in Leo’s utterance of (8.2) here is a referential reading.
B. A reader’s intuition about the natural reading of the definite description in this case is the result of interpretation in an interpretive context.

C. Given our definition of an interpretive context, assumption B in the case, and the stipulated context for (8.2), the interpretive context for Leo’s utterance in (8.2) includes the following assumptions:
   * Leo and Gray are typical adults.
   * Leo is making an assertion.
   * Smith was murdered by a single individual.
   * Jones murdered Smith.
   * All assumptions which we attribute to adults generally, unless the interpreter has good reason on balance to believe that they are not mutually manifest to Leo and Gray.

D. The stipulated context for Leo’s utterance of (8.2) is the only evidence available to the interpreter in interpreting the utterance.

E. The set of interpretations that the interpreter has good reason to believe Leo would judge to be available to Gray in is the same as the set of interpretations the interpreter has good reason to believe are available to Gray in this case.

F. The accessibility ordering the interpreter has reason to believe Leo would judge to be Gray’s ordering is the same as the accessibility ordering that the interpreter has reason to believe is Gray’s ordering.

G. Per (ORIC), the optimally relevant interpretation in an interpretive context is the most relevant of those which the interpreter has reason to believe the speaker would judge to be available to the hearer. Given assumption E about the case, this means that the optimally relevant interpretation of (8.2) will be the most relevant of those the interpreter has reason to believe are available to Gray.

As noted in assumption A, the natural reading of the description in Leo’s utterance of (8.2) is a referential reading on which the description refers to Jones. According to the account of definite descriptions we are attempting to test here, this reading results from the selection of a *de re* concept of Jones as the enrichment of the content encoded by the description during pragmatic processing. To show that the referential reading of the description is the optimally relevant reading in the interpretive context, we will use the strategy outlined in chapter 7. The first step in this strategy is to determine which interpretations of the description the interpreter has reason to believe are available to Gray. I will argue in what follows that a *de re* concept of Jones is the
only concept available as an enrichment of the content encoded by the description, and hence is
the only available interpretation of it.

Let us begin by looking at whether a *de re* concept is available to Gray as an enrichment
of the description’s encoded content prior to any *ad hoc* concept construction. In order for a *de re*
concept to satisfy both clauses of (CA) prior to any *ad hoc* concept construction, it must be
mutually manifest to Leo and Gray that the concept is retrievable for Gray and, per clause (2) of
(COM), that the referent of the concept is uniquely identified by the content encoded by the
description. Since the stipulated context tells us it is mutually manifest to Leo and Gray that they
are acquainted with Jones, the interpreter has good reason to believe that a *de re* concept of Jones
is retrievable for Gray. The stipulated context also tells us that it is mutually manifest to Leo and
Gray that Jones murdered Smith and that only one individual murdered Smith. The interpreter
thus has good reason to believe it is mutually manifest to Leo and Gray that the content encoded
by the description – *MURDERER OF SMITH* – uniquely identifies Jones, who is the referent of the
*de re* concept JONES. The stipulated context does not give the interpreter good reason to believe
the referent of any other *de re* concept is uniquely identified by this content. Thus, the *de re*
concept of JONES is the only *de re* concept the interpreter has reason to believe is available to
Gray prior to any *ad hoc* concept construction.

The referential interpretation of the description in Leo’s utterance of (8.2) resulting from
the selection of the *de re* concept of Jones would yield cognitive effects in the context – *Jones is
not normal*, for example, or *Jones may not be competent to stand trial*. The referential
interpretation thus has relevance in the context. Since there do not appear to be any specific
expectations of relevance in the context, this means that none of the conditions for triggering *ad*
hoc concept construction detailed in (ACC) would be satisfied in the case. If any other interpretations of the description in (8.2) are to be such that the interpreter has reason to believe they are available to Gray, they must be available prior to any ad hoc concept construction.

