Negation:

Implications for Theories of Natural Language

Ana von Klopp

Degree of Doctor of Philosophy

University of Edinburgh

1993

Declaration

I declare that this thesis has been composed by myself and that the research reported therein has been conducted by myself unless otherwise indicated.

Ana von Klopp

Abstract

Natural language negation has been the object of numerous studies, but still no satisfactory description has yet been given. Current theories tend to be based on the negation in predicate logic, which they extend or extrapolate to increase the coverage. However, they incur a number of empirical inadequacies, and for some uses of negation, such as in yes/no-questions, it is difficult to envisage how a truth functional description could be modified to accommodate them.

This thesis claims that the failure of previous theories stem from their reliance on the common assumption that sentences refer to facts, which is linked to a view of human communication as being a matter of transferring knowledge in the form of propositions from one person to another. As a result of assuming that uttering a sentence is equivalent to making a reference to a fact, it has been thought that uttering a sentence with a negation amounts to a claim that the information of the corresponding affirmative sentence is wrong. Negation has been viewed as applying to information which is conveyed by sentence.

An alternative to the hypothesis that sentences are used to refer to facts is to assume that they are used to make small changes in someone's representation of the world—not in terms of adding and deleting propositions, but in terms of slight reorganisation of entities and concepts, and the links between them. If sentence meaning is thought of as instructions for how information about the world should be organised, it is possible to give a simple, unified characterisation of negation as being *about* a representation of information, not *part* of it. It is shown that this approach has great advantages in terms of descriptive adequacy and explanatory potential. For instance, it is possible to explain certain features of the use and interpretation of multiple negations and scalar expressions. The description proposed in this thesis also lends itself to an account of why speakers would want to use sentences with negation, a question that it is difficult to answer against the background of the traditional approach.

Acknowledgements

At the 1991 European Summer School in Language, Logic, and Information, Jay Atlas told me about the dangers of wasting one's youth on negation. I was foolish enough not to listen, and this is the result.

There are several entities in my representation of the world which I associate with this thesis. First of all I would like to thank my supervisor, Elisabet Engdahl, for her help, advice, and support. She has been instrumental in helping me organise the argumentation, which used to be a hopeless case of too many thoughts and too little structure. I also want to thank Kevin Humphreys, who has contributed enormously by fighting just about every theoretical point, and by moving my adverbials. Jon Oberlander entered the scene as my second supervisor a few weeks before I intended to submit, and told me that most of it was wrong. Unfortunately I could see his point, which meant another month chained to my desk.

Many other people at the Centre for Cognitive Science have been helpful or supportive, in particular Lex Holt, Alex Lascarides, Catrin Rhys, Sophie Cormack, David Beaver, Carl Vogel, Bethan Davies, Jeremy Olsen, Betty Hughes, and Sheila Glasbey. I would also like to thank my other friends, who have put up with my moaning (especially about said extra month of hard labour), distracted me from writing, persuaded me that cycling is more fun than writing a thesis, and used negation in such a way that I had to try something else.

Finally, I do not know whether I should thank or curse Marierose Blomgren for opening my eyes to linguistics...

I ain't takin' no for an answer.
Oh no?
I ain't takin' it for a question either.

B.C., J. Hart, 1986.

Contents

Chapter 1

Introduction

1.1 Why Negation?

Natural language negation is so common that it is often difficult to ignore while studying any communicative aspects of language. Certain language processing operations have even been argued to be affected by whether they occur in sentences with or without negation, e.g. focussing and determination of the speaker's expectations (Anscombre and Ducrot, 1983; Moxey and Sanford, to appear), and the interpretation of referential expressions (Kamp and Reyle, in press). Negation appears to be a relatively basic feature of natural language, but despite this, no satisfactory description of it has yet been given.

Because the function of negation is significant in the study of other issues, it has often been necessary to act as if an adequate description of it existed. The approximation that has been chosen is often that of the truth functional negation of predicate logic. This, however, is inadequate in many cases—so inadequate that it would be necessary to describe some uses of natural language negation in a completely different way. But there is little evidence that *"not"* has more than one function, so this is not a very appealing solution. If the aim is to provide a unified

account of negation, then it must be quite different from a truth function. In view of how much language research relies on there being an account of negation, it is extremely important that a more satisfying characterisation is developed.

The issue of describing negation is not only interesting from an instrumental point of view. The fact that it has been so difficult to formulate a satisfactory description in the traditional disciplines of semantics and pragmatics may mean that it might be necessary to revise some of the assumptions which form the basis for them. The challenge of characterising negation should therefore also provide an opportunity to consider some of the fundamental aspects of studying meaning in language.

1.2 Background

Despite being the most common approximation of natural language negation, the bivalent truth function of predicate logic is inadequate as a description. However, it would not have been adopted unless there was some reason to assume that it could be suitable. Before embarking on the task of proposing an alternative, it will be useful to examine why this description is so commonly used, despite its shortcomings.

1.2.1 Language and Communication

Natural language has developed as a tool for communication and therefore should be expected to be highly adapted for these purposes. This statement is a truism,[1] but the conclusions that are drawn from it can be quite distinct, depending on what type of information is taken to be the object of communication. For the pur-

[1] Or maybe almost a truism—Relevance theory takes a different view, which will be discussed in chapter 3.

pose of characterising natural language, a number of assumptions are commonly made, and they are worth drawing attention to.

Shannon and Weaver (1949) suggested a model of communication which has been extremely influential. Information is defined as the ability of making non-random choices, and communication is the transfer of this ability from a source to a destination. Schematically, some information is selected at the source for transfer to the destination. But it cannot be sent directly—it must be coded into a signal, which can be transmitted. The signal passes through a channel where it is possibly distorted, and then reaches a receiver at the other end, where it is decoded, ideally into a copy of the original information, which then reaches the destination.

The Shannon-Weaver model was not intended primarily for human verbal communication, but for information transfer in general, with a telegraph as the prototype. It is nevertheless often used for the purpose of describing human communication, or rather, for describing the role of language in communication. In this case, the source and transmitter are both one human agent, who decides to convey some information. The information is converted into a signal (in this case sentences of the natural language) which is uttered[2] and another human agent (the receiver/destination) hears it, decodes it, and, if all goes according to plan, constructs a copy of it.

Assuming that this model is a useful way of thinking about the function of language in communication, the following questions instantly arise: firstly, what is the nature of the information that sentences correspond to, and secondly, what part does this play in the more complex process of human communication?

These are issues which have been considered numerous times, including, of course, prior to the formulation of the Shannon-Weaver model. In the case of the former question, one particular answer tends to be given: "thoughts". "Thoughts" may

[2]The terms "utter", "utterance", and "speaker" etc. will be used as generic terms for the production of natural language signals (written or spoken). "Write" etc. will only be used in the discussion of specific data from written texts.

seem like a rather vague concept, which could denote various types of "mental objects", but for the purpose of describing meaning in natural language, a thought is considered to be something rather precise (Dowty et al., 1981, p.144):

> Frege referred to the sense of a sentence as the "thought" (*Gedanke*) expressed by the sentence, what is now often called the *proposition* expressed by the sentence.

"Thoughts" are the objects of communication in for example the Relevance framework (Sperber and Wilson, 1986) and Discourse Representation Theory (Kamp and Reyle, in press), who also define them as propositions, or a set of conditions. Lyons suggests that restricting language meaning this way makes sense for the purpose of studying communication (1977, p.32):

> The narrowing consists in the restriction of the term [communication] to the intentional transmission of information by some means of some established signaling system; and initially at least, we will restrict the term still further, to the transmission of factual, or propositional, information.

This may seem like a rather drastic restriction, but it is nevertheless a common one to make, and as a result, it is assumed that language is often used in communication to convey factual descriptions. The precise details of how these factual descriptions can be represented differ between various approaches to meaning: "propositions" in truth conditional semantics, "infons" in Situation Semantics (Devlin, 1991), etc.[3]

The assumption that at least a subset of sentences of natural language can be appropriately thought of as corresponding to factual descriptions is rarely questioned, and it is common for those who study interactional properties of language to accept it, and to consider factual descriptions to be input to the more complex

[3] "Factual description" (sometimes only "description") will be used as a theory neutral term.

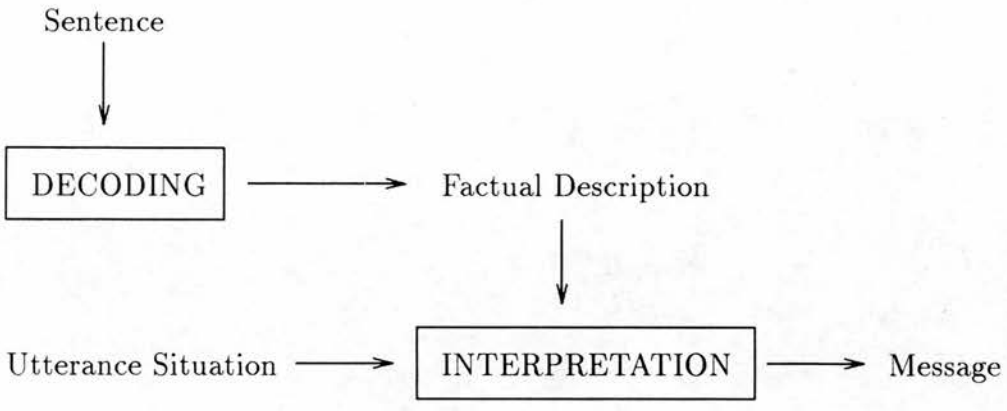

Figure 1.1: Schematic model of the interpretation of an utterance.

interpretation system. This is the view of pragmatics taken in for example Levinson (1983). Although Brown and Yule (1983) suggest that the importance of the type of communication in which factual descriptions are conveyed is greatly exaggerated, they do not dispute that it is a useful way of characterising some linguistic activity. For instance, they suggest that written language is used mainly for this purpose.

If this view of the function of language in (some) communication is adopted, then in such circumstances, the task of generating an utterance is equivalent to deciding what factual description to convey, and then converting it into words. The task of interpreting one is to reconstruct the factual description, and then decide what the speaker's purpose in mentioning it was, whether the description is correct, etc.

This suggests a distinction between what can be called the MESSAGE[4] (what the speaker's intention behind making the utterance is), a factual description chosen to perform the function of carrying the message, and the signal (the sentence) which corresponds to the factual description. These two steps can be taken to indicate that there are two different aspects to studying the communicative function of language: the relation between sentences and factual descriptions, and the relation between the use of factual descriptions and speakers' intentions (fi-

[4] "Message" is not used here in the same way as Shannon and Weaver (1949) use it.

gure 1.1). The distinction between decoding (which only takes sentences as its input) and interpretation (which takes the output from the decoding as input, together with knowledge and contextual factors) represents one way of seeing the semantics/pragmatics distinction.[5]

The distinction was originally made by Morris (1938 and 1946) and Carnap (1942), although it should be pointed out that for Carnap in particular, it does not correspond exactly to this outline. Significantly, both Morris and Carnap viewed the study of natural language as necessarily involving pragmatics. It is however quite common to treat fragments of language as if the relation between factual descriptions and sentences can be explained without reference to the language user, i.e. semantically—Montague Semantics (Dowty et al., 1981) is a paradigm example. For others, such as the proponents of Radical Pragmatics (e.g. Atlas, 1989), and Relevance theory, sentences correspond to factual descriptions, but the correspondence cannot be determined without reference to their use.

1.2.2 Negation

If the meaning of a sentence is a factual description, it makes sense to talk about it as being true or false with respect to some discourse universe. This is the basis for truth conditional semantics. Dowty et al. write (1981, p. 4):

> A truth conditional theory of semantics is one which adheres to the following dictum: To know the meaning of a (declarative) sentence is to know what the world would have to be like for the sentence to be true. Put another way, to give the meaning of a sentence is to specify its truth conditions, i.e., to give necessary and sufficient conditions for the truth of that sentence.

[5] I use the phrase "decode a sentence" to mean the construction of a representation of the information of a sentence (token) when it has been used in a discourse. This is not the same thing as constructing an interpretation (figuring out why the speaker wanted the addressee to construct such a representation).

One common assumption about negation is that it indicates that there is something wrong with the sentence it occurs in. What is wrong often seems to be that the information it corresponds to is incompatible with the discourse universe. If the information in sentences is taken to be equivalent to factual descriptions, which can be true or false, then it follows that negation can often be thought of as reversing their truth values: it is a truth function.

There are however some problems with this description. For instance, there are many examples of natural language negation where it is used in a way which is inconsistent with this characterisation. Some of them are incompatible to the point that it is nearly impossible to envisage how the description could be extended to cover them, as will be shown in chapter 2. Adopting the truth function of predicate logic as a description of natural language negation makes it necessary to postulate that it is ambiguous, for which there is little linguistic motivation (this issue will be discussed in chapters 2 and 4).

Another problem arises, not at the level of correspondence between factual descriptions and sentences, but at the level of messages, in the sense it was used above. Sentences with negation are often recognised as less "informative" than those without (this debate is reviewed in Horn, 1989), but speakers still use them. Why should a speaker want to convey that a factual description is false with respect to the world? It will be shown in chapter 3 that the "obvious" answer, that they are trying to correct a previous assertion or assumption that it is true, is too simplistic. This is not a trivial issue—it means that there is no obvious or generally applicable procedure for describing the interpretation of sentences with negation, in the sense of retrieving the message. All it contributes in the context of studying natural language and communication is at the level of encoding/decoding factual descriptions.

1.3 Goals of this Investigation

The main goal of this thesis is to formulate a characterisation of natural language negation, which has the following properties: it is descriptively adequate, in that it covers all data, and explanatorily adequate, so that it can explain restrictions on its use, and allow for an explanation of why speakers use sentences with negation—i.e. it should be suitable for the generation and interpretation processes.

It is also hoped that this discussion may be useful in the wider perspective of describing the function of language in communication. The issues that are of particular interest are the assumptions that some sentences convey user-independent information, and that the information they convey corresponds to factual descriptions.

1.4 Outline of the Chapters

The first task of this thesis will be to introduce a body of data which will serve as a basis for the rest of the discussion. In particular, it will be shown why it is unlikely that a truth functional description could be extended to cover the use of natural language negation in general. In chapter 3 various approaches to describing language and communication will be examined. The assumptions about language and communication outlined above will be discussed in more detail.

Chapter 4 investigates possible evidence for considering *"not"* to have more than one function. It also draws attention to some interesting restrictions on it, which will be used to formulate some basic requirements on a representational framework which allows for a more plausible description of negation. A conceptual framework which meets these requirements will be proposed in chapter 5. Using the interaction between *"but"* and negation, a hypothesis about the function of negation will be formulated. This hypothesis will be examined in chapter 6, in

which its descriptive and explanatory potential will be assessed. Finally, chapter 7 concludes the discussion by relating the proposal to the other frameworks that were discussed in chapter 3 and discusses the implications for theories of language in general.

Chapter 2

Using Negation

2.1 Introduction

The aim of this chapter is to give an overview of the various uses of natural language negation, and to set out the goals concerning what a theory of it should be expected to cover. The data introduced here will be used as a basis for the discussion in the rest of this thesis.

The starting point of this discussion will be the by now uncontroversial observation that the traditional account of natural language negation as a straightforward bivalent truth function is inadequate for explaining at least some of its uses. There are many problems with it, the most outstanding one being that it is incompatible with a lot of data: speakers sometimes use sentences with negations when the corresponding sentence without a negation would be true. Section 2.2 contains an overview of data of how negation is used, including examples of varying degrees of incompatibility with the truth functional description.

Despite the discrepancies between negation in natural language and the negation of predicate logic, the traditional account is often retained for the subset of data

with which it appears to be compatible, with the consequence that *"not"* must be ambiguous between that and some other function. Section 2.3 considers whether there is any known cross-linguistic evidence which supports such a claim.

2.2 Truth Functional Operator and Natural Negation

2.2.1 Negation as a Bivalent Truth Function

The traditional semantic account of negation suggests that it corresponds to a bivalent truth function which maps the truth value false to true and vice versa.

> **Negation as a Truth Function**
> Let S be a sentence in a natural language, whose semantic representation is r. Let NEG-S be a sentence identical to S except that it also contains a main clause negation. Then NEG-S is true if and only if r is false. The representation of NEG-S is ¬r, where ¬ is a bivalent truth function.

According to this description, the only function of negation is to reverse the truth values of sentences: a sentence with a negation is true if and only if the corresponding sentence without a negation is false. For instance, the (a)-sentence in each of the examples 2.1 and 2.2 below is true if and only if the (b)-sentence in the same example is false.

(2.1) a. He's not in.
b. He's in.

(2.2) a. Florence is not irritatingly well-adjusted.

b. Florence is irritatingly well-adjusted.

It is usually assumed that the representation of a sentence can be derived more or less directly from it. This is particularly the case with e.g. Montague Semantics (Dowty et al., 1981), in which rules to construct the semantic representations from the syntactic form are spelt out explicitly, but it tends to be assumed informally in other frameworks too.[1] For this reason, it is common practice to say that a sentence of the natural language is true when its representation is true. By uttering the sentence in 2.1(a) the speaker thus asserts that it is false that whoever *"he"*[2] refers to is at wherever *"in"* is at the time of the utterance. Likewise, uttering the sentence in example 2.2(a) amounts to asserting that it is false that Florence is irritatingly well-adjusted.

It seems straightforward to give skeletal truth conditions for the sentence in example 2.1(a), even without knowing what person and place *"he"* and *"in"* refer to. Say that *"he"* refers to some male person P, and *"in"* to some location L, and that t is a time when the sentence is uttered. Then a one-sentence truth condition can be formulated:

P is not at L at t

Establishing the truth conditions for the sentence in example 2.2(a) is more complicated. This sentence could be used truthfully about at least two different and incompatible situations. *"Florence is not irritatingly well-adjusted"* would be true both in worlds where Florence is not well-adjusted (in which case it does not matter whether this is irritating or not), and worlds where she is well-adjusted, but where this is not irritating. While the corresponding positive sentence only describes one type of situation (where Florence is irritatingly well-adjusted) the

[1] Proponents of Radical Pragmatics and Relevance theory, as pointed out in chapter 1, would however disagree.

[2] Quoted italics will be used throughout to indicate that the word or phrase is used as signifying the linguistic expression.

negative one can describe different types.[3] The negation can apply either to the modifier *"irritatingly"* or to the predication *"well-adjusted"*. Therefore it is less obvious how to go about it with the sentence in example 2.2(a). Rather than having truth conditions which are of the form of a proposition (or a set of propositions), which are to be satisfied, its truth conditions must consist of a disjunction of propositions:

It is not the case that Florence is well-adjusted **or**
It is not irritating that Florence is well-adjusted

This suggests that there is an asymmetry between sentences with negation, and those without. It is easier to establish truth conditions for the latter than for the former. Unless a sentence without a negation contains a disjunction, or is semantically ambiguous, its truth conditions are a non-disjunctive set of propositions. To see this, consider a sentence S which has three truth conditions {p,q,r}. It is enough that any single one of these does not obtain for S to be false. So S is false and NEG-S (S with a main clause negation) is true e.g. if {¬p,q,r} or {p,¬q,r} obtain. But {¬p,q,r} and {p,¬q,r} cannot simultaneously hold, so there are at least two distinct, and incompatible, types of situation in which NEG-S would be true.

The truth conditions of a sentence with a negation are indirect ones, in the sense that they must be thought of in terms of the truth conditions of the corresponding positive sentence, and consist of a disjunctive list of the negations of its truth conditions. Only for the most simple negative sentences can non-disjunctive truth conditions (that pick out only one type of context) be given. Natural language

[3]Someone who is used to thinking about truth conditions as simple to establish may at first wonder if the complex conditions on the use of this sentence stem from presuppositions (see section 2.2.2) rather than its "meaning". However, *"irritatingly"* does not presuppose, but implies, the proposition that follows it. *"Florence is irritatingly well-adjusted"* can hardly be said to lack a truth value if she is not well-adjusted, and *"No, she isn't"* would be natural reply in that case.

sentences are typically not that simple, and there may be many conditions associated with their appropriate use, meaning that in general, a specific interpretation of a sentence with a negation can only be given with reference to a specific situation which it is used about.

It may be that it is sometimes appropriate to think of some of these conditions as propositions that must be true in the world that the sentence is used to describe. When natural language negation is used to indicate that some condition of this type does not obtain, it is compatible with the truth functional account. But expressions of natural language have other types of conditions as well, and it will be shown below that negation can be used when such conditions fail too.

2.2.2 Presupposition Canceling Negation

It was noted above that the traditional account of negation as a truth function means that negation is considered to indicate that some condition on the discourse universe associated with the corresponding positive sentence does not obtain. It was hinted that negation is not restricted to applying to this this sort of condition, but before considering such cases, it will be worthwhile to examine one commonly suggested objection to this theory.

The objection in question is derived from an intuition that while it makes sense to talk about some sentences as having a truth value of either true or false relative to the world, for some sentences and some worlds this does not apply. The intuition can be thought of as correlating the truth value of a sentence with whether it can be satisfactorily confirmed by saying *"yes"*, in which case it is true, or denied by saying *"no"*, in which case it is false. If neither of these responses would be adequate, it is either considered to have a third truth value (e.g. Seuren, 1985) or to have no truth value at all (Fodor, 1979).

For instance, although the information conveyed by the sentence *"Steve has stop-*

ped smoking" could, and perhaps should, be objected to, if Steve never smoked in the first place, to merely say *"no"* would usually imply that Steve has smoked. To say *"yes"* also implies that Steve has smoked, so neither *"yes"* nor *"no"* are accurate replies. If the accuracy of a *"yes"* or a *"no"* reply is considered to be the diagnostic of truth and falsity respectively, then the sentence is neither true nor false. The implication, that Steve has smoked, from both these answers is termed a PRESUPPOSITION.

Although *"no"* might not be an appropriate reply in case Steve has never smoked, it is possible to object to that sentence with a negation, at least if the information that Steve never smoked is somehow provided too, as illustrated in example 2.3.

(2.3) Steve hasn't stopped smoking. He never smoked in the first place.

In this example, the negation is said to apply to a sentence S (*"Steve has stopped smoking"*) which is neither true nor false, to form a new sentence NEG-S, which is true. Since the sentence is used to object to the information that Steve has stopped smoking on the grounds that its presupposition (that Steve has smoked) does not hold, this usage of negation is often termed PRESUPPOSITION CANCELING (as opposed to PRESUPPOSITION PRESERVING). Some authors (e.g. Seuren, 1985) argue that the presupposition canceling use is semantically different from a negation which maps falsity into truth, such as the one in the example below.

(2.4) Steve hasn't stopped smoking. He couldn't go without more than one day.

If the conditions of truth and falsity outlined above are accepted, then the negations in examples 2.3 and 2.4 cannot be described as the same truth function. This type of example is sometimes used to suggest that there is a weakness in the description of negation as a mapping between truth and falsity. However, it

is usually not considered to discredit the truth functional account of negation in general. Instead, it is taken as evidence that *"not"* is semantically ambiguous between one operator which maps falsity to truth, and one which maps a third truth value to truth, which Seuren (1985) call minimal and radical negation respectively.

The alleged ambiguity obviously hinges on the definition of truth and falsity described above. The truth conditions of *"Steve has stopped smoking"* include that he has smoked, so the issue is whether this is also a truth condition of *"Steve hasn't stopped smoking"*. It is not for the *"Steve hasn't stopped smoking"* in example 2.4, so it remains to decide if it is for the *"Steve hasn't stopped smoking"* in example 2.3, i.e. if they are semantically different. It all depends on whether *"Steve has stopped smoking"* is considered false if Steve never has smoked, or if it is neither true nor false because it would be inappropriate to reply by only saying *"no"*.[4]

The sentence *"Steve hasn't stopped smoking"* is verified by different, and incompatible, contexts in examples 2.3 and 2.4. This on its own, however, does not prove that the negations are different, since, as pointed out in section 2.2.1, sentences

[4]Presupposition failure is not, as it were, defined by whether *"yes"* or *"no"* applies, but this is a test which I believe correlates with the examples of presuppositions that the proponents of these theories provide. They tend to be defined either as conditions both for the truth and the falsity of a sentence, or as conditions for the truth both of a sentence and the same sentence with a presupposition preserving negation. In both these cases, it is assumed that some sentences can be shown to be neither true nor false, but the test for this is not provided.

Atlas (1989) argues that the purported correlation with whether *"yes"* and *"no"* are appropriate replies is "the wrong kind of evidence", citing Kripke (recounted in Quine, 1974). Essentially, the argument is that although an informant may be puzzled by a statement such as *"Steve has stopped smoking"* if Steve never smoked, and consider *"no"* an appropriate reply, there are other sentences where similar preconditions fail where such a reply would be natural. Kripke's example is *"the present king of France will invade us"*—definite noun phrases are supposed to presuppose the existence of a single entity which they refer to—to which a *"no"* reply is perfectly acceptable.

There is one "test" which is sometimes mentioned in connection with presuppositions. However, it is not a "test" of truth values, but of type of negation (presupposition preserving or cancelling). The idea is that by paraphrasing a sentence NEG-S into *"it is not the case that* S*"*, the presuppositions of S are not preserved (assuming that they exist of course). Atlas (1989) points out that the purported weaker interpretation of the paraphrase is hardly recognised by normal language users, but only exists because of logicians' prejudices. If the third truth value and the presupposition canceling negation always co-occur, and no independent tests for either of them are at hand, then the theory is non-demonstrable. Burton-Roberts (1989) argues on these grounds that the theory of semantic ambiguity cannot be falsified.

with negation cannot generally be expected to have non-disjunctive truth conditions. The difference is in other words not as significant as it may seem at first, and it is questionable whether the proposed negations could be comprehensively defined and distinguished.

It is likely that the reason why these negations have been perceived as different is the selection of data to illustrate "ordinary" presupposition preserving negation. The function of "ordinary" negation is usually explained with reference to comparatively simple sentences, e.g. *"John doesn't love Mary"*, for which it may have appeared possible to give a single truth condition. But as noted above, natural language negation can be rather more complex while still being compatible with the bivalent truth function.

2.2.3 Truth Conditions vs Utterance Conditions

The examples of sentences with negation that have been considered so far illustrate one way natural language negation can be used. This use has often been thought of as semantic, in the sense of being independent of the situation where it is used. It is taken to indicate that some truth condition of the corresponding sentence without the negation fails to obtain. The use of natural language negation is however not restricted to expressing that a sentence does not accurately describe the world. It can also be used when some condition on the *use* of the sentence, as opposed to on the world it describes, does not obtain, as illustrated by the following example:

(2.5) A and B live in Edinburgh. B is going to Glasgow to practice with his band. A is not a native speaker of English.

A: Do you want me to help you to carry the stuff to the station?

B: No, it's OK, thanks.
A: Are you bringing the (guitar) amp?
B: I'm not BRINGING it, I'm TAKING it.

In English, the choice of *"take"* and *"bring"* depends on where the speaker and the addressee are at the time of the utterance—in general, if either of them are at the target location of the transporting action then *"bring"* should be used, otherwise *take*.[5] A is native in Swedish, which has one single phrase corresponding to both *"take"* and *"bring"*. When B says *"I'm not bringing it"* the negation does not indicate that some truth condition on the action of B transporting the amplifier from Edinburgh to Glasgow fails to obtain. Although the negation could be said to apply to a condition associated with the phrase *"bring"*, it is a condition on the utterance situation and not on the discourse universe. The sentence *"are you bringing the amp?"* would have been appropriate if A or B had been in Glasgow at the time of the utterance, in which case the sentence *"I'm not bringing it"* would have been false. The use of *"bring"* is therefore not inappropriate with respect to the situation that the sentence purports to describe, but with respect to the circumstances of the utterance.[6]

[5]It appears that there is some individual, or perhaps regional, variation in how strong this distinction is.

[6]In general, the choice of *"take"* and *"bring"* and the analogous pair *"come"* and *"go"* depends on the position of the agents in the discourse universe, and if one location has been designated as the setting of the discourse. Consider the following sentence:

John's mother came/went to see him in the hospital.

"Came" would be more natural if the previous discourse had been about something which happened in the hospital, and *"went"* if it was set somewhere else. Certain locations can also be strongly associated with an agent, so that *"come"* can be used even if no other relevant agent is at the location at the time of either the utterance or the action, e.g. *"home"*:

When I came home yesterday...

To choose between *"take"* and *"bring"* in English does not amount to making a statement, or implying something, about the world. Instead, the speaker is complying with constraints on lexical items with regard to the utterance situation. Example 2.5 illustrates how one person fails to comply with such a constraint, and how this failure is indicated by another person using a negation. Although conditions on the utterance situation can theoretically be expressed as conditions on the sentence, they are clearly not independent of the language user.

2.2.4 Non-Optimal Descriptions

A truth functional description of negation may be compatible with some of its uses, but as seen in section 2.2.3, this is not always the case. In this section some data will be considered where the negation applies to how the sentence relates to the discourse universe, but in contrast to the examples in section 2.2.1 and 2.2.2, there does not appear to be any truth conditions which fail to obtain.

This typically happens when an agent has uttered a sentence which another agent considers to be inadequate given the circumstances. The sentence may be "true" relative to the world, but the second agent can still use a negation to indicate that the choice of that particular sentence is not considered optimal. Often, this use of negation occurs when a part of a representation can be perceived as a position on some scale, such as the scale of natural numbers, or a scale of how positive the feelings an agent has for something or someone are. e.g.:

(2.6) A: Does he have a good bike too?
B: He doesn't have *a* good bike, he has three.

(2.7) A: Thérèse seems to like Carol.
B: She doesn't like her, she loves her.

If someone has three good bikes, they certainly have one,[7] and *"love"* is usually taken to imply *"like"*. If the second clauses of B's utterances are true, the speaker should be committed to the representation of the sentences that were rejected. Such apparently contradictory utterances are, however, used in felicitous utterances, and their addressees can make sense of them. Although the addressee might first be led down the garden path (due to a precocious assimilation of the first sentence), they do not end up assuming that the speaker has contradicted themselves.

Use of negation in sentences that would theoretically be true without it is not confined to scalar expressions, although such examples have perhaps received more attention than others (e.g. Grice, 1975; Gazdar, 1979; Levinson, 1983; and Horn, 1989). But this phenomenon is not limited to scalar expressions, as shown in example 2.8. For instance, if it is true that someone did something, then it is also true that they managed to do it.

(2.8) A: Did you manage to put the wheel back on?
B: What do you mean, manage? I didn't manage anything. It was dead easy.

Assuming that agents who make utterances like these are not being inconsistent, it seems that the negation does not indicate that some truth condition fails to obtain. It cannot be described as applying to some condition on the utterance situation either. The issue is the appropriateness of the information: someone who uses negation in this way is typically trying to convey that some description is not considered optimal.

The B speakers' utterances in example 2.6 to 2.8 are intended to change the A speakers' representations, in that the rejected descriptions and the preferred ones

[7]One could think of other ways of representing these sentences such that they do not imply each other, but it is commonly assumed that it ought to be done this way (e.g. Horn, 1989). See chapter 3, section 3.4.1.

are not identical. If one wishes to see this in terms of truth conditions, then those of the A speakers' sentences are a subset of those of the B speakers' preferred ones (at least in the former two examples). This means that it might be tempting to suggest some quasi-semantic (pragmatic) solution which keeps the negation truth functional, even though it cannot be as a mapping between truth and falsity of the explicit information. Possible approaches here are to postulate that scalar expressions automatically assert their own maximality (i.e. the representation of *"a"* in example 2.6 should be 'EXACTLY ONE', or that the speaker has left out a *"just"* or an *"only"*, which should be part of the semantic representation even though it is not realised linguistically.

The first approach leads to inconsistencies with sentences such as as B's reply in example 2.9, which would be contradictory if *"he has a good bike"* asserted that the person in question has only one bike.

(2.9) A: Does he have a good bike too?
B: He certainly has a good bike, in fact he has three.

Thus it will not do to claim that a sentence with an expression corresponding to the position N on a scale always means that N is the upper bound. The claim would have to be weakened, so that such expressions sometimes mean 'AT LEAST N' and sometimes 'EXACTLY N'. This means that the problem is shifted somewhere else: instead of explaining why negation can be used in two apparently different ways, the task is seen as one of determining which of two senses a different type of expression has.

The second approach differs from the first one in that the scalar (or other) expression is not considered ambiguous, but instead the 'AT LEAST' and 'EXACTLY' part of the interpretation is part of the representation of the sentence, although the speaker chose not to realise it linguistically. That is, in example 2.6, the B speaker meant that HE DOES NOT HAVE EXACTLY ONE BIKE', but decided to leave out

the *"just", "exactly", "only"* or other expression which would have conveyed this sense. Since this is a common phenomenon, it would appear that in general, when one wants to reject one description in favour of another one which implies it, it is not necessary to explicitly realise the modifier. But why is it then, that although *"X arranged Y"* implies that *"X knows about Y"*, it would sound odd if someone used the B' sentence in example 2.10?

(2.10) A: Does the head know about it?
B: He doesn't just know about it, he arranged it.
B': *He doesn't know about it, he arranged it.

If it is only possible to leave out maximality indicators in certain contexts, then it is clearly necessary to specify what conditions must be satisfied. Stating that speakers sometimes leave out maximality markers which have to be inserted into the semantic representation is merely a post hoc measure, which is applied if a negation cannot be given a truth functional interpretation otherwise.

Examples 2.6 to 2.8 are interesting, because in the circumstances where they would be appropriate, the negated sentences do not violate neither truth conditions nor utterance conditions of the type discussed in the previous section. It could be argued that accepting the sentence which was negated would be misleading, and that to this extent it is in breech of some form of utterance conditions, but they are not of the same type as the direct ones considered above.

2.2.5 Objections to Linguistic Realisations

Horn (1985) drew attention to the fact that negation can be used when an agent wants to point out what they perceive to be a linguistic inaccuracy in an utterance made by another agent. One example including a linguistic error has already been considered in section 2.2.3, but that differs from the ones that Horn discusses by

violating some constraint associated with the expression that was used. In Horn's examples, the mistakes, or perhaps misjudgements, are ineffectual, in that they do not result in the word or phrase in question having any truth or utterance conditions that are distinct from what the correcting speaker prefers. Examples 2.11 (from Horn, 1989) and 2.12 contain objections to a gender mistake and a mispronunciation respectively:

(2.11) A: Est-ce que tu as coupé le viande? (pronounced with an English accent)
Is it that you have cut the{masculine} meat?
'Have you cut the meat?'
B: Non, je n'ai pas 'cou-pay luh vee-and', j'ai coupé la viande.
No, I not have not cut the{masculine} meat, I have cut the{feminine} meat.
No, I haven't "cut the meat", I have cut the meat.

(2.12) A: Vill du ha en äppel?
Want you have "an apple" {wrong gender & mispronounced}?
'Do you want "an apple"?'
B: Jag vill inte ha en äppel, men ett *äpple* vore inte så dumt.
I want not have "an apple" {wrong gender & mispronounced} but an apple were not so bad.
'I don't want "an apple", but I wouldn't mind an apple.'

The mistakes in these examples do not mean that the relevant word is automatically understood as conceptually different from what the speaker intended, as was the case with the example in section 2.2.3. No new, or different, conditions are directly associated with the A-utterance, that are not also associated with the same utterance, had it been realised properly. Accordingly, Horn suggests that the negation does not apply to information at all in these cases.

Not only non-native speakers' failed attempts can be objected to in this way, but so can also other linguistic expressions which are deemed "inappropriate". This

might be for instance because the register, focus, or connotations of a phrase are considered unsuitable (examples from Horn 1989):

(2.13) Grandpa isn't feeling lousy, Johnny, he is just a tad indisposed.

(2.14) I'm not his sister, he's my brother.

2.2.6 Dealing with Conditions

The intuition that a negation indicates that there is something wrong with the corresponding positive sentence is equivalent to saying that there is some condition on the use of the sentence that is not met. It is reasonably easy to specify what the conditions are when it comes to incompatible information and constraints on the utterance situation. When the information is correct, but not optimal, the task becomes more difficult; even more so when the grounds for using the negation is the linguistic realisation.

Although the examples that have been discussed so far differ in the aspect of the utterance the negation seems to apply to, they seem to have this important property in common: the negation indicates that something is not quite right. Because of this, it may seem that one way of dealing with uses of negation which do not conform to the ordinary truth functional description would be to adapt the semantic framework. By allowing linguistic or pragmatic operators into the semantic representation, the various conditions on the utterance situation, and on linguistic wellformedness and appropriateness are elevated to truth conditions; and a negation which applies to such a condition is truth functional. This approach is favoured by e.g. Seuren (1991), for at least some of the examples above: the idea is that by uttering the phrase *"le viande"*, one simultaneously asserts that

"viande" is a masculine noun in French (or at least that the expressions used is the correct one).