So, are there any type concepts available to Gray as enrichments of the content encoded by the description in the absence of ad hoc concept construction? It seems to be that there are not. For a concept to be available as an enrichment, remember, it must satisfy both conditions of (CA). For a type concept to satisfy clause (1) of (CA) prior to any ad hoc concept construction in this case, then, it would have to be mutually manifest to both Leo and Gray that the concept is retrievable for Gray. And, given clause (3) of (COM), for a type concept to satisfy clause (2) of (CA), it would have to be mutually manifest to both Leo and Gray that the referent of the concept is a kind whose members map onto the extension of the encoded content MURDERER OF SMITH. Now, as we saw in the attributive case, it is uncertain whether or not there even is such a kind as murderer of Smith. Given the unclear nature of kinds, however, it may ultimately be the case that there is a kind murderer of Smith. Since murderer of Smith is not intuitively a kind in the way that doctor or tax collector are, however, it is highly unlikely that we attribute to adults generally the assumption that there is such a kind. Since there is nothing else in the stipulated context which would indicate that it is mutually manifest to Leo and Gray that murderer of Smith is a kind, the interpreter does not have good reason to believe that it is mutually manifest to Leo and Gray that there is a type concept referring to such a kind which is retrievable for Gray. In other words, the interpreter does not have good reason to believe it is mutually manifest to Leo and Gray that there is a type concept which satisfies both of the conditions laid out in (CA) prior to ad hoc concept construction.
If the foregoing is correct, the interpreter lacks good reason to believe a type concept is available to Gray as an enrichment of the content encoded by the description in Leo’s utterance of (8.2). Let us turn, then, to the question of whether a descriptive concept is available prior to any ad hoc concept construction. For a descriptive concept to satisfy clause (2) of (CA) prior to ad hoc concept construction, it must be mutually manifest to Leo and Gray that the encoded content MURDERER OF SMITH constitutes the descriptive content of the concept. The descriptive concept MURDERER-OF-SMITH satisfies this condition. In order for MURDERER-OF-SMITH to satisfy clause (1) of (CA), it must be mutually manifest to Leo and Gray that this concept is retrievable for Gray. But I contend that the stipulated context does not give the interpreter good reason to believe that it is mutually manifest to Leo and Gray that the descriptive concept MURDERER-OF-SMITH is retrievable for Gray at the time of Leo’s utterance.

For the concept MURDERER-OF-SMITH to be retrievable for Gray at the time of Leo’s utterance, it would have to be the case that Gray had constructed the concept at some point prior to the time of the utterance and retained it in memory until that time. Let us look first at whether the stipulated context gives the interpreter good reason to believe that Gray has constructed the concept MURDERER-OF-SMITH prior to the time of Leo’s utterance. It is stipulated in the case that the assumptions Smith was murdered by a single individual and Jones murdered Smith are mutually manifest to Leo and Gray. As we saw in the attributive case, however, this is not the same as saying that it is mutually manifest that these assumptions or any concepts which might be associated with them have actually been mentally represented by Gray prior to Leo’s utterance. An assumption is manifest to an individual, remember, when that individual is capable of mentally representing the assumption and taking it to be true or probably true. The mere fact
that the assumptions *Smith was murdered by a single individual* and *Jones murdered Smith* are mutually manifest, then, tells us nothing about whether Gray has constructed the concept *MURDERER-OF-SMITH* at some point prior to Leo’s utterance.

In fact, even if the stipulated context indicated that it was mutually manifest to Gray and Leo that Gray *believed* the assumptions *Smith was murdered by a single individual* and *Jones murdered Smith*, it would still fail to give the interpreter good reason on balance to believe that it is mutually manifest to Gray and Leo that Gray had constructed the descriptive concept prior to Leo’s utterance. The context tells us nothing about the case surrounding Smith’s murder or about Leo and Gray’s sentiments about the case. But whether or not Gray has constructed a descriptive concept of Smith’s murderer depends entirely on how Gray came across the relevant information about the case and how he would be disposed to react cognitively to that information. Because the stipulated context fails to provide the interpreter with any information on that score, it fails to provide the interpreter with good reason on balance to believe that it is mutually manifest to Gray and Leo that Gray has constructed the descriptive concept *MURDERER-OF-SMITH* prior to Leo’s utterance.

If the interpreter lacks good reason to believe it is mutually manifest that Gray has constructed the descriptive concept *MURDERER-OF-SMITH* at some point in the past, then he or she lacks good reason to believe it is mutually manifest that the concept is retrievable for Gray at the time of Leo’s utterance. But even were we to grant that it is mutually manifest that Gray had constructed the concept at some point in the past, this might still not be enough to give the interpreter good reason to believe that it is mutually manifest to Leo and Gray that *MURDERER-OF-SMITH* is retrievable for Gray at the time of Leo’s utterance. For the concept to be retrievable
for Gray at the time of Leo’s utterance, after all, it must be the case both that Gray had constructed the concept at some point prior to that time and that Gray had retained the descriptive concept in memory from the time of construction until the time of Leo’s utterance. Now, retaining a concept in memory requires the dedication of limited cognitive resources. Memory is finite, after all. It seems plausible, then, to suppose that, in general, cognitive resources will not be dedicated to the storage of a concept in memory unless there is reason to think that retaining that concept in memory will facilitate the generation of cognitive effects in the foreseeable future. But as far as I can tell, the stipulated context does not give us any reason to think it is mutually manifest to Gray and Leo that MURDERER-OF-SMITH is the sort of concept that would appear to Gray to be likely enough to facilitate the generation of cognitive effects to warrant retaining it in memory. It would seem, then, that the interpreter lacks good reason to believe that it is mutually manifest to Leo and Gray that Gray has previously constructed MURDERER-OF-SMITH and retained it in memory until the time of Leo’s utterance. Thus, the interpreter lacks good reason on balance to believe that the descriptive concept is retrievable for Gray at the time of Leo’s utterance.

If the foregoing is correct, then the only concept that is available as an enrichment of the content encoded by the description in Leo’s utterance prior to any ad hoc concept construction is the de re concept of Jones. According to our account, the referential reading of the description results from the selection of a de re concept. Since, as we have seen, the interpretation of the utterance involving this concept is relevant, and since there are no specific expectations of relevance in the case, ad hoc concept construction is not triggered. Thus, the de re concept of Jones is the only concept which is available as an enrichment in this case. By default, then, the
referential reading of the description is the most relevant of those interpretations the interpreter has reason to believe are available to the hearer. According to assumption G about the case, the optimally relevant interpretation of (8.2) in the interpretive context will be the most relevant of those the interpreter has reason to believe are available to Gray. The referential interpretation is thus the optimally relevant interpretation of Leo’s utterance of (8.2) in the interpretive context. The prediction principle laid out in (PP) tells us that in an interpretive context, the natural interpretation of an utterance is the optimally relevant one. Since the reader is interpreting Leo’s utterance in an interpretive context, our account thus predicts that the reader’s natural interpretation of the description in Leo’s utterance of (8.2) is the referential interpretation. This prediction lines up with our intuitive reading of (8.2).

8.4 The Kind-Referring Case

Let us suppose it is mutually manifest to Alfred and Arthur that they are both typical adults. Arthur utters (8.3). Suppose the tone of Arthur’s utterance makes it manifest that he is making an assertion.

(8.3) The panda is territorial.