Natural language negation can be used by speakers when they want to indicate that some feature of a sentence is inappropriate given the circumstances. The different features are however not easily classified into distinct categories. The layout of this chapter has perhaps made it seem like this can be done: that there are truth conditions, utterance conditions, conditions on the linguistic realisation etc. In reality the boundaries between these are vague: the anomalous features cannot always be easily classified (this will be discussed in chapter 4, section 4.2). Although it is questionable whether it is a good idea to treat all different aspects of an utterance as truth conditions, this approach has the merit of reflecting that there are no obvious subcategories of negation. It would, however, be difficult to find suitable operators for some of the examples discussed above.

The notion that negation applies to some feature of a sentence or an utterance is a common hypothesis, which is nearly omnipresent, even in accounts such as Horn's (1989) which does not consider all negation truth functional. Why does it seem natural to take this type of approach? The answer is, because the set of data used as examples is limited to such sentences that would typically be used to indicate that "something is not quite right" with the corresponding positive sentence. It is usually data of affirmative sentences with a main clause negation. In the next section, some examples will be considered where the truth functional approach fails completely, even if it is weakened to seeing negation merely as an indication that the sentence is not optimal.

2.2.7 Yes/No-Questions

While negation in declarative sentences has received a lot of attention, its function in other types of sentences has not. Consequently most of what has been written

about negation is conceived to fit this type of data. It is however worthwhile considering other types as well, as the following examples illustrate:

(2.15) a. Is Manchester rainy?
b. Isn't Manchester rainy?

If someone who considers Manchester to be rainy was asked these questions, and answered them sincerely, the answer should in both cases be *"yes"*. If the negation in 2.15(b)[8] is interpreted truth functionally, the whole utterance will be understood as a request to confirm or deny that the proposition that 'MANCHESTER IS RAINY' is false, which is not how this type of question is used. Not even the weakest version of negation as a bivalent operator, according to which negation only indicates that something is not right, applies in this case—an agent who treats an utterance of the question *"isn't Manchester rainy?"* as the speaker seeking confirmation or denial that there is something wrong with 'MANCHESTER IS RAINY' would produce an anomalous answer. This holds for intonationally unmarked yes/no-questions in general. When a speaker requests a confirmation or a rejection of a fact, the status of the fact in the question is not affected by whether the speaker uses a negation or not.

(2.16) Isn't the weather in Scotland dreadful?

In English, this usage of negation is only common in demands for the assessment of descriptions. In Swedish, on the other hand, it is also found in requests to perform an action. The two sentences in example 2.17 are equivalent if one wants to ask an addressee to open a door. The (b)-sentence is not perceived as the speaker asking for the second time, or being irritated, as the corresponding English question could be.

[8]This example was brought to my attention by Tom Wachtel.

(2.17) a. Kan du öppna dörren?
Can you open door-the?
'Can you open the door?'

b. Kan du inte öppna dörren?
Can you not open door-the?
'Can't you open the door? '

Negation used like this may possibly add a degree of politeness. This is the effect it has in the French and Irish[9] examples below. The Irish b example would be perceived as more polite in a formal conversation. In a more informal one, it may indicate that the speaker is becoming irritated.

(2.18) T'as pas un clop?
You have not a fag?
'You don't have a fag? '

(2.19) a. An bhfreagróidh tú ceist domsa?
part. answer-will you question for-me?
'Will you answer a question for me?'

b. Nach bhfreagróidh tú ceist domsa?
NEG answer-will you question for-me?
'Will you not answer a question for me?'

The sentences in examples 2.17 to 2.19 are typically used to ask the addressee to perform an action, rather than to assess the status of of the information. It is nevertheless interesting that the negation is irrelevant for the interpretation of the request. That is, a speaker who uses the sentence in example 2.17(b) is not asking their addressee to not open the door. In fact, it is not possible to use this type of construction with a negation to ask an addressee not to perform an action in Swedish—it would be necessary to use an expression corresponding

[9]I am grateful to Breanndán Ó Nuallain for providing this example.

do *"refrain"* instead.[10] Unless one is prepared to claim that negation used in the main-clause position in yes/no-questions is different from negation used in declarative sentences, the position that negation is essentially a truth function is clearly untenable.

2.2.8 Summary

This section has introduced a body of data that will serve as a basis for discussion throughout this thesis. Negation has been shown to be a versatile expression, which is not limited to indicating that some truth condition of the corresponding positive sentence does not obtain. On the contrary, it can apply to a wide range of properties of utterances of sentences, many of which it would be difficult to formulate in terms of a truth conditional semantics. In order to extend the coverage of the traditional account, it would be necessary to allow the representations to contain also predicates that are about the situation in which the sentence is uttered and about formal properties of language.

While some of the data made it seem likely that the function of natural language negation is to indicate inappropriate features of sentences, this interpretation simply is not available for the use of negation in yes/no-questions. If such a question with a main clause negation were interpreted as a request for the confirmation that there is something wrong with the corresponding positive description, the wrong answer would be produced (except if the negation is heavily stressed). A main clause negation in a yes/no-question should not normally appear in the re-

[10]It is possible to get this reading out of a corresponding English utterance, with the negation in a different position and with contrastive stress on *"not"*:

Can you *not* open the door, please?

Swedish also has a choice of syntactic position for the negation, but there is only one possible reading in both cases. It is worth noting that the English example above would probably be interpreted as meaning more than a mere request to leave the door shut, such as the speaker being rather irritated etc. As opposed to the ordinary request, a *"yes"* or a *"no"* would not be suitable replies to indicate that one accepts/rejects the order. Instead, it would be necessary to reply with *"OK"* or similar.

presentation of it, if this is thought of as a factual description whose status is unknown. This suggest that the intuition that negation indicates that there is something anomalous with the sentence it occurs in is, after all, not as reasonable as it may appear.

2.3 Cross-Linguistic Considerations

Despite the fact that the differences between various uses of natural language negation are slight, and gradual, it has sometimes been suggested to be semantically ambiguous, on the grounds that no single description will fit all cases. This section contains a brief discussion of the distinctions made in some languages with more than one morpheme corresponding to the English *"not"*, for the purpose of investigating whether the differences between them could be used to justify viewing *"not"* as ambiguous.

Why should cross-linguistic evidence matter? One reason to consider it is that most, if not all, accounts of *"not"* that have been proposed so far either have exceptions or suggest that negation is ambiguous, while there is no evidence from English that it is not a uniform phenomenon.[11] This is for instance the case with the presupposition preserving and cancelling functions, outlined in section 2.2.2 above. One is required to decide whether the sentence *"Steve hasn't stopped smoking"* in example 2.3 is different from the one in example 2.4:

(2.3) Steve hasn't stopped smoking. He never smoked in the first place.

(2.4) Steve hasn't stopped smoking. He couldn't go without more than one day.

[11]The arguments about possible multifunctionality (as opposed to ambiguity) of negation in English will be examined in chapter 4.

Since they are linguistically identical, declaring them different requires an act of faith—unless some other language distinguished them lexically.

Semantic ambiguity (when two different concepts happen to be denoted by identical linguistic strings), can usually be verified by translation into another language. If one word has two meanings, and this fact is purely incidental, then it would be expected that other languages, at least if they belong to different families, have different words for the two. An example of this is the English word *"cry"*: other languages tend not to have one word which signifies both 'WEEP' and 'SHOUT'.

The translation test for ambiguity is however hardly a failsafe one, since languages often differ in the specificity of the concepts they have lexicalised. This point is made in Atlas (1989). To demonstrate it with yet another example, there is no single Swedish word corresponding to the English *"cut"*. Instead, there are different verbs for cutting with a knife, cutting with an axe, cutting with a pair of scissors, etc. If the translation test was the be-all-and-end-all of semantic ambiguity, then *"cut"* would be ambiguous between these functions, rather than a general label for all of them. A native English speaker would probably not feel that is the case. Similarly, a Swedish speaker would probably not consider the phrase *"ta med sej"* to be ambiguous between the concepts labeled as *"take"* and *"bring"* in English.

Any evidence produced by the translation test is therefore likely to be influenced by the concepts one is used to having specific labels for, and perhaps, on a different level, by the theory that one is trying to promote. Nevertheless, it is interesting to see what categories different linguistic communities have come to label, so even if it cannot be considered hard evidence that *"not"* is ambiguous, it is surely worthwhile to consider what categorisations that are made or, equally significantly, not made. This does not purport to be a comprehensive review of such categories, but some points are worth making.

There are many languages into which the English *"not"* does not translate uni-

formly as one particular morpheme (or combination of morphemes). To take a familiar example, there may be specific morphemes for certain categories of negated objects, so that a morpheme meaning "nobody" is preferred to a phrase corresponding to "not anybody". Negation in French, for instance, is in the form of a particle *"ne"* (which is usually dropped in spoken language), followed by a comparatively specific negative morpheme, so there is no single word for *"not"*:

(2.20) a. On n'a rien vu.
We not have nothing seen.
'We didn't see anything.'

b. On ne l'a pas vu.
We not it have not seen.
'We didn't see it.'

c. On n'a vu personne.
We not have seen nobody.
'We didn't see anyone.'

d. On ne l'a jamais vu.
We not it have never seen.
'We have never seen it.'

The difference between *"not"* and negation in French is similar to the one between *"cut"* and the Swedish equivalents: it is a case of more finely grained descriptions in one of the languages, rather than of ambiguity in the other. Still, there would be more substance to a claim that negation is ambiguous between two particular functions (e.g. those that it is said to have in example 2.3 and example 2.4 respectively) if it can be shown that some language actually has different morphemes corresponding to them. As for the presupposition preserving and canceling negations, Gazdar (1979) notes that no language is known to make this distinction morphologically, which makes it hard to believe that they are indeed distinct.

One use of negation that constitutes a counterexample to virtually every account that has been suggested so far is that in yes/no-questions. This use in particular, it would be desirable to discount as a separate category, but no data available data supports this (although this could be due to lack of reasons to look for it

among those who have investigated such things). Likely candidates would have been for instance Scots and Irish Gaelic, which have different negation morphemes in declaratives and questions. However, the one (or the ones) used in yes/no-questions are also used in question-word-questions (as well as in relative clauses), while it is only in the former that the truth functional interpretation fails.

In fact, even if a language which consistently used a different negation morpheme in yes/no-questions could be found, the problem would not be solved. It must also be explained why the ordinary truth functional negation does not usually occur in yes/no-questions, and why one would have to use some complicated paraphrase using an embedded declarative if one really seeks the verification of a negative fact:[12]

(2.21) Can you confirm that it is not Thursday today?

The lack of a formal negation in the most obvious semantic representation of a yes/no-question is not an incident or an idiomatic peculiarity—it is impossible to construct a question with a main clause negation such that it would be interpreted as a request to verify a negative fact. This might reflect that there seldom is any reason to seek negative information, but that does not make it any less interesting for the purposes of representing natural language negation.

While no evidence has been found to support a categorisation of negation which includes one sense which corresponds to the truth function, it is worth drawing attention to a distinction that appears to be made in several unrelated languages: they have special morphemes and/or constructions for rejecting a focussed element. In English, which permits little syntactic freedom, such negation is typically expressed using intonation or an *it*-cleft.

[12]Even so, it is not obvious that one can be completely certain of what e.g. a *"yes"* reply would mean.

(2.22) In France they don't *eat* their food, they enjoy it.

(2.23) It wasn't the derailleur that was bent, it was the dropouts.

I will refer to such use as CONTRASTIVE NEGATION for the time being.[13] Certain languages, e.g. Welsh and Yoruba (cited by Payne, 1985; a further few are mentioned by Horn, 1989) have specific forms for it. That is, the English sentence *"we didn't go to the Canaries last year"* in examples 2.24(a) and (b) would not be translated to the same sentence in such a language. The Welsh equivalents are given in example 2.25.[14]

(2.24) a. We didn't go to the Canaries last year. We didn't go anywhere.

b. We didn't go to the Canaries last year. It was the year before.

(2.25) a. Aethon ni **ddim** i'r Ynysoedd y Canary y llynedd. Aethon ni i le'm byd.
Went we **not** to-the Islands of Canary last-year went we to no-place-at-all.
'We didn't go to the Canaries last year. We didn't go anywhere.'

b. **Nid** y llynedd aethon ni i'r Ynysoedd y Canary, ond y flwyddyn cynt.
Not{contrastive} the last-year went we to-the islands of Canary, but the year before.
'We didn't go to the Canaries last year, but the year before.'

[13]It has also been labelled "focussing" by Atlas 1990. I prefer the term "contrastive" for reasons that should become clear later on.

[14]I am grateful to Geraint Jones for helping me with the Welsh data. The accounts of negation in Welsh differ quite a lot, which probably is due partially to a conflict between written and spoken language, and partially to regional differences. *Contemporary Welsh Grammar* (1976) lists *"nid"* as a literary form, but it seems to be in use at least in some parts of Wales. The alternative would be using *"dim"* (of which *"ddim"* is a form) in the same construction.

It is interesting that contrastive negation is sometimes distinguished morphologically, and as will be seen in chapter 4, this contrastive use is sometimes reflected in other morphological distinctions. Apart from this, there is little evidence for any conceptually different negations, which is striking, given that many languages have more than one word or phrase which corresponds to the English *"not"*.

2.4 Summary

The purpose of this chapter was to describe the versatility of the phenomenon that is natural language negation, and to discuss what is required of a description of it. Although natural language negation has often been described as corresponding to the truth function of predicate logic, it was shown that it would be difficult to give a unified description of it based on that characterisation. There is no cross-linguistic evidence to suggest that *"not"* is semantically ambiguous, so postulating that it has more than one meaning is undesirable. The challenge for a theory of negation is to account for all the examples above.

Chapter 3

Approaches to Describing Language and Negation

3.1 Introduction

This chapter examines some different approaches to describing the meaning or information carried by sentences, as well as some recently proposed descriptions of natural language negation. I will begin by comparing some ideas about the nature of sentence meaning, which increasingly differ from the traditional conception of meaning as truth. The frameworks that will be considered are Relevance theory (Sperber and Wilson 1986), Radical Pragmatics (Atlas, 1989), and Argumentation theory (Anscombre and Ducrot, 1983).

With this background, some different approaches to the description of negation will be considered. I will concentrate on four approaches: negation as a pragmatic operator; Horn's (1989) proposal that it has more than one function; Atlas' (1981) and Kempson's (1986) approach to keeping negation unambiguously truth functional; and Anscombre's and Ducrot's (1983) account which suggests that it has

argumentative properties. These approaches were chosen because they represent distinct ways of accounting for the behaviour of natural language negation, and because their coverage of data is comparatively large.

Finally I will consider a different, but related, issue, namely how to account for the interpretations of what has been called scalar expressions. This problem has been addressed by some of those whose accounts of negation are reviewed, and examining this part of their theories will help explicate the differences between their approaches.

3.2 Approaches to Describing Language

3.2.1 Meaning and Context

I have deliberately left out until now the sentence which has probably received the greatest amount of attention in the literature about negation. It is, of course,

(3.1) The king of France is not bald. (There is no king of France.)

This particular sentence has been given numerous explanations, and anyone who writes about negation is almost obliged to include it among the set of data they discuss. But what is interesting about this sentence is perhaps not so much the function of the negation in it, as the study of the sentence itself.

It is difficult to imagine in what sort of context this sentence would be uttered. Presumably it would be used to reject an earlier assertion that the king of France is indeed bald. But why did the other speaker say that? Did they genuinely believe that there is a king of France, and if so, on what did they base the judgement that he is bald? Was it a jocular reference to president Mitterand, or to Jacques

Delors? Was it a reference to one of the two people who claim to be the legitimate heir to the French throne?

On the view that sentences have a meaning which can be derived more or less directly from their linguistic form, there is nothing incongruent in considering the sentence above without knowing when or if it was used. If meaning is independent of the language user (the Morris/Carnap definition of semantics) and a subset of sentences of natural language can be given semantic representations, then that subset has informational properties which can be discussed without reference to use.

Semantic representations can be given for expressions of language which do not reflect any communicative aspects except at a user-independent level of referring to factual representations. But it is generally accepted that not all expressions of natural language fulfil this criterion. Pronouns and deixis is one exception—they need to be evaluated with respect to how they are used (although it has been argued that some of them can be defined with respect to the context only, so that no specific reference to the user need be made, cf. Montague Semantics, Dowty et al., 1981). But many types of expressions, including those sentences which are taken to correspond to propositions, and connectives such as *"and"* and *"not"*, are often considered to be semantic, and therefore user-independent.

Some theories of language would disagree with the assumption that sentences have a meaning (in the specific sense of a factual description) which can be derived independently of their use. Sentences are taken to correspond to factual descriptions, but their representations can only be computed with respect to a situation where they are used. This is the approach taken in Relevance theory (Sperber and Wilson, 1986) and Radical Pragmatics (e.g. Atlas, 1989). Further, more far-reaching criticisms of the traditional role of language in communication have been made by other scholars, who reject the assumption that sentences correspond to factual descriptions at all. This perspective is taken in Argumentation theory (Anscombre and Ducrot, 1983).

Those who have adopted (truth conditional) semantics as a description of meaning in language in general do not express, or even view, this as making a point about the role of language in communication. It does however imply a commitment to a number of non-trivial philosophical and practical assumptions about the nature of language and its use, some of which I will try to point out below.

3.2.2 Relevance theory

If sentences are considered to correspond to factual descriptions, then communication is viewed as a subroutine in reasoning—a way of providing input to the deductive system. To many, this consequence of adopting it as a representational framework would seem too strong. However, one recent proposal, Relevance theory, prefers to strengthen these claims.

Sperber and Wilson (1986) argue that language must not be seen in the light of communication. In their eyes, a language is merely "a grammar governed representational system" (p.173) created for cognitive, and not for communicative, purposes. Although it would be possible to define language as "semantically interpreted well-formed formulas used for communication," this would be scientifically unjustifiable, since there is no natural connection between language and communication. Language is a necessary tool for information processing and memorising, but its use for communication is completely incidental. In fact, they go so far as claiming that (Sperber and Wilson, 1986, p.173):

> the fact that humans have developed languages which can be used to communicate is interesting, but it tells us nothing about the essential nature of language. The originality of the human species is exactly to have found this curious additional use of something which many other species also possess, as the originality of elephants is to have found that they can use their noses for the curious additional purpose of picking things up. In both cases, the result has been that something widely found in other species has undergone remarkable adaptation and development because of the new uses it has been put to. However,

> it is as strange for humans to conclude that the essential purpose of language is for communication as it would be for elephants to conclude that the essential purpose of noses is for picking things up.

In other words, a language is a tool for reasoning. Human natural language is, in addition, used as a tool for communication, in which case its function is to *describe* reasoning.

According to Sperber and Wilson, the purpose of communication is that the agents involved share representations of the topic in the end. This is achieved by presenting stimuli, which the recipient ideally processes in such a way that their representation becomes similar to that of whoever produced it. The role of language in this is to provide a means for producing complex stimuli in the form of coded representations, which would be impossible if all the agent could do was to draw attention to particular states. To some extent, these claims probably represent consequences of accepting that sentences of natural language correspond to propositions, but what is interesting in this case is that Sperber and Wilson accept it prior to choosing a representational framework.

Although Relevance theory ultimately derives a semantic representation from sentences, it differs from the ordinary approach in how the representations are obtained. Conventionally, there is considered to be a direct mapping from each sentence of the natural language to a proposition (or set of propositions, if the sentence is ambiguous). According to Sperber and Wilson, when an agent hears an utterance of a sentence, they retrieve a skeleton proposition corresponding to the information represented by it (this is possible because there is a direct mapping for some words). Certain words, such as pronouns, will be represented as variables in the skeleton proposition, but they contribute constraints on the type of context that it can be interpreted in. The proposition is completed by finding a context which satisfies the constraints. The variables are instantiated, and it is made sure that the full proposition is consistent with it. After this, the proposition is assessed, so as to decide if it contributes something new to the agent's representation. If

this fails, then other bindings are tried until it does (the information has become "relevant").

3.2.3 Radical Pragmatics

Radical Pragmatics is the name of a collection of papers edited by Cole (1981). Although the various articles in it share the feature that they question that meaning can be derived directly from sentences, the alternative proposals are quite distinct from each other. This section introduces one particular view of Radical Pragmatics, that of Atlas (1989). Radical Pragmatics shares with Relevance theory the idea that the meaning of a sentence only can be derived with reference to how it is used. In Atlas' terms, sentences are SENSE-GENERAL. They do not have specific meanings—only when a sentence is used in an utterance can its meaning be specified. There are, however, some significant differences from Relevance theory. For instance, language is viewed as a tool for communication, as opposed to a tool for reasoning.[1] This is an important theoretical difference, in that it forms part of the justification for the concept of sense-generality, although it is less critical for the application of the accounts.

By considering language to be a tool for communication, Atlas argues that meaning must be defined in terms of *use* (assertability) rather than *truth*. The meaning-as-truth position defines the meaning of sentences as being completely independent of the language user, and what they know. But this is unlikely, given that language is acquired through use,[2] and he argues that assertability captures ordinary language users intuitions about meaning better than truth does.

These different approaches to describing meaning are the basis for the distinction

[1]It should be said, though, that Radical Pragmatics is not presented as a theory of language in communication, but rather as a theory of meaning in language.

[2]The meaning-as-truth position seems to necessitate the view that natural language is isomorphic to some intrinsic capacity such as Fodor's concept of a language of thought (Fodor, 1975). This is the view of language taken in the Relevance framework.

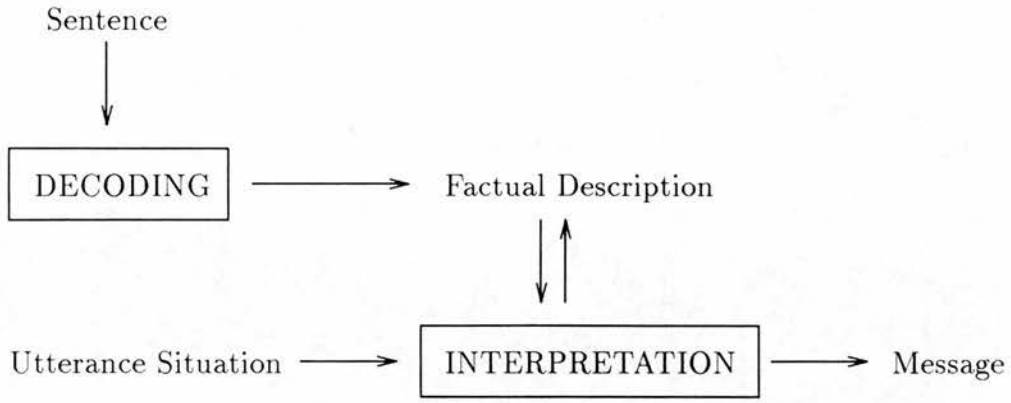

Figure 3.1: Schematic model of the interpretation of an utterance: Radical Pragmatics and Relevance Theory

between sense-generality and ambiguity. Sense-generality means that there are several types of context where a sentence would be appropriate. This is reflected by the fact that language users often can name different situations where an utterance of a sentence would be appropriate. But naming situations in which a sentence could be used is not equivalent to giving disjunctive truth conditions. Atlas gives the following example: the sentence *"I'm going to the game"* can be used for instance to convey that the speaker is going to watch a baseball match or a football match, but it does not *mean* that the speaker is going to watch baseball *or* football (or some other game). It is non-specific and would be suitable on both occasions. He writes (1989, p. 146):

> Understanding an utterance is not simply knowing a logical form that the context SELECTS from the meanings of an ambiguous sentence. Understanding an utterance is knowing a proposition that in the context the hearer CONSTRUCTS from the meaning of a univocal, sense-general sentence. The general meaning of the sentence is made specific in the hearer's understanding of the utterance.

Considering this in the view of the model of utterance interpretation in chapter 1, figure 1.1, this means that although sentences in principle refer to factual descriptions, the actual description intended cannot be derived directly by decoding the sentence. Rather it is determined by feedback from the interpretation process (as illustrated in figure 3.1). Atlas does not specify what the sense-general repre-

sentations are like, except for stating that they cannot be a logical form (1989, p. 131):

> Formal languages as we know them are not designed to represent this aspect of grammatical meaning.

3.2.4 Data or Manipulation

The approaches to describing meaning in language which have been discussed so far have one thing in common: they agree that the act of uttering a sentence[3] is equivalent to conveying a factual description. The last of the theoretical frameworks that will be discussed here, Argumentation theory, is based on a different assumption. The amount of conceptual rethinking that is required for this is considerable, so it will be useful to begin with an attempt to explicate what the alternative view of sentence meaning is.

Reddy (1979) argues that thinking about communication in natural language as the transfer of thought (the CONDUIT METAPHOR) is a mistake, and a dangerous one at that. But it is difficult to avoid this perspective, as much, if not most, of language about language is metaphoric and biased towards this view, e.g. *"put ideas into words"*. He offers an alternative metaphor for the role of language in communication—the toolmakers paradigm (henceforth TMP).

The setting of the metaphor is a wheel-shaped area which is divided into partitions, with one person living in each of them. The environments in the partitions are similar, but not identical, and the people try to make the best of their situation by constructing various tools so that they can cultivate the soil, etc. They can only communicate with each other by passing blueprints of tools through a letterbox in the middle, in the hope that the recipient has the materials to construct the

[3]The phrase "the act of uttering a sentence" is used in order to distinguish it from the purpose of uttering the sentence, which is to convey the message in the terminology used above.

tool, and that it will be of use for them.

When they receive a blueprint, they try to build a copy of the tool in question, and to work out what it is meant to be used for. Since their environments are different, the tools they build are usually not identical to the original ones, and they often use them for other purposes than what the person who made the blueprint intended.

Green discusses the TMP in her pragmatics textbook (1989, p.11):

> Reddy argued that the conduit metaphor is not only misleading, but harmfully, perniciously so. For our purposes, however, it is enough to assume that there is more to understanding utterances than parsing them and deriving representations of their propositional meanings in terms of intensional logic or model theory (cf. Dowty et al., 1981). It is necessary also to make inferences about what the utterer believes about what the addressee believes, and about what effect the utterer intends the utterance to have.

There are at least two ways of interpreting Reddy's argument. Using the terminology from chapter 1, it could either be about the interpretation process, or about the decoding of sentences. I am not entirely sure which one he intended, but it seems to me that Green's description of it is biased towards the former, to the exclusion of the other one.

Green's description of Reddy's argument can be interpreted as being about the necessity of using pragmatics for the understanding of utterances, in the traditional way that was outlined in chapter 1. That is, sentences are decoded into factual descriptions, which serve as input to the interpretation mechanisms that yield the message. The TMP is understood as saying that *interpretation* is not equivalent to *decoding*; the message is not equivalent to the factual description.

However, it is also possible to interpret the TMP as being about the decoding process. The position that the interpretation of utterances requires more than the derivation of propositions is after all hardly controversial, but Reddy mentions

several times how difficult it is to explain what he means, since it requires a fundamental rethinking about basic concepts. It is because of these statements that I think that he might have meant a more far-reaching change than what is perhaps reflected in Green's paragraph above.

The metaphor of blueprints makes it natural to assume that Reddy wanted us to think about language as instructions. Now, there already is an approach to the study of language which is based on this type of conception of its function: Speech Act theory (Austin, 1962; Searle, 1969). In this framework, to understand an utterance it is not sufficient to know what it refers to, but one must also determine what the speaker was trying to achieve by making this reference. Utterances should be interpreted as instructions to *behave* in a certain way.

Reading the TMP at another level, the blueprints need not be seen as instructions about behaviour, but could be seen as instructions about how to represent information. The *fact* that the blueprint was passed on can be viewed as an instruction about behaviour (the equivalent of a speech act): this tool may be useful. To interpret the toolmaker's intention behind sending the blueprint at this level is to find out how to *use* the tool once it has be constructed. But the blueprints as such are also sets of instructions, which are decoded by *building* the tool from what is available in the environment.

Decoding and interpretation can be viewed as trying to follow instructions at different levels of the representation: achieving small organisational changes on the one hand (building the tool), and determining the speaker's intentions on the other (using it). This is in contrast with the traditional view of the meaning of sentences, which using this metaphor would be equivalent to saying that the blueprints represent tools, rather than being instructions of how to build them.

If the decoding of sentences[4] is viewed as the process of working out what changes

[4]Recall that this phrase is used in the sense of constructing a representation of sentence tokens that have been uttered, see chapter 1, note 5.

the speaker wanted the addressee to make, it is not just the interpretation of utterances which depends on taking other agents' representations into account—the actual decoding process makes the same requirements. I will refer to the two approaches to describing meaning of language as the DATA HYPOTHESIS (= sentences correspond to factual descriptions) and the MANIPULATION HYPOTHESIS (= sentences correspond to sets of instructions as to how some representation should be changed) respectively.

This issue is not often discussed in the Anglo-Saxon philosophical tradition, where the data hypothesis tends to be taken for granted. Sperber and Wilson are an exception, in that they state explicitly that they assume that the function of language is to represent facts. However, they do not consider any alternatives. This is in contrast with some of the French tradition, e.g. Anscombre and Ducrot (1986), Roulet et al. (1985), where these distinctions are debated, and some version of the manipulation hypothesis often is preferred.

It may seem that the distinction between a set of instructions about the organisation of someone's representation and proposition is inconsequential, since the result of following the instructions should be equivalent to a propositional representation. This would make it similar to the use of context in deriving a propositional representation of sentences in Radical Pragmatics and Relevance theory. But there are non-trivial differences, and that is what this thesis is about.

3.2.5 Argumentation Theory

Adopting a data oriented approach to describing natural language meaning basically is to suggest that speakers only provide propositional input to their addressees—they cannot influence their addressees' processing of information unless they mention explicitly what particular inference rules should be applied. In contrast with this approach, Anscombre's and Ducrot's Argumentation theory emphasises the

function of language in guiding the addressee's processing and changing their representations. Their central theory (e.g. Anscombre and Ducrot 1983, Anscombre and Ducrot, 1986) is that coherent communication does not depend on any intrinsic logical relation between the facts referred to in utterances, but on how the speaker presents the information. They argue that the argumentative function of utterances is more important than the facts they refer to when it comes to understanding why certain expressions are chosen.

This claim is based on observations of communication where the informational content and the argumentative properties of sequences of sentences are at odds with each other. For instance, there are situations in which, from the logical point of view, speakers appear to be contradicting themselves. In example 3.2, B simultaneously supports two facts that are mutually exclusive, but the utterance is nevertheless perfectly functional as coherent argumentation.[5]

(3.2) A: Is dinner ready yet?
B: Yeah, almost.

In this example, the speaker first accepts that dinner is ready, and then adds that it is almost ready, which amounts to simultaneously accepting that dinner is ready and that it is not.[6] In contrast, references to facts that are logically compatible and relevant can be combined into less than coherent utterances, if an unsuitable modifier is chosen. Example 3.3 illustrates how the use of the quantifier *"few"* is odd when followed by another fact modified by *"nearly"*, but not when followed by *"only"*.[7]

[5]It is worth noting that one is often not even aware of the potential contradiction in an utterance of this type, as opposed to for instance the examples in chapter 2, section 2.2.4.

[6]Atlas (1984) argues that *"almost"* does not implicate *"not"*, but only *"not quite"*. However, if *"ready"* is an achievement predicate, in which case *"not quite ready"* entails *"not ready"*. Anscombre's and Ducrot's account of *"almost"/"not quite"* is quite different, see page 52.

[7]The issue is not that 20% is too many to be *"few"*, because *"only 20%"* might well be a larger number than *"nearly 20%"*.

(3.3) a. ?Few drivers exceed the 70mph speed limit (nearly 20%).
b. Few drivers exceed the 70mph speed limit (only 20%).

As for joining sentences, there may be several possibilities which are truth functionally equivalent. Despite this, one might be acceptable while the other one is not. A speaker who wants to recommend the film "The Doors" on the grounds that the acting is good, while admitting that the direction was not perfect, could not use the sentence in example 3.4(a). The (b)-sentence, which refers to the same facts is, on the other hand, fine.

(3.4) a. ?You should go and see "The Doors", it is badly directed, and extremely well acted.
b. You should go and see "The Doors", it is badly directed, but extremely well acted.

The reason why these examples make good, or bad, utterances cannot be deduced from their "informational content". Anscombre and Ducrot (1983, p.79, my translation) write:

> Linguists and philosophers have always been struck by the the possibilities of "reasoning"—used in a very vague sense—offered by language. But they have generally chosen to reduce this activity to an orderly presentation of logical relations, i.e. relations depending on the truth value of sentences are used: incompatibility, implication..., etc. Even though "reasoning" is considered linguistically relevant, its reduction to logic (understood as a truth functional system) is ... unacceptable.

Anscombre and Ducrot offer an explanation of the examples above in terms of what they label ARGUMENTATIVE ORIENTATION. Argumentative orientation is a property of utterances with respect to a context. If an utterance is presented as supporting a certain conclusion, then it is oriented towards that conclusion. The

---

Atlas (personal communication) does not agree that (3.3a) is odd. Even if he were right, it is beyond question that the (b)-sentence is better, which cannot be explained in terms of logical properties of *"few"*, *"nearly"*, and *"only"*.

speaker can use certain linguistic expressions to indicate that some information is used as an argument for something. I will return to their definition of argumentative orientation below, but first it will be useful to consider intuitively what this means for the examples discussed above.

In example 3.2, speaker B can simultaneously confirm that the dinner is ready and commit himself or herself to a description of the dinner according to which it is not ready (but not far off) because both these situations would justify A taking the same kind of action, e.g. setting the table. *"Almost"* as a description of the state of the readiness of the dinner is used to indicate that the addressee should behave in the same way as if it was ready. This can be contrasted with an expression such as *"not quite"* which would steer the addressee in the opposite direction, i.e. towards acting as if the dinner was not ready. *"Yes, not quite"* would be an odd answer (cf. Sadock, 1981).[8]. For Anscombre and Ducrot, there is no relation (by for instance implicature) between *"almost"* and *"not quite"*. They are both used to refer to the same type of state of affairs, but they are used to achieve different reactions.

An agent who used the sentence in example 3.3 would appear to assert that the set of drivers who break the law is insignificant, but following it by the qualification *"nearly 20%"* suggests that the set is close to some particular size, as if this size is significant. Although the number of drivers referred to by *"few drivers"* and *"nearly 20% of the drivers"* may be compatible (depending on one's views of driving behaviour), the ways it is referred to in the sentence appear to support opposite conclusions.

The sentence in example 3.4(b) could be used by speaker to recommend a film to someone. If a film is well acted, that usually counts as an advantage for it, whereas the opposite holds for being badly directed. This means that *"the*

[8] *"Yes, but not quite"* can be an acceptable combination in principle (although it can hardly be used, at least not to the same effect, in example 3.2). This is because of the *"but"*, see chapter 4, section 4.3 and chapter 5, section 5.3.

film is well acted" and *"the film is badly directed"* have the opposite orientations with respect to a recommendation to see it. To join two sentences with *"and"*, they must have the same orientation relative to the context which prompted the speaker to use them. Joining them with *"but"* requires them to have the opposite orientation. So if the speaker considered both the direction and the acting to be good, then only *"and"* would be suitable:

(3.5) a. You should go and see "The Doors", it is well directed, and extremely well acted.
b. ?You should go and see "The Doors", it is well directed, but extremely well acted.

Argumentative orientation reflects what Anscombre and Ducrot consider to be the most important feature of utterances: they have a direction with respect to the topic of the discourse in which they are uttered. When a speaker makes an utterance which refers to some fact p, they are not primarily trying to convey p to an addressee, but to steer the addressee in a certain direction in the discussion (which may be towards accepting p, but need not be).

Argumentative orientation has been defined in several ways throughout their works, moving away from a logical definition towards a speaker dependent one. The earlier definitions suggested that when agents argue a point, they apply rules to facts in order to show that other facts can be deduced from the present ones. The definition in Anscombre and Ducrot (1983) links the concept to particular agents' representations of information. It uses a different view of "rules", in which the logical ones have been replaced by (possibly agent-specific) probabilistic connections. They label them TOPOI[9], loosely based on the Aristotelian notion. Using it, a description of argumentative orientation can be given.

[9]The existence of the topoi is a claim about a feature of human knowledge representation, which they suggest is most closely approximated by probabilistic rules. The topoi are open to criticism that they are not rigorous enough, but rigour in this case imposes structure that is not necessarily empirically justified. The topoi are intended to refer to a feature of cognitive processing, not a logical system which approximates it. There is nothing intrinsically wrong with holding a discussion at that level.

Argumentative Orientation
An utterance u_1 representing the information p_1 is oriented towards another utterance u_2 representing the information p_2 if u_1 presents p_1 in such a way that it licences an application of a topos from p_1 to p_2.[10]

The way this is formulated shows an important difference from data oriented accounts: identifying the argumentative orientation of an utterance is not equivalent to recognising *which* inference rule the speaker used, but *that* they used one.