We will make the following assumptions about the case:

A. The natural interpretation of the definite description in Arthur’s utterance here is the kind-referring reading.
B. A reader’s intuition about the natural interpretation of the description in this case is the result of interpreting the utterance in an interpretive context.
C. Given our definition of an interpretive context, assumption B about the case, and the stipulated context for (8.3), the interpretive context for Arthur’s utterance of (8.3) includes the assumption that Arthur and Alfred are adults and all assumptions we attribute to adults generally, unless the interpreter has good reason on balance to believe that they are not mutually manifest to Arthur and Alfred.
D. The stipulated context for Arthur’s utterance of (8.3) is the only evidence available to
the interpreter in interpreting the utterance.
E. The set of assumptions that the interpreter has good reason to believe Arthur would
judge to be available to Alfred in is the same as the set of assumptions the interpreter
has good reason to believe are available to Alfred in this case.
F. The accessibility ordering the interpreter has reason to believe Arthur would judge to
be Alfred’s ordering is the same as the accessibility ordering that the interpreter has
reason to believe is Alfred’s ordering.
G. Per (ORIC), the optimally relevant interpretation in an interpretive context is the most
relevant of those which the interpreter has reason to believe the speaker would judge
to be available to the hearer. Given assumption E about the case, this means that the
optimally relevant interpretation of (8.3) will be the most relevant of those the
interpreter has reason to believe are available to Alfred.

As noted in assumption A, the natural reading of Arthur’s utterance of (8.3) is the kind-referring
reading. According to the theory of descriptions on offer here, the kind-referring reading of a
description results from the selection of a type concept – in this case, the type concept $T_PANDA$.
As in the previous cases, we will attempt to show that the kind-referring reading is the optimally
relevant interpretation of the description using the strategy outlined in chapter 7. This would
normally involve the completion of four steps: (I) determining which interpretations of the
description are available, (II) showing that the available interpretations of the description are
independent of the interpretation of the predicate, (III) showing that the available interpretations
have the same contextual effects, and (IV) showing that the kind-referring reading has the lowest
processing cost of the available interpretations. In what follows, however, I will argue that the
type concept $T_PANDA$ is the only available enrichment of the description ‘the panda’ as it occurs
in Arthur’s utterance of (8.3).

Let us begin with the question of whether or not the type concept $T_PANDA$ is available as
an enrichment of the content encoded by the description prior to any ad hoc concept
construction. According to clause (2) of (CA), for a concept to be available as an enrichment of
encoded content, it must be mutually manifest that the interpretation of the encoded content as that concept is compatible with that encoded content. Clause (3) of (COM) tells us that for the interpretation of a description as a type concept to be compatible with the content encoded by that description, it must be the case that (a) the extension of the encoded conceptual content or of content derived from it via conceptual broadening or narrowing contains all and only members of the kind which is the referent of the type concept and (b) the referent is a kind. The type concept \( T_{\text{PANDA}} \) refers to the kind panda (Ailuropoda melanoleuca). The extension of the kind concept encoded in the description (PANDA) consists of all and only the members of the kind panda.\(^{190} \)

We would assume that this is manifest to anyone who possessed the concepts in question. Thus, as long as the interpreter has reason to believe that it is mutually manifest to Arthur and Alfred that the concepts PANDA and \( T_{\text{PANDA}} \) are retrievable for Alfred, it will be mutually manifest that \( T_{\text{PANDA}} \) satisfies both of the compatibility requirements on concept availability laid out in clause (3) of (COM), and thereby satisfies condition (2) of (CA) prior to any ad hoc concept construction. We attribute to adults generally the assumption that the typical adult is in possession of the kind concept PANDA. The interpreter thus has good reason to believe it is