Anscombre and Ducrot argue that decoding an utterance is not about retrieving a proposition, but about retrieving instructions about how some information should be incorporated into the representation. Truth conditional semantics fails as a representation, because its very purpose is to isolate propositions and their truth value from context. Deduction schemas cannot be used to describe natural language communication precisely because in the former, premises must not be interconnected, and all references to the same fact must have the same value. This is typically not the case in natural language, where referring to the same information in different ways can influence the addressee to process the information in radically different manner.

3.2.6 Comparing Approaches

The data hypothesis corresponds to a model of communication in which speakers convey factual descriptions to their addressees, which the latter process using some form of inference rules. It invites the question of how the addressees select what rules to apply—speakers typically do not convey enough premises for a valid deduction. But if the input is in the form of apparently unconnected propositions, then the addressees are clearly going to have to reason with them to make sense of the communication. This problem has been discussed by e.g. Sadock (1977), who suggests that conventional abbreviations of deduction schemata are used. He calls

[10] $p_1 = p_2$ is not excluded.

these "modus brevis". When Anscombre and Ducrot would say that a speaker has presented p as an argument for q, indicating *that* they consider there to be an applicable rule, a logical approach to communication prefers it as the speaker having presented p, omitted the rule that p $\rightarrow$ q, and concluded q. Understanding the utterance is a matter of *identifying* the rule p $\rightarrow$ q.

A typical utterance which lacks a connecting premise is given below:

(3.6) I walked in today. The front tyre of the grey bike was flat.

If this utterance only conveys two propositions (which by the nature of propositions must be unconnected), then the addressee will need to find some rule which connects them. If Sadock's modus brevis are adequate as reasoning schemata, then there is a (single) rule which connects the two propositions referred to by the sentences. One rule which would comply with this requirement is

i. The Speaker's (S) grey bike has a flat tyre $\rightarrow$ S walks to the office.

There are some important comments to be made about this. To begin with, it is "obvious" that the propositions should be ordered this way (q $\rightarrow$ p rather than p $\rightarrow$ q), although it is very unlikely that anyone would actually have internalised a rule like (i). This is because of contextual knowledge, but such knowledge need not correspond to rules as usually conceived of. The default situation with respect to example 3.6 may well be that S fixes flat tyres instantly, so that this situation is an exception. It would not make the utterance less valid.

Not only is the addressee not likely to have internalised the rule given in (i), but the same holds for the speaker—it is too arbitrary. If the speaker was asked to actually describe the reasoning "used" to arrive at the conclusion to walk to their office, it would in all likelihood be significantly more complex than the simple rule given above. More premises, including for instance the following would be needed:

ii. If one of the bikes can be used, S usually commutes by bike.

iii. If S does not commute by bike, then S walks.

iv. If a bike has a flat tyre and someone fixes it then the bike does not have a flat tyre.

v. S could not be bothered to fix a flat tyre at the relevant time.

But what about the addressee? Although they may be able to recognise some of these premises, they are unlikely to be able to reconstitute the whole reasoning behind the action. This fact does however not prevent the understanding of the utterance—the recognition that there is some connection which corresponds to (i) is quite enough, and the speaker can hardly have intended to convey more. To this extent, the modus brevis hypothesis is acceptable, but the way it was formulated indicates that it should be viewed as a logical deduction rather than simply expressing that there is a reasoning behind it.

It follows from the data hypothesis that understanding a sentence such as the one in example 3.6 is equivalent to either applying the suitable rules for deduction, or postulating some if they are not already available. This is spelt out explicitly in Relevance theory. This is where Argumentation theory differs, as it does not describe argumentation in terms of factual constraints. Consider another example:

(3.7) It's not worth going via Pat's, { I'm not sure / I doubt / it's not sure } he'll be in.

A data oriented theory of meaning would view the interpretation of these sentences as a matter of recognising that the speaker has used a rule, in this case one which says that if one is going to visit someone, they need to be home. The reason why this rule should be used is the fact referred to in the second clause (the probability of Pat being home is low), together with the fact that going to Pat's place is a salient topic. This explains why the sentence in example 3.8 appears ill-formed.

(3.8) ?It's not worth going via Pat's, it's possible he'll be in.

Argumentation theory gives quite a different account of the interpretation of these sentences. The reason why a topos from 'VISIT PAT' to 'PAT IS HOME' is chosen is the formulation of the sentence.[11] The phrases *"I'm not sure"* etc. have the effect of weakening the argumentative strength of the expressions they modify to the extent that its argumentative orientation is reversed. The phrase *"it is not worth it"*, modifying the first clause, suggest that the conclusion 'VISIT PAT' should not be accepted. It is because of the reversal of the orientation of the presented argument in conjunction with the rejection of the conclusion that a topos from 'VISIT PAT' to 'PAT IS HOME' is selected for the interpretation.

The advantage of this approach becomes clear when the sentence in example 3.8 is reconsidered. There is in fact nothing wrong with it. The reason why it seemed odd is that example 3.7 has just been considered, and the way it is expressed leads us to consider a connection from Pat's being home to the possibility of visiting him. Example 3.8 on the other hand, by using *"it is possible"* rather than *"I'm not sure"* suggests that 'PAT IS IN' is oriented against 'GO TO PAT'S'. i.e. a different topos should be used. This is because the phrase *"it is possible"* weakens the argumentative strength of the expression it modifies, but contrary to those in example 3.7 it does not reverse the argumentative orientation.

When first considering example 3.8 above, one is still influenced by the topos selected for the previous example, which despite being out of context apparently is interpreted as being about the same individuals and actions. The sentence in the latter example appears odd: not because of any intrinsic contradictions in the factual descriptions, but because of context that was established by the selection of the expressions of uncertainty. This illustrates how much power linguistic expressions have when it comes to influencing the addressee's selection of context.

[11] Normally it would be contextual knowledge and the formulation of the sentence, but since not much context is given here, one basically has to rely on the latter.

and that factual information is not the only input to reasoning. This influence cannot be ignored as a feature of the interpretation of sentences, and Argumentation theory therefore has an advantage over the semantics based accounts in that it recognises argumentative orientation as an important part of lexical meaning.[12]

3.3 Some Approaches to Describing Negation

It was noted in chapter 2 that it does not seem possible to account for negation in general as a feature of factual descriptions which are expressed by sentences. If semantics fails, a legitimate question would be if it be possible to give a pragmatic account instead, i.e. if it can be described in terms of the interaction between the participants of the communication. This will be discussed in section 3.3.1.

The fact that there are discrepancies between natural language negation and the negation of predicate logic has been recognised for a long time, and there have been many attempts to account for them. Sections 3.3.2–3.3.4 review three comparatively recent proposals. The discussion here is not intended as a comprehensive account of theories of negation, and perhaps the selected accounts are not representative of the issues that have been debated through the years.[13] However, they illustrate three quite different ways of dealing with the task. The main reason why they were selected is that they are formulated to cover a large range of data, which until recently has not been widely considered, and that they represent the different approaches to describing language outlined in section 3.2.

[12]Grice's conventional implicature can perhaps be said to reflect the same phenomena, although it has not been used or discussed on the same scale.

[13]For instance, I have chosen not to review the negation/presupposition debate. For a more comprehensive review of theories of negation, see Horn (1989).

3.3.1 Negation as a Pragmatic Operator?

When semantics fails to account for some feature of natural language, it is common to designate it pragmatic instead. In chapter 1, semantics was described as the relation between sentences of natural language and the factual descriptions they express. Pragmatics, on the other hand, is supposed to explain how the factual descriptions are integrated into the addressee's representation. If negation is not semantic, but pragmatic, it must be used to indicate what the addressee should do with the factual description, instead of simply applying to the description as such. The aim of this section is to consider whether negation could be described in these terms.

Although nobody, to my knowledge, has proposed a purely pragmatic description of negation, several people (e.g. Allwood, 1972 and Givón, 1978) have suggested that its use is pragmatically restricted. Their reason for this is that they feel that sentences with negation have a different status in the discourse from those without. But this cannot be explained in terms of the logical characterisation of negation, since it is a feature of the bivalent negation of predicate logic that there is no difference in status between a sentence S without a negation and a sentence NEG-S which has one. Allwood and Givón consider the meaning of natural language negation to be equivalent to the bivalent truth function, but they argue that it is not sufficiently described unless restrictions on its use are specified. In order to completely account for natural language negation, one must turn to pragmatics.

Both Allwood and Givón treat their pragmatic restrictions as complements to a logical description, but given that the latter has been shown to be untenable as a general account, the pragmatic restrictions will be treated as if they were *the* description of negation for the present examination. The proposed restrictions range from Allwood's relatively weak one, to Givón's much stronger statement. According to Allwood, a sentence with a negation is only used when the corresponding sentence without a negation would have been expected to be an accurate

representation of what the speaker is talking about. Givón claims that sentences with negation are only used to correct erroneous beliefs on behalf of the addressee.

Givón's position is too strong, and is not supported by data on negation. Even though some dubious cases, such as when a speaker makes a linguistic mistake, could be construed as manifestations of erroneous beliefs, this is not true for negation in general. This can be seen in the Norwegian example below (the italics are mine):

(3.9) The interviewee is talking to a journalist about how he used to deal hash, but now only sells larger quantities to other dealers:

> Nå etter at jeg sluttet å selge sjelv, har jeg bare noen få personer å forholde meg til. Og færre folk betyr mindre risiko for å bli avslørt. Men ennå er det noen som kommer på dørene og som ennå ikke har forstått *at jeg ikke selger lenger.*
>
> 'Now after I've stopped dealing myself, I only need to see a few people. And fewer people means less risk for being busted. But there's still the odd person who comes to the door, and who still hasn't understood that *I don't deal anymore.*'

That the interviewee does not sell anymore is known to the journalist at the time of the utterance of the underlined sentence. Although this example could be characterised as drawing attention to someone's erroneous beliefs, it is not the addressee (the journalist) who is the offender, but the people who turn up at the ex-dealer's door (a similar example is cited in Atlas, 1980).

Example 3.10 proves even more problematic: it can hardly be argued that the person who wrote the instructions in example 3.10 thought that the reader would believe that the cable ends would fray if the cable end covers are taken off:

(3.10) If you can't take out the cable at the lever end, which is probably unlikely, then you'll have to undo the bolts on the calipers. Mark the cable positions before you do, and you may have to take off any cable end covers, but be careful not to fray the ends.

If anything, the writer seems to assume that the reader might fail to consider a case in which the cables fray. It is also impossible to reconcile the notion of erroneous beliefs with the use of negation in yes/no-questions, as described in chapter 2, section 2.2.6.

Allwood's proposal, being weaker than Givón's, is compatible with much of the data of main clause negation in declarative sentences. To some extent it actually seems compatible with the use of negation in yes/no-questions (although Allwood's simultaneous assertion that negation is truth functional is not): when a speaker uses a negation in a yes/no-question, it could well be that they indicate that they are of the opinion that it would have been reasonable to assume that the information were accurate. Moreover, it does not have the same problems with the negation in example 3.10 as Givón's approach has, since it does not refer to the addressee's beliefs. By simply saying that negation applies to something that could have been expected (by anyone), cases where the speaker is the only one who thinks that some state of affairs may obtain are not excluded. This means that the approach is consistent with example 3.10, where the speaker, but not necessarily the addressee, is aware of the possibilities that the cables may fray.

It would clearly be necessary to define the notion of "expectancy", which is likely to involve more than a truth conditional semantics. As far as the representation of information is concerned, the use of defaults or probabilistic rules is implied. Furthermore, it would be necessary to distinguish appeal to different agents' beliefs. In other words, quite extensive changes to the representational framework are needed.

Turning to the use of negation in relative clauses, it becomes obvious that both the notion of erroneous beliefs and that of expected situations are too restrictive. Both Allwood's and Givón's proposals would predict that the negation in example 3.11 is misused.

(3.11) B has just opened the fridge to take some milk out. A sits at the table.

A: Can you get me some jam while you're at it?
B: Sure. Which one do you want?
A: Apricot. It's the one that doesn't have a label.

A description of negation which claims that it indicates that a default rule does not apply, or that someone holds an erroneous belief, would fail for this example. Firstly, although jam jars may be considered to have labels by default, it is clear that A knows that the default is overridden in this case. Secondly, given that A knows that B has already opened the fridge, and quite possibly has seen the jar, the default may for all that B knows be overridden for B as well. Now it could perhaps be argued that A uses the negation because the jar could have had a label in principle (consistent with Allwood's account), but I disagree. The negation is used in order to distinguish a particular jar from the others—it is used to contrast, not to correct or reject. The important feature is not that the jar *could* have had a label, but that it *lacks* a property that a set of similar objects has, and that this distinguishes it from that set.

Nor does the use of the negation in example 3.11 depend on there being an accepted default in principle. To see this, consider another example:

(3.12) A and B have just arrived at a party. A does not know anyone, but has been told that Ian, who A has heard of, will be there. B knows most of the people, including Ian.

A: Who is Ian?
B: It's the one who doesn't read Sartre.

It is hardly a default (in Britain at least) that people read Sartre, but if most of the people in the room look very serious and/or wear black rimmed glasses, while Ian is a sporty type, B may well succeed in referrring to him. In making this utterance, B is effectively commenting on the appearance of the others as well as identifying Ian, The expression *"the one who doesn't read Sartre"* is not used

to deny an expectation of Ian, but to contrast two sets of individuals by some selected identifying property. A pragmatic account of negation which specifies that it is used to object to an erroneous belief or previous expectation does not capture this use.

3.3.2 Horn: Multifunctionality

Horn's (1985, 1989) takes a fairly conventional approach to describing language, assuming that sentences have semantic representations which are independent of their use. Because he considers the representations of sentences to be close to their linguistic surface form, a lot of the data discussed in chapter 2 cannot be covered by the truth functional description of negation. Despite this, he argues that some uses of natural language negation are truth functional, and hence that it must be ambiguous. So when negation appears to object to a linguistic realisation, occurs in a sentence that would be true without it, or is used to cancel a presupposition (as described in chapter 2, section 2.2.2), then it is of a different type, which Horn claims can be characterised as a pragmatic operator. He does not give any formal descriptions of the two, but the following informal ones can be found in the text:[14]

> "Apparent sentence negation represents either a descriptive truth-functional operator, taking a proposition **p** into a proposition **not-p** (or a predicate **P** into a predicate **not-P**) or a metalinguistic operator, which can be glossed 'I object to **U**', where **U** is crucially a linguistic utterance or utterance type rather than an abstract proposition." (1989, p.377)

Horn considers this ambiguity to be a pragmatic one, rather than a case of two different meanings. That is, the metalinguistic negation is in principle performing a similar function as the descriptive one, but at a different level, with the negation applying to the sentence as if it denoted an utterance of it, rather than the fact

[14]His categorisation of negation is based on a distinction made by Ducrot (1972, 1973) (see below), from which the labels are borrowed, but differs from it, as will be shown in section 3.3.5.

it refers to. He suggests that this is common in natural language, and that the situation with other logical operators is analogous. To illustrate this, he uses the following example of a conditional.

(3.13) If I may say so, you look beautiful tonight.

In order to make sense of the conditional as a material implication, it is necessary to consider the sentence *"you look beautiful tonight"* to refer to the act of uttering it, rather than the fact it normally would be taken to represent.

Although he argues that the dual function of negation is best described as a pragmatic ambiguity, the parallel between the two is called into question by some of his other assertions. For instance, he writes that metalinguistic negation can be glossed as "I object to [an utterance]", while elsewhere he makes it clear that descriptive negation cannot be (analogously) characterised as "I object to [information]."[15] If this were the case, then natural language negation would in general correspond to a speech act of objecting. Horn is quite adamant that this is not the case in his criticism of Givón's (1978) position (summarised above). His reason for rejecting this characterisation of description negation is that it leads to undesirable interpretations of sentences like the ones in examples 3.14 and 3.15.

(3.14) Either he isn't going to the opera tonight or he's going to miss the first act.

[15]A characterisation which is close to pragmatic ambiguity has been suggested by van der Sandt, using what he calls an "echo-operator", which applies to sentences in order to denote utterances of them. The truth functional negation can operate outside the echo-operator, which allows it to apply to any (informative) aspect of an utterance. This basically means that the negation can indicate that some utterance condition of the sentence fails to obtain, in the same way as it would indicate that a truth condition fails to obtain, which is a way of formalising pragmatic ambiguity.

Horn (personal communication) however considers this approach to fail for objections to linguistic realisation, and therefore that metalinguistic negation requires a descriptions which is *different* from that of descriptive negation. But that means that the descriptions of the two negations would not be analogous, which they ought to be if negation were pragmatic ambiguous in the same way as the other phenomena that Horn discusses.

(3.15) I promise not to come.

Clearly these sentences would not be used in the sense of 'EITHER I OBJECT TO HIS GOING TO THE OPERA/THE ASSUMPTION THAT HE IS AT THE OPERA TONIGHT...' or 'I PROMISE TO OBJECT TO THE ASSUMPTION THAT I WILL COME'.

There is also other evidence that he considers the ambiguity to be more than pragmatic. He gives what he calls three "diagnostics" of the different types of negation, which are other linguistic phenomena that he argues correlate with them (the diagnostics will be discussed in chapter 4). One of them is the selection of *"but"*-morpheme in languages that have two different ones, which according to Horn covary with his two categories of negation. But the difference between the two BUTs is not one of pragmatic ambiguity (which is typically not lexicalised), so the purported correlation, if it is significant, also indicates that the difference between the two negations is not merely one of pragmatic ambiguity.

3.3.3 Atlas and Kempson: the Return of the Truth Function

Despite differences in their general theories, Atlas (1977, 1981, 1989) and Kempson (1986) give relatively similar accounts of natural language negation. In contrast to Horn, they do not assign any significance to ambiguity in the explanation of it, but consider most of its uses to be adequately described by the bivalent truth function. Kempson acknowledges that there is metalinguistic use of negation, but reserves this description for what could be called "typical" objections to a linguistic realisation,[16] such as objections to e.g. pronunciation. Atlas does not discuss this type of data, and for him, natural language negation is invariably to be interpreted as a truth function.

[16]See chapter 4, section 4.2.3 for a discussion of negation as applying to linguistic realisation.

The idea behind these accounts is that if the reason why a negation is used is tht the information of the sentence in which it occurs is less than accurate with respect to the world, then the negation *can* be characterised as truth functional, and therefore *should* be. This includes not only the instances of negation that Horn would consider descriptive, but also those where the negation is presupposition canceling, or objects to the suitability of e.g. a scalar expression.

As shown in chapter 2, there are many acceptable utterances with negation where a straightforwardly derived propositional representation would be contradictory or inconsistent with the world, if it is taken to correspond to the ordinary truth function. In order to avoid these problems, propositional representations of utterances cannot simply be derived from the sentences, but a more complicated approach must be adopted.

Atlas describes the process using the concept of sense-generality (as discussed in section 3.2.3). Only when a sentence is uttered can a representation be constructed. For a sentence such as the one below, this means that *"eggs"* is sense-unspecified for 'EGGS ONLY' and 'EGGS (AS A PART)'.

(3.16) John didn't cook eggs, he cooked broccoli and eggs.

If this sentence was uttered, the first occurrence of *"eggs"* would be interpreted as ('EGGS ONLY'), as its representation would otherwise be contradictory.

Kempson's relevance based account is similar, except that she makes slightly stronger claims about the status of the different readings. Certain expressions may have two possible interpretations in a sentence, one "literal", and one strengthened by relevance criteria (i.e. required by the context in order to make the sentence consistent with the context and maximally informative). She only gives one example to illustrate this, but argues that it generalises to other cases:

(3.17) Mark didn't have three biscuits, he had four.

The literal meaning of *"three"* is, according to her, 'AT LEAST THREE', so the literal meaning of the first clause is that Mark had less than three biscuits. However, that is not consistent with the context (he had four), so *"three"* is strengthened to mean 'EXACTLY THREE'.

These two accounts shift the burden of explanation from the description of negation to the representation of the other expressions used in sentences. The former is kept relatively simple, in that negation is only considered to have one function, but is still so specific that it requires a lot of adjustment to the construction of representations.

3.3.4 Anscombre and Ducrot: Argumentative Properties

The last of the accounts of negation to be examined here is the one by Anscombre and Ducrot. The discussion of it will however be rather tentative, as they have not written specifically on negation for some time, while the framework in which the account is situated has developed considerably. Because of this, it is almost necessary to look back at how they were first described in order to understand the motivation for the way they are characterised in Anscombre and Ducrot (1977 and 1983).

They consider negation to have two functions: DESCRIPTIVE and POLEMIC. The first time a distinction similar to the current one is mentioned is in Ducrot (1973).[17] The difference between them is one of function, rather than of meaning (p.123, my translation):

[17]Ducrot (1972) also discusses two categories of negation, then labelled DESCRIPTIVE and METALINGUISTIC (hence Horn's terms) which more or less correspond to presupposition preserving and presupposition canceling negation as described in chapter 2, section 2.2.2. Horn's two negations are basically modelled on those in Ducrot (1973) rather than on this early account, but he prefers the term "metalinguistic".

> "we distinguish two types of negation. One polemic negation which corresponds to a speech act of negation, and appears as a refutation of a corresponding positive utterance (it is always, in this case, a phrasal negation). And on the other hand, a descriptive negation, which is an affirmation of a negative content without reference to an antithetical affirmation (we can have either phrasal or predicate negation here)."

The descriptive negation is considered to be semantically equivalent to the bivalent truth function. The polemic one is characterised as a speech act: it indicates that some speech act that has been or could be carried out using the same sentence without the negation in the context would be inappropriate. In Ducrot (1973), descriptive negation can be used polemically, as long as it is primarily used to present a fact, and does not have as immediate function to contradict. This is in contrast with Anscombre and Ducrot (1977), in which a descriptive negation must make no reference to an earlier assumption at all. But this stipulation still leaves a grey area of uses which could be considered either polemic (in a weak sense) or descriptive, such as a negative reply to a yes/no-question, e.g. *"Is Pierre French?"*.

However, their characterisation of negation is not complete with this, but it is also considered to have argumentative properties. Interestingly, it is descriptive negation which receives most attention in this aspect, as opposed to the majority of the other accounts of negation which tend to assume that the truth function is sufficient as a description of negation in those cases where it is compatible at all.

Descriptive negation is considered to be subject to two argumentative constraints: the LAW OF NEGATION and the LAW OF INVERSION (Anscombre and Ducrot 1983). The law of negation states that if an utterance of a sentence S would be considered an argument for a conclusion C in a particular context, then an utterance of NEG-S is an argument against C. This is possible since the argument relation is not defined in terms of factual constraints between situations, but in terms of the significance of uttering a sentence. If the argument relation corresponded to a logical relation, e.g. $s \rightarrow c$, then $\neg c$ could not be inferred from $\neg s$.

But the argument relation is about how an actual utterance referring to s, or ¬s, would be used. That is, if the purpose of the current conversation is to say something about c, it is known that s → c, and one speaker goes to the trouble of conveying ¬s (by saying NEG-S), it is quite easy to accept that the speaker would be trying to say that c is not likely, although that does not follow logically. That this should not primarily be viewed as a question of making the addressee perform certain deductions, in the sense of realising that their reasons to believe something has been undermined. Rather, it is a case of the addressee accepting that the conclusion does not obtain on the grounds that the speaker seems to think that it does not. The law of negation is about taking negative sentences as evidence for what the speaker thinks, and accept a representation on authority so to speak, not about taking sentences as input to reasoning.

However, the process is more subtle than this. It is not simply about drawing attention to certain propositions. Anscombre and Ducrot point out that two equivalent propositions, one of which is expressed using a negation and the other one not can have radically different effects. They use the following utterance, made by a French minister of state, to illustrate this:

(3.18) Les 3/4 des travailleurs touchés par des suppressions d'emploi ne connaîtront pas une situation de chômage.
The 3/4 of workers touched by [determiner] suppressions of work [negation particle] will-know not a situation of unemployment.
'Three quarters of the workers affected by redundancies will not experience unemployment.'

This is presented as an optimistic statement, as an argument for believing that the politicians are doing fine (at least if unemployment is considered undesirable). Contrast this with the effect it would have had if the minister had used the sentence in example 3.19 instead.

(3.19) Le 1/4 des travailleurs touchés par des suppressions d'emploi connaîtront une situation de chômage.
The 1/4 of workers touched by [determiner] suppressions of work will-know a situation of unemployment.
'One quarter of the workers affected by redundancies will experience unemployment.'

Although the propositions expressed by the sentences in examples 3.18 and 3.19 have the same truth conditions, they are clearly not equivalent for argumentative purposes. The action of negation with respect to argumentative orientation therefore is more complicated, and does not simply derive from the fact that the speaker found it relevant to draw attention to the situation. Anscombre and Ducrot suggest that descriptive negation has the property of reversing the argumentative orientation of a predication, so that two sentences S and NEG-S' have the opposite orientation if they have the same predicate and the subjects are of the same type.

There is one case where the law of negation is not supposed to hold for descriptive negation, and that is when the phrase it applies to contains a numerical expression. That is, if one considers a pound to be a cheap price for a cinema ticket, then both the sentences in example 3.20 would constitute arguments for going to see a film at that price.

(3.20) a. It costs a pound.
b. It doesn't cost a pound.

In Ducrot (1973), this was just considered to be an exception. In Anscombre and Ducrot (1983) they claim that they have different argumentative orientation with respect to some aspect of cinema ticket buying. The idea is that the (b)-sentence would be used if someone would be likely to assume that it was more expensive (using their notion of polyphony, see below), so the two sentences would have a different argumentative orientation for them. However, since they allow the notion

of argumentative orientation to apply also to potential use of sentences, both (a) and (b) must be considered to have the same orientation in some cases.[18]

The law of inversion says that if S and S' can be used as arguments for a conclusion C and S is stronger than S', then using NEG-S' amounts to making a stronger argument against C than using NEG-S.[19] This is supposed to reflect the different effects of the sentences in example 3.21 in some suitable context. Consider for instance a situation where two people are at a lake and one of them asks the other, who is already in the water, what it is like. Say that the person who is in the water interprets this as a question about whether the other person should go for a swim too, and uses a topos which correlates a good swim with the water having a pleasant temperature.

(3.21) a. It's freezing.
b. It's cold.
c. It's not cold.
d. It's not freezing.

[18]Thinking of sentences as potential utterances indicates that they consider them to have an independent meaning. They distance themselves from that assumption in Anscombre and Ducrot (1986). However, it is necessary to have a notion of potential utterance if one wants to talk of descriptive negation as reversing the argumentative orientation—this would not be possible unless it can be determined what argumentative orientation the corresponding sentence without negation would have in the same context.

[19]The notion of argumentative strength is not defined in Anscombre and Ducrot (1983). Moeschler (1989) suggests a definition in terms of sets of conclusions that the argument supports, e.g. S is stronger than S' if the set of conclusions C supported by S is a superset of the set C' supported by S'. For Anscombre and Ducrot, however, argumentative strength is a linguistic primitive, which cannot be defined, in that certain expressions are ordered with respect to their strength without reference to any factual situation (they can be used to modify a reference to a fact to signal how strong an argument it should be considered, see section 3.4). But references to different facts used as arguments can also be of different strength, in which case the strength must depend on the facts rather than on any intrinsic linguistic property.

The formulation of the law of inversion in Anscombre and Ducrot (1983) is actually considerably stronger than the version I have given here. It says that if S is an argument for C and S' is an argument for C', and the latter is stronger than the former, then the reverse is true for the negated sentences with respect to the opposite conclusions. They do not discuss any cases of where C and C' are different, which is why I have chosen the formulation given here. It is difficult to conceive of how the argumentative strength of different arguments with respect to different conclusions could be compared. If one tries, the law does not appear to hold anyway.

Anscombre and Ducrot argue for expressions such as *"cold"*, *"freezing"* etc., the difference between them is not what temperature they indicate, but their argumentative strength. If *"freezing"* is used rather than *"cold"*, then the speaker is trying to make a relatively strong argument against going for a swim. When these expressions are negated (with a descriptive negation), the ordering of them is reversed so that *"not cold"* is a stronger argument *for* going for a swim than *"not freezing"*.

It is noteworthy that Anscombre and Ducrot have stated the law of inversion as an argumentative one. If it were the case that *"freezing"* implied *"cold"*, and the negation is truth functional (which they accept that it is), then the negative phrases would denote different subsets of the temperature scale, with *"not cold"* situated in the warmer subset of *"not freezing"*. If the pleasantness of swimming is linked to the temperature of the water, the law of inversion should not need to be expressed as a particular argumentative constraint, but would follow from the truth functional description of the negation and the law of negation. But as noted, Anscombre and Ducrot (1983) do not consider expressions such as *"freezing"* and *"cold"* to imply each other. They simply have different argumentative strength (why they chose this approach will be explained in section 3.4), so the law of inversion is needed.[20]

Polemic negation is not bound by the law of negation. Ducrot (1973) writes that it may or may not reverse the argumentative orientation. Anscombre and Ducrot (1983) have very little to say about it, but they (perhaps unintentionally) strengthen the earlier description of it to require that it applies to something which has been explicitly or implicitly referred to, as opposed to *could* be referred to, and insist that it must be followed by a rectification.

The last property that Anscombre and Ducrot assign to negation is that it is

[20]Even if the arguments in question could be ordered according to some logical rule, the reason for the effect of negation would be the same as above, i.e. a reversal of the argumentative strength. But of course the varying argumentative strength might be a consequence of a factual constraint.

POLYPHONIC, i.e. if a speaker uses a sentence with a negation, they attribute the belief that the corresponding sentence without the negation would be appropriate to some other agent. It does not have to be an actual agent, but can be an imagined one.[21] For some reason, they are quite adamant that this only applies to descriptive negation.

Finally, there is an important aspect of the characterisation of descriptive negation that is problematic, and must be mentioned. Although it is defined functionally (as asserting negative information), it is nevertheless thought of a corresponding to the truth function of predicate logic. The problem with this is that in the last chapter of Anscombre and Ducrot (1983), and in particular in Anscombre and Ducrot (1986), they move away from propositional representations. Instead, they write that "the meaning of a phrase is the set of topoi whose application it legitimises in the situation where it is uttered" (1986, p.88), which makes it appear that their idea of of meaning is akin to a spread of activation in a network. But it is not obvious how the idea of a truth functional negation translates to this type of representation.

3.3.5 Comparing Negations

This section evaluates how the three accounts described in sections 3.3.2–3.3.4 fare with respect to the data discussed here and in chapter 2. The only concern here will be coverage: whether the categories as such can be justified will be discussed in chapter 4.

To begin with, they all assume that if a sentence with a negation is appropriately used, the negation indicates that there is something wrong with the sentence it occurs in—either that one of its truth conditions fails to obtain, or that there is some inadequacy with respect to the circumstances of its utterance. This means

[21] This resembles a formal description of Allwood's notion of expectancy. Some problems with this was noted in section 3.3.1.

that none of them is compatible with how negation is used in yes/no-questions. The are no obvious extensions or changes to them that would remedy this.

They also have individual problems. One of them is related to the way non-truth-functional negation is accounted for by Horn and by Anscombre and Ducrot. Both metalinguistic and polemic negation are defined as objections to a previous utterance. But speakers do not have to object to any feature of a previous statement to use a negation in such a way that it cannot be straightforwardly accounted for as truth functional:

(3.22) A: Who do you think will win?
B: David Gibson I suppose.
A: Sarah Phillips will get the ladies' prize.
B: Yeah, you don't expect her to win, you know she will.

B does not refer to any previous assertion that Sarah can be expected to win. Anscombre and Ducrot could in principle revert to Ducrot's earlier characterisation of polemic negation, and claim that it refers to a potential utterance on behalf of somebody else. But that would miss the point, because this utterance is not a rejection of someone's assumptions but a comparison of the probability that David will win with the probability that Sarah will.

It is possible to achieve observational adequacy by weakening "previous utterance" to "potential utterance", but while this ensures that the above example is not ruled out, it is unclear what it achieves from an explanatory point of view. If, for instance, this description of metalinguistic negation was used in a natural language generation system, when should the program postulate a potential utterance? And for that sake, when should an ordinary language user postulate one? A possible answer would be when there is some reason to believe that some agent could utter it because they would believe it to convey appropriate information. But that means redefining it in terms of information, and no longer simply in terms of linguistic activity, as the speaker who uses it must consider the *reasoning* that

lead to the inadequate statement (otherwise they would not know that it could be made).

It has been pointed out several times that it can be appropriate to use a negation to indicate that some representation is inaccurate even when the addressee is not thought to believe that it is. This has typically only been illustrated with data where either a straightforward propositional representation of the sentence is consistent with the world, or where there is a third party who could be assumed to hold the non-optimal belief. One such example (3.9 was discussed in section 3.3.1. Another one is given in Atlas (1980): a speaker can begin his talk by saying that *"I shall not speak for more than 45 minutes"*even if he or she is known to give short talks and there is no convention to speak for longer. Significantly, in both these examples, a simple representation of the sentence with a logical negation is accurate with respect to the situation.

But example 3.22 can be used to make an even stronger statement: it includes a sentence of the type that is typically used to illustrate metalinguistic negation, as it could seemingly be inferred that if one knows that something will happen, one also expects it. In this case, however, neither the addressee nor some third party holds the limited view that Sarah is only expected to win. The negation contrasts the two states of 'KNOWING' and 'EXPECTING', but does not object to a (possible) assertion of the validity of one of them with respect to Sarah's winning. The reason for drawing attention to the contrast is that the rejected description has been asserted (and accepted) about somebody else. That someone holds a less than optimal belief is thus not necessary factor for the use of the type of negative sentences where the corresponding positive sentence could be considered true either.

Atlas and Kempson do not have this particular problem, since they do not explain this type of usage by reference to a previous utterance. However, there are problems with their approach as well. If a sentence such as the one in example 3.22

is used, its representation is considered to contain a maximality marker of some sort, e.g. 'YOU DON'T JUST EXPECT...' . It is not necessary to realise the maximality marker linguistically in this example, as the addressee can be relied on to adopt an appropriate interpretation of 'EXPECT' (which is general with respect to wether the agent only expects or is in some stronger epistemological state which implies expecting). But then it must be explained why maximality markers cannot always be left out. If their reasoning is right, then a similar interpretation procedure should be available for B's reply in example 3.23, as 'KNOW' presumably is general with respect to whether the agent just knows about the existence of the object, or if they know because they caused it to exist. But in this example it would be odd to leave out the maximality marker, as in the B' reply.

(3.23) A: Does the head know about it?
B: He doesn't just know about it, he arranged it.
B': *He doesn't know about it, he arranged it.

Another problem is the choice of possible interpretations as such. Recall that Atlas suggested that the correct interpretation of the sentence in example 3.16 (repeated here) is obtained by choosing specific representations of the sense-general word *"eggs"*, namely 'EGGS ONLY' for the first occurrence, and 'EGGS (AS A PART)' for the second one.

(3.16) John didn't cook eggs, he cooked broccoli and eggs.

Why is it that these particular specifications are chosen? For instance, the word *"eggs"* must also be general with respect to colour, i.e. it could be used both about brown and white eggs. Sense-generality does not explain why *"John did not cook only eggs only, but eggs and broccoli"* seems to be a more reasonable type of alternative formulation than *"John didn't cook white eggs, he cooked brown eggs and broccoli"*. The latter one would be exactly as likely on Atlas's account

of sense-generality. But it is implausible that anyone would want to make that type of corresponding without explicitly mentioning both colours, i.e. it is unlikely that anyone would use the sentence in example 3.24, which is parallel to that in example 3.16 in order to distinguish between brown and white ones.[22]

(3.24) ?John didn't cook eggs, he cooked brown eggs.

If it is the maximality that is the relevant feature in example 3.16 (as suggested by Atlas), then it must be explained why this is such a central feature, and indeed differs from other types of properties with respect to which a word is general. Sense-generality is for this reason at best a description of a property of linguistic labelling. It does not reflect the cognitive procedure which lead to the interpretation of example 3.16.

Horn, Kempson, and Atlas lack a possibility to describe the sentences in examples 3.18 and 3.19 (repeated here) as different.

(3.18) Les 3/4 des travailleurs touchés par des suppressions d'emploi ne connaîtront pas une situation de chômage.
The 3/4 of workers touched by [determiner] suppressions of work [negation particle] will-know not a situation of unemployment.
'Three quarters of the workers affected by redundancies will not experience unemployment.'

(3.19) Le 1/4 des travailleurs touchés par des suppressions d'emploi connaîtront une situation de chômage.
The 1/4 of workers touched by [determiner] suppressions of work will-know a situation of unemployment.
'One quarter of the workers affected by redundancies will experience unemployment.'

[22]My intution is that the sentence could only be used if the speaker considered there to be a quality difference associated with the colours.

This is inevitable on at least Kempson's account—if two propositions that are logically equivalent are fed into the deductive mechanism, the agent cannot possibly arrive at different conclusions. Horn does not offer any solution. Atlas has argued elsewhere (Atlas and Levinson, 1981) that sentences that have identical truth conditions do not need to have identical representations (and can therefore have distinct meanings), but it still remains to explain *why* these sentences are so different.

It should be said however that Anscombre and Ducrot do not *explain* this either, they merely attribute to negation the property of causing this effect. But at least their framework permits the formulation of this type of feature, and treats them as significant. If the purpose of utterances are to convey propositions, and negation can be described exhaustively as a truth function, then it is not possible to distinguish between two utterances which convey what are obviously equivalent propositions.

Despite the shortcomings of the pragmatic descriptions of negation discussed in section 3.3.1, they have an important advantage over a truth functional account of negation, in that they furnish an explanation of why negation should be used at all. By describing negation as indicating that someone's expectations are not met, a reason for using it can be inferred—to make sure that no participant has a substandard representation of the discourse universe. Unless it is accompanied by a plausible account of why speakers would want to convey negative facts, the truth functional description of negation is of little use for the generation of utterances, as well as leading to problems with their interpretation.

This issue is relevant if one wants to use the description of negation in e.g. a natural language interface to a database. If the purpose of the system is to supply information to the user, should it also supply negative facts? Are there any circumstances where negative sentences would be appropriate? When faced with this question, it is likely that even the most ardent semanticist would find themselves taking a similar position to that of Givón: negative sentences are useful

for indicating that the addressee has an erroneous belief. That is, it is in the end considered to have a rather significant pragmatic component to its description.

This assumption might work acceptably well in a simple system, but it does not capture the use of negation in general, as shown above. It would be desirable that a theory of negation be explanatory in the sense that it has some intrinsic property which explains why speakers use it. On this account, a truth functional description does not have much to offer.

In contrast, Anscombre's and Ducrot's approach lends itself more easily to these concerns. Polemic negation could be hypothesised to be used to object to an erroneous belief, while descriptive negation would typically be expected to be used when a speaker thinks their addressee may draw some unwanted inference—it could prevent the addressee from using the information referred to by the corresponding positive utterance as input to some reasoning process. Viewed from this perspective, Anscombre's and Ducrot's polemic negation would correspond to the usage described by Givón , and their descriptive negation approximately to Allwood's restriction that the corresponding sentence without a negation must be expected to be accurate (minus those cases where it is used for explicit objection). It is, in other words possible to make predictions about when a speaker would want to use a negation (whether these predictions are accurate, and explanatory, is of course a different matter).

Anscombre and Ducrot argue that language can be used for argumentation at an autonomous level, in that some expressions have functions in the discourse which are not derived from any informational property. Because of this, their position is radically different from Horn's and Kempson's, who both emphasise the relation between language and reasoning in a fairly Gricean spirit (it is difficult to place Atlas with respect to this dimension, as he is less specific about how the interpretations are arrived at). Relevance theory and Radical Pragmatics are closer to Argumentation theory than the traditional semantic approach is in some aspects, but they are still committed to sentences as corresponding to factual

descriptions.

With Relevance theory and Radical Pragmatics somewhere between the other two, it should come as a surprise that when it comes to negation, Anscombre's and Ducrot's account is closer to Horn's than to Atlas' and Kempson's (ignoring the argumentative properties for the moment). The reason for this, apart from the fact that Horn modelled his account on Ducrot's, is that the non-argumentative properties that Anscombre and Ducrot attribute to negation were specified a long time ago, when they were assigning much more importance to factual constraints. In many ways, this part of the specification has outlived itself, and comes very close to being inconsistent with how Argumentation Theory is described in Anscombre and Ducrot (1986), as pointed out in the end of the previous section.

It is hardly necessary for them to distinguish between two types of negation anymore. Given that they do not consider it utterances to be interpreted as (sets of) propositions, there is no reason to single out a negation whose main characteristic is that it corresponds to a truth function. As for the functional definitions of the two, Argumentation theory does not use any features which depend on whether the utterance is used as an objection or not.[23] The only significant property of the negations that is used is their, effect on the argumentative orientation of the sentence. But as noted previously, the definition of descriptive negation as reversing the argumentative orientation hinges on the possibility of considering "potential utterances", which is not really consistent with their general approach. Moreover, as suggested by example 3.22, descriptive negation (which is defined as not objecting) does not always reverse argumentative orientation, and polemic negation may either reverse or preserve it, so it would be legitimate to ask whether there is any reason to distinguish the two. The argumentative orientation of utterances with negation could depend on something else.

[23]They probably used to consider this to be a property of the two BUTs though, see chapter 4.

3.4 Scalar Expressions—A Case Study

3.4.1 The Problem

Sometimes it seems natural to think of a set of linguistic expressions as being ordered. One such set of expressions are *"freezing"*, *"cold"* and *"fresh"*. They relate to a certain quality (temperature), and if *"freezing"* is used about some object, it suggests that the object has more of that quality than it would have if *"cold"* had been used. Such words are often called scalar expressions, as they appear to denote positions on a scale. A few examples are given in example 3.25, together with some expressions that could belong to the relevant set.

(3.25)	a.	The water is cold.	<fresh,cold,freezing>
	b.	Pat has three children.	<one,two,three,...>
	c.	It's probable.	<possible,probable,certain>
	d.	Some students submit their PhDs in time.	<some,most,all>
	e.	She is pleased.	<satisfied,pleased,delighted>
	f.	He is unhappy.	<unhappy,depressed>

How to interpret scalar expressions in natural language sentences is one of the more frequently debated issues in pragmatics. The reason for this is that there appears to be a discrepancy between the literal meaning of the sentence, and what it is usually taken to mean when it is used. If uttering a sentence is equivalent to conveying that the proposition it refers to is true, then a sentence like *"Pat has three children"* would be correctly used if the number of children belonging to Pat is at least three.[24] However, it seems that in the type of contexts where such a

[24] At least this is the common position in the literature on the topic. There are however several ways of representing such sentences. For instance, if it is treated as a statement about the cardinality of the set of children belonging to Pat, then it would not be true if she in fact had four. In mathematical text, it is common to use the numbers on their own as doing this,

sentence would be used, it would be natural to interpret it as a statement that Pat has exactly three children.

The same holds for the other examples: assuming that it is possible to specify exactly what *"cold"* means, an utterance of example 3.25(a) would entitle the addressee to assume that the water was not such that it could be called *"freezing"*, but just cold. In other words, there is a discrepancy between the chosen primary meaning and the preferred interpretation.

As such, scalar expressions are not of primary interest to this thesis. However, the way they have been accounted for by some of the proponents of the different theories of negation examined in section 3.3 highlights and may clarify the differences between their approaches to describing language. Horn, Kempson and Anscombre and Ducrot all propose quite different accounts of scalar expressions, and their interaction with negation. The discussion of their accounts has so far been quite terse, and it is hoped that by using this case study as an illustration, the difference between them can be more easily appreciated.

3.4.2 Scalar Expressions and Negation

Scalar expressions pose an interesting challenge to a truth functional description of negation. It was noted above that the interpretation of the sentence in example 3.25(b) is often more specific than its truth conditions. For the corresponding sentence with a negation however, the default interpretation and its truth conditions seem to coincide, i.e. the sentence is interpreted as meaning that Pat has less than three children.

---

i.e. *"there are three elements which fulfil this condition"* does not mean 'THERE ARE AT LEAST THREE ELEMENTS WHICH FULFIL THIS CONDITION'. However, it will be assumed for the sake of the argument that the sentence *"Pat has three children"* has a propositional representation which is true when Pat has at least three children, as this is the representation that linguists have chosen.

(3.26) Pat doesn't have three children.

However, that is not the only possible interpretation of that sentence. If it was followed by an assertion that Pat has more than three children, that interpretation obviously is not available.

(3.27) Pat doesn't have three children, she has four.

In this case, the meaning is quite different from the truth conditions normally associated with the sentence. When interpreting an utterance of this type, one may first be lead down the garden path, and assume a different representation than the speaker intended. But after the utterance there is no question of what the speaker meant, and the addressee is not left with the impression that the speaker said something contradictory. Given this, it seems that one of the following hypotheses about the representation of sentences must be abandoned.

H I. *"Not"* corresponds uniformly to the bivalent negation of predicate calculus.

H II. A set of expressions which can be taken to denote positions on the same scale can be ordered by implication.[25]

H III. The meaning of natural language sentences can be derived from the linguistic form.

The three accounts which will be discussed here differ in which hypotheses they reject: for Horn (1989) it is the first one, for Kempson (1986) the second and the third, and for Anscombre and Ducrot, all of them.

[25]In fact, it is enough that *some* scales have this property, since it suffices that negation can apply to *a* logically ordered scale in the way it does in example 3.27 for it to be non-truth-functional.

3.4.3 Scalar Implicature: Reasoning with the Expressions

According to Horn (1989), scalar expressions should be interpreted as assigning a "lower bound"[26] to how much of a certain quality, such as coldness, applies to an object. He writes about numbers that (1989, p. 214):

> Cardinals like 3 are lower-bounded by their literal or conventional meaning; hence [Pat has three children] means (is true iff) Pat has at least three children.

What he means by this is that by saying *"the water is cold"*, one has effectively only conveyed that it is at least cold (assuming that what one conveys is what one means, which, to judge from the quote, Horn considers to be equivalent to the truth conditions of the sentence. But when an addressee hears the sentence, they are usually permitted to infer that the speaker has communicated an upper limit of the quality as well, so that the water is in fact not any colder than what could be described as exactly *"cold"*, i.e. *"cold"* applies but *"freezing"* does not.

Horn favours a Gricean account of the process by which this happens (Grice, 1975). Essentially, the "exactly" reading is arrived at by appealing to some imperatives governing communication which stipulate that speakers should make their contributions as informative as required by the context, and that they should only convey information which they believe to be true. A well known example of such rules are Grice's (1975) conversational maxims, the relevant ones being the maxim of quantity (first submaxim), and the maxim of quality:[27]

The Maxim of Quantity

[26] It is unfortunate to use "lower bound" in this sense, as it is usually taken to mean that no entity less than the lower bound can be included in a set. For Horn however, if *"freezing"* has been used about an object, then *"cold"* also applies.

[27] Horn uses his own version of these.

1. Make your contribution as informative as is required (for the current purposes of the exchange).

The Maxim of Quality: try to make your contribution one that is true.

1. Do not say what you believe to be false.
2. Do not say that for which you lack evidence.

Horn proposes that while a scalar expression *e* means 'AT LEAST E', 'EXACTLY E' can be derived using conversational implicatures,[28] which arise as a result of assuming that other agents obey the maxims because they are being cooperative. The 'EXACTLY THREE' reading of example 3.25(b) (*"Pat has three children"*) is arrived at by reasoning in approximately the following steps:

i. It would have been possible to make a stronger statement (i.e. *"Pat has four children"*).

ii. If the speaker had believed that it is true that Pat has four children and that it is in the addressee's interest to know that, they would have used that sentence.

iii. For all the speaker knows, Pat does not have more than three kids.

He argues that his account generalises to "all weak scalar operators, including cardinal numbers and evaluative or gradable adjectives such as *good*" (p.213).

Although the maxims appear to refer only to interactive behaviour (i.e. how the speaker's linguistic behaviour should be from the perspective of being cooperative), there are also references to *how* language means—the account of scalar expressions depends on there being some kind of ordering of them. This is because of the notion of "informativeness" in the maxim of quantity: it must be possible

[28]Actually, he writes (p. 212) "context-dependent generalised conversational implicature" which is a contradiction. Generalised conversational implicatures are by definition not context dependent. Jay Atlas (personal communication) suggest that he meant generalised conversational implicature despite the reference to context.

to order the scalar expressions according to how informative they are. On Horn's account, an expression is more informative than another if it implies it.

For some types of scalar expressions it is quite easy to accept such an order, but for others it is less obvious. Some expressions that exhibit the type of behaviour associated with scales cannot be associated with some universally applicable scale available, as illustrated by the sentence below:

(3.28) Most photographers were inarticulate if not subhuman.

It is not generally the case that *"subhuman"* and *"inarticulate"* are thought of as positions on the same scale, or even on a scale at all. Horn (1989) suggests that they can however be ordered that way with respect to a certain context (this approach was originally suggested by Fauconnier (1975), who labels such local orders pragmatic scales). The expressions are considered to imply each other given the appropriate background constraints.

It remains now to explain the behaviour of scalar expressions with negation. The issue is how to obtain for instance the two different readings 'PAT HAS LESS THAN THREE CHILDREN' and 'THE NUMBER OF PAT'S CHILDREN IS NOT THREE' respectively for the two examples below.

(3.26) Pat doesn't have three children.

(3.27) Pat doesn't have three children, she has four.

Horn's solution is to reject the hypothesis that negation unambiguously corresponds to the bivalent truth function. That means that the negation in the former example is interpreted as a truth function, and hence *"not three"* as meaning

'NOT AT LEAST THREE' (= less than three).[29] The first clause of the sentence in example 3.27 may first be interpreted in the same way, but when this is found to be inconsistent with the second clause, the negation is reinterpreted as a metalinguistic negation. This means that it is considered to be an objection to a previous utterance, so the first clause does not have an ordinary semantic representation. Its representation must contain a metalinguistic operator and something which denotes an utterance.

3.4.4 Relevance: Reasoning with the Context

For Kempson, the intuition that negation corresponds to the bivalent truth function is so strong that she prefers to reject another one of the three hypotheses listed on page 83. She agrees with Horn that a different analysis of negation is needed for the cases where it appears to apply to e.g. the pronunciation of a word, but not for example such as 3.27. She writes (1986, p.88):

> I conclude that though a metalinguistic analysis is available for these paradoxical negation cases, the problems they present cannot be pushed aside by only providing a metalinguistic analysis; they can, and naturally are interpreted as straightforward cases of descriptive negation.

Instead, she rejects the hypothesis that natural language sentences have representations that can be derived from their linguistic form (H III above). She also partially rejects H II, that scalar expressions can be ordered. The expressions denote positions on a scale, and those positions can be ordered universally or locally, but the expressions as such do not imply each other.

[29]He does not explain exactly why *"three"* does not get interpreted as 'EXACTLY THREE' in this example. The latter is a stronger statement, but there is no reason why one should prefer a stronger interpretation per se, as the maxim of quantity is not defined in terms of maximal strength, but in terms of strength relative to the context. It could be argued that the addressee should select the 'LESS THAN THREE' interpretation because that is stronger than 'NOT EXACTLY THREE', but the maxims could equally well used to argue the opposite point ("if the speaker had meant the stronger statement, he or she would have used it").

According to her, a scalar expression e denoting the quantity E receives either of two readings, which are 'AT LEAST E' and 'EXACTLY E'. She exemplifies the process of how these readings are assigned using the following example:

(3.29) A: How many children do you have?
B: I have two children.

If the context of these utterances is one in which B wants to claim a state benefit, which happens to be a fixed amount for all families which have two or more children, then the 'AT LEAST TWO' reading will be selected. This is because all the "interesting" inferences that can drawn from the number of children that B has can be drawn from this weak statement. It wouldn't matter if B had four or five children, as long as there are definitely two of them. If, on the other hand, the dialogue was uttered among friends, while talking about families in general, the 'EXACTLY TWO' reading should be selected. Kempson explains this by arguing that the information that one can infer about someone's family life depends on the number of children they have. If A is not allowed to assume that B has exactly two children, A cannot infer enough of B's domestic situation for the utterance to be informative.

Despite insisting earlier that scalar expressions are not ambiguous between two readings, and that neither interpretation should be chosen at the decoding stage, Kempson considers them to have different status. The "at least" reading is the "lexically stipulated meaning", whereas the "exactly" reading is the one obtained by applying relevance criteria (1986, p. 97). From this it appears that the two readings are after all not alternatives, but that just as in Horn's account, the "at least" reading is the actual meaning of scalar predicates, which is potentially restricted by the context.

The reference to the "at least" interpretation as being more basic is made in relation to the explanation of the interaction of scalar predicates with negation.

The idea that the "exactly" reading is more difficult to arrive at is used to explain why sentences such as the one in example 3.26 is usually interpreted as meaning 'LESS THAN THREE CHILDREN'. Kempson writes (1986, p. 97):

> The explanation of why, in negative sentences, the 'at least' reading is normally predominant follows directly from the principle of relevance. In order to be able to interpret an utterance of 'Mark didn't eat three biscuits' as compatible with 'Mark ate three biscuits' [sic] one has to be able to process both the linguistic meaning of the sentence and its narrowed 'exactly 3' interpretation, simultaneously, and yet separately, as it were, ignoring the lexically stipulated meaning in favour of the relevance restricted one. Other things being equal then, this predicts that a negative sentence containing a numeral [sic] construed as within the scope of the negative element will be interpreted as 'at least n'.

If scalar expressions have a conventional meaning which can be restricted, then it makes sense to claim that the "at least" reading is more common, because the cognitive cost (as measured in "increased processing complication") of picking that one will be lower. This is because the agent does not need to carry out the processing that leads to the relevance restricted reading. However, this predicts that the more common reading of the corresponding sentence without the negation is the more expensive one.[30]

3.4.5 Argumentative Scales: Determining the Reasoning

Anscombre's and Ducrot's (1983) account of scalar expressions is quite different from the previous two, in that they consider the interpretations of sentences that contain them to derive from argumentative properties instead of informational ones. Rather than using scales which are ordered by logical properties of the relevant concepts, they consider the expressions to be ordered for communicative

[30] A similar contradiction arises with the use of Grice's maxims—it does not seem possible to formulate any rules of this type which will generate the right interpretation for both sentences with and without negation. This was noted e.g. by Atlas and Levinson (1981).

purposes. That is, the most, and sometimes only, significant property of many expressions which appear to be ordered is that they have different argumentative strength with respect to a quality. This means that the perceived order of e.g. <fresh,cold,freezing> reflects the impact of the expressions when used in argumentation, rather than denoting a position on a scale corresponding to increasing or decreasing quantities of some external quality. The claim is not that such orderings do not have any relation to external properties, only that they are not isomorphic to one. For instance, the expressions <fresh,cold,freezing> are used to make arguments of various strength with respect to coldness, not to correspond to decreasing temperatures.

It might seem that the distinction between argumentative ordering and logical ordering is unnecessary. Horn (1989) dismisses argumentative strength as being equivalent to an external, implication based ordering, since he considers that an argumentative ordering always must depend on an external, implication-based one. However, the distinction is non-trivial, and there are some advantages of adopting Anscombre's and Ducrot's account when it comes to explaining the behaviour of certain expressions.

One set of expressions that are often thought of as being of the same type and related by implication are <permitted,compulsory>. If the sentence below is uttered about an action, one tends to assume that it is forbidden.

(3.30) It is not permitted.

This could be explained logically, if it is assumed that for all actions, any agent stands in either of three relations to them: either they must perform it, they must not, or they have a choice. Then if 'COMPULSORY' implies 'PERMITTED', denying that something is 'PERMITTED' also excludes that any predicate which implies it could apply.[31] Hence the sentence in example 3.30 excludes the possibilities that

[31]I will try to adhere as strictly as I can to distinguishing between expressions (e.g. *"permit-*

the agent has a choice whether to perform the action or must do it, and this is why an utterance of the sentence typically would be interpreted as suggesting that the action is forbidden.

Anscombre and Ducrot give a different explanation. They argue that by predicating about an action that it is *"permitted"* or *"compulsory"*, one is typically making an argument for performing it.[32] The latter expression is stronger than the former, so with a descriptive negation (it must be assumed to be descriptive in the absence of an utterance that it could be considered to object to) the sentence will count as a strong argument against performing the action (by the law of inversion).[33]

This is only a comparative judgement, but presumably, by making the same assumption as for the logical explanation above (that the agent must, mustn't, or has a choice whether to perform the action), the effect of using *"not permitted"* in a discourse can be similar to that of using *"forbidden"*. That is, the exclusion of the agent having a choice whether to perform the action is made by the actual assertion; and the exclusion of the agent having to perform the action is achieved by assuming that the negation is descriptive, finding the relevant topoi, and applying the law of inversion. The equating of that expression with an assertion that the action is forbidden must however be done by reference to the statuses of performing the action as such.

---

ted") from representations of words ('PERMITTED') in the discussion in this section. This is because of the fact that neither Kempson, nor Anscombre and Ducrot consider there to be a direct correspondence between the two. For Horn this does not matter; if two concepts imply each other, the corresponding words do as well.

[32] To be more precise, the words can be used in this sense given the right topoi. Without access to a particular context, a standard, or minimal context is used, in which people do what is compulsory. In a more particular context, for instance talking about Anne, who likes to break rules, the observation that something is compulsory could be an argument for believing that she avoids doing it. However, without access to such a context, one tends to assume a general one, in which people are law abiding. The semantically defined scales would obviously have to be sensitive to exceptions like this too.

[33] Note that for Anscombre and Ducrot this is not a property of reasoning with the predicates, but a property of using the expressions. It is not a result of *"not permitted"* undermining the reasons to perform the action, but a general property of negating an expression, captured by the law of inversion.

Which one of the approaches that is chosen does not seem to matter in this case, but consider the following example:

(3.31) It is not optional.

"Optional" and *"permitted"* seem to denote similar states of affairs (where the agent has a choice as to whether to perform an action or not). If something is 'OPTIONAL', it should be 'PERMITTED' too. Despite this, the sentence in example 3.31 tends to be interpreted differently from the one in example 3.30—it suggests that the action is compulsory. If this is to be explained logically, in the same way as the sentence in example 3.30, then 'FORBIDDEN' must imply 'OPTIONAL'. But this is counterintuitive, in particular since it means that 'FORBIDDEN' would imply 'PERMITTED' too.

By characterising <optional, forbidden> as an argumentative scale instead, the default interpretation can be accounted for. While <permitted,compulsory> are typically predicated of actions as an argument for performing them, these two tend to be used against performing them. This can be seen by considering the combinations in example 3.32, where the status of the action of giving money is presented as being in opposition with what someone actually did.[34]

(3.32)
a. ?Although giving money was permitted, she did.
b. Although giving money was permitted, she didn't.
c. ?Although giving money was compulsory, she did.
d. Although giving money was compulsory, she didn't.
e. Although giving money was optional, she did.
f. ?Although giving money was optional, she didn't.

[34] *"x although y"* indicates that *"y"* is an argument against *"x"*. Note that some of the examples which are preceded by a ? would be acceptable given a different topos in a different context (such as the one mentioned in footnote 32).

g. Although giving money was forbidden, she did.

h. ?Although giving money was forbidden, she didn't.

Presenting the fact that an action is *"permitted"* or *"compulsory"* as an argument against performing it makes for odd utterances (given a context where people are assumed to comply with regulations). The corresponding sentences using the expressions *"optional"* or *"forbidden"* are however quite natural. Anscombre and Ducrot argue that although it would be absurd to claim that *"forbidden"* implies *"optional"* (or that their representations do), they support the same type of conclusions (have the same argumentative orientation), and *"forbidden"* is stronger than *"optional"*. The negation in example 3.31, just as in example 3.30, reverses the orientation, and hence *"not optional"* is interpreted as a strong argument *for* performing the action.

Anscombre and Ducrot do not discuss examples such as *"Pat has three children"* above, but a plausible account of it in Argumentation theory would be to assume that agents will make as strong arguments as possible (while ideally being truthful), with respect to the topos used in the context.

As for the interaction between scalar expressions and negation, their account is less convincing. Similarly to Horn, they propose that this is a different type of negation, and that speakers only use it when they are objecting to a previous utterance. But there is no need to do this, as using a descriptive negation could not lead to a factual contradiction anyway, by their assumption that linguistic expressions do not imply one another. If the negation were descriptive, there would obviously be a clash in argumentative orientation, but that is only by their stipulation that the two negations are different. As pointed out in section 3.3.5, that claim is not supported by data.

The consequence of their approach is that the negation used in example 3.26 is conceptually different from the one used in example 3.27, despite the only "real" difference being the argumentative orientation. The law of inversion obviously

does not apply in example 3.27, but it seems drastic to solve this by redefining the negation pragmatically. A more plausible approach would be to generalise the description of negation and reformulate the law.

3.4.6 Summary

Horn and Kempson consider sentences with scalar expressions to have two different interpretations: one in which the expression is taken to have an "at least" reading and one in which it has an "exactly" reading. They both treat scalar expressions as if the problem is whether the "exactly" reading should apply or not. The "at least" reading is assumed by both of them to be more primary than the other one.

Why, if one is allowed to infer the "exactly" reading from all the sentences in example 3.25, is this not considered to be their meaning? To someone who is not a linguist or a logician this must seem odd—it would be more natural to maintain that the "exactly" reading is the primary one, and that the "at least" one is permitted in retrospect, if it should turn out that the water is freezing or Pat has four kids. But the problem with considering the "exactly" reading primary is that the following sentences would be contradictory.

(3.33) a. The water is cold, in fact freezing.
b. Pat has three children, in fact four.
c. She's pleased with her job offer. In fact absolutely thrilled.

They are not inconsistent on the "at least" reading, so this is a reason to prefer it. However, adopting the "at least" reading as primary can cause exactly the same type of problems as the adoption of the "exactly" reading would for the examples above. Consider the following example:

(3.34) Morten wants to run a marathon in three hours.

This sentence should clearly be interpreted as Morten aiming for a time less than three hours. It is not possible to attribute this reversal of the order to any semantic feature of the sentence. The only candidate for causing this would be the preposition *"in"*, which would then be assumed to mean 'LESS THAN'.[35] But this position is not tenable either, as illustrated by the following sentence:

(3.35) Doing ten miles in more than thirty minutes is pretty bad.

This sentence does not mean that only exactly thirty minutes is a bad time, but rather that any time from thirty minutes upwards is, so *"in"* cannot mean 'LESS THAN'. The interpretation of the sentence in example 3.34 cannot be attributed to the linguistic realisation of the sentence, but it is obtained from using world knowledge about how time is important in fixed distance sporting events.

Another example of why it would be a disadvantage to consider the "at least" reading primary is that some numbers can be used in yet another sense:

(3.36) Lisa has twenty rabbits.

References to certain numbers *"n"* are often used to mean 'APPROXIMATELY N'. This sentence could be used to say that Lisa has a lot of rabbits, somewhere around 20, which is so many that one more or less does not matter. The sentence could be used to make an acceptable utterance even if Lisa only has 19 rabbits.

These examples show that there are more than two possible interpretations of at least numerical modifiers. None of these can be considered more basic ('literal')

[35] For this to work, the default reading of a scalar expression *e* has to be 'E', rather than 'AT LEAST E', or else the combination *"in e hours"* could only mean 'EXACTLY E HOURS'. I am not sure why Kempson chose to call it the "at least" reading.

than the others, as it is not possible to define all the readings in terms of a single one of them.

The problem with an approach such as Horn's is that although it takes context into account when there is no obvious order of the descriptions, cases where there seem to be default interpretations are considered to derive from reasoning with the expressions as such exclusively. But this means that e.g. *"three"* in example 3.34 and *"Pat has three children"* will be interpreted in the same way, which is obviously not appropriate. The context cannot be ignored.

The Relevance account is more flexible, since the specific interpretations are attributed to context rather than to the expressions (at least in principle—in practice Kempson's account is quite similar to Horn's). It would however be necessary to permit other interpretations, apart from the 'AT LEAST' and 'EXACTLY' ones.

However, both these accounts treat scalar expressions as if their use can be described in terms of entailments or contextually motivated implications. It was demonstrated in section 3.4.5 that this cannot explain the seemingly scalar behaviour of expressions such as *"forbidden"* and *"optional"*. The failure of the data based approaches for these expressions suggest that argumentative properties may indeed be cause of our intuitions that some expressions can be ordered.

3.5 Summary

This chapter has reviewed some different approaches to describing natural language, in terms of the general assumptions that are made about communication, and at a more specific level, how they account for negation, and the use and interpretation of scalar expressions. A distinction was made between data oriented approaches and manipulation based ones. The former assume that the function of language in communication is to refer to factual descriptions, which speakers

convey to addressees, whereas the latter views the function of language as a tool for making the addressee process information in a certain way.

Although this is perhaps not a intrinsic feature of data oriented approaches, the ones discussed here all attribute a great deal of specificity to the meaning of sentences when they are uttered. This is because uttering a sentence is seen as equivalent to conveying a proposition which has a truth value. This makes it necessary to attribute very detailed specifications to what is in principle vague expressions.

Both Radical Pragmatics and Relevance theory acknowledge that sentences are general, and only acquire a specific interpretation when they are used. It would be natural to assume that the sentences remained general even when they are used, and that the specific interpretation only is a property of the agent's representation of the world. But because certain expressions, such as negation, are considered to have a logical function in the language, it is necessary that the rest of the sentences (what negation applies to) correspond to a logical form as well. The consequence of this is that an addressee's processing of a sentence is effectively considered to be fed back into it.

As opposed to the data oriented accounts, Argumentation theory does not view communication as a matter of conveying propositions. Instead, utterances are used to constrain the addressee's representation. Like Relevance theory and Radical Pragmatics, the meaning of a sentence depends on the context where it is interpreted. However, they make the additional claim that certain expressions of language directly influence which type of context should be chosen, without requiring the addressee to select the particular interpretation by reasoning. Argumentation is about indicating *that* there is a reasoning.

Argumentation theory has a lot to offer, for instance the possibility of viewing what is obviously acceptable discourse as coherent although the facts that the sentences correspond to are incompatible. In can also help explain the function

of connectives. What it did not offer was a convincing account of negation. It is however possible to see how a different characterisation could be given using a manipulation based approach to language, and this potential will be explored in the following chapters.

Chapter 4

Negation and its Linguistic Environment

4.1 Introduction

This chapter investigates how negation interacts with other linguistic factors. Firstly, it will be considered whether there is any motivation for postulating that negation has two different functions from the perspective of requirements on the linguistic environment. If it could be shown that there are two different sets of constraints imposed by *"not"*, then that could support a claim that negation has two distinct functions.

Secondly, even if functional ambiguity cannot be supported, constraints on the linguistic environment are interesting with respect to the description of negation as such, and the rest of the chapter will examine various relevant correlations. Some of the relevant phenomena have already been discussed in relation to negation, such as its interaction with *"but"* (Anscombre and Ducrot, 1977; Horn, 1989). Other interesting data have not been widely taken into account. The goal of

analysing these constraints is to specify a set of requirements on a representational framework such that a characterisation of negation can be given.

4.2 A Well Founded Split?

4.2.1 Does Negation Have More Than One Function?

If it is accepted that natural language negation is ambiguous (pragmatically or otherwise), the two core functions are usually taken to be the reversal of truth values and the indication of erroneous linguistic realisation. But where the precise boundary between the functions is drawn varies between different accounts. For instance, both Horn (1989) and Kempson (1986) acknowledge a truth functional and a metalinguistic use of negation but they differ on which category some of the data belongs to. Horn distinguishes between truth functional and non-truth-functional use of negation, which means that a distinction is made between sentences in which negated information is incompatible and sentences in which negated information is non-optimal. Kempson prefers to draw the line between negation which applies to information, and negation that applies to language, i.e. as long as it is the information and not the language that is non-optimal, there should be no metalinguistic interpretation of the negation. Horn cites various linguistic evidence for his distinction, while Kempson relies on the difference between information and language at the representational level. This section examines whether their positions are justified.

4.2.2 Incompatibility vs Non-Optimality

4.2.2.1 Preamble

It was shown in chapter 2 that natural language negation can be used to indicate a wide range of inaccuracies of sentences. Ignoring for the time being those inaccuracies that derive from the linguistic realisation, there are at least two cases where some information can be negated. It may be that the information is incompatible with the speaker's representation of the discourse universe (illustrated by example 4.1), or that it is not the optimal characterisation of the situation in question (example 4.2).

(4.1) A: I thought Megan was back in New Zealand.
B: She's not. I saw her in Negociants the other day.

(4.2) Graeme Obree comments on his choice of gears for a hilly cycle race:

I wasn't a bit overgeared, it was a big bit.

Horn (1989), as well as Anscombre and Ducrot (1977, 1983), considers this distinction relevant for the description of negation: the *"not"* in the former example is different from that in the latter. This section considers whether there is any linguistic evidence in support of this distinction.

A couple of points will need clarifying prior to commencing this discussion. Firstly, Horn sometimes suggests that the difference between the two negations is one of pragmatic ambiguity. Now, if this were the case, then an examination of possible differences in linguistic environments would be of little use, as the data that has been cited as instances of pragmatic ambiguity typically does not exhibit distinct linguistic characteristics correlated with the two uses. It is the concept of pragmatic ambiguity as such that would need defending. As for Horn's two

negations, they are not analogously defined, and if they have different meanings and exhibit different linguistic behaviour, then it is an ambiguity of some other type. This, and similar points have been made in several replies to Horn (1985), e.g. by Burton-Roberts (1989).

Horn primarily offers his metalinguistic negation as an account of such uses of negation that are not compatible with the truth functional description. He argues that the division is justified, on the grounds that this type of negation has distinct requirements on the linguistic context. For the purpose of evaluating this claim, it will be assumed that the two functions are both defined at the same level, as compatible and incompatible with the truth function respectively. This is true to Horn's spirit, if not exactly to his words.

4.2.2.2 Is Non-Optimality Linguistically Marked?

Since erroneous information is certainly non-optimal, it could be argued that a negation which indicates incompatibility is a special case of negation which indicates non-optimality. If compatibility with the bivalent truth function is the only criterion of definition, then it would be necessary to find some independent justification for considering incompatibility negation a separate category, rather than representing a typical or default interpretation. Identifying a special incompatibility negation by appealing to an intuition that some negation is truth functional, i.e. indicates incompatibility only, will not do, since that "intuition" in all likelihood is a learned one—an ordinary language user does not think of linguistic expressions in terms of truth values.

As noted above, Horn does not use compatibility with the truth function in both his descriptions, but the distinction he makes is analogous. He argues that it is supported by an observation that when speakers use *"not"* metalinguistically, this must be signaled by immediately providing an alternative statement (and using a special intonation). This claim is put forward, perhaps with even more insistence,

by Anscombre and Ducrot (1986). What this means is that for the negation in example 4.2 to be interpreted as indicating that Obree did not have low enough gears, it was necessary that he added that he was very much overgeared, and he also had to use contrastive intonation on the first occurrence of *"bit"*.

Horn suggests that this requirement indicates that such uses of negation are different from "ordinary truth functional negation", and that moreover, it shows that negation indicating non-optimality is of the same type as a negation used to object to an erroneous linguistic realisation, since the same constraint holds for such uses. For instance, the speaker who corrects the choice of gender for *viande* must add a sentence containing the correct gender afterwards:

(4.3) Non, je n'ai pas 'cou-pay luh vee-and', j'ai coupé la viande.
No, I not have not cut the{masculine} meat, I have cut the{feminine} meat.
'No, I haven't cut the{masculine} meat, I've cut the{feminine} meat.'

For Horn, as well as for Anscombre and Ducrot, a rectifying sentence is necessary in order for the addressee to be able to reinterpret the negation as metalinguistic. Horn writes (1989, p.374), with respect to uses of negation that are "irreducible to the ordinary truth functional operator" (e.g. objections to scalar implicature, lexical choice, and linguistic realisation) that for sentences containing them,

> a felicitous utterance involves contrastive intonation [and it must be] followed by a continuation in which the offending item is replaced by the correct item in the appropriate lexical, morphological and phonetic garb—a RECTIFICATION to borrow the label of Anscombre and Ducrot (1977).

Anscombre and Ducrot, as mentioned above, insist that this is always necessary for a metalinguistic negation, but Horn allows for some exceptions. He suggests later, on p.403, that

> the rectification might not occur overtly, in which case the 'incompleteness' of what I am calling the metalinguistic understanding of *not* will still be 'made explicit through intonation'.

It is not clear whether this passage refers to such instances of metalinguistic negation where the negation can in principle be interpreted as truth functional, or to any instance, since he otherwise suggests quite strongly that a rectification is needed. If the rectification can be left out, the 'requirement' for one can hardly be considered a distinctive feature of metalinguistic negation.

A relevant question to ask here is who it is that needs the rectifying sentence. Is it the addressee of the utterance, whose general interpretative activities are performed with reference to a context and background knowledge, or is it the analyst, who constructs sentences and pretends to be interpreting them without?

It is virtually impossible to avoid using world knowledge when interpreting sentences, even if they *are* considered without reference to a particular context. For instance, one tends to prefer to interpret the first of the two sentences in example 4.4 as if the person referred to as *"her"* is seen through the binoculars, and the second one as if she is carrying a gun.

(4.4) a. I saw her with my binoculars.
b. I saw her with a gun.

This happens despite no context being given beforehand, and there is no syntactic justification for assuming that the preposition phrase modifies a different constituent in these two cases. That is, in interpreting the sentences, contexts are selected using knowledge about the objects referred to. The contextualisation process cannot be switched off—there is no such thing as a null context. The examples above with the limited information they provide illustrate that even for such restricted examples, it is knowledge that determines the interpretation, rather than the linguistic form.

Given this, it is quite conceivable that addressees are able to select a non-truth functional interpretation of a negation without the speaker having to provide any linguistic clues, as long as they use a context in which that interpretation is the one which makes sense. The interpretation of a non-optimality negation could be achieved by accommodating the information into a representation, just as that of an incompatibility negation would be interpreted.