\[ ^{190} \text{It might be objected that 'panda' is ambiguous, potentially picking out the giant panda (Ailuropoda melanoleuca) or the red panda (Ailurus fulgens). In a conversation whose participants were mutually manifestly well-educated in the relevant areas of biology, this would be the case. In such a scenario, individual concepts compatible with either the giant panda or the red panda would have to be considered. Since the interpreter only has reason to believe that it is mutually manifest to Alfred and Arthur that they are typical adults, however, the interpreter does not have reason to believe that it is mutually manifest to the conversational participants that the kind concept RED PANDA is retrievable for Alfred, as we do not attribute to adults generally the assumption that other adults possess the a concept of the red panda. On the contrary, most of us only know of the giant panda, and take the word 'panda' to refer to this kind. Prior to any ad hoc concept construction, then, a concept of the red panda would fail to satisfy clause (1) of (CA), and hence would not be available. Since, as I will argue, ad hoc concept construction is not triggered in this case, the concept RED PANDA would be unavailable tout court in this case. Even were ad hoc concept construction triggered, the stipulated context provides the interpreter with no reason to believe the concept would be derivable for Alfred in this case. For the sake of clarity and brevity, then, we will thus proceed on the assumption that the conceptual content encoded by the description 'the panda' in Arthur’s utterance of (8.3) is a kind concept of the giant panda. We will use 'PANDA' to indicate this concept, and 'panda' to pick out the kind giant panda.} \]
mutually manifest to Alfred and Arthur that PANDA is retrievable for Alfred. If the interpreter also has reason to believe it is mutually manifest that $^\top$PANDA is retrievable for Alfred – that is, if the interpreter has reason to believe that $^\top$PANDA satisfies clause (1) of (CA) prior to ad hoc concept construction – then clause (2) will also be satisfied, and the interpreter will have good reason to believe the concept is available as an enrichment of the description’s encoded content prior to any ad hoc concept construction.

Now, it seems plausible to suppose that we generally assume that adults possess type concepts for any kind concept of a well-established kind that they possess. Since we generally assume that adults possess the kind concept PANDA, and panda is a well-established kind, both PANDA and $^\top$PANDA will satisfy clause (1) of (CA).

One might be concerned about the claim that type concepts like $^\top$PANDA are generally assumed to be possessed by adults, however. While it is certainly the case that we generally assume that adults possess the kind concept PANDA, one may not be as comfortable with the notion that we generally assume that adults possess a concept of the kind as an individual. It seems to me, however, that the claim that we generally assume that adults possess the kind concept PANDA ultimately commits us to the notion that adults are generally assumed to possess the type concept $^\top$PANDA.

There would seem to be a close, if somewhat complex, relationship between a natural kind concept and the associated type concept. Sadly, we do not have time to fully explore that relationship here. What seems clear to me, however, is that the occurrence of the natural kind concept in an interpretation or assumption will provide access to at least one deductive elimination rule stored in the logical entry for the concept which involves the type concept. The
reverse will hold for occurrences of the type concept in an assumption or interpretation. At the very least, there would have to be such a rule to explain the intuitive validity of the argument in (8.4).

\[(8.4)\]

1) The panda is a mammal.
2) \(\alpha\) is a panda.
\(.\quad \therefore \alpha\) is a mammal.

Reaching the conclusion in (8.4) requires some bridge between the type concept in premise (1) and the natural kind concept in premise (2). Perhaps the rules in (8.5) might do, assuming that there is some rule in the logical entry for the type concept which allows the move from a claim about a type to the claim about a token. \(R\), here, is the instantiation relation.

\[(8.5)\]

\[\begin{array}{c|c|c}
R(x, \text{\(^\text{PANDA}\))} & \text{\(\text{PANDA}(x)\)} & R(x, \text{\(^\text{PANDA}\))} \\
\hline
\text{\(\text{PANDA}(x)\)} & & \\
\end{array}\]

Together, these rules exemplify the truth that \(x\) is a panda if and only if \(x\) is of the type panda.

I do not know if the rules suggested in (8.5) are the exact rules located in the logical entries for \(\text{PANDA}\) and \(^\text{PANDA}\). Other similar rules may better explain the semantic and inferential relationship between these two concepts. What I think is clear, however, is that some rules like those in (8.5) must be present in the logical entries for \(\text{PANDA}\) and \(^\text{PANDA}\). Not only do such rules appear to be necessary to explain the validity of arguments like that in (8.4), but it seems fair to me to say that someone who did not believe that \(x\) is a panda if and only if \(x\) is of the type panda could not truly be said to be in possession of either the kind concept \(\text{PANDA}\) or the type concept \(^\text{PANDA}\). This would seem to indicate just the sort of connection between the
concepts and the rule which Sperber and Wilson take to hold between concepts and the rules found in their logical entries.