What Horn proposes is, however, that this is not the case: in order to recognise a metalinguistic negation, agents do not use their representations, but rely on the recognition of certain linguistic clues. In order to use this linguistic markedness as an argument for considering metalinguistic negation different from descriptive, it would be necessary to show that the former is always marked this way, while the latter never is. But as it happens, contrastive intonation, and rectifications is by no means restricted to cases where the information that the negation applies to is non-optimal without being incompatible. This is illustrated by examples 4.5, 4.6 and 4.7.

(4.5) A: Tim behaves so strangely you start wondering if he's in love with her.
B: He's not in LOVE, he HATES her.

(4.6) A: So Carol's going out with Abby.
B: She isn't going out with HER, she's going out with THÉRÈSE!

(4.7) Den kvällen gjorde [Aron] den stora förändringen. Han skrev om alltsammans rakt av utan att bry sig om huruvida det blev en helhet eller inte, om språket var perfekt eller om det blev en enda rappakalja. *Att skriva var inte längre ett tvång utan en befrielse.* Sida efter sida skrev han i rasande fart.[1]

'That night Aron made the big change. He rewrote everything straight off without worrying whether it held together or not, if the language was perfect or if it was a great big mess. *Writing was no longer a duty, but a liberation.* He wrote page after page at breakneck speed.'

These examples are all contrastive in the sense that the rejection of one item is immediately followed by the presentation of a preferred one. They nevertheless contain a negation which applies to some unit of information which, given the context, is incompatible with the world, i.e. the negation is compatible with the bivalent truth function. This shows that rectifications are not confined to cases where the negation applies to information which is in principle correct but non-optimal. Rather, they occur whenever the speaker considers it useful to contrast two items, be they compatible or not (contrastive use of negation will be discussed further in section 4.3 and in chapter 5 and 6).

Conversely, it is not necessary that a negation which cannot be given a straightforward interpretation as a truth function must be followed by a rectification. For instance, consider example 4.8.

(4.8) At a cycling club meeting, and A wants to know the latest results. It is generally accepted that the person who won is a better cyclist than everyone else in the club.

A: How did you get on last week?
B: Dave managed to win the Road Club ten {mile time trial}.
C: He didn't *manage* to win.

It is quite possible for those present to interpret B's utterance as a statement that *"manage"* does not accurately describe Dave's effort to win, without B having to

[1]From *Passionsspelet* by Jonas Gardell, 1991, Norstedts Förlag AB, Stockholm. The italics are mine.

add that *"it was easy for him"* or some sentence to that effect. The interpretation does not depend on the speaker providing a linguistic recitification, but on the availability of the information required to process the utterance to the addressee.

Analogically, an objection to a linguistic realisation need not be followed by a preferred version either, if the speaker thinks that the addressee can recognise their mistake anyway. Consider a situation where both A and B know that A tends to confuse *take* and *bring* (as discussed in chapter 2, section 2.2.3), and makes mistakes regularly, which B usually corrects.

(4.9) A and B are discussing what camping equipment B should take on a trip to Italy.

A: What are you bringing?
B: I'm not bringing anything.

It is not necessary for B to linguistically mark the negation as metalinguistic, by adding that A should have used *"take"*. A is able to interpret the sentence as an objection to the choice of *"bring"* anyway. In the dialogue where the sentences in example 4.9 occurred, *"bringing"* was not marked intonationally either.

When a sentence like *"he didn't manage to win"* is interpreted "out of context", what one does is to assign a likely context to it. Additional knowledge about the person in question might mean that there are constraints on what a likely context would be. Additional information provided by the speaker (e.g. *"it was easy for him"*) may have the same effect, but it is not the utterance of the information as such that makes the addressee interpret the negation as not signaling incompatibility, it is the fact it refers to.

In other words, the addressee's ability to interpret a negation as doing something other than indicating that some information is incompatible with the context does not depend on linguistic clues given by the speaker. It depends on which context is the most natural one to interpret the utterance in. It may be necessary

that the speaker indicates what context they intend linguistically, by adding more information in a following sentence, but this is by no means always so. The task of interpreting a negation as applying to different types of inaccuracies is not one of recognising a particular linguistic form, but one of interpreting utterances in a context.

4.2.2.3 Horn's Negation Diagnostics

Linguistic markedness does not justify a split between an incompatibility negation and a non-optimality one . However, Horn also cites some other linguistic factors which he suggests covary with these negations to a greater or lesser extent. These are whether the negation can be incorporated prefixally, what type of polarity item it induces, and what type of BUT-morpheme it would be followed by. Significantly, Horn terms these factors "diagnostics"—the state of one of them may decide of what type the preceding negation is, but not necessarily—indicating that he does not consider there to be a direct correlation between them and the two negations. This means that they are of limited interest to a discussion of whether Horn's negations are well founded, and will as such only be briefly described. The exception to this is the BUT-morphemes which are relevant for other reasons.

Prefixal Negation

According to Horn, a word of the form *NegativePrefix+stem*, e.g. *"unhappy"*, *"improbable"*, *"disoriented"*, and *"non-entity"* can be analysed as ¬STEM, and can be alternatively expressed as *"not (stem)"*, in which case the negation is descriptive. The reverse is only true if the negation is descriptive. This means that if the sentence in example 4.10a is suitable in a discourse, then so is the (b)-sentence.

(4.10) a. She's unhappy.

b. She's not happy.

The latter cannot be replaced by the former if the negation is used metalinguistically:

(4.11) a. She's not happy, she's ecstatic.
b. *She's unhappy, she's ecstatic.

If a negation can be incorporated prefixally, it is a descriptive one, otherwise it may or may not be metalinguistic. This is because of several factors, some of which are discussed by Horn himself.

A prefixal negation only applies to the word that carries the prefix. A *"not"* which is descriptive in Horn's sense need not do that, and for some types of sentence, it may indeed be impossible to obtain a reading in which the negation only applies to one word. He gives the following examples:

(4.12) a. She was not fortunate enough to lose her husband.
b. She was unfortunate enough to lose her husband.

The freestanding negation cannot take scope over the word *fortunate* only, and hence the (a)-sentence cannot be used in place of the b-sentence.

Certain negative prefixed words such as *"uneasy"* can simply not be analysed as ¬EASY. . And even when a pair of expressions *NegativePrefix+stem* and *"not (stem)"* seem interchangeable when considered out of context, this is not necessarily the case in a discourse, as illustrated in example 4.13 (further evidence for this will be discussed in section 4.3).

(4.13) A knocks on the bedroom door:

A: Can I come in?
B: No, I'm not decent.
B': *No, I'm indecent.

Only a subclass of words can be prefixed. Not all words that can be negated accept a negative prefix—they primarily occur with adjectives, and to a lesser extent with nouns, in which the affixed noun often seems to denote the lack of a property or characteristic:

(4.14) A Renault 4 is a non-car.

Rather than functioning as an incompatibility negation in general, a negative prefix denotes the absence of a characteristic. Moreover, Horn points out that negatively affixed words tend to have negative connotations: we have for instance *uncivilised, unclean, untrue* but not *unbarbarian, undirty, unfalse*, which means that the class of words that take negative prefixes is even further restricted.[2] Given this, prefixal incorporation cannot be cited as evidence for the existence of a specific incompatibility negation. At most, it defines a subset of such negation.

Polarity Items

Polarity items are words and phrases which occur only in either affirmative syntactic environments (declaratives) or non-affirmative ones (e.g. after a negation, in questions and in conditionals). The ones that are found in the former are POSITIVE POLARITY ITEMS (PPIs) while the ones that are found in the latter are NEGATIVE POLARITY ITEMS (NPIs). Examples of PPIs are *"would rather"*, *"some"*, *"still"*, *"already"*, while *"can be bothered"*, *"any"*, *"ever"* and *"yet"* are NPIs. The idea is

[2]This is a rather interesting issue with respect to the representation of information, but it is beyond the scope of this thesis.

that the (b)-sentences in example 4.15 and 4.16 are not acceptable, unless in very specific circumstances.

(4.15) a. I have already done that.
b. ?I haven't already done that.

(4.16) a. I haven't done that yet.
b. ?I have done that yet.

Several authors, e.g. Baker (1970) have observed that PPIs can occur in non-affirmative contexts, e.g. after a single negative element, as illustrated by the sentence in example 4.17.

(4.17) I haven't already finished it, it was supposed to be handed in weeks ago.

Baker suggests that this happens when the speaker's purpose is to emphatically deny a previous speaker's assertion "word by word". It is not clear whether he meant that this was the only situation, or just one that merited particular interest. Horn chooses the narrow approach, suggesting that word by word denial is the only time a PPI can occur after a negation, and that a descriptive negation must be followed by an NPI. A polarity item in a sentence containing a metalinguistic negation, however, might be either negative or positive, meaning that the situation is similar to that of negative prefixes: the correlation with his categories of negation is only partial. It could have been a relevant observation if the choice of polarity item after a metalinguistic negation was free, but it is not. The sentences in examples 4.18 and 4.19 contain metalinguistic negation, but in one case a PPI is required and in the other an NPI.

(4.18) A: Chris solved some of the problems.
B: She didn't solve some of the problems, she solved all of them.
B': ?She didn't solve any of the problems, she solved all of them.

(4.19) A: Did you manage to put the wheel back on?
B: I didn't manage anything. It was dead easy.
B': ?I didn't manage something. It was dead easy.

Clearly there is some factor which determines whether a PPI or an NPI should be selected, but this is not whether the negation is compatible with the bivalent truth function or not. Horn himself suggests that the reason why PPIs occur after metalinguistic negation is not the nature of the negation, but that they are scope-sensitive. The scope of a metalinguistic negation is typically, but not always, different from that of a descriptive one, and it is this feature rather than the negation as such which determines the choice of polarity item.[3] While differences in scope seem like a plausible explanation, other research suggests that the issue is more complex than that. For instance, examples 4.20 and 4.21, illustrate how in one case a PPI is preferred and in the other an NPI, despite the syntactic environments being identical.

(4.20) a. Did they at least say something?
b. ?Did they at least say anything?

(4.21) a. ?Has he at least cooked dinner already?
b. Has he at least cooked dinner yet?

[3]Since Horn argues that metalinguistic negation takes its scope over the entire sentence, the polarity item in these two sentences ought to be of the same type. If scope really is a deciding factor, Horn must be wrong about the scope of metalinguistic negation—rather than having a wide scope over the entire sentence, it would be limited to the particular word or phrase that the negation applies to, i.e. *"some"* in example 4.18 and *"manage"* in example 4.19.

The pairs *"some"/"any"* and *"already"/"yet"* are often cited as typical examples of polarity items, and are considered similar. But it has been pointed out, e.g. by Zwarts (1991), that there seem to be different categories of NPI, depending on on what type of expressions that is needed to license them (e.g. an overt negation might be required or an expression like *"few"* might suffice). Lakoff (1969) and Baker (1970) cite examples which suggest that the difference between *"some"* and *"any"* is one of meaning rather than of conforming to a syntactic environment. As in the case of negative prefixes, it must be concluded that although the selection of polarity items depends on something, it is not whether a negation that precedes it indicates incompatibility.

"But"

The last diagnostic proposed by Horn is the selection of BUT-morpheme. This one differs from the previous two, in that he suggests that there is not only a tendency, but a significant correlation with his two negations. The background to this claim is that several languages (e.g. Spanish, German, Swedish and Finnish) have two morphemes corresponding to the English *"but"*, and they are not interchangeable. The difference between the two was described in some detail by Anscombre and Ducrot (1977) in terms of Argumentation theory, and will be examined in section 4.3. They label the two morphemes PA and SN, from the Spanish *Pero* and *SiNo* and the German *Aber* and *SoNdern* (I will follow Horn, and refer to them as BUT_{pa} and BUT_{sn}). Anscombre and Ducrot argue that a BUT_{sn} is always preceded by a metalinguistic negation, and that if there is a negation before a BUT_{pa}, then it is a descriptive one.

Their categories of negation are, as indicated in chapter 3, functional, or at least the definitions in Anscombre and Ducrot (1977) are. This means that they differ from Horn's descriptions, since at least one of his is descriptive and not functional. However, Horn argues that Anscombre's and Ducrot's theory carries over to his descriptions, with a slight modification: a negation before a BUT_{pa} is always

descriptive, while a negation preceding a BUT_{sn} is "typically understood as metalinguistic". As it happens, neither Anscombre and Ducrot nor Horn are right about these correlations, but the interaction between the BUTs and negation is an interesting area, which will be explored further later in this thesis.

4.2.2.4 Summary

There is no reason to distinguish between two negations which indicate incompatibility and non-optimality on the grounds that they have different requirements on the linguistic environment. The interpretation of sentences with negation seems to be driven by their accommodation into the addressee's representation, rather than, as Horn argues, being a matter of recognising certain linguistic clues. If the speaker considers the addressee to have sufficient background knowledge to interpret a negation as indicating incompatibility, then the speaker does not have to provide any specific hints that this is how it is intended. Similarly, the speaker may or may not want to specify how an incompatibility negation should be treated. The only reason to draw a line between these two is that one is convinced that some natural language negation is best characterised as corresponding to the bivalent truth function.

4.2.3 Information vs Language

Kempson's (1986) account of negation differs from Horn's in that she does not consider it ambiguous with respect to different types of informational inaccuracies. Her account is in accord with the idea that context determines the interpretation of sentences, not the linguistic form. But she also maintains that negation is truth functional—this is not one potential interpretation, but the only one, which the representation of the sentence is enriched to fit. This procedure is available when the reason for using the negation is to mark some information as inaccurate, but

since she does not advocate the inclusion of metalinguistic predicates in the logical form, it is not an option when the negation appears to be applying to a linguistic realisation. For Kempson, natural language negation therefore either corresponds to a truth functional operator or to a marker of linguistic inappropriateness.

The distinction between a negation which applies to information, and one which applies to linguistic realisations at first blush might seem so clear that there is no reason to object to it—language and information are distinct categories, so two functions of negation based on these categories would be easily defined and justified. This is a pragmatic ambiguity that is perhaps not too hard to accept: negation reinterpreted as applying to language rather than information. Theoretically, it should be unproblematic to identify any use of negation as belonging to one category or the other, and with the type of sentences that are usually employed to illustrate the difference (exemplified in 4.22 and 4.23), one can easily believe that this is so.

(4.22) John does not love Mary.

(4.23) Claire de Lune is not by [d^{h}ə'b^{y}uːsiː], it's by [dəby'si].

In practice, however, the difference between linguistic realisation and informational inaccuracy is not as clear cut as it may seem from these examples. Consider example 4.24 (which was discussed in more detail in chapter 2):

(4.24) A and B live in Edinburgh. B is going to Glasgow to practice with his band. A is not a native speaker of English.

A: Do you want me to help you to carry the stuff to the station?

B: No, it's OK, thanks.
A: Are you bringing the {guitar} amp?
B: I'm not BRINGING it, I'm TAKING it.

To recapitulate the circumstances of this example, A is native in Swedish, which has only one word corresponding to the English words *"take"* and *"bring"*. As far as A is concerned, there is only one relevant concept in this context, which roughly corresponds to 'agent transports object'. For B, on the other hand, there are two concepts, which are informationally distinct, and in this case depend on the position of A and B with respect to the target location (Glasgow) of the transporting event. If either of them is at the target location at the time of the utterance or at the time of the event, then *"bring"* should be used, otherwise *"take"*.

There is no reason for B to assume that A actually thinks that they are in Glasgow at the time of the utterance, so the problem with A's utterance is not the information that A refers to, but A's linguistic realisation of it: what B considers inadequate is not A's beliefs, but A's utterance. As in example 4.23, it is the expression chosen to refer to some information that is inadequate. At the same time, the reason why this particular utterance is perceived as inadequate is that the linguistic realisation that A chose conventionally is used in a state of affairs which does not obtain. From this aspect it has more in common with example 4.22: the negation applies to some aspect of information.

Is B objecting to *"bring"* as if it were informationally inaccurate, or as if it were a linguistic mistake? It seems difficult to decide if only this example is considered, but some other observations can shed light on the problem. As it happens, recognising that a sentence has been badly realised is not sufficient for negating it. This can be seen by considering the situation described in example 4.25, in which B cannot identify what A is referring to with the expression ['wudu].

(4.25) B: Is there anything interesting in it? ("it" refers to a magazine)
A: There's an article about ['wudu].
B: ['wudu]?
A: Oh shit, I don't remember, is it 'v' or 'w'? You know, the religion.
B: Oh, you mean Voodoo.

B can, however, identify [wudu] as a non-signifying expression, and given that B knows that A is not native in English, B can safely assume that A has mispronounced something. Now if it were true that the function of some uses of negation is to indicate an erroneous linguistic realisation, then B should be able to object to A's utterance of *"there is an article about* ['wudu]*"* by uttering the sentence in example 4.26.

(4.26) It's not about ['wudu].

However, an utterance of this sentence could hardly be considered acceptable in this context. B could only use this sentence if B had understood what A intended by ['wudu] in the first place. An inadequate linguistic realisation as such is therefore not enough to allow a speaker to use a negation. The speaker must know what the offending speaker meant.

Another relevant case is when a speaker uses a negation to object to a linguistic realisation of a word when it has been mispronounced in such a way that it sounds like another word (example 4.27, A is a child).

(4.27) B: What pudding do you want?
A (reading from the packet): Raspberry [maus]!
B: You don't want raspberry [maus]. That would be horrible. You want raspberry [muːs].

Speaker A mispronounces the word *"mousse"* in such a way that it sounds like another word, *"mouse"*. A did not intend to refer to the animal, but B treats A's utterance as if A did, which is made clear by B's use of the sentence *"that would be horrible"*. Given this, the negation used by B must be operating on information. But what would the status of the negation be if B had left out the utterance of *"that would be horrible"*? If B had not made explicit that she was considering the interpretation usually associated with [maus], then her utterance would have

been analogous to the standard examples of "objections to linguistic realisation", as exemplified by 4.23. In my opinion, B would not have been using negation in a different way in that case. It would be hard to justify, and one would have to address problems such as what would happen if B left out *"that would be horrible"* because she thought that the reference to mice was obvious, etc.

Agents cannot use negation as a single means of objecting to the linguistic realisation of a sentence unless they can identify what they offending speaker intended. In view of this, the distinction between objecting to language and objecting to information is not as clear cut as it may seem from the examples usually employed to illustrate it. Instead of considering such use of negation to apply to the linguistic realisation, it should be seen as the speaker treating the badly chosen expression as if the offending speaker had used it to refer to some inaccurate information. Whether or not the expression conventionally refers is not relevant—the correcting speaker is treating it as if it did, which means contrasting either a real, but unintended object, or a pseudo-object (referred to by the bad expression) with a preferred object (referred to by the preferred expression).

Linguistic and informational mistakes must be treated on the same conceptual level, but not by introducing metalinguistic predicates into the representation of information. If this approach were taken, a metalinguistic predicate which failed to apply would be sufficient to allow a negation to be used but as has been shown, this would make incorrect predictions of how negation can be used. A metalinguistic negation is hence not just a shorthand for a longer expression which includes predicates expressing well-formedness. It is an ordinary negation which applies to an item in the agent's representation which has been added because of the expression, even if the agent does not consider it to correspond to anything particular in the world they are talking about.

4.2.4 Consequences for a Theory of Negation

The assumption that negation is truth functional is sometimes justified by researchers by saying that they have an intuition that this is so. This reflects the following observation: there is a "default" interpretation of sentences with negation, which is that the negation signals that something the sentence refers to is incompatible with the context that an utterance of it is to be interpreted in. This is something that probably everyone who studies negation agrees on. In all likelihood, it is an accurate observation, and it seems that a reasonable requirement of a theory of natural language negation is that it predicts this interpretation in the absence of more specific knowledge which suggests something else.

This thesis by no means wishes to take objection with this observation. What is questionable, however, is the extrapolation to the claim that natural language negation is truth functional, which is arrived at by making two additional assumptions. The first one is that what was described above as "something is incompatible with a context" can be equated with "the corresponding sentence without a negation is false" (i.e. that a truth functional account of reasoning with incompatibility can be given). The second one is that the incompatibility interpretation, now transformed into a falsity interpretation, is not just *one* interpretation, but the very meaning of (some) negation.

It is interesting to compare this attitude to the characterisation of negation with how that of other types of expression have been approached. One relevant example is scalar expressions (chapter 3): in order to avoid inconsistency, they are assumed to have a weak meaning, which tends to be strengthened in certain contexts in the absence of a counterindication. It is curious that negation has not received a similar treatment, but that the specific interpretation is considered primary.

What I have sought to show in this section is that the incompatibility interpretation of negation is just that: one way of accommodating the information in an utterance into the addressee's representation. If a *"not"* is interpreted in a

different way, it is not because it represents a different type, or use, of negation, but because the context and background knowledge suggest that some other understanding is better.

4.3 'BUT' and Negation

4.3.1 Background

This section introduces the two BUTs, and presents the descriptions of them proposed by Anscombre and Ducrot (1977). The purpose of this is twofold: I want to show that the connection between the two BUTs and the categories of negation proposed by Anscombre and Ducrot and by Horn does not exist, but I also want to give a thorough introduction to the BUTs themselves, as it will be argued in section 4.4.1 that they provide some important pointers to the understanding of negation. As there is a lot to be said for Anscombre's and Ducrot's understanding of the two BUTs, their descriptions will be an good starting point.

The issue of describing *"but"*, in particular in comparison with *"and"*, has received a lot of attention (e.g. Dascal and Katriel, 1977; Elhadad and McKeown, 1988; and Moeschler, 1989). It has sometimes been suggested to be ambiguous, e.g. by Lakoff (1971), but it is worth noting that the purported ambiguity concerns whether *"but"* in Anscombre's and Ducrot's BUT_{pa} sense is ambiguous. This claim has been refuted by e.g. Kempson (1986) and Blakemore (1987). I will not discuss these ambiguity proposals, suffice to say that the relevant data are all covered by Anscombre's and Ducrot's description of BUT_{pa}. The distinction between the two BUTs considered here is a different one.[4]

[4]I will not use the term "ambiguity", as I am referring to the lexicalisation of different discourse markers, rather than a semantic ambiguity. The BUTs are maybe better viewed as two functions than as two senses. The exact characterisation of the nature of the different is not relevant for the present discussion.

Prior to examining the formal descriptions, an informal account of the difference between them will be given, since it may not be an obvious one to speakers of a language that does not distinguish them morphologically. BUT$_{pa}$ is the more familiar one: indeed a lot of the writings about *"but"* acknowledge only this function (e.g. Lakoff, 1971; Cohen, 1984; Kempson, 1986; and Blakemore, 1987). It can be thought of as concessive. A BUT$_{pa}$ is used in order to prevent the addressee from drawing an unwanted conclusion from a statement that the speaker accepts, as illustrated in examples 4.28.

(4.28) a. The river has flooded but the bridge is still there.
b. The river hasn't flooded, but it is not far off.

"But" used in the BUT$_{sn}$ sense is less common in English, in particular in spoken language. When translating a sentence with a BUT$_{sn}$ into English, it often sounds natural not to use an overt *"but"* at all, and it might even be necessary to do so. BUT$_{sn}$ is used to contrast an item which the speaker considers should not be part of an accurate representation, with one that should be. It must be preceded by a negation, and can therefore be viewed as an optional marker of contrastive negation; the type of negation that Horn labels linguistically marked (section 4.2.2). This means that it basically has the same properties as the morphologically distinct contrastive negation in the languages that have such (chapter 2, section 2.3).

(4.29) a. It's not semantic, but pragmatic.
it's pragmatic.
b. Pedro doesn't smoke cigarettes, but cigars.
he smokes cigars.

In languages which distinguish the two BUTs, the morphemes are associated with different preferences with respect to their syntactic environment: BUT$_{sn}$ is typically, although not necessarily, followed by a reduced clause (this will be discussed

in detail below). Anscombre and Ducrot argue that the French *"mais"* has syntactic constraints which correspond to these preferences, so that the distinction is present, but syntactically rather than morphologically. Horn (1985) argues that the same holds for *"but"* in English. It is by no means a failsafe test, but it is often the case that when an English *"but"* is followed by a reduced clause, it is used in the BUT_{sn} sense.

In the sentences in (4.29), it is possible to use an overt *"but"*. In some circumstances, however, a BUT_{sn}-morpheme can be used in the languages which distinguish them phonetically, while it is not possible to use a *"but"* in English. For instance, it seems more natural not to use an overt *"but"* if the two items are obviously mutually exclusive. Using a *"but"* in the sentence in example 4.30(a) does not seem very natural (if it is the properties, as opposed to the words that are contrasted). The corresponding Swedish sentences are equally natural.

(4.30) a. It wasn't easy, {?but hard. / it was hard.}

b. Det var inte lätt, {utan svårt. / det var svårt.}
It was not easy, {BUT_{sn} hard. / it was hard.}

Moreover, there are syntactic environments which do not permit a BUT_{sn} reading in English, only a BUT_{pa} one, where this is notthe case if there is a morphological difference. English would have to use a *"so"* or an *"instead"* instead in these cases:

(4.31) a. Jag orkade inte tala med honom, utan åt frukost på Caféva.
I had-the-energy not talk to him, BUT_{sn} ate breakfast in Caféva.
I didn't have the energy to talk to him, so instead I had breakfast in Caféva.

b. I didn't have the energy to talk to him, but had breakfast in Caféva.

It is difficult to interpret the English sentence in such a way that the speaker having breakfast in the café was something that was done to avoid the person in question. Despite the ellipsis, one tends to prefer a BUT_{pa} reading, where the speaker went to the café despite having to talk to the person there. In other words, it is sometimes necessary to use a different word than *"but"* in order to achieve the BUT_{sn} effect: *"but"* does not entirely cover the use of the two morphemes in languages that have both.[5]

The two BUT-morphemes are crucially not interchangeable. From a discourse perspective, if one was used, the other one could not have been used to the same effect. This is a communicative distinction, not a semantic one. The two BUTs could sometimes be used to link references to the same information, as illustrated by the Swedish sentences in example 4.32

(4.32) a. Han är inte spanjor, men han är argentinare.
He is not Spaniard, BUT_{pa} he is Argentinian.
'He isn't Spanish, but he is Argentinian.'

b. Han är inte spanjor, utan argentinare.
He is not Spaniard, BUT_{sn} Argentinian.
'He isn't Spanish, he's Argentinian.'

Although both these sentences could theoretically be used to make a well-formed utterance in similar circumstances, they differ in the effect they would have on the addressee. To use the first sentence would suggest that the speaker accepts that assuming that the individual in question is Spanish would have been a good guess. The second sentence suggests that this assumption is unjustified.

[5]It is possible that the purported correlation with two different negations is partially a result of assuming that it does.

4.3.2 The BUTs

The descriptions of BUT_{pa} and BUT_{sn} given by Anscombre and Ducrot (1977) are quite complex. I have made some changes to the presentations, in particular to that of BUT_{sn}, for the sake of clarity. In some cases, I have chosen to expand on their observations, as I felt that a more thorough explanation was needed. The only part of it that will be examined critically in this section is the purported restrictions on a preceding negation. Some other aspects of the descriptions will be discussed in section 4.4.1. It is important to bear in mind that the references to descriptive and metalinguistic negation made in the article are to the early descriptions, as given in chapter 3 section 3.3.4, not the ones in Anscombre and Ducrot (1983) which are more similar to Horn's.

4.3.2.1 BUT_{pa}

Although their conception of negation has changed, Anscombre and Ducrot (1983) still use the original description of BUT_{pa}, and it has been used by others too, e.g. Moeschler (1989), and Elhadad (Elhadad and McKeown, 1988; Elhadad, 1990).

Description of BUT_{pa}

Given an utterance of the form P BUT_{pa} Q,

1. P is presented as a possible argument for a conclusion C.
2. Q is presented as an argument for $\neg$C.[6]
3. Q is a stronger argument for $\neg$C than P is for C.
   The whole utterance of P BUT_{pa} Q is thus oriented towards $\neg$C.

BUT_{pa} and Negation

Given an utterance of the form NEG-R BUT_{pa} Q (i.e. the clause preceding the *"but"* contains a negation),

1. The negation must be descriptive.
2. If R and Q can be thought of as belonging to a graded scale, then R must be stronger than Q.

The first two conditions of the description taken together mean that P and Q must have opposite argumentative orientations with respect to a theme. This requirement is an interesting one: it means that the speaker must have access to some context which links the information expressed by the two clauses. Accessing such a context is not equivalent to relating the information expressed by the two clauses through logical properties. It has to be one which is provided by the immediately preceding discourse. Consider the following two Swedish examples (adapted from Anscombre and Ducrot)[7].

(4.33) A: Carlos pratar spanska väldigt bra. Är han spanjor?
Carlos talks Spanish very well. Is he Spaniard?
'Carlos speaks Spanish really well. Is he Spanish?'

B: Nej, men han är argentinare.
No, BUT$_{pa}$ he is Argentinian.
'No, but he is Argentinian.'

(4.34) A: Juan måste känna till Spanien väl.
Juan must know about Spain well.
'Juan must know Spain well.'

B: *Nej, för han är inte spanjor, men han är argentinare.
No, for he is not Spaniard, BUT$_{pa}$ he is Argentinian.
'No, because he's not Spanish, but he's Argentinian.'

In example 4.33, it is possible to use BUT$_{pa}$, because a context (that Carlos speaks Spanish) has been provided which would be justified both by Carlos being Spanish and by his being Argentinian. In example 4.34 however, the preceding context

[6]Anscombre and Ducrot and use ¬C in the article where the description first appeared. This is perhaps unfortunate, since the argument relation is defined as holding between utterances and not between an utterance and a belief. But the original notation will be used here, as the intention is to discuss the descriptions in the 1977 article. It might be more true to Anscombre's and Ducrot's later writings to think of it as Q being an argument against C, or as P and Q being opposed with respect to whether C should be accepted.

[7]I chose to translate these examples since Anscombre and Ducrot gave them in French, which does not have two morphemes.

does not provide such a standard (being Argentinian is not a reason for knowing Spain), and therefore BUT$_{pa}$ cannot be used. Although the facts contrasted in (4.33) and (4.34) are the same, only the former allows for the use of BUT$_{pa}$. This shows that the possibility of using a BUT$_{pa}$ cannot be judged simply by considering the sentences, but depends on the discourse context. As a general rule, it does not make sense to study a sentence with a BUT$_{pa}$ in isolation, because there simply will not be enough information to decide whether it is well-formed or not.

The claims about negation preceding the BUT$_{pa}$ are intended to explain why example 4.35a may be acceptable while 4.35b can never be:[8]

(4.35) a. No es cierto pero es probable.
Not is-it certain, BUT$_{pa}$ is-it probable.
'It's not certain, but it's probable.'

b. *No es probable, pero es cierto.
Not is-it probable, BUT$_{pa}$ is-it certain.
'It's not probable, but it's certain.'

4.3.2.2 BUT$_{sn}$

Restrictions on the Linguistic Context of BUT$_{sn}$:

i. BUT$_{sn}$ only occurs in utterances of the type NEG-P BUT$_{sn}$ Q, i.e. the preceding clause must contain an overt "autonomous" negation.

ii. Both the NEG-P clause and the BUT$_{sn}$ Q clause must be uttered by the same speaker.

[8]The corresponding examples with "sino", i.e. BUT$_{sn}$, are both possible utterances:

No es cierto, sino probable.
Not is-it certain BUT$_{sn}$ probable.
'It is not certain, but probable.'

No es probable, sino cierto.
Not is-it probable BUT$_{sn}$ certain.
'It is not probable, but certain.'

Description of BUT$_{sn}$

i. An utterance of the type NEG-P BUT$_{sn}$ Q serves to present Q as a reason to reject P.

ii. P and Q must be alternative descriptions of the same type of fact.

BUT$_{sn}$ and Negation

i. The negation preceding the BUT$_{sn}$ must be polemic.

There are no restrictions on the linguistic context of BUT$_{pa}$ comparable with those on BUT$_{sn}$. As for the requirement of an overt negation, a sentence of the form P BUT$_{pa}$ Q may have an overt negation in either, neither, or both, of P and Q. The "autonomous overt negation" before a BUT$_{sn}$ can be either a *"not"* or a negative particle such as *"no"*, *"nothing"*, *"never"*, *"nobody"* and *"none"*, as illustrated by the Swedish and German examples below.[9]

(4.36) Jag såg ingenting, utan hörde honom bara.
I saw nothing, BUT$_{sn}$ heard him only.
'I saw nothing, but only heard him/I didn't see anything, I only heard him.'

(4.37) Ich habe keine Pferde gesehen, sondern nur Kühe.
I have no horses seen, BUT$_{sn}$ only cows.
'I saw no horses, but only cows/I didn't see any horses, only cows.'

Negative prefixes and "negative verbs" (such as *deny*) on their own (i.e. when not in combination with an overt negation) do not admit the use of BUT$_{sn}$. Compare the following Swedish examples:

[9]The English translations tend to sound better using *"not anything"* rather than *"nothing"* etc., as well as hardly permitting the use of an overt *"but"*, which illustrates the point made earlier that *"but"* does not fully cover the BUT$_{sn}$ function.

(4.38) a. Det är inte praktiskt, utan omständigt.
It is not practical, BUT$_{sn}$ complicated.
'It's not practical, but complicated.'[10]

b. * Det är opraktiskt, utan omständigt.
It is impractical, BUT$_{sn}$ complicated.

(4.39) a. Han sa inte att han skulle komma utan antydde bara.
He said not that he would come BUT$_{sn}$ hinted only.
'He didn't say that he'd come, he only hinted.'

b. *Han förnekade att han skulle komma, utan sa att de skulle stanna hemma.
He denied that he would come, BUT$_{sn}$ said that they would stay home.

The second restriction on the linguistic context of BUT$_{sn}$, that both NEG-P and Q must be uttered by the same speaker, is an interesting difference between the two BUTs. It is common to follow on from another speaker's utterance with a BUT$_{pa}$ clause, as illustrated in the Swedish example 4.40.

(4.40) A: Jag vill ha ett utkast till det kapitlet till torsdag nästa vecka.
I want-to have a draft of the chapter-the by Thursday next week.
'I want a draft version of that chapter by Thursday next week.'

B: Men jag måste bli färdig med artikeln.
BUT$_{pa}$ I must be finished with article-the.
'But I've got to finish my paper.'

[10]While this sentence is not perfect in English (*"It's complicated"* would be better), it was chosen because it is one of the relatively few sentences where an overt *"but"* can be used in the BUT$_{sn}$ sense in English, when the two contrasted descriptions are obviously incompatible. In Swedish and German, there is no corresponding restriction, and BUT$_{sn}$ can be used for any contrast whatever the nature of the contrasted items.

The reason why translation with an overt *"but"* was chosen is that it is definitely not acceptable after *"it's impractical"*, while *"it's complicated"* would be. But an utterance of *"it's impractical, it's complicated"* would have the nature of providing two pieces of evidence, rather than of contrasting two different descriptions, and hence is not equivalent. To see this, consider how unnatural it would seem to insert a *"rather"* or *"instead"* between the two sentences, which would be quite acceptable if an overt negation had been used (this is probably the nearest linguistic analogy in English to the requirement of a negation before a BUT$_{sn}$.). Negation in combination with markers of contrast provides an important counterexample to the theory that an incompatibility indicating negation is a feature of information, which could equally well be expressed by a negative prefix (section 4.2.2.2).

Corresponding use of BUT_{sn} is not possible. Speakers cannot continue from other people's utterances with a BUT_{sn} clause, at least not unless they take it over and "treat it at theirs" (as in utterances made jointly by Huey, Dewey and Louie to Donald Duck in Walt Disney's comics). The utterances (in Swedish) in example 4.41 below would thus be unacceptable if the speakers did not act as if it were a single joint utterance.

(4.41) H: Vi vill inte ha...
We want not have ...
'We don't want...'

D: ...karameller...
'...sweets ...'

L: ...utan glass!
...BUT_{sn} ice-cream.
'...but ice-cream.'

Anscombre and Ducrot describe the function of BUT_{sn} as signaling that the description referred to by the second (usually reduced) clause excludes that of the first clause. However, this characterisation, like that of BUT_{pa} must be seen in relation to the speakers purposes, and cannot simply be determined by static properties of the information. It is not necessary that the two pieces of information are logically incompatible to allow for the use of BUT_{sn}, nor is it sufficient if they are.

The second part of the description of the function of BUT_{sn} is rather vague, stating that the juxtaposed clauses must be descriptions of the same fact. It is difficult to express this constraint more precisely, but it is a very interesting phenomenon which deserves more attention than it has received so far. Anscombre and Ducrot phrase it as (1977, p.25, my translation):

> the utterance of Q must be a characterisation—held incompatible with the one given by P—of the same fact that P purports to characterise (my translation).

This is a slightly different claim from the one associated with BUT$_{pa}$—the clauses joined by a BUT$_{sn}$ are not required to be arguments for or against believing something, but express different instantiations of a certain parameter.