Now, it seems plausible to hold that any concept occurring in a deductive rule which is stored in the logical entry for a concept present in an individual’s mental lexicon will itself be a concept present in the individual’s mental lexicon. It is not clear to me how a concept could end up in a deductive rule stored in some other concept’s logical entry if it were not already part of the individual’s mental lexicon, nor how a deductive rule involving a concept which was not stored in the individual’s memory could be used an inference process. Another way of putting this same claim would be to say that if possession of concept $A$ requires possession of concept $B$ than any individual who possesses $A$ also possesses $B$. Not only is this claim straightforwardly true, but it seems to me that it is also a claim which is manifest to adults generally. And if this is true, then the fact that we generally assume that adults possess the kind concept $\text{PANDA}$ means that we also generally assume that adults possess the type concept $\text{T}\text{PANDA}$. Thus, the interpreter has good reason to believe that it is mutually manifest to Arthur and Alfred that $\text{T}\text{PANDA}$ is retrievable for Alfred and hence satisfies clause (1) of (CA). As we saw earlier, as long as the interpreter had good reason to believe $\text{T}\text{PANDA}$ satisfied clause (1) of (CA), he or she would also have good reason to believe clause (2) is satisfied, and hence would have good reason to believe that $\text{T}\text{PANDA}$ is available to Alfred as an enrichment of the content encoded by the description ‘the panda’ in Arthur’s utterance of (8.3) prior to any ad hoc concept construction.

We have just seen that the type concept $\text{T}\text{PANDA}$ is available to Alfred as an enrichment of the content encoded by ‘the panda’ prior to any ad hoc concept construction. According to our account of definition descriptions, selection of this concept during pragmatic processing yields
the kind-referring reading of the description. A kind-referring reading of the description in
Arthur’s utterance of (8.3) would result in a full interpretation of the utterance which would have
relevance, yielding, for example, the cognitive effects *the panda marks its territory* or *a panda
will defend its territory*. Since there do not appear to be any specific expectations of relevance
for Arthur’s utterance of (8.3), what this means is that none of the triggering conditions in (ACC)
hold and *ad hoc* concept construction will not be triggered in this case. The only available
enrichments of the content encoded by the description ‘the panda’ in Arthur’s utterance here,
then, will be available in the absence of any *ad hoc* concept construction.

Since *ad hoc* concept construction is not triggered in this case, \( ^T \text{PANDA} \) is the only type
concept available as an enrichment of the content encoded by the description. According to
clause 2 of (CA), remember, in order for a concept to be available as an enrichment of encoded
content, remember, it must be the case that the interpretation of the encoded content as that
concept must be compatible with the encoded content in question. According to clause (3) of
(COM), for an interpretation of a description as a type concept to be compatible with the content
encoded by that description, the extension of the encoded conceptual content or content derived
from it via conceptual narrowing or broadening must contain all and only tokens of the type
referred to by that type concept. Since *ad hoc* concept construction is not triggered in this case,
the only type concept interpretations of the descriptive content which will be available are those
that are retrievable and compatible with the encoded content – \( \text{PANDA} \). The only type concept
which would appear to fit that bill is \( ^T \text{PANDA} \).