The statement that the two clauses must be descriptions of the same type of fact is meant to capture the difference in acceptability between the B-sentence and the B'-sentence in example 4.42.

(4.42) A: Al-Helal har stängt på onsdagar.
Al-Helal have closed on Wednesdays.
'Al-Helal is closed on Wednesdays.'

B: Det är inte onsdag idag, utan (det är) tisdag.
It is not Wednesday today, BUT$_{sn}$ (it is) Tuesday.
'It isn't Wednesday today, but Tuesday.'

B': * Al-Helal har inte stängt idag, utan det är tisdag.
Al-Helal have not closed today, BUT$_{sn}$ it is Tuesday.
'It is not the case that Al-Helal is closed today, but that it is Tuesday.'[11]

Only the B-sentence is acceptable as a reply to A's utterance, and it is virtually impossible to think of any context at all that would license the B'-sentence. This is because it is difficult to think of *"it is Tuesday"* and *"Al-Helal is closed"* as descriptions of the same type of fact.

While it could be thought that the problem with the sentence in the B'-reply is that there is no ellipsis in the second clause of the Swedish sentence, this is not what makes the sentence odd. According to Anscombre and Ducrot, language text books have often suggested that the only difference between the two BUTs is their requirements on syntactic context, which they note is incorrect. They do not elaborate this point further, but it is worthwhile explaining why the difference cannot be explained merely in terms of syntactic context.

[11]The translation of the B'-sentence in example 4.42 is admittedly a bit odd, since *"It is not the case..."* is hardly part of ordinary English. The reason it was chosen is that *"Al-Helal is not open today but it is Tuesday"* could only have a BUT$_{pa}$ reading in English.

In at least Swedish and German, a BUT$_{sn}$ can be followed by a full clause. For instance, the Swedish sentence in example 4.43 is a possible utterance.

(4.43) Karo bet inte Fido, utan Fido bet Karo.
Karo bit not Fido, BUT$_{sn}$ Fido bit Karo.
'Karo didn't bite Fido, Fido bit Karo.'

There is often a preference for ellipsis after a BUT$_{sn}$, if the speaker has a choice, as well as for the two clauses to have the same syntactic subject. These syntactic preferences do not, however, in general entail any constraints on the type of information which can be contrasted. The preferred syntax can often be achieved by selecting an appropriate syntactic structure (typically an *it*-cleft which also has a contrasting function, see von Klopp and Humphreys, forthcoming). That is, although example 4.43 is acceptable, Swedish speakers would prefer the sentence in example 4.44, which has an ellipsis in the second clause, and a dummy subject "common" to both of them.

(4.44) Det var inte Karo som bet Fido, utan Fido som bet Karo.
It was not Karo who bit Fido, BUT$_{sn}$ Fido who bit Karo.
'It wasn't Karo who bit Fido, but Fido who bit Karo.'

Given the choice of syntactic structure, it is usually possible to refer to most information in such a way that an ellipsis can be made. So it is not the syntactic context which constrains what information a BUT$_{sn}$ can link. In particular it is not the case that it cannot contrast two states of affairs because they would have to be referred to by a full sentence.

4.3.3 The Connection Between the BUTs and Negation

The functions of the BUTs will be discussed in more detail below (section 4.4.1), but for the moment the discussion will be limited to the purported connection

between the BUTs and two, possibly different, types of negation.

From Anscombre's and Ducrot's (1977) descriptions it emerged that a BUT_{sn} is always preceded by a negation, while a BUT_{pa} may or may not be. They claim that the negation before a BUT_{sn} is always polemic, while a negation before a BUT_{pa} is descriptive. Horn (1989, p.413) writes that

> the negation which (optionally) figures in the concessive PA constructions is necessarily descriptive, while the negation required by the SN environments is typically understood as metalinguistic.

Horn claims about this statement that it is "consistent with Anscombre and Ducrot's thesis", which it is not, if by "typically understood" Horn means "not always". The reason for Horn's use of "typically understood" is probably that some of the examples of BUT_{sn} that he cites contain a negation which is in principle compatible with the bivalent truth function. As noted previously, this is a conflict for Horn—mostly he stresses truth functionality as the defining criterion, but sometimes he uses the definition of metalinguistic negation, i.e. whether the it is used to object to a previous utterance.

$\textbf{BUT}_{sn}$

Starting with BUT_{sn}, its function is to indicate that the speaker wishes to mark some description as unsuitable and suggest another one. It is difficult to see why someone would want to do this unless the former description was likely to be adopted by someone. Given this, it can be understood why Anscombre and Ducrot and Horn assumed that BUT_{sn} only occurs in utterances which reject a previous assertion—the most likely situation where a speaker may want to replace a non-optimal description is if it has been explicitly communicated. In this case, the BUT_{sn} would be used for an objection. Since both Horn and Anscombre and Ducrot want to distinguish between two negations, one of which objects to previous utterances, the negation preceding a BUT_{sn} is considered to have a rejecting

function as well.

Although BUT$_{sn}$ can occur in contrastive utterances used to object to some previous description, the objection part is not essential to the use of BUT$_{sn}$. One feature that should count against such a characterisation is that it is so common in written text, which would not normally be expected to contain objections to previous sentences. Consider example 4.7 again:

(4.7) Den kvällen gjorde [Aron] den stora förändringen. Han skrev om alltsammans rakt av utan att bry sig om huruvida det blev en helhet eller inte, om språket var perfekt eller om det blev en enda rappakalja. *Att skriva var inte längre ett tvång utan en befrielse.* Sida efter sida skrev han i rasande fart.

'That night Aron made the big change. He rewrote everything straight off without worrying whether it held together or not, if the language was perfect or if it was a great big mess. *Writing was not a duty anymore, but a liberation.* He wrote page after page at breakneck speed.'

In this case the writer contrasts Aron's previous experience of writing with his present one. It has not been suggested previously in the text that Aron felt that writing was a duty after he changed his approach to the task. It would be difficult to construe it as an objection to a previous utterance—if there was a sentence which suggested that Aron found writing to be a duty under any circumstances, and the writer sincerely objected to its use, then he could have changed it.

The important property of BUT$_{sn}$ is not that it objects to a previous utterance, but that it contrasts two items in the representation. It is a mistake to equate the reasons an agent might have to take a certain action with the action itself. Rather than viewing the combination of the negation and BUT$_{sn}$ as an objection to a previous linguistic action, it would be more natural to consider it as an instruction to the reader to do something with their representation. The speaker's ideas of the addressee's representation is a *result* of previous linguistic activity,

but that does not mean that they are equivalent.

Defining the function of polemic negation as objecting to any information expressed, implied or suggested by the previous discourse would make the connection between it and BUT$_{sn}$ trivial, but it does not entail that there is such a thing as a polemic or metalinguistic negation in the first place. Even if BUT$_{sn}$ were considered to express correction, this is no reason to assume that the negation preceding it has the exact same function. The fact that a sentence which contains a BUT$_{sn}$ typically also contains a negation in the rejection part definitely says something about how negation can be *used*, but it does not prove that it is of a different type.

Negation used before a BUT$_{sn}$ exhibits the same variation as it does in other contexts. It may be compatible with a truth functional interpretation or not, and it can apply to information or to a linguistic realisation:

(4.45) a. Vi gick inte till stationen, utan sprang av bara helvete.
We walked not to station-the, BUT$_{sn}$ ran of only hell.
'We didn't walk to the station, but ran like hell.'

b. Jag har inte två månader på mej, utan bara en.
I have not two months on me, BUT$_{sn}$ only one.
'I don't have two months, but just one.'

c. Han har inte två cyklar, utan tre.
He has not two bikes, BUT$_{sn}$ three.
'He doesn't have two bikes, but three.'

d. Jag såg inte Lisa, utan Magnus.
I saw not Lisa, BUT$_{sn}$ Magnus.
'I didn't see Lisa, but Magnus.'

e. Jag såg inte Lisa, utan Lisa och Magnus.
I saw not Lisa, BUT$_{sn}$ Lisa and Magnus.
'I didn't see Lisa, but Lisa and Magnus.'

f. Jag vill inte ha en äppel, utan ett äpple.
I want not have "an apple"{mispronounced & wrong gender}, 'I BUT$_{sn}$ an apple.
don't want an "an apple", but an apple.'

This shows that BUT_{sn} does not only occur with non-optimality negation.

BUT_{pa}

The connection between BUT_{pa} and descriptive negation is also questionable. BUT_{pa} was described as concessive: in an utterance of P BUT_{pa} Q, the state of affairs referred to by P supports a state of affairs C. The state referred to by Q, on the other hand, indicates that C does not obtain. There are several possible reasons why a speaker could want to make such a concession. For instance, some description of an object may have been accepted in a communication, and C would usually be a likely consequence. It does not obtain on this occasion, and one of the participants wants to prevent the other(s) from inferring or otherwise accepting it. In example 4.46, B considers there to be an association between the river flooding and the bridge being washed away. A supports the information that the river has overflown[12], which B accepts, but B wants to prevent A from inferring that the bridge is gone.

(4.46) A: Looks like the river has overflown.
B: It has, but the bridge is still there.

In example 4.47, the writer introduces some new information (*"you may have to take off any cable end covers"*) and wants to make sure that the reader prevents the state of affairs in which the cables fray from obtaining.

(4.47) If you can't take out the cable at the lever end, which is probably unlikely, then you'll have to undo the bolts on the calipers. Mark the cable positions before you do, and you may have to take off any cable end covers, but be careful not to fray the ends.

[12]It may be in the reader's interest to know that Collins English dictionary supports both *"overflown"* and *"overflowed"*.

These examples were deliberately chosen because the speaker (or writer) agrees with the addressee, or introduces new information to them. But there is no reason, given the description of BUT$_{pa}$, why it should be restricted to contexts in which the information in the P-clause is compatible with the addressee's representation. And, naturally, there is no such restriction either. In example 4.48, B refutes the information provided by A, showing that a BUT$_{pa}$ certainly can follow a clause which is used to signal that an agent disagrees with what another agent has communicated.

(4.48) A: Johan kommer aldrig hit nuförtiden.
Johan comes never here nowadays.
'Johan never comes here anymore.'

B: Det gör han visst[13], men du är aldrig här när han kommer.
That does he so, but you are never here when he comes.
'Yes he does / 'He does so } but you're never here when he comes.'

Disagreement with some statement can of course often be expressed by a clause containing a negation, which would then be polemic or metalinguistic. The negation in example 4.49 is polemic on Ducrot's account, and metalinguistic on Horn's (although perhaps Horn would say that it is simultaneously descriptive).

(4.49) A: Så du har hela frysen full me äppelmos nu?
So you have entire freezer-the full of apple-mash now?
'So you've got the whole freezer full of apple sauce now?'

B: Du de vart inte så mycke äpplen i år men de vart en jäkla massa plommon.
You it was/gave not so much apples in year BUT$_{pa}$ it was/gave a devil's lot of plums.
'Well, there weren't a lot of apples this year, but there were a fair few plums.'

[13] *Visst* is a particle marking adversity, which is used similarly to *"at all"*, except that it is not a polarity item. In some dialects of English, *"so"* or *"too"* is used in this way in non-negative sentences.

Another example of BUT$_{pa}$ following a polemic/metalinguistic negation is B's utterance in (4.33) on page 125. Anscombre and Ducrot actually recognise this type of example, and cite (4.50) as a potential counterexample.

(4.50) A: C'est probable qu'il va faire mauvais.
It is probable that it will make bad.
'The weather will probably be bad.'

B: Non, ce n'est pas probable, mais ça reste possible.
No, it is not probable BUT$_{pa}$ it stays possible.
'No, it's not probable, but it's still possible.'

They argue that the *"ne... pas..."* in B's utterance is a descriptive negation, on the grounds that the *"non"* preceding it is the only polemic element, thus allowing the descriptive *"ne... pas..."* to combine with the BUT$_{pa}$. The problem with this explanation is that B's utterance would be just as acceptable without the *"non"*, which forces the negation to take on the polemic function.

A solution based on illocutionary speech acts only being performed once has no real foundation. It is common to attempt to convey some information, and then rephrase the utterance to convey the same information again. This is not limited to cases where the speaker thinks they failed to convey the information the first time, but is common as a rhetorical device. The same holds for rejecting information. To argue, as Anscombre and Ducrot do, that the act of rejecting information can be carried out only once means that the sentence *"ce n'est pas possible"* in example 4.50 would have different functions depending on whether B used a *"non"* before it or not. It also leads to difficulties when the word *"non"* is used twice (which is common in French): what would the second negation in example 4.51 be considered to do if only the first one could carry the rejection function?

(4.51) Non, non, ce n'est pas probable, mais ça reste possible.

The negation in examples 4.49 and 4.50 must be polemic and hence cannot be descriptive on Anscombre's and Ducrot's (1977) account.

Horn could probably claim that the *"not"* in these examples are simultaneously metalinguistic and descriptive on his account, thus avoiding the problem so far. But BUT$_{pa}$ can be used in utterances with negations that cannot be descriptive on Horn's account either:

(4.52) A and B are two native Swedish speakers living in Britain. A translates the English *"do you take sugar"* literally into Swedish.

A: Tar du socker?
Take you sugar.
'Do you steal sugar?'

B: Jag TAR inte socker, men jag brukar ha lite i kaffet!
I take not sugar BUT$_{pa}$ I do-usually have little in coffee-the.
'I don't STEAL sugar, but I normally have a bit in my coffee.'

(4.53) A is a non-native Swedish speaker, who pronounces *"ett äpple"* like the English *"an apple"*:

A: Vill du ha en äppel?
Want you have "an apple" {mispronounced}?
'Do you want "an apple" {mispronounced}?'

B: Nej, men ett ÄPPLE vore inte så dumt.
No BUT$_{pa}$ ett äpple were not so bad.
'No, but I wouldn't mind an apple.'

Both these negations are metalinguistic on Horn's account, demonstrating that BUT$_{pa}$ is not always used with a descriptive negation.

4.3.4 Summary

This section examined whether there is a covariation between the two BUTs and some categorisation of negation. Anscombre's and Ducrot's intuitions about the

BUTs having distinct functions were found to be basically sound. These functions are drawing attention to a contrast between two objects or descriptions, while selecting only one of them (BUT_{sn}), and preventing an inference or consequence from some accepted fact (BUT_{pa}). The descriptive and metalinguistic/polemic negations, if defined in terms of the presentation of new information and the objection to a previous utterance, are viewed as speech acts. The descriptions of the BUTs do not entail that either of them should be used exclusively tied to one of these speech acts, as they are defined at a different level of description, and the data does not support such a connection.

It was also considered whether the BUTs would combine with different types of negation with respect to whether it can be defined as a truth function or not. It was shown that *"not"* preceding both the BUTs span the same wide range of uses as negation in general.

Even if it had been possible to show that a negation before one of the BUTs tended to have a particular function, that would not have proved that it was a different type of negation—it would only have shown that negation *could* be used for that effect. It would still have remained to show that it was a different negation, as opposed to a specific use of the ordinary one.

4.4 The Behaviour of Negation

In certain circumstances, the behaviour of negation, or the interaction between negation and other linguistic expressions is distinctly at odds with what one would expect given the conventional view of it as corresponding to the bivalent truth function. This section discusses three important features of natural language negation, picking up and expanding on previous threads: the interaction between negation and BUT_{sn}, how one must know what one is negating, and finally, negation in yes/no-questions.

4.4.1 More on BUT$_{sn}$ and Negation

When the BUTs were discussed in section 4.3, only the nature of a possible preceding negation was paid attention to. This was because it was the only connection between them and negation considered both by Anscombre and Ducrot (1977) and by Horn (1989). This section considers another interesting requirement on negation by BUT$_{sn}$, which is analogous to one already mentioned by Anscombre and Ducrot.

Anscombre and Ducrot note that BUT$_{sn}$ needs an overt preceding negation, i.e. that in a sentence of the form P BUT$_{sn}$ Q, there must be a *"not"*, *"no"*, *"nothing"*, *"none"* or similar in the P-clause. This accounts for the difference in acceptability between the two Swedish sentences in example 4.38.

(4.38) a. Det är inte praktiskt, utan omständigt.
It is not practical, BUT$_{sn}$ complicated.
'It's not practical, but complicated.'

b. *Det är opraktiskt, utan omständigt.
It is impractical, BUT$_{sn}$ complicated.

But equally interesting, the inverse holds for the Q-clause: an overt main clause negation is usually *not* permitted. That is, similar to the pair in example 4.38, there are pairs like the German one in example 4.54.

(4.54) a. *Es ist nicht schwer, sondern nicht möglich.
It is not difficult, BUT$_{sn}$ not possible.
'It isn't difficult, but not possible.'

b. Es ist nicht schwer, sondern unmöglich.
It is not difficult, BUT$_{sn}$ impossible.
'It isn't difficult, but impossible.'

If this restriction is considered in the view of the function of BUT$_{sn}$, the result is rather startling. The function of BUT$_{sn}$ is to reject one description in favour of

another one, but a main clause negation is not premitted in this presentation of what would conventionally be considered new information. This is a significant observation in the context of Ducrot's and Horn's proposals about negation: if there were such a thing as a descriptive negation, which is used only when there is no reference to a previous utterance, then the clause following a BUT_{sn} would be a typical syntactic context where one would expect to find one. But negation cannot occur in this context, despite negative prefixes being allowed. This is particularly problematic for Horn, who suggested that prefixal incorporation is a test of descriptive negation.

The fact that negation is restricted this way suggests that no function of it can be exhaustively described as a property of information—if it could, then *"not possible"* and *"impossible"* should be interchangeable in this context. It is possible that the function of BUT_{sn}, rejection and presentation of a description, may be central to a negation being necessary and forbidden in the two clauses respectively. While the claim that negation objects to erroneous beliefs was noted to be too strong, it seems that its function after all may have to be thought of in terms of considering someone's representation.

4.4.2 Do You Know What You Are Talking About?

In section 4.2.3, while considering whether it was possible for negation to apply to the linguistic realisation of an utterance (as opposed to being used to object to it), example 4.25 was employed to illustrate that a speaker cannot negate a sentence (as in example 4.26) unless they know what the speaker who made the mistake refers to:

(4.25) B: Is there anything interesting in it? ("it" refers to a magazine)

A: There's an article about [wudu].
B: [wudu]?
A: Oh shit, I don't remember, is it 'v' or 'w'? You know, the religion.
B: Oh, you mean Voodoo.

(4.26) It's not about [wudu].

It was suggested that in order to use a negation in an utterance which refers to information asserted by someone else, the speaker must be able to identify the objects, individuals, actions, etc. that the information concerns.

This is not only relevant to mispronunciations, but also to other types of data. To see this, it will be useful to consider the example below again, as well as some related ones.

(4.55) The king of France is not bald, there is no king of France.

There are basically three types of accounts of this sentence:

1. The definite article *"the"* can be viewed as a logical operator, which specifies that one, and only one, referent exists. If one of these conditions fail, then the sentence is false. This was the approach taken by Russell (*On Denoting*, 1967), and a similar one is probably necessary if negation is to be kept truth functional as on Kempson's account.

2. The expression *"the king of France"* lacks a referent, so the sentence *"the king of France is bald"* is neither true nor false. The negation in example 4.55 indicates that the sentence has a third truth value because of presupposition failure (Seuren, 1985).

3. The expression *"the king of France"* lacks a referent, so the sentence *"the king of France is bald"* was badly used. The negation in example 4.55 is metalinguistic, and indicates that the speaker objects to the fact that someone has used it (e.g. Anscombre and Ducrot, Horn).

Although the details differ, all three approaches are based on the assumption that the negation indicates referential failure. There are however some problems with this. Firstly, although it may seem that this is what the *"not"* in example 4.55 does, negation cannot in general indicate referential failure in the Russellian, one-and-only-one sense. If the problem is that there are several possible referents of the definite expression, *"not"* cannot be used.

(4.56) A: Lisa's teacher is mad.
B: Which one?
B': *Lisa's teacher isn't mad, she's got more than one teacher.

The B'-sentence would not constitute an acceptable reply to A's utterance. This is inconsistent with the logical explanation of *"the"* as a logical quantifier, and so rules it out. For the other two, the sentence *"Lisa's teacher is mad"* is presumably ambiguous, and it is possible that the difficulties with using the negation can be explained as originating from this fact. But excess of referents is not the only problematic case. Consider this dialogue:

(4.57) A: Was the tape that Justin gave you any good?
B: What tape?
B': Justin didn't give me any tape.
B": *The tape that Justin gave me wasn't any good, he didn't give me any tape.

While the first two replies in example 4.57 come across as natural, the third one does not.[14] In view of examples 4.25, 4.56, and 4.57, it seems that a sentence

[14]Atlas (personal communication) argues that this sentence would be less bad without the contraction *"wasn't"*. Whether this is so, contracting *"is not"* in example 4.55 does not seem

cannot be negated unless the speaker can identify the entities (objects, individuals, relations, descriptions, etc.) that the sentence refers to. The B speakers in these examples do not know what [wudu] is meant to refer to, which one of Lisa's teachers the A speaker intends, or what a referent of *"the tape* would be.

These examples are in fact typical—referential failure usually cannot be indicated by a negation. Instead, it is the sentence in example 4.55 which must be accounted for as an exception.[15] The reason why it seems possible for the negation to indicate referential failure in that sentence is that we have some background knowledge about heads of state, how they tend to be the only person in their position, etc. Especially with some habituation, we can reason with "the king of France" as a hypothetical, but specific, individual. B in example 4.57 does not have this option—she has no knowledge of a referent of *"the tape"*.

It would be impractical to deny *"the"* the function of picking out a particular referent, or set of referents. The fact that one referent is singled out is, in addition, definitely part of the information that the utterance of such a sentence conveys. But a negation cannot apply to this referential information. An alternative way of thinking about negation, which avoids this problem, is to distinguish between the information *in* a sentence and the representation that the sentence refers to in the agent's mind. A referring expression is an instruction to use a certain object in the representation. Negation applies to representations of the world, and not to the information carried by the sentence.

---

to have any effect, although these examples are parallel.

[15]Although there is no standard account of definite noun phrases in DRT, it seems likely that *"The king of France is not bald"* would fail to have a proper representation on the grounds that the expression *"the king of France"* lacks a referent (see chapter 6). This means that contrary to most accounts, DRT would not consider the sentence well-formed with respect to a world where there is no king of France.

4.4.3 Yes/No-Questions

The most compelling argument against considering negation truth functional is probably its behaviour in yes/no-questions (this issue was first raised in chapter 2, section 2.2.7). It is not only an argument against the truth functional account, but also provides further evidence that negation should not be seen as applying to the information of the sentence.

(4.58) a. Is that Alan Peiper?
b. Isn't that Alan Peiper?

As for the factual verification requested by utterances these two sentences, they are equivalent: both can be used to ask for a confirmation of whether a designated person has the property of being Alan Peiper. To answer *"yes"* or *"no"* would amount to making the same statement whichever of the questions were asked. As observed in chapter 2, the two sentences are not completely equivalent, in that the second one would probably be interpreted as the speaker is quite sure that the person in question is Alan Peiper. This is less likely if the (a)-sentence is used. This makes it seem that negation is used about representations, rather than being a feature of them.

4.4.4 Some Properties of Natural Language Negation

This section used some data which has not been widely taken into consideration so that two hypotheses about natural language negation can now be formulated. The first one is that a negation does not apply to the information *in* a sentence, but to the parts of a representation of the topic that the sentence *refers* to. This is supported by two observations. Firstly, negation cannot apply to referential failure as a feature of a sentence. It can be used by a speaker to draw attention to referential failure under certain circumstances, i.e. when the speaker can

conceptualise a hypothetical entity which would have been a referent. But lack of referent on its own does not license the use of a negation. Secondly, using a negative yes/no-question is not equivalent to asking for a verification that a description is inaccurate. The function of the negation seems to be to indicate something about an agent's representation, and not of the description as such.

The seconatlas claims that noone did not attend the party has two descriptive negations.d hypothesis is that using a negation cannot be viewed as conveying the status of a proposition—instead, it should be seen as an attempt to perform an operation on some agent's representation. The reason for making this assumption is that the restrictions on negation in combination with BUT_{sn} suggest that a negation cannot be used in an expression which is used to convey what according to the speaker is a suitable description.

4.5 Summary

This chapter has examined the use and interpretation of negation in relation to various other linguistic phenomena, markedness, contrast and referential failure, to name a few. The purpose of this was twofold: to decide whether negation should be considered to have more than one function, and to gather data which provide indications as to what is required of a representational framework in which it can be characterised.

No support was found for considering negation to have more than one function in terms of there being any special linguistic constraints associated with some particular use of it. A more realistic approach would be to argue that it is the available contextual and background knowledge which determines what function the negation has. Speakers may help their addressee to select a certain interpretation by linguistic means, but the speaker uses e.g. a rectification in order to suggest one particular interpretation, it is not the fact that a rectification was used

that had this effect (i.e. the interpretation is not a consequence of the linguistic form). Rather, it is the changes to the addressee's representation(s) that result from interpreting the rectification that constrains the interpreting of the sentence with the negation.

It was also noted that there is no clear-cut boundary between a negation which applies to information, and one which applies to linguistic realisation. In fact there can be no such thing as a negation whose *function* is to indicate that the speaker dislikes a linguistic realisation. Speakers can *use* negation for this purpose, but no function of it can be *defined* that way, without causing erroneous predictions. Negation should be viewed exclusively as applying to information. In those cases where a speaker appears to be using negation about a linguistic realisation, the speaker is treating the linguistic mistake as if it denoted some erroneous fact asserted by the addressee.

Using some observations about the interpretation and use of negation, two hypotheses about its function could be formulated: that it applies to representations of the discourse universe rather than to the information of utterances, and that it is used to manipulate someone's representation. This suggests that significantly more representational capacity is required for the purpose of characterising negation than has usually been assumed. The next chapter will discuss these requirements in more detail.

Chapter 5

Towards a Characterisation of Negation

5.1 Introduction

The purpose of this chapter is to propose a characterisation of natural language negation, which adequately accounts for all the data which has been discussed in this thesis. In chapter 4 it was argued that negation should be viewed as a uniform phenomenon, which has a function, rather than a meaning. The reason for this is that at least in some cases, it would lead to an inadequate analysis to assume that there is some operator corresponding to the negation at the level of the factual descriptions that sentences are conventionally taken to express. For this, and various other reasons, it was suggested that negation should be viewed as being *about* an agent's description, rather than being part of it.

The aim here is to define, in general terms, what properties are required of a representational framework to permit the formulation of a description of negation. It is not within the scope of this thesis to try and suggest appropriate amendments

to any existing framework. Rather, the characterisation will be given in theory neutral terms, and the terminology and illustrations that are used are employed mainly for reasons of transparency.

Despite the fact that a large body of data has been considered, there are still not many indications as to what negation actually does. Instead of speculating further, the approach taken here will be to begin by discussing another problem, although one that has been suggested to be partially related: the description of the two BUTs. The idea is that by providing a set of properties of information that permits a description of what the BUTs are used for, these properties could also be used for the description of negation. I will attempt such a description in section 5.4.

5.2 Perspective in the Representation of Information

The claim that negation is not part of, but about, someone's representation of the discourse universe has some important consequences for how the relation between sentences and a representation of the world should be viewed. It is not possible to think of sentences as factual descriptions of the world. In order to characterise negation it is necessary to consider not only the representation of the person who utters it, but also how the speaker thinks that the other participants of the communication represent the topic.

Although it is common for linguists to discuss language as if the informational aspects of many sentences can be described in terms of a single representation, the idea of maintaining separate representations for the participants of a communication is not new. Multiple representations have been used extensively in Artificial Intelligence and are often seen as necessary both for the production and

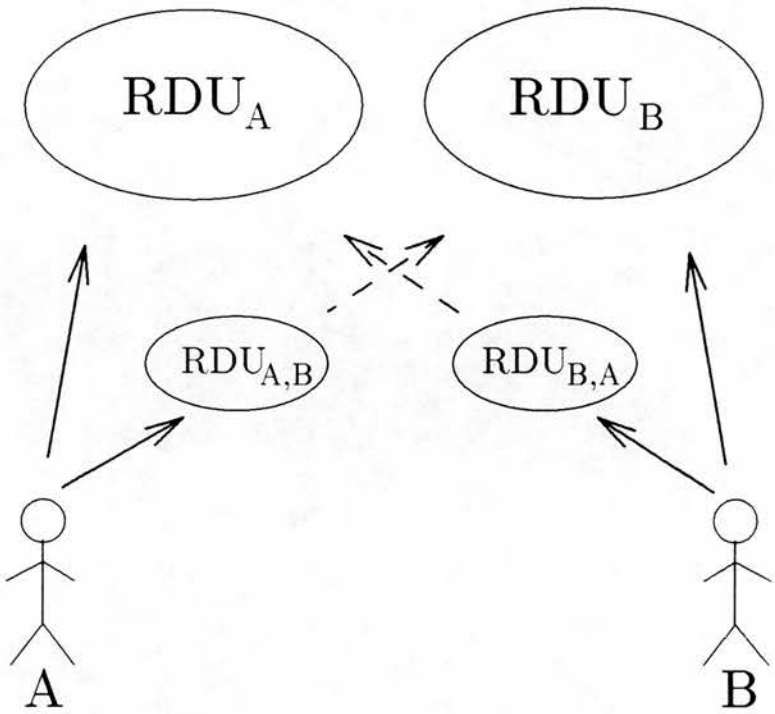

Figure 5.1: RDUs in a two-person communication.

interpretation of utterances (e.g. Cohen and Perrault, 1979).

The basic assumption of these frameworks is that each participant of a communication have their own representation of the topic, which they may believe to be accurate or which they have just chosen for the present purposes. They also have a representation of how the other participants in the communication represent the topic, either individually or collectively, but these secondary representations are not identical to the primary ones of the people concerned.

For the present purposes, the following notation will be used: the different representations, or perspectives, of the topic will be labeled Representations of the Discourse Universe (RDUs), and they will be indexed for who holds them, and, if they are secondary representations, for whom they are held (as exemplified in figure 5.1). One reason for choosing this notation is that it is more convenient for talking about the objects *in*, and the structure *of*, someone's representations, which would be difficult using a more conventional BELIEVE(AGENT,PROPOSITION) notation. There are also other reasons for avoiding a commitment to some par-

ticular framework, in terms of the assumption of what the nature of these representations are (beliefs or temporary constructions).[1] This will discussed in section 5.3.

Multiple representations are often used to allow for the definition of speech acts, at the level of adding and removing propositions in an agent's representation (cf. Cohen and Perrault, 1979). The type of acts (or intentions) that can be formulated this way are e.g. "inform", "request information" and "deny". Using the RDU-notation, an act such as "inform" may be defined as having the preconditions that a proposition p $\in$ RDU$_{speaker}$ and p $\notin$ RDU$_{speaker,addressee}$. The speaker's goal is that p $\in$ RDU$_{speaker,addressee}$, and the description of the act may specify that this is typically achieved by uttering a declarative sentence which expresses p.

This means that multiple representations are typically used for a fairly high-level planning of the communication, where the goals are defined in terms of adding and possibly removing propositions in them. This is not quite how they will be used here. They will not be considered to contain propositions, but structured representations of information. Using the idea that utterances are made in order to manipulate someone's representations, the important use of them will be to reflect how agents have chosen to structure the concepts in their representations, and what they want to do with this structure.

The RDUs will be seen as representations that agents use for a specific communication. The RDU$_{agent}$ need not be identical to the agent's beliefs about the topic under discussion. Consider for instance a physics professor when she writes a paper for a scientific journal, lectures undergraduates, or gives a series of talks to children on television. In all these cases, she is likely to use different representations, in such a way that what she says is (hopefully) intelligible to the audience. When talking to the children, she might for instance behave as if Newtonian phy-

[1]This is not to be taken as an assertion that a framework based on e.g. a belief operator necessarily represents beliefs, but as it will be argued that the distinction is important, I prefer to use a notation which does not carry the connotations of representing knowledge and/or belief.

sics applies globally. It is a good enough approximation and it works for all the cases she will talk about, so she simplifies and pretends that it always works. The way she talks about physics to her different audiences reflects the way she chooses to represent it on each particular occasion.

Simplification is of course not the only reason why RDUs may differ from the beliefs of the agents who hold them. They may have been asked to argue some arbitrary point in a French conversation class, or by a political party, or they may be lying or joking. The fact that people are able to talk coherently in such situations means that they must be able to represent the information somehow, even if it is not a representation that they actually believe is accurate.

5.3 Representing Information

If a perspectival framework such as the one outlined above is combined with a representation of information in the form of a list of propositions or equivalent, then the level of manipulations that can be expressed are of the same type as the speech acts mentioned above: they permit us to talk about propositions being added or removed from someone's representation. For the purpose of describing negation, this would allow for a characterisation of e.g. "denial of an erroneous belief", which was suggested by Givón (1978) to be its function. However, it was shown in chapter 2 that this was too crude a description, so it will be necessary to use a representation of information which permits more sophisticated operations. In other words, propositions cannot be treated as the basic unit.

Data on the use of negation gives precious few hints to what its actual function is. A common assumption is that it indicates that there is something wrong with a sentence, but this was shown not to apply to e.g. yes/no-questions. A notion of denying an expectation (*pace* Allwood, 1927) may be more suitable, but it is unclear how it should be defined. A consequence of this lack of pointers to what

its function is, is that theorising about negation typically has a taste of vagueness or speculation. Rather than speculating further using the insufficient source of data on negation on its own, the problem of deciding what the relevant features of the representation of information are will be approached from a different angle: by considering what features of a representational framework that would be needed to characterise the two BUTs. Given that BUT_{sn} always occurs together with a negation, and that BUT_{pa} sometimes does, it is possible that the representational properties that are needed to understand these two could also be used for the characterisation of negation.

5.3.1 Constraints associated with BUT_{pa}

5.3.1.1 Communication Matters

How to describe the two BUTs is an issue which has already been considered several times in this thesis. The discussion has so far been centered around the account given in Anscombre and Ducrot (1977), which in the case of BUT_{pa} only refers to argumentative properties, and not to representational ones (the description of BUT_{sn} on the other hand relies on features of the organisation of information). Argumentative orientation was later (1983) defined in terms of topoi, which can be thought of as a type of probabilistic rule. This section will try to link the argumentative properties to representational ones, and examine what is required of the rules.

Anscombre and Ducrot (1977) pointed out that it is not possible to characterise the BUTs in terms of a single representation, i.e. as conveying a factual constraint. If this were the case, then both the replies in the examples below should be acceptable, as they both connect the same factual descriptions. But because of the context from the preceding dialogue, in each case only one of them is suitable. This was taken to suggest that it is necessary to take the addressee's representation

into account.

(5.1) A: Carlos pratar spanska väldigt bra. Är han spanjor?
Carlos talks Spanish very well. Is he Spaniard?
'Carlos speaks Spanish really well. Is he Spanish?'

B: Nej, men han är argentinare.
No, BUT$_{pa}$ he is Argentinian.
'No, but he is Argentinian.'

B': ?Han är inte spanjor, utan argentinare.
He is not Spaniard BUT$_{sn}$ Argentinian.
'He isn't Spanish, he's Argentinian.'

(5.2) A: Juan måste känna till Spanien väl.
Juan must know about Spain well.
'Juan must know Spain well.'

B: *Nej, för han är inte spanjor, men han är argentinare.
No, for he is not Spaniard, BUT$_{pa}$ he is Argentinian.
'No, because he's not Spanish, but he is Argentinian.'

B': Han är inte spanjor, utan argentinare.
He is not Spaniard BUT$_{sn}$ Argentinian.
'He isn't Spanish, he's Argentinian.'

By linking two clauses P and Q with a BUT$_{pa}$, P is presented as an argument against a conclusion which Q is an argument for (or vice versa). Since it is not in B's interest to argue for some conclusion which both the fact that Juan is Argentinian and the assumption that he is Spanish would support, a BUT$_{pa}$ is not suitable.

5.3.1.2 What are the Rules?

The fact that BUT$_{pa}$ must be described in terms of the interaction between participants of the communication does not mean that there is no factual background to the argumentation, only that it is not sufficient for a description. It will be

useful to consider what type of factual constraints an agent may consider when using a BUT_{pa}.

It is clear that the argument relation cannot be thought of as corresponding to material implication. The "obvious" translation into propositional logic (with P and Q expressing the propositions p and q respectively) would be to say that there is a third proposition c such that $p \rightarrow c$ and $q \rightarrow \neg c$. However, with the usual interpretation of material implication, P BUT_{pa} Q would lead to a contradiction ($c \wedge \neg c$). The link between at least p and c must be something different from implication, since this link is always overridden. It would have to be a default implication or a probabilistic rule.

Considering some examples, it becomes clear that the relation between q and c is not typically one of material implication either, and the notion of default is often too strong to capture the sort of links that are referred to. Consider the following two examples:

(5.3) A: Victor speaks Spanish really well. Is he Spanish?
B: No, but he's Argentinian.

(5.4) A: Victor speaks Spanish really well. Is he Spanish?
B: No, but he's married to a Spanish woman.

Assume that they can be explained in terms of defaults that are overridden. Then we have for example 5.3:

X is not Spanish $\overset{default}{\longrightarrow}$ X does not speak Spanish.
X is Argentinian $\overset{default}{\longrightarrow}$ X speaks Spanish.

and for example 5.4:

X is not Spanish $\overset{default}{\longrightarrow}$ X does not speak Spanish.
X is married to a Spaniard $\overset{default}{\longrightarrow}$ X speaks Spanish.

To begin with the links that were used for the second clauses of the examples, it is probably reasonable to assume that the speaker can appeal to an accepted generalisation that Argentinian people speak Spanish. It is however questionable if the same can be said to hold for people who are married to Spaniards (at least if they are of a different nationality and live outside Spain). So is the speaker really trying to make the addressee access a general rule of this form, or is the speaker in fact conveying specific knowledge about Victor? The question is whether on the one hand, the speaker uses the *"but"* and the addressee separately interpretations the sentence by using their own internalised rules, or, on the other hand, if the speaker uses it in order that the addressee can infer that a particular individual has some property (which may be more or less likely).

Interpreting *"but"* has often been viewed statically (semantically), as a question of finding the right rules; whether the sentence can be understood or not depends on whether the the addressee's representation contains the required rule or not. Alternatively, it could be viewed dynamically, as the speaker telling the addressee how to structure the knowledge, possibly by making a connection which is only appropriate in the particular context where the sentence is uttered.

The appeal to default rules for the first clause also raises some questions. The notion of a generalisation is in principle more plausible with respect to the state of affairs referred to by this clause, as the *"but"* often appears to be used in order to prevent the addressee from making what could be construed as a reasonable inference. However, the fact that some generalisation or default rule is involved does not necessarily mean that the relation between the fact referred to by the clause and the linking context is in the form of a default, at least not where the clause contains a negation, as in examples 5.3 and 5.4. It is easy to accept a default rule from being Spanish to speaking Spanish, but should the same be said

to hold between not being Spanish and not speaking Spanish? This is where it is appealing to switch from properties of information to properties of argumentation.

Anscombre and Ducrot suggest that if a sentence S can be used to make an argument for C in a context (for instance because S refers to a fact s which implies C by default) then uttering NEG-S would be an argument against C in the same context (the "law of negation"). The law of negation is not a claim about propositions, but about the significance of referring to a state of affairs in a discourse—it cannot be concluded from $s \rightarrow C$ that $\neg s \rightarrow \neg C$. However, if a fact s could be used in a particular discourse to support a certain other fact C, and someone takes the trouble to point out that s does not obtain, it does seem reasonable that the utterance should be taken as an argument against C unless the speaker indicates otherwise.[2]

In principle, it would be possible to attribute the link between 'NOT BEING SPANISH' and 'NOT SPEAKING SPANISH' to a default rule. But it is not a rule that one would assume that people would have internalised. The alternative, which is not based on the generalisations or rules as such, but on their applicability under the circumstances offers a more satisfactory explanation.

5.3.1.3 More than Simple Generalisations

The previous section dealt with whether BUT_{pa} can be defined in terms of the properties of the states referred to by the clauses it conjoins. It was suggested that rather than characterising BUT_{pa} directly in terms of probabilistic rules that are overridden, it would be more realistic to use argumentative properties, which rely on, but are not equivalent to, such rules. This section further discusses the rules and generalisations that arguments can be based on.

The examples considered above used BUT_{pa} in order to contrast descriptions of an

[2]As shown in chapter 4, this phenomenon is stronger than simply depending on the fact that the sentence is uttered.

individual. The data of BUT_{pa} that is studied is often of this type: the *"but"* is used to indicate that some entity does not have a property that it would usually be expected to have. The type of knowledge that is required is declarative; it relates to properties of classes of objects. Another such example is given below:

(5.5) A: But I thought he had a PhD.
B: No, he's a lecturer, but he actually only has an MA.

There are however other types of examples, where more complex representations are required. It may for instance be necessary to consider properties of series of actions, as illustrated by the following example:

(5.6) The waiter came, but he didn't bring a menu. Later we learnt that there was usually one, or sometimes two main courses, and a more or less fixed set of starters—you just go in, sit down, tell them which starter you want and then get served.

It is not a default property of waiters that they bring menus when they go to a table. However, it is something that they often do the first time they go to a table after a customer has sat down at it (in certain types of restaurant). As in the case of the lecturer in example 5.5 above, the restaurant in example 5.6 could be said to deviate from some expectations, but the latter example differs from the former in that the deviation cannot simply be attributed to the waiter, but is a property of the waiter's behaviour at a certain stage of a routine.

It has been suggested that this type of knowledge, which relates to procedures rather than objects can be represented in the form of plans in which subgoals are executed, e.g. STRIPS (Fikes, 1971), or scripts (Schank and Abelson, 1977). Examples such as example 5.6 led Jayez (1989) to attempt a description of BUT_{pa} in terms of a representation which is inspired by AI planning systems. The significance of this slight digression is that the generalisations that speakers have in mind when using a BUT_{pa} are not always easily captured by simple rules.

5.3.1.4 Who Infers?

In the examples discussed so far, the BUT$_{pa}$ has been viewed as having the function of indicating that some state of affairs C which could be assumed to obtain if the clause preceding the *"but"* is accepted as accurate, does in fact not obtain. For many of the examples, it is natural to assume that the speaker is using the BUT$_{pa}$ sentence in order to prevent the addressee from inferring that C obtains. However, using a BUT$_{pa}$ is not always equivalent to trying to avert a misrepresentation. Consider the following example:

(5.7) If you can't take out the cable at the lever end, which is probably unlikely, then you'll have to undo the bolts on the calipers. Mark the cable positions before you do, and you may have to take off any cable end covers, but be careful not to fray the ends.

In this case, the speaker is not trying to prevent the addressee from *inferring* that the cables will fray. If the writer had considered the addressee to be aware of the problem of the cables possibly fraying, he would probably not have expressed it in this way.[3] So this is not a case of preventing an inference, but rather of trying to make the addressee take actions such that a situation where she will have to add to her representation that the cables are frayed does not obtain. Although the *"but"* is used to prevent a certain description from entering the addressee's representation (as in the previous examples), in this case it is not a question of avoiding a misrepresentation, but of avoiding a state of affairs which would make the description accurate.

5.3.1.5 Summary

Describing BUT$_{pa}$ is an intricate problem, and this exposition has hardly paved the way for a straightforward characterisation of it. It should be clear that it

[3]In such circumstances, if the fraying were to be mentioned, *"and"* would probably be a better choice.

has argumentative properties, which cannot be ignored, but it is not trivial how these can be defined. Although the notion of argumentative orientation might be intuitively acceptable, it hardly solves the problem unless it is clearly stated *how* they reflect topoi or rules, since the data examined above seems to suggest that there is no direct correspondence between the argument relation and some type of rules.

The nature of the rules or links that are used for these purposes is by no means a settled matter either, as demonstrated in the discussion above. It would have been elegant to be able to formulate a rule which says that speakers use sentences of the type P BUT_{pa} Q when they want to make a point of rejecting the conclusion C (the linking context), while accepting P, which describes a fact p, if the probability of C given that p exceeds a certain threshold (for an utterance of the type ¬P BUT_{pa} Q it could be preferrable to refer to a rule which uses p rather than ¬p, see section 5.3.1.2). However, this does not seem to correctly describe what is going on.

Consider example 5.7 again. It is not necessary that the cables usually fray for the writer to warn the reader that it might happen when the cable covers are taken off. It does not have to happen most of the time. In fact it is not even necessary that it has ever happened, as long as the writer thinks that it is a possible, and significant, consequence of removing the cable ends.

It would be possible to use a BUT_{pa} in example 5.7 even if the probability of the ends fraying is low. Using the same probability threshold in another situation, however, may not work. Assume that the same person who wrote the text has knowledge about an individual called Stephan who regularly goes hill walking in the Scottish Highlands. This being Scotland, the probability that it rains when Stephan goes hill-walking is much higher than the probability that the cable ends will fray. Now assume that the writer wanted to report that last time Stephan went hill-walking, it was not raining. If the choice of BUT_{pa} only depended on the probabilities of the two correlations exceeding a certain threshold, then it would be

predicted that the sentence in example 5.8a is better than that in example 5.8b.

(5.8) a. Stephan went hill walking on Sunday, but it wasn't raining.
b. Stephan went hill walking on Sunday, and it wasn't raining.

The linking context in this example is that it did not rain (analogically to the cables not fraying being a linking context in example 5.7).[4] One could speculate about the relative importance of links, or perhaps the difference in preference is due to the fact that the weather being dry is usually seen as advantageous, as opposed to the cables fraying. Clearly, a lot of research remains to be done.

5.3.2 Constraints associated with BUT$_{sn}$

BUT$_{sn}$ is used to contrast two description of the same type. As in the case of BUT$_{pa}$, the reasons for considering two descriptions to be of the same type may, but does not have to be a globally applicable constraint. The following three examples illustrate how properties and actions can be contrasted for increasingly context specific reasons.

(5.9) Han är inte spanjor, utan argentinare.
He is not Spaniard, BUT$_{sn}$ Argentinian.
'He isn't Spanish, he's Argentinian.'

(5.10) Det fanns ingen meny, utan några oaptitliga referenser till vegan bönröra var nedkrafsade på en griffeltavla.
It was/existed no menu, BUT$_{sn}$ some unappetizing references to vegan bean-gunk were scribbled on a blackboard.
'There was no menu, instead a couple of unappetizing references to vegan bean gunk were scribbled on a blackboard.'

[4]This means that the Q-clause refers to the conclusion, rather than to some reason to accept it. In both these examples, the BUT$_{pa}$ is used for a DIRECT OPPOSITION, as opposed to an INDIRECT one, in Anscombre's and Ducrot's terminology.

(5.11) Börja inte med såsen, utan med grönsakerna.
Start not with sauce-the, BUT$_{sn}$ with vegetables-the.
'Don't start with the sauce, start with the vegetables.'

While BUT$_{pa}$ is used to indicate that some states of affairs which the speaker considers to be correlated should not be taken to stand in their ordinary relationship on a particular occasion, BUT$_{sn}$ is used for quite different purposes. As with BUT$_{pa}$, there is a requirement that the two contrasted entities have something in common, but when a BUT$_{sn}$ is used to reject a description, the speaker typically indicates that the reasoning that would lead to the adoption of it is invalid.

(5.12) A: Måns har varit och grävt bland ringblommorna igen.
Måns has been and dug among the calendulas again.
'Måns has been digging in the calendulas again.

B: Det var inte Måns, utan den där katten som inte har någon svans.
It was not Måns, BUT$_{sn}$ the there cat that not has any tail.
'It wasn't Måns, it was the cat that doesn't have a tail.'

B simply rejects the association of Måns with the damage done to the flowerbed. There is no element of concession that this was a valid assumption, the way that a BUT$_{pa}$ indicates that some states of affairs would normally be linked.

5.3.3 Comparing BUTs

There is an important general point to be made about the BUTs and the descriptions that they link. Most accounts of the BUTs (BUT$_{pa}$ mainly) attempt to describe them as indicating that some rule does not apply on a particular occasion. However, given an utterance of the type P BUT$_{pa}$ Q, the links between the state of affairs referred to by P and Q (as obtaining or not obtaining) and the linking context C need not be of the type that one would think of as a commonly accepted rule or generalisation. Consequently, the reason for using a *"but"* is not necessarily to

draw the addressee's attention to a known rule, but it may be used to convey *that* the speaker considers there to be a link. Still, most accounts of BUT_{pa} emphasise their reference to rules.

Kempson (1986) and Blakemore (1987) offer explanations of the function of BUT_{pa}, in which it is essentially seen as a constraint on the context in which the sentence is to be interpreted. Kempson proposes that the function of BUT_{pa} can be described in terms of defaults (p.94):

> the contrastive element in *but* is a lexically specified instruction to the hearer to consider only particular kinds of context, namely those in which some contextual implication drawn off the propositions expressed by the first sentence conflicts with some implication to be drawn off the two sentences taken together.

She argues that part of the linguistic meaning of BUT_{pa} consists of a contextualisation as described below (she does not indicate what the other part would be). It should be said that she specifies this as an instruction as to how a BUT_{pa} should be *interpreted.* She does not argue that this description can be used for the generation of sentences. This is what she suggests (p.94, the symbol "–" corresponds to the negation of predicate logic, the symbols R and R_m are rendered as in the original):

> In interpreting 'P but Q' where P and Q are propositional variables, construct a context C_m for P such that:
>
> C_m & P $\rightarrow$ R
>
> and also construct a context C_n for (P &) Q such that:
>
> C_n & (P &) Q $\rightarrow$ – R_m

Blakemore's (1987) account of BUT_{pa}, like Kempson's, is set in the Relevance framework, but hers is different and is closer to the argumentation based description. As opposed to Kempson, she argues that *"but"* (in the BUT_{pa} sense) does not have

a linguistic meaning, but only affects the "pragmatic interpretation" of utterances, by which she means how the information is assimilated into the addressee's representation. Specifically, she argues that *"but"* does not have *"and"* as part of its meaning. She does not give the particulars of a representation of BUT_{pa}, but writes that it (1987,p.130)

> constrains the interpretation of the proposition it introduces so that its relevance must be understood to lie in its effect on the interpretation of the proposition in the preceding clause. More specifically, the hearer is instructed to process the proposition *but* introduces in a context in which she can derive a proposition logically inconsistent with one assumed to have been derived from the proposition expressed by the utterance of the first clause.

While all these accounts emphasise the importance of using context in the interpretation, they attribute the interpretation of *"but"* to the use of logical rules. It could be thought that this does not matter from the point of interpretation, but one effect of referring to deduction is that it is difficult to explain how we are able to "understand" an utterance with a BUT_{pa} without being able to access some very specific contexts which permit the deduction of two incompatible propositions. Consider for instance the following example, discussed by Kempson:

(5.13) My father's coming to stay in our house, but he's going to be out of the house all day.

On seeing this sentence outside a context, we know that a speaker could use it to indicate that they consider there to be an opposition between the two states of affairs that it refers to. If this sentence was uttered in a context, the addressee would probably have the background knowledge to be able to infer whether the speaker thinks it is good or bad that the father is going to be out of the house—it does not follow from the sentence above. Assuming it is a good thing, it is not necessary to know exactly why that is—it could be intended to be interpreted in

a general way, and there is no reason to assume that an addressee would need to know all the (possibly unpleasant) details.

However, the deductivist approach requires not that the addressee makes a general mental note about the state of the speaker's well-being with respect to the fatherly visit, but that the addressee finds a particular proposition that is implied by the father being around, unless he is out of the house. The particular proposition that Kempson proposes (without actually describing a context) is that the speaker would have to keep the children quiet. But are such specific interpretations really required, and moreover, are they not bordering on the unrealistic?

Another problem is that many utterances with a BUT_{pa} fall well short of providing enough material for a successful deduction until a contradiction is reached. This was pointed out in section 5.3.1.2, with respect to the example in which an agent is questioned whether Carlos proficiency in Spanish is due to him being a Spaniard. The reply *"No, but he's Argentinian"* would have to oppose the two propositions that Carlos is not Spanish and that Carlos is Argentinian in such a way that a *logical* contradiction can be derived. The prospect of adding rules such that propositions like these can lead to logical contradictions seems rather daunting.

In the discussion of example 5.3 (repeated below), it was suggested that if it is to be explained in terms of default or probability rules, it would be necessary that agents come equipped with a rule that says that if someone is not Spanish, then they (probably) do not speak Spanish.

(5.3) A: Victor speaks Spanish really well. Is he Spanish?
B: No, but he's Argentinian.

It was argued that an argumentation based account of this example would be more realistic. Such an account would still use e.g. probabilistic rules, but although the argument relation depends on there being a reasoning behind it, it is not isomorphic to the reasoning. For the example above, asserting that someone is

not Spanish can be an argument agains assuming that they speak Spanish, by virtue of the fact that there is a generalisation that Spaniards speak Spanish, and the fact that the speaker found it worthwhile to point out that the individual is not Spanish in a context in which Spanish speaking is relevant. That is, the argument relation depends on the generalisation, but does not correspond to it.

This effectively means that the speaker is considered to have imposed a structure on the information which is specific to the communication. This structure is different from the agents knowledge, as inferences can be made that are based on linguistic actions, and not only on rules governing information. In other words, it is suggested that agents use a type of intermediate representation for the purpose of carrying out individual communications which are based on, but not equivalent to, their knowledge representation. This is why it is desirable to use intermediate representations, which are specific to a particular communication, rather than reflecting a general, fixed, representation of agents' knowledge: argumentation and knowledge representation are different, and something is needed to bridge the gap between them.

5.3.4 BUTs and Structured Information

It will be useful to look at some of the examples of the BUTs with respect to the information structures that they may correspond to. Let us reconsider examples 5.1 and 5.2.

(5.1) A: Carlos pratar spanska väldigt bra. Är han spanjor?
Carlos talks Spanish very well. Is he Spaniard?
'Carlos speaks Spanish really well. Is he Spanish?'

B: Nej, men han är argentinare.
No, BUT_{pa} he is Argentinian.
'No, but he is Argentinian.'

B': ?Han är inte spanjor, utan argentinare.
He is not Spaniard BUT$_{sn}$ Argentinian.
'He isn't Spanish, he's Argentinian.'

(5.2) A: Juan måste känna till Spanien väl.
Juan must know about Spain well.
'Juan must know Spain well.'

B: *Nej, för han är inte spanjor, men han är argentinare.
No, for he is not Spaniard, BUT$_{pa}$ he is Argentinian.
'No, because he's not Spanish, but he is Argentinian.'

B': Han är inte spanjor, utan argentinare.
He is not Spaniard BUT$_{sn}$ Argentinian.
'He isn't Spanish, he's Argentinian.'

The setting initiated by A in example 5.1 is that Carlos speaks Spanish, which A tries to relate to Carlos being Spanish. That is, A first associates 'CARLOS' with 'SPEAKING SPANISH', and then tentatively also with 'BEING SPANISH'. On hearing this utterance, B attempts to reconstruct A's representation (RDU$_{B,A}$), by making the associations that A has indicated in it, and also assumes that A considers the descriptions 'SPEAKING SPANISH' and 'BEING SPANISH' to be linked figure 5.2. The full lines in the figure represent the associations that B thinks that A has made (the one between 'CARLOS' and 'BEING SPANISH' is only tentative, which is indicated with a question mark). The dotted line represents a link between two descriptions, that is, B assumes that A thinks that if one applies, then so does the other.[5]

B also constructs a corresponding structure in RDU$_B$. 'CARLOS' is associated with 'SPEAKING SPANISH', and with 'BEING ARGENTINIAN', and these descriptions are correlated. B accepts that 'BEING SPANISH' is correlated with 'SPEAKING SPANISH' too. Presumably B could think about a lot of other descriptions that

[5]The notation is used for the sake of transparency, to illustrate the "association" and "link" relations. The choice of entities and their labels should not be taken as a claim about the exact nature of representations.

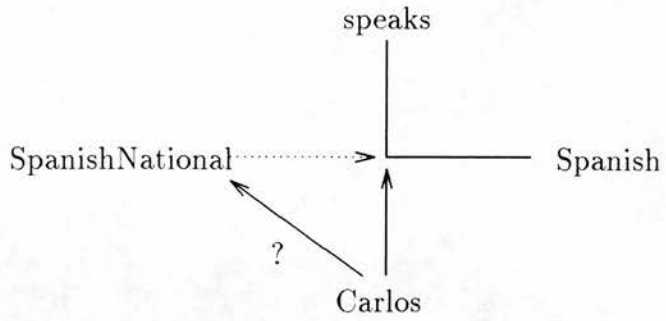

Figure 5.2: $\text{RDU}_{B,A}$ after A's utterance in example 5.1

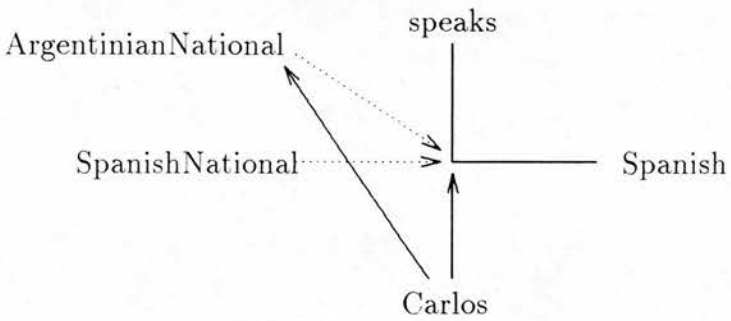

Figure 5.3: RDU_B after A's utterance in example 5.1

would be linked to the property of speaking Spanish too, such as being a Hispanic American, being married to someone Spanish-speaking, having studied Spanish, etc. But although B probably knows that there are other possible reasons, it is unlikely that B will use that knowledge in this context. A possible representation of RDU_B is given in figure 5.3.

The RDU_B and the $\text{RDU}_{B,A}$ are different. B judges that RDU_B is the more adequate one, and therefore wants A not to associate 'CARLOS' with 'BEING SPANISH'. But as an acknowledgement that the first association was accurate ('CARLOS' with 'SPEAKING SPANISH', B provides an alternative association: 'CARLOS' with 'BEING ARGENTINIAN' (figure 5.4). The BUT_{pa} indicates that both the rejected and the preferred descriptions are linked to the first one.

In example 5.2, A associates 'JUAN' with 'KNOWS SPAIN' for some reason that does not show in the discourse fragment. B associates the relevant entities in $\text{RDU}_{B,A}$. From B's reply it appears that B thought that the reason for A's as-

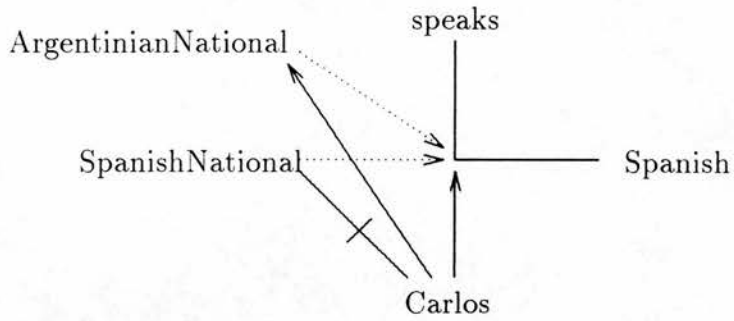

Figure 5.4: Intended changes to $\text{RDU}_{B,A}$ after B's utterance in example 5.1

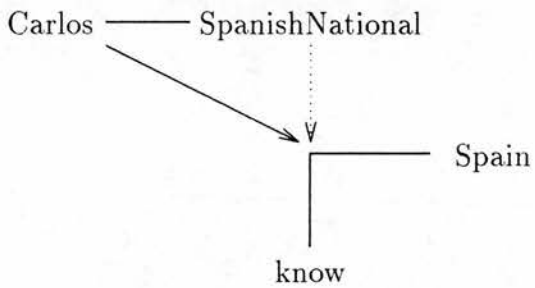

Figure 5.5: $\text{RDU}_{B,A}$ after A's utterance in example 5.2

sumption derives from another assumption, that Juan is Spanish (figure 5.5).

B disagrees with both these associations, on the grounds that Juan is Argentinian. Being Argentinian and being Spanish are descriptions of the same type, (for whatever reason B might consider—possibly because they are both nationalities, or because they are both correlated with Spanish speaking, if that is relevant enough for B), and JUAN should only be associated with the former (figure 5.6—the dashed line indicates that two entities are of the same type. T is used as a placeholder for whatever type B considered.). That is, B is effectively trying to cut the association between 'JUAN' and 'BEING SPANISH' in A's representation.

Figure 5.6: RDU_B after A's utterance in example 5.2

Figure 5.7: Intended changes to $\text{RDU}_{B,A}$ after B's utterance in example 5.2

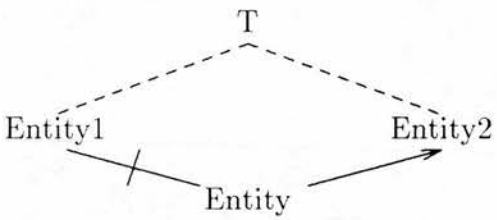

Figure 5.8: BUT_{sn}

and associate it with 'BEING ARGENTINIAN' instead (figure 5.7).

The fact that there is no linking description in this context is why BUT_{pa} cannot be used. As for using a BUT_{sn} in example 5.1, it is possible, but it requires the speaker to ignore the linking description and only treat the nationality descriptions as being of the same type, i.e. B does not acknowledge the link from 'SPEAKING SPANISH' to 'BEING SPANISH'—A is simply wrong and has not made a reasonable guess.

From these examples, it can tentatively be assumed that the information structures referred to by BUT_{sn} is of the type illustrated figure 5.8, and that one typical structure referred to by BUT_{pa} is as in figure 5.9. The first figure illustrates how a BUT_{sn} indicates that an entity in the RDU has been or could be associated with some other entity (Entity1 in the figure), but this should be avoided, and instead it should be associated with another one (Entity2) of the same type. It is worth recalling that "entity" is used in a very weak sense, meaning an object, a relation, an association of an object and a relation, etc.

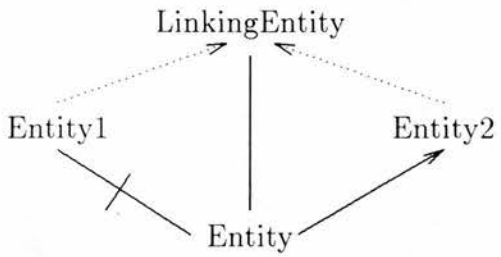

Figure 5.9: BUT_{pa}

The second figure illustrates one of the possible structures that could be manipulated by using a BUT_{pa}. The one in the figure corresponds to a sentence of the type $\neg$-PBUT_{pa}Q where there is an indirect opposition between some entities referred to by P and Q. This is the most common type of sentence that has been discussed here, which is why I have chosen to illustrate that. The extension to direct oppositions and inhibitory links between entities can easily be imagined.[6] This suggests that the two BUTs operate on quite different levels of the representation: while BUT_{sn} is about the associations of entities as such, BUT_{pa} is about the links between e.g. properties which determine whether the association of an entity with one property means that it should be associated with another property as well.

5.4 The Function of Negation

This chapter has sought to suggest a conceptual framework such that what has previously been considered to be distinct pragmatic and semantic features can be integrated. By using representations of the discourse universe which are not considered to correspond to general knowledge or beliefs, but whose structure largely are motivated by the particular context, it is possible to characterise connectives

[6] A direct opposition would have been used if B had replied *"No, but he does speak Spanish well"* in example 5.1). Most of the examples of BUT_{pa} discussed here had it link two descriptions whose applicability increase the probability of one another. But it can also be used when to descriptions tend to, or should exclude each other (inhibitory links):

He'd had four pints but he still insisted on driving home.

such as the BUTs as being about someone's representation, rather than corresponding to some state in it. With this background, it is time to return to negation.

In the discussion of BUT_{sn} and negation in chapter 4, it was noted that BUT_{sn} requires a negation in the clause that precedes it and will not allow one in the one that follows it. Taken together with the function of BUT_{sn}, which is to reject the association of an entity with a description and instead associate it with another one, it could be hypothesised that negation indicates that some entity should not be associated with a certain description.

This would explain the confrontational nature of much use of negation, since there are particularly good reasons to reject an unwanted association if someone has indicated that they have made it. As a feature of representations, negation would be used to point out that an association that has been, or could be made should not be. It would not be expected to be used unless the association of the entity with the description could be made, either because the association seems likely (rejection), or because the description applies to a similar entity (contrast). Chapter 6 will use this hypothesis to give an alternative account of the examples that have been discussed in this thesis.

Chapter 6

Examination of the Hypothesis about Negation

6.1 Introduction

The previous chapter suggested that natural language negation can be characterised as indicating that two entities in an agent's representation should not be associated. This chapter will explain how and why this description differs from other ones, by comparing how they account for the type of data that the traditional accounts seem to handle well, as well as the data that were problematic for them. It will also be shown how the characterisation of negation proposed here can be used in conjunction with agent-specific context-motivated representations of the discourse universe to give a comparatively simple account of the use of scalar expressions. Finally, the consequences of using the present approach for the interpretation and generation of utterances with negation will be discussed.

6.2 Negation and the Representation of Information

The characterisation of natural language negation proposed in chapter 5 makes the assumption that sentences do not have representations that are isomorphic to some part of an agent's representation. Instead, it was suggested that they contain information of (at least) the following two types:

1. Expressions that refer to entities (concepts) in a representation (individuals, objects, actions, properties, etc.).

2. Expressions which suggest how these entities should be organised.

Steps in this direction have already been taken to account for certain types of linguistic expression. An example of this is the account of referential expressions in Discourse Representation Theory (Kamp and Reyle, in press). DRT acknowledges the need to look at individuals and objects as separate entities. However, properties and actions, etc. are treated in a more traditional way, which means that they are seen as conditions on the objects. Because of this, their approach to negation is essentially equivalent to the classical semantic approach,[1] with a few exceptions.[2]

DRT basically views the task of interpreting sentences as one of finding a set of referents and making sure that a certain set of conditions holds for them.

---

[1] Apart from the verification process, which is more complex.

[2] This is why it has not been considered on its own. The main notable advantage over other semantic accounts is that it would probably not consider

The king of France is not bald.

to have a proper representation in the absence of a king of France, so referential failure could not indicated by negation (I say "probably" because there are no rules for definite noun phrases in Kamp and Reyle, in press).

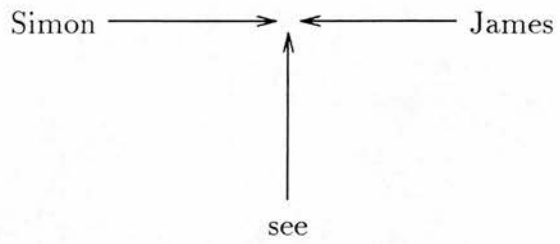

Figure 6.1: The associations made for the sentence in example 6.1

(6.1) Simon saw James.

A sentence such as this is interpreted as being about two entities (say u and v), which have the properties of being Simon and James respectively. It also states another condition on u and v, that u sees v. The sentence is true if two objects u and v exist in the discourse universe such that simon(u), james(v) and see(u,v) is true about them.

For the purpose of characterising negation, it is necessary to take this type of reasoning further: not only objects are entities that we reason about, but so are properties, actions, etc. All these need to be viewed as separate entities which may be associated or not. For instance, the sentence in example 6.1 may be interpreted as conveying that there is a relation between two objects, and that the relation is one of seeing. This can be represented as in figure 6.1, in the same graphical notation as in the previous chapter.[3] It was argued that a speaker who uses a negation indicates that some association between two entities should be rejected. So if the sentence in example 6.2 is used, then (at least) one of the associations in figure 6.1 should not be made.

(6.2) Simon didn't see James.

[3]The reason for using this notation is, as stated in chapter 5, that it is more transparent for the present purposes, not to make any theoretical claims about representations, except that not only objects must be viewed as separate entities. In practice, it would be necessary to account for argument structure, etc., but these issue will be ignored here, for the sake of simplicity.

But which particular one the speaker intends cannot be determined simply by looking at the sentence. It could for instance be about who Simon saw, who saw James, or what the relation between Simon and James is. The addressee should usually be able to determine this from the context, or perhaps the speaker would use intonation to mark one entity, or supply more information which decides it:

(6.3) a. Simon didn't see James. He saw Peter though.
b. Simon didn't see James, but Peter did.
c. Simon didn't see James, but he heard him.

Without a context, it can only be inferred that some association should be rejected, not which one, and because of this it appears equivalent to a representation such as ¬SEE(SIMON,JAMES). However, I claim that in a discourse, addressees are usually able to identify which association the speaker intends, and that this is an important part of decoding a sentence (token) with a negation. This cannot be captured by a simplistic representation like ¬SEE(SIMON,JAMES). It will therefore be necessary to use a more powerful framework, such that all the associations between the relevant concepts can be singled out. There is still a case for viewing negation as being about a representation, rather than part of it, and this is what the following two sections are about.

6.3 Why Negation is Not a Property of Sentences

6.3.1 Yes/No-Questions Revisited

In chapter 4, section 4.4, three issues were discussed which were considered to constitute strong evidence against traditional accounts of negation, as well as

providing important constraints on what type of representation could be used: the interaction between negation and BUT_{sn}, how negation cannot apply to referential failure, and its behaviour in yes/no-questions. The two former were used more or less directly for the formulation of the hypothesis of the function of negation—as evidence that negation applies to (structured) representations—while the third one was only used indirectly, as evidence that negation should not be viewed as an operator which is part of factual descriptions conveyed by sentences.

As shown in chapter 2, section 2.2.7, the problem with negation in yes/no-questions is that whether they include a negation or not does not affect the type of answer that should be given. For instance, the following examples have been discussed:

(6.4) a. Is Manchester rainy?
b. Isn't Manchester rainy?

(6.5) a. Is that Alan Peiper?
b. Isn't that Alan Peiper?

When answering questions such as these, whether a *"yes"* or a *"no"* should be used is not affected by whether the question contains a negation or not, i.e. if a person who replies thinks that Manchester is rainy, *"yes"* would be an appropriate answer to both of the questions in example 6.4. The problem with viewing negation as part of the factual description is that the opposite prediction is made. The sentence in example 6.4(a) would be used to ask for a confirmation that Manchester is rainy, while the (b) sentence would be used to ask for a confirmation that it is not.

If negation is viewed as an indication that two entities in an agents RDU should not be associated, then this observation can be accounted for quite easily. Consider a person A who has just spotted a person in a crowd, and thinks that it might be

Figure 6.2: RDU_A before uttering one of the sentences in example 6.5

Alan Peiper. A's representation (RDU_A) now contains an entity P corresponding to the person, and the property of being Alan Peiper, and A wants to know if they should be associated (figure 6.2). Whether A uses a positive sentence (asks if the association should be made) or one with a negation (asks if it should not be made) they draw attention to a possible association of the two entities. It would be possible to hypothesise that they use a negation if they have already made the link and want to verify that it is correct. A *"yes"* has the function of accepting an association, whereas an *"no"* rejects it (in the same way as *"not"* does). In this case the distinction between being about and being part of a representation matters. *"Not"*, *"yes"* and *"no"* are all about the entities in the representation, and because of this cannot apply to each other.

6.3.2 Double Negation

In section 6.2 it was claimed that interpreting a sentence with a negation typically involves determining which particular association is being questioned. In English, this usually cannot be decided directly from the sentence, as the syntactic position of negation is relatively fixed. Speakers can often rely on their addressees' contextual knowledge to allow them to understand which association is intended.[4] If the

[4]In fact, this is frequently capitalised on, in English and also in other languages, by putting the negation in the preferred (main clause) position, even though it would be possible to have it in a position closer to the references to the entities whose association is rejected. This phenomenon was named neg-raising by Horn (1975, cited in 1989). The first of the two sentences below sounds more natural than the second one, although the second one puts it closer to the expressions denoting the relevant entities.

1. I don't think he has come.

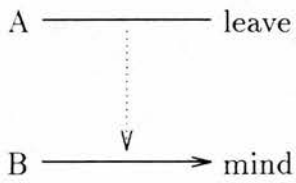

Figure 6.3: RDU$_A$ for example 6.6—which association does B mean?

speaker does not think that the addressee is able to determine which association is relevant, it can be marked linguistically: contrastive intonation, BUT$_{sn}$, *it*-clefts (see von Klopp and Humphreys, in preparation), etc., or they can provide more information about the context.

However, this does not mean that it never happens that a participant in a communication is unsure of which associated entities that a speaker means, or that it never matters. On the contrary, there are many situations which can lead to uncertainty about which association the speaker intends. One typical case of when it is unclear which link the speaker intended arises with questions such as the one illustrated in example 6.6.

(6.6) A: Do you mind if I leave?
B: No.

A *"no"* answer to this type of question often leaves the A speaker feeling not quite sure whether the reply was meant to indicate that B does not want to be associated with having objections to A leaving, or if B wants A not to leave (figure 6.3). But interestingly there is not the same uncertainty if the question includes a negation, such as the one in this example:

(6.7) A: Do you mind if I don't go?
B: No.

2. I think he hasn't come.

In this case, there is only one likely interpretation, that B should not be associated with minding. With a truth functional negation which reverses truth values, this is difficult to explain. It is possible to get the right interpretation of example 6.7 by treating the conditional as a material implication, but then there will only be one possible interpretation of example 6.6. Conversely, if (6.6) is allowed to have two different interpretations, then (6.7) would too.