If the foregoing is correct, then the interpreter only has reason to believe that one type
concept – \( ^T \text{PANDA} \) – is available as an enrichment of the content encoded by the description ‘the
panda’ in Arthur’s utterance of (8.3). Let us now turn to the question of whether the interpreter has reason to believe a \textit{de re} concept is available to Alfred. According to clause (2) of (CA), remember, in order for a \textit{de re} concept to be available, it must be mutually manifest that the interpretation of the encoded content as that \textit{de re} concept is compatible with the encoded content in question. Given clause (2) of (COM) and the fact that \textit{ad hoc} concept construction would not be triggered in this case, for a \textit{de re} concept interpretation of the content encoded by ‘the panda’ in Arthur’s utterance to be compatible with that content, the referent of the \textit{de re} concept would have to be uniquely identified by the encoded content \textsc{panda}. But this content is insufficient to uniquely identify any individual. Thus, the interpreter lacks good reason to believe that any \textit{de re} concept is such that it is mutually manifest that it satisfies clause (2) of (COM). As a result, the interpreter lacks good reason to believe that any \textit{de re} concept is available as an enrichment of the encoded content in this case.

Finally, let us examine whether there are any descriptive concepts available as enrichments of the content encoded by ‘the panda’ here. To be available as an enrichment, any descriptive concept must satisfy (CA). To do that, it must be mutually manifest that the concept satisfies clause (1) of (COM). Since \textit{ad hoc} concept construction is not triggered in this case, this means it must be the case that (a) the encoded conceptual content constitutes the descriptive content of the concept, and (b) the referent of the concept, when there is one, uniquely satisfies the content of the concept. The encoded content here is \textsc{panda}, but the stipulated context gives the interpreter no reason to think there is any individual such that it is mutually manifest to Alfred and Arthur that this individual would uniquely satisfy the content \textsc{panda}. Hence, the
The interpreter lacks good reason to believe that any descriptive concept is available to Alfred as an enrichment of ‘the panda’ as it occurs in Arthur’s utterance of (8.3).

If the foregoing is correct, then the interpreter only has good reason to believe that one concept – $^T\text{PANDA}$ – is available to Alfred as an enrichment of the content encoded by the description ‘the panda’ as it occurs in Arthur’s utterance. According to the relevance theoretic account of descriptions we are currently attempting to test, the kind-referring interpretation of a description results from the selection of a type concept like $^T\text{PANDA}$ as an enrichment of the description’s encoded content. As we have seen, selection of $^T\text{PANDA}$ as the enrichment of the content encoded by the description would yield an interpretation which has relevance in the context. Since it is the only interpretation the interpreter has reason to believe is available in this case, the kind-referring interpretation is by default the most relevant interpretation of those the interpreter has reason to believe are available to the hearer. According to assumption (G) about the case, the optimally relevant interpretation of (8.3) will be the most relevant of those the interpreter has reason to believe are available to Alfred. The kind-referring interpretation is the optimally relevant interpretation of Arthur’s utterance of (8.3) in the interpretive context. The prediction principle laid out in (PP) tells us that in an interpretive context, the natural interpretation of an utterance is the optimally relevant one. Since the reader is interpreting Arthur’s utterance in an interpretive context, our account thus predicts that the reader’s natural interpretation of the description in Arthur’s utterance of (8.3) is the kind-referring interpretation. This prediction lines up with our intuitive reading of (8.3).
8.5 Conclusion

The most worrisome objection facing our relevance theoretic account of the English definite article is the charge that it is untestable, that it could not be used to generate predictions as to the natural interpretations of definite descriptions in example utterances. The goal of this chapter was to overcome this objection. With the help of theoretical machinery put forward in chapter 7, we have used our relevance theoretic account to generate predictions of the natural interpretations of standard examples of the different readings of definite descriptions. In §8.2 we looked at a standard example of an attributively used description and demonstrated that our relevance theoretic account predicts that the natural reading would be the attributive reading. In §8.3 we showed that our relevance theoretic account predicts that the referential reading would be the natural reading of a standard example of a referentially used description. And in §8.4 we used the account to accurately predict the natural reading of a generic kind-referring use of a definite description. We have thus shown not only that our relevance theoretic account can be used to generate predictions, but also that it makes accurate predictions about standard examples of the different uses of definite descriptions.
Bibliography


296


297