However, if negation is about representations, it cannot be about another rejection (cf. *"yes"* and *"no"* as replies to negative yes/no-questions in section 6.3.1). So in example 6.7 there is only one association that can be rejected by B's reply, i.e. the one between B and 'MIND', as the one between 'A' and 'STAY' is already rejected in the condition. This is an interesting feature, which leads to an issue which has not yet been considered in this thesis, namely double negation.[5]

In natural language, two negations do not cancel each other out, like the bivalent negation of classical logic. The two sentences in examples 6.8 and 6.9 are not interchangeable; definitely not from the perspective of when they would be appropriate in a discourse, and hardly from a detached "meaning" perspective either. A double negation would not occur unless the sentence which contains it is used as a reply to an earlier utterance which had a single negation, and in such a case, a sentence without a negation could not be used to the same effect.

(6.8) a. I'm not not smoking.
b. I'm smoking.

(6.9) a. I didn't not like it.
b. I liked it.

[5]By double negation I mean a *"not"* combined with another *"not"*, *"never"* or *"no"*. A negation and a negative prefix do not count as a double negation, as negative prefixes have a different function from the free-standing morphemes (see chapter 4, section 4.3.2).

There are many questions surrounding the use of double negation that still need to be clarified. I cannot offer a solution to the problem, but I will sketch a possible explanation based on the characterisation of negation proposed here.

It will be useful to start with some cross-linguistic observations. One notable syntactic restriction on negation in many languages is that it needs to occur with a verb, or a "verbal expression".[6] This means that a double negation can only be used where there are two expressions in the sentence which license it. In English, given that most verbs require the dummy verb *"do"* when negated, this is usually a trivial requirement. But if sentences like those in examples 6.8 and 6.9 are to be translated into e.g. German or Swedish, a reformulation is needed unless they are expressed with some auxiliary verb. To begin with *"I'm not not smoking"*, this cannot be translated directly into either German (6.10a) or Swedish (6.11a). A sentence which could be used in one type of context where the English one could be appropriate needs to have more material in it.[7]

(6.10) a. *Ich rauche nicht nicht.
I smoke not not.

b. Ich bin nicht Nichtraucher.
I am not Non-smoker.
'I'm not a non-smoker.'

(6.11) a. *Jag röker inte inte.
I smoke not not.

b. Jag har inte slutat röka.
I have not stopped smoking.
'I haven't stopped smoking.'

[6]Adjective and preposition phrases may count as verbs for these purposes, in the Germanic languages, as well as any "real" verb. In for example Spanish, however, the requirement seems to be that not only is there a "real" verb, but that it has tense too.

[7]The "translations" were suggested by native speakers as phrases that could be used in the type of situation where they imagined that the English ones would be used.

As for *"I didn't not like it"*, past tense tends to be expressed with the perfect construction in German, which means that there are two verbs, and hence that it can be translated directly. In Swedish this is however not the case.

(6.12) a. Ich habe es nicht nicht gemocht.
I have it not not liked.
'I don't not like it.'

(6.13) a. *Jag tyckte inte inte om den.
I thought not not of it.
'I liked it not not.'

b. Det är inte så att jag inte tycker om den.
It is not so that I not think of it.
'I don't not like it.'/'It's not the case that I don't like it.'

In order to translate a sentence which lacks two verbal expressions into Swedish or German, it is thus necessary to add some linguistic material which will typically amount to saying that the agent is making a point out of avoiding that the description in question applies. Obviously, I would not want to make too strong an argument based on syntactic evidence, but it is nevertheless interesting that it is often not possible to translate a double negation directly into other languages. The fact that it often needs to be expressed along the lines of not wanting to avoid something could indicate that a double negation rejects two associations: one leading to two other entities whose association is also rejected. But I accept that it is not always clear what these associations would be.

To briefly reconnect to the accounts of negation discussed in chapter 3, it is clear that if a negation which applies to information is truth functional in the ordinary way, then the (a) and (b)-sentences in example 6.8 and 6.9 should be equivalent. In accordance with the Gricean tradition, it could perhaps be argued that there must be a reason for a speaker to use a longer expression, since it flouts the

third maxim of manner ("be brief"). This reason could be that the sentence is a shorthand for one of the longer formulations that would be necessary in Swedish and German.

Horn and Anscombre and Ducrot would consider the first *"not"* of a double negation metalinguistic/polemic, objecting to the previous utterance, rather than being used in the ordinary truth functional way. It may seem a fair assumption that double negation is only used to object to the fact that a negation has been used, so stipulating that double negation is only permitted when at least one of them is metalinguistic/polemic might appear to solve the problem of why they do not cancel out. However, even if the assumption about the use of double negation were justified, these two solutions, whether based on the manner maxim or on not permitting a descriptive negation, are only solutions in the sense that they can be *formulated* in the theories. It has not been *explained* why it seems impossible to interpret a double negation as two descriptive negations (in the sense of two truth functions that cancel each other out).[8] Using the characterisation of negation proposed here, it is possible to reach a little further, as negation is prevented from applying to the same association twice.

[8]A sentence with a double negation (NEG-NEG-S) and the corresponding sentence without a negation (S) can never be equivalent from a discourse perspective. This does not mean that they could not both theoretically be used with reference to the same state of affairs. However, they would not suggest the same changes to the addressee's representation(s): S would indicate that one or more associations should be made, while NEG-NEG-S would indicate that some entity should not be associated with a rejected association between two others. That is, the difference between two sentences such as *"I like it"* and *"I don't not like it"* is that with the first sentence, the speaker associates (or accepts the associations of) three entities, while with the second one, the speaker merely removes the associations—he or she has effectively not said anything about the entities. The addressee may feel justified in reconstructing some associations between the entities because of the context, so that their representation ends up identical to what it would have been like if the corresponding sentence without negations had been used. But that is a feature of the *interpretation*, and not of the decoding, of the sentence (as defined in chapter 1).

6.3.3 Equivalent Propositions—Different Utterances

Anscombre and Ducrot (1983) discuss two sentences that correspond to equivalent propositions, but which have different connotations when used in a discourse.

(6.14) Les 3/4 des travailleurs touchés par des suppressions d'emploi ne connaîtront pas une situation de chômage.
The 3/4 of workers touched by [determiner] suppressions of work [negation particle] will-know not a situation of unemployment.
'Three quarters of the workers affected by redundancies will not experience unemployment.'

(6.15) Le 1/4 des travailleurs touchés par des suppressions d'emploi connaîtront une situation de chômage.
The 1/4 of workers touched by [determiner] suppressions of work will-know a situation of unemployment.
'One quarter of the workers affected by redundancies will experience unemployment.'

They suggested that the difference between these two is due to the negation reversing the argumentative orientation that the predicate has when applied to a certain type of object. As pointed out in chapter 4, this may be descriptively accurate, but it hardly explains the phenomenon. Moxey and Sanford (to appear) discuss a related example:

> A salesman is trying to sell a car to a customer. The customer asks a question about reliability, and the salesman can give one of the following replies: A few/Few/Not many of our cars break down within two years of purchase.

Asking subjects which quantifier would be best for the salesman to use, they reliably ranked them as *"few"* being the best one, and *"a few"* as the worst, with *"not many"* in between. Why are *"few"* and *"not many"* not equivalent?

Moxey and Sanford suggest that using *"not many"* may suggest that the customer expected a lot, while *"few"* does not carry this assumption.

On the account proposed here, to use the sentence in example 6.15 is to try to associate an entity (a group of people) with a state (that of being unemployed). The speaker is presenting this as a new connection, and because unemployment for most people is considered to be linked to undesirable consequences, this means that the addressee is making a new connection to something that should be avoided.

By using the sentence in example 6.14, the speaker suggests that an association between an entity (a different group of people) and a state (being unemployed) that has been made or is likely to be made should be rejected. The speaker is effectively treating the situation as if at least someone would have expected the people in question to become unemployed, but that it did not happen. Because of the problems linked with unemployment, this presents the situation as if things are better than expected, and the government is doing well. Similar reasoning applies to the car salesman example.

6.4 Scalar Expressions Revisited

In chapter 5 it was argued that there are significant advantages in using context-specific representations of knowledge as an intermediate stage between communication and reasoning. This was motivated by the apparent appeal to links between descriptions which seem to be motivated by particular contexts rather than regular co-occurrences. This section suggests that there are further reasons for adopting this approach, by showing that it permits a simple account of the use of scalar expressions.

The problem of explaining scalar expressions, seen from the perspective of truth

conditional semantics, is that it seems that there is a default interpretation of sentences such as the one in example 6.16.

(6.16) Pat doesn't have three children.

This example is taken to (typically) mean that Pat has less than three children. In order to achieve this semantically, it is necessary to assume that having four children implies having three, etc., so that all predictions of numbers larger than N are automatically excluded by the rejection of the number N. This type of reasoning is extended to other predicates which appear to behave in a similar way.

However, speakers can, and do, use sentences of this type without excluding the predications that there would be if the reasoning above was followed:

(6.17) Pat doesn't have three children, she has four.

Sentences such as the one in this example are sometimes considered unusual. In the sort of contexts where one may want to talk about how many children someone has, it is probable that the addressee is first lead down the garden path, in that they may initially assume a representation which they will have to change on hearing the second clause.[9] But the addressee can nevertheless make sense of the utterance, and can recognise the use of negation as a valid one, although perhaps unexpected.

Using the concept of an intermediate context-specific representation, I suggest that the reason why the sentence in example 6.17 appears odd is not that it violates any rules for how negation is normally used, but that the speaker has

[9]This is at least what is assumed by e.g. Kempson (1986) and Horn (1985).

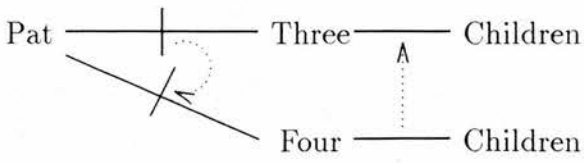

Figure 6.4: One interpretation of *"Pat doesn't have three children"*.

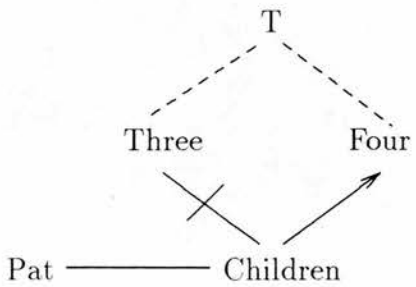

Figure 6.5: *"Pat doesn't have three children, she has four."*

chosen to organise their representation in a slightly different way from what is normally expected *in that type of context.*

The negation in example 6.17 indicates that two entities should not be associated, as with any negation. But when it comes to talking about how many children someone has, it is common to view the cardinality as referring to subsequent supersets. That is perhaps because usually when people talk about how many children they have, parents with four children have all the problems of parents with three children plus presumably a few more, so that having three children is naturally seen as part of having four. If an agent agrees with that, and uses a representation which reflects that, then if a speaker indicates that Pat should not be associated with having three children, this would mean that it can be inferred that she does not have four either (figure 6.4). A speaker who uses the sentence in example 6.17 has however chosen a different organisation, where the numerical modifiers are seen as alternative descriptions of Pat's children (figure 6.5).

The fact that one tends to prefer one organisation over the other might make it appear that the latter is less available, and perhaps unnatural. But it is worth noting that even if the speaker wants to refer to the number three, but this is not the exact one, they will usually indicate the right number and which organisation

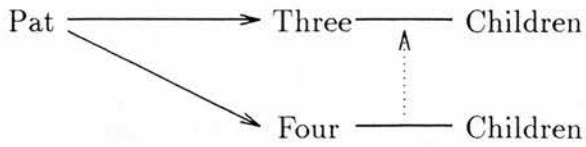

Figure 6.6: *"Pat has three children, four even."*

they have chosen whether they use the "normal" organisation of the information or not. If Pat is associated explicitly with having three children, and then with having four, it is almost necessary to linguistically mark that it is the organisation illustrated in figure 6.6 that is chosen, for instance by using *"even"* or *"in fact"*.

(6.18) a. She has three kids, four even.
b. She has three kids, in fact she has four.

Simply lining up the two descriptions—as in example 6.19—would make a strange utterance.

(6.19) ?She has three kids, she has four.

When talking about the two organisations of the information with respect to the examples above, one of them was referred to as the "normal" one. It is however important to realise that the reason why this one seemed more normal is not the numbers as such, but the type of context one imagines. In other contexts an organisation of the type illustrated in figure 6.4 may not be natural at all.

One example of a different context which does not invite this structure is when discussing mathematics. Say that the participants of a conversation are discussing the equation $x^2 - 3x - 4 = 0$ and that one of them has mistakenly suggested that $x = 3$ is a solution. Now if another participant utters the sentence in example 6.20. they will typically not be interpreted as having claimed that no number larger than three is a solution.

(6.20) Three is not a solution.[10]

Although numbers are ordered, the order is not applicable as licensing inferences about certain numbers from other ones in this context. So while the participants certainly recognise that numbers can be ordered in principle, they are not seen as ordered in the sense of larger numbers implying smaller ones when the topic is solutions to equations. This demonstrates how important it is to distinguish between having the capability of being ordered according to some criteria, and the property of being ordered for some particular purpose.

Other types of context may license yet another organisation of numerical modifiers, as illustrated by example 6.21(b).

(6.21) a. John wants to have three children.
b. Morten wants to run a marathon in three hours.

The (a)-sentence tends to be understood as John not being content if he has less than three children, while the (b)-sentence tends to be interpreted the opposite way—a shorter time than three hours would certainly be expected to please Morten. The order used in the latter example is the opposite from that used in the former, and it is the normal one for discussing satisfactory performance in a fixed distance sporting event.

The use and interpretation of scalar expressions has often been treated as if they depended at least in part on reasoning with intrinsic features of the expressions as such (e.g. Horn, 1989; Kempson, 1986, see quote on page 89). A simpler and more plausible account can be given by appealing to representations whose organisation is motivated by context-specific constraints instead.

[10]On Atlas' view (1990, quoted with permission) this should be written as *"3 is not a solution"*, as he considers *"3"* to mean the numeral, while *"three"* is semantically general. There is no similar distinction in spoken language, so even if the written sentence is anomalous, the discussion above holds for [θ r i:].

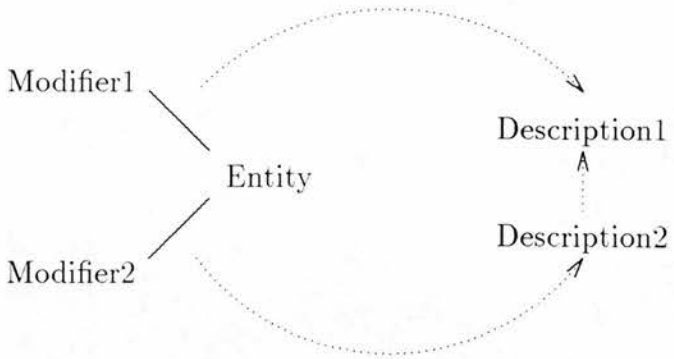

Figure 6.7: When scalar expressions appear to be ordered, that is because they reflect a context-dependent order of a set of modifiers which they refer to. The modifiers are ordered with respect to association with some entity which in turn is linked to some state that can be quantitatively assessed.

It should be pointed out, however, that the "links" are not equivalent to (pragmatic) implication caused by contextual factors, as suggested by Fauconnier (1976) and Horn (1989). The reason why associations of scalar descriptions with certain types of entities appear to be ordered is that they in turned are linked to some separate phenomenon that can be perceived as ordered in terms of increasing or decreasing quantities of a quality. The dependence on an external motivating context is illustrated in figure 6.7—the previous illustrations were simplifications, ignoring the licensing context.

The need for a licensing context is acknowledged by Fauconnier, but he considers it to give rise to a logical ordering of the scalar descriptions also. Anscombre and Ducrot (1983), in their reply to Fauconnier, argue that although the external order justifies an ordering of the scalar expressions, this secondary ordering is not equivalent to a logical one.

Expressions such as *"cold"* and *"freezing"*, are ordered with respect to the type of arguments that a speaker may want to make with reference to the quality of being cold (to take a naive view). They both denote quantities of the same quality, but they are not ordered in the sense that they correspond to different quantities. They have sometimes been suggested to be ordered relatively, so that one implies the other, although order is obviously by stipulation rather than ob-

servation. The argumentative analysis is that the order is not there by virtue of implication (which has to be imposed anyway), but that it feels natural to impose the analysis because of the argumentative properties—using *"freezing"* amounts assigning more importance to the cold, in a context where the "quantity of cold" is correlated with the unpleasantness of taking a swim. The order is there so that we can make arguments of various strenghts, not so that we can indicate that the actual "quantity of cold" is larger than if we had used another term. For that, one would use a temperature, or a precise description of what the effects of the temperature are.

Anscombre and Ducrot base their claim that the ordering of scalar expressions is not defined in terms of factual constraints on the behaviour of *"optional"* and *"forbidden"* (discussed in chapter 3, section 3.4.5). They behave like ordinary scalar expressions, in that *"not optional"* seems to suggest *"compulsory"*, analogically to how for instance *"not permitted"* suggests *"forbidden"*. But while the latter inference can theoretically be explained as being due to *"compulsory"* implying *"permitted* (if it is assumed that for any action there are three possibilities: the agent has to perform the action, must not perform it, or has a choice.), an analogical analysis is not available in the case of *"forbidden"* and *"optional"*.

It is actually possible to describe Anscombre's and Ducrot's notion of an argumentative ordering in terms of a representation without equating it with an inferential link. This is most easily seen if the expressions are considered with respect to figure 6.7. The expressions *"optional"* and *"forbidden"* are used to label modifiers which correspond to approximately 'HAS A CHOICE and 'MUST NOT' respectively. The way to understand Anscombre's and Ducrot's claim about the ordering being argumentative and not factual is to distinguish between the entities in the RDU and the labels that are used to designate them.

Consider a context where the modifier entities 'HAS A CHOICE' and 'MUST NOT' are ordered in such a way that the latter is stronger than the former (this may be because they are used in the context of persuading someone not to do something).

Because the modifiers are ordered in this way, the labels chosen to refer to them are selected from a set that reflects this order. In other circumstances, where they had been ordered the opposite way, *"permitted"* would have been better than *"optional"*.

That is, it must be acknowledged firstly that the order is not intrinsic to the entities in the RDUs as such, but depend on a licensing context, and secondly, that the choice of labels for referring to the ordered entities may well be determined by how the entities as such are considered to be ordered. This is why argumentative ordering can be a property of expressions, while reflecting a context-motivated ordering of some concepts.

6.5 Why Use a Negation?

The study of the type of linguistic phenomena that have been discussed in this thesis has typically been approached as a task of describing how sentences containing them are interpreted. The aim of this work was however to provide a characterisation of natural language negation which can be used for the production of sentences as well.

It was pointed out in section 6.2 that viewing negation as a truth function, without a more expressive notation than e.g. ¬RELATION(X,Y), means that the representation of declarative sentences with negation only contains an indication *that* something is wrong with it. However, it was argued that understanding negative sentences typically requires the agent to determine *what* it is wrong, or in the terminology used here, which particular association should be rejected. The characterisation given here therefore means that the interpretation of a sentence with a negation requires more than the addition of a negative proposition to the representation.

When it comes to the generation of utterances, the difference from the traditional descriptions is significant. It was observed in chapter 2 that it is difficult to explain why speakers would use a sentence with a negation if the truth functional account of it were correct, unless it is accompanied by a theory which specifies when it is suitable to convey negative facts. When negation is characterised as has been suggested here, however, it is comparatively easy to predict in what situation a negation would be suitable.

The idea is that negation is used to indicate that an association has been, or is likely to be made in someone's representation. Each participant of a communication is continuously modelling their own and others' RDUs, and there will arise situations where one of them is likely to contain an association which one participant, A say, does not think or is not certain should be there. To give a possible reason for A to assume this, another participant may have explicitly mentioned that they have made an association, or some association may have been mentioned from which the undesired one could be inferred. Another possible reason for using a negation is when it makes sense to distinguish two sets of objects on the grounds that one set has a property that the other does not. Using the characterisation proposed here for such cases should be fairly straightforward.

6.6 Summary

The appeal of the characterisation of negation suggested in chapter 5 is that it is extremely simple, while still managing to describe natural language negation as a uniform phenomenon. All that is required is to distinguish between entities in the representation and their organisation on the one hand, and talking about this organisation on the other.

This chapter has tried to clarify how this differs from previous approaches, and to demonstrate how it can account for the data that is problematic for them, as well

as for the data that they are compatible with. It was also noted that a number of other observations can be explained in a more satisfactory way.

Chapter 7

Conclusion

7.1 Achievements

The most serious obstacle in the way of progress towards an adequate theory of natural language negation is that it has such a long tradition of being done in a certain way. It is clear that the truth function is inappropriate as a description of many uses of negation, but because it fits into a well-established theoretical framework, it is appealing to keep this description anyway. Instead of focussing on negation in general, the effort has been concentrated on accounting for the exceptions—the academic equivalent of treating the symptoms while ignoring the cause.

The characterisation of negation that I have proposed in this thesis is not a theory of exceptions. What is offered is a new, and simple, description of the basic phenomenon, which also covers the uses that the traditional approach has difficulties in accommodating. In order to formulate this description, it was necessary to revise some assumptions that are commonly made about the function of language in communication. The main points of this can be summarised as follows:

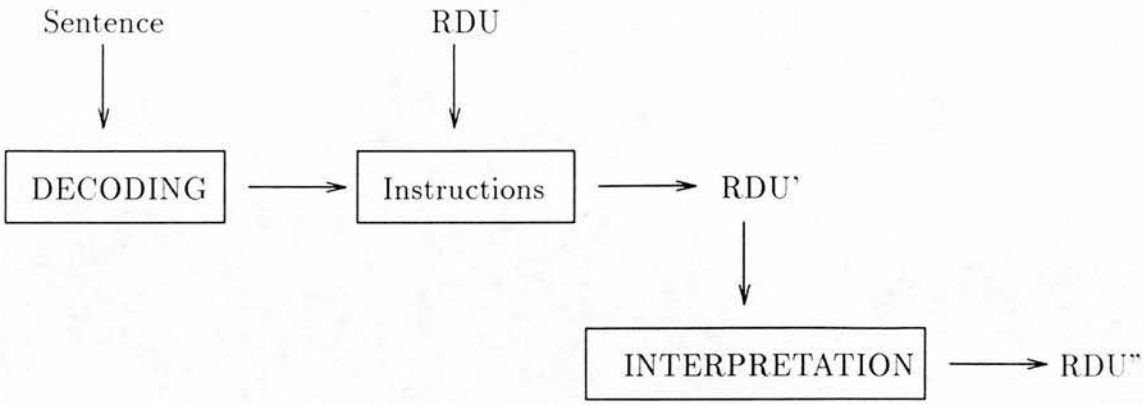

Figure 7.1: Schematic model of the interpretation of an utterance.

- Sentences do not represent factual descriptions (propositions, infons, etc.). Instead, they correspond to descriptions of the structure of a representation of the discourse universe. Sentences are decoded by carrying out changes to the representation. The changes are not equivalent to the result of performing them. Sentences do not have *meanings* in the traditional sense of corresponding to factual descriptions, but *functions*.

- The representations that sentences are about contain entities (individuals, objects, actions, descriptions, etc.—things that can be talked *about*), corresponding to entities in the discourse universe which are organised into a certain structure. The entities are structured and connected.

- The expressions that make up sentences can have (at least) two different functions: they can identify entities in the representations, or they can be used to say something about how some entities should be organised.

This means that the understanding of sentences involves the following steps (figure 7.1): the decoding of a sentence yields a set of instructions to find some entities in a representation of the discourse universe and to change it in some way. The representation is changed, and the agent then tries to determine how the change is significant, which may lead to further changes. With this background, negation can be characterised as indicating that two entities in a representation

of a discourse universe should not be associated.

The step from viewing negation as *being part of* representations of sentences to *applying to* representations of the discourse universe is a simple one, but it nevertheless has some very important consequences. It permits a unified characterisation of how natural language negation is used, and it also works at an explanatory level for many related phenomena, as shown in chapter 6. It also makes it easier to understand why speakers would want to use sentences with negation in a communication.

7.2 Relation to Other Frameworks

In making the assumptions about the representation of sentences that were summarised above, I have committed myself to a particular view of the function of language in communication. It will be useful to compare this with the other theories that have been discussed in this work.

Briefly, the main points of Relevance Theory, or at least its programme, are that sentences do not have meanings, in the sense of corresponding to factual descriptions—utterances do. A sentence only receives a representation when an addressee processes it with respect to a context. Despite the emphasis on the importance of information processing, speakers are not considered to be able to influence it beyond providing simple stimuli in the form of propositions, and logical constraints on them. As for knowledge representation and the function of language, these are intimately linked, in that language is seen primarily as a tool for information processing. The latter can be adequately modelled using natural deduction, and language is used to code propositions which function as input to the deduction process. Natural language expressions such as *"not"* are considered to be truth functional, which means that utterances must express highly specified information, so that the sentence can be assigned a proposition with a truth value.

Atlas' version of Radical Pragmatics differs from Relevance theory in that language is viewed as a tool for communication, rather than directly reflecting some intrinsic reasoning facility. However, the function of language in communication is still seen as providing a means for referring to factual descriptions. Despite sentences being general, utterances are considered to have specific representations, which can be determined in the context of their use—a parallel assumption to the one made in Relevance theory.

The idea that sentences are sense-general and can only be considered to refer to specific facts when they are used suggests that the specificity of meaning in language should be questioned. If sentences do not have a specific meaning on their own, why should they be considered to contain that information when because they are uttered? If it is the speakers processing of the sentence which constructs the meaning of it, then it was never part of the sentence in the first place. The consequence of considering negation in natural language to correspond to a truth function is however that the sentence must *contain* the specific meaning when it is uttered. If the negation in the sentence is a truth function, then the sentence must have a propositional representation that has a truth value, which means that other expressions in the sentence must be considered to express for instance non-overt maximality markers. This may only happen when the sentence is uttered, but it is still necessary that it happens, as the various readings are explained as properties of the expressions, and not as properties of the structure of the representation. In other words, sentences are supposed to carry the very information that the addressee finally arrives at by interpreting them.

Argumentation theory also emphasises the importance of the addressee's processing of information, but there the similarities with Relevance theory and Radical Pragmatics end. For Anscombre and Ducrot, language is a tool for communication, and not for information processing, and communication is not equivalent to conveying factual descriptions. They consider communication to be about steering the addressee's information processing in a certain way.

In Relevance theory guiding someone's processing of information is seen as a question of making them apply the intended deduction rule. In contrast, Anscombre and Ducrot argue that this consists in making the addressee organise their information in a specific way, independently of whether they actually perform the reasoning involved in arriving at this organisation. Natural language contains expressions which have the function of suggesting a certain organisation.

However, Anscombre and Ducrot are mainly interested in the formal (or conventional) attributes of these expressions, to the point that they sometimes seem to suggest that e.g. argumentative orientation is not a property of information at all (it is a "primitive"). But clearly, references to facts can also be used for the purpose of trying to change someone's representation, so function is not limited to specific expressions which do not have any other properties. The reason why certain expressions appear to have a conventional argumentative function is that they reflect properties of information structuring. For a better understanding of them, it is necessary to appeal to a representation of information at the level of processing where they apply. The conceptual framework suggested in this thesis can be viewed as an attempt to explain how conventional and non-conventional features of argumentation can be linked.

7.3 Applications

If the reasoning in this thesis is correct, there are many expressions and features of natural language whose main function is to suggest an organisational change in someone's representation of the discourse universe. I believe that the conceptual framework developed in this thesis can be useful as a general approach to the characterisation of their properties. For example, syntactic variants such as clefts can be thought of this way (von Klopp and Humphreys, in preparation). As the framework was formulated in order to allow for a characterisation of the BUTs,

it is likely that other natural language connectives could be characterised in a similar way. Although no completely satisfactory characterisation of BUT_{pa} could be given, at least some of its properties could be described. Significant work has already been done on similar connectives in Argumentation theory, so this framework could be used to ground the descriptions in a more representation-based account.

One issue where this approach should offer a substantial advantage is the description of the functions of *"yes"*, *"no"* and comparable expressions. It has previously only been possible to give relatively informal characterisations of their functions in different languages (e.g. "accepting/rejecting a factual description" in English, "agreeing/disagreeing" in Russian), whereas it should be relatively straightforward to give precise formulations using this framework.

7.4 Wovon Man Nicht Sprechen Kann...

If retrieving information from sentences consisted of deriving a representation from the linguistic form, the task of describing meaning in language would at least be straightforward, if not easy—the data would be readily available, making it comparable to studying syntax. However, I have argued that this is not a suitable way of viewing the information conveyed by sentences. Instead, it must be described in terms of a representation of the relevant discourse universe. But this makes the task a lot more difficult, as knowledge representation, in the sense of human mental representations, is poorly understood.

At the same time, language is one of the few sources of pointers to how this is done, which means that we are faced with what appears to be a conflict of enormous proportions: to study meaning in language, we must understand knowledge representation better, and one of the few ways to achieve such an understanding is through studying how language is used. Not only this: we are also forced to

hold any discussion about it using the very object of the study. The solution, if it is one, must be to tread carefully, and attempt to rid oneself of prejudices of how representation should be done, a near-impossible task.

Bibliography

Allwood, J. (1972) Negation and the strength of presupposition. Logical Grammar Reports 2, Department of Linguistics, University of Göteborg.

Anscombre, J.- C. and O. Ducrot (1977) Deux mais en francais? *Lingua* **43**, 23–40.

Anscombre, J.- C. and O. Ducrot (1983) *L'Argumentation dans la Langue*. Brussels: Pierre Mardaga.

Anscombre, J.- C. and O. Ducrot (1986) Argumentation et Informativité. In *De la Métaphysique à la Rhétorique*, pp. 79–94. Brussels: Editions de l'Université de Bruxelles.

Atlas, J. D. (1977) Negation, ambiguity and presupposition. *Linguistics and Philosophy* **1**, 321–336.

Atlas, J. D. (1980) A Note on the Confusion of Semantic and Pragmatic Aspects of Negation. *Linguistics and Philosophy* **3**, 411–414.

Atlas, J. D. (1981) Is *not* logical? In *11th International Symposium on Multiple-Valued Logic*, pp. 124–128, Oklahoma City, Oklahoma.

Atlas, J. D. (1984) Comparative Adjectives and Adverbials of Degree. *Linguistics and Philosophy* **7**, 347–77.

Atlas, J. D. (1989) *Philosophy Without Ambiguity: A Logico-Linguistic Essay.* Oxford: Clarendon Press.

Atlas, J. D. (1990) Implicature and Logical Form: The Semantics-Pragmatics Interface. Lecture notes from a course given at the Second European Summer School in Language, Logic and Information.

Atlas, J. D. and S. C. Levinson (1981) It-clefts, informativeness, and logical form. In P. Cole, ed., *Radical Pragmatics*, pp. 1–59. New York: Academic Press.

Austin, J. L. (1962) *How To Do Things With Words.* Oxford: Clarendon Press.

Baker, C. L. (1970) Double negatives. *Linguistic Inquiry* **1**, 169–186.

Blakemore, D. (1987) *Semantic Constraints on Relevance.* Oxford: Basil Blackwell.

Brown, D. and a'i Feibion Cyf (1976) *Gramadeg Cymraeg Cyfoes.* Y Bontfaen Morgannwg. (*Contemporary Welsh Grammar*).

Brown, G. and G. Yule (1983) *Discourse Analysis.* Cambridge: Cambridge University Press.

Burton-Roberts, N. (1989) *The Limits To Debate: A Revised Theory of Semantic Presupposition.* Cambridge: Cambridge University Press.

Carnap, R. (1942) *Introduction to Semantics.* Cambridge, Mass.: MIT Press.

Cohen, P. R. and C. R. Perrault (1979) Elements of a plan-based theory of speech acts. *Cognitive Science* **3**(3), 177–212.

Cohen, R. (1984) A computational theory of the function of clue words in argument understanding. In *Proceedings of the 10th International Conference on Computational Linguistics and the 22nd Annual Meeting of the Association for Computational Linguistics*, pp. 251–258, Stanford University, Stanford, Ca.

Dascal, M. and T. Katriel (1977) Between semantics and pragmatics: The two types of "but"—Hebrew "aval" and "ela". *Theoretical Linguistics* **4**, 143–72.

Devlin, K. (1991) *Logic and Information.* Cambridge: Cambridge University Press.

Dowty, D. R., R. E. Wall and S. Peters (1981) *Introduction to Montague Semantics.* Dordrecht: D. Reidel.

Ducrot, O. (1972) *Dire et ne pas dire.* Paris: Hermann.

Ducrot, O. (1973) *La preuve et le dire.* Paris: Maison Mame.

Elhadad, M. (1991) A procedure for the selection of connectives. Technical Report CUCS-419-89, Department of Computer Science, Columbia University, New York.

Elhadad, M. and K. R. McKeown (1991) What do you need to produce a "but"? Technical Report CUCS-334-88, Computer Science, Columbia University, New York.

Fauconnier, G. (1975) Polarity and the scale principle. In R. E. Grossman, L. J. San and T. J. Vance, eds., *Papers from the Eleventh Regional Meeting of the Chicago Linguistics Society*, pp. 188–199, Chicago.

Fauconnier, G. (1976) Remarque sur la théorie des phénomènes scalaires. *Semantikos* **1**, 13–36.

Fauconnier, G. (1984) *Espaces Mentaux: Aspects de la Construction du Sens dans les Langues Naturelles.* Paris: Minuit.

Fikes, R. E. and N. J. Nilsson (1971) STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence* **2**, 189–208.

Fodor, J. D. (1979) In defense of the truth value gap. In *Syntax and Semantics*, Vol. 11: *Presupposition*, pp. 199–223. New York: Academic Press.

Fodor, J. A. (1975) *The Language of Thought.* New York: Thomas Crowell.

Gazdar, G. (1979) *Pragmatics, Implicature, Presupposition, and Logical Form.* New York: Academic Press.

Givón, T. (1978) Negation in language: Pragmatics, function, ontology. In P. Cole, ed., *Syntax and Semantics*, Vol. 9: *Pragmatics*, pp. 69–112. New York: Academic Press.

Green, G. M. (1989) *Pragmatics and Natural Language Understanding.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Grice, H. P. (1975) Logic and conversation. In P. Cole and J. L. Morgan, eds., *Syntax and Semantics*, Vol. 3: *Speech Acts*, pp. 41–58. New York: Academic Press.

Horn, L. (1985) Metalinguistic negation and pragmatic ambiguity. *Language* **61**, 121–174.

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

Horn, L. (1975) Neg-raising predicates: Towards an explanation. In R. E. Grossman, L. J. San and T. J. Vance, eds., *Papers from the Eleventh Regional Meeting of the Chicago Linguistics Society*, pp. 279–94, Chicago.

Jayez, J. (1989) *L'inférence en langue naturelle.* Paris: Hermès.

Kamp, H. and U. Reyle (in press) *From Discourse to Logic.* Dordrecht: D. Reidel.

Kempson, R. M. (1975) *Presupposition and the Delimitation of Semantics.* Cambridge: Cambridge University Press.

Kempson, R. M. (1986) Ambiguity and the semantics-pragmatics distinction. In C. Travis, ed., *Meaning and Interpretation*, pp. 77–103. Oxford: Basil Blackwell.

von Klopp, A. and K. Humphreys (in preparation) Syntactic Choice:: The Relevance of Pragmatics.

Lakoff, R. (1969) Some reasons why there can't be any *some-any* rule. *Language* No. 45, 608–615.

Lakoff, R. (1971) *If's, And's* and *But's* about conjunction. In C. J. Fillmore and D. T. Langedoen, eds., *Studies in Linguistic Semantics.* New York: Holt, Rinehart and Winston.

Lang, E. (1984) *The Semantics of Coordination.* Amsterdam: John Benjamins Publishing Co.

Levinson, S. C. (1983) *Pragmatics.* Cambridge: Cambridge University Press.

Lyons, J. (1977) *Semantics.* Cambridge: Cambridge University Press.

Moeschler, J. (1989) *Modélisation de dialogue: Représentation de l'inférence argumentative.* Paris: Hermès.

Morris, C. (1946) *Signs, Language, and Behaviour.* Englewood Cliffs, N.J.: Prentice-Hall.

Morris, C. (1971) *Writings on the General Theory of Signs.* The Hague: Mouton.

Morris, C. W. (1938) Foundations of the theory of signs. In O. Neurath, R. Carnap and C. W. Morris, eds., *International Encyclopedia of Unified Science*, Vol. 1, pp. 77–138. Chicago: The University of Chicago Press.

Moxey, L. M. and A. J. Sanford (to appear) Prior expectations and the interpretation of natural language quantifiers. *European Journal of Cognitive Psychology.*

Payne, J. (1985) Negation. In T. Shopen, ed., *Language Typology and Syntactic Description*, Vol. 1: *Clause Structure*, pp. 197–242. Cambridge: Cambridge University Press.

Reddy, M. J. (1979) The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony, ed., *Metaphor and Thought.* Cambridge: Cambridge University Press.

Roulet, E., A. Auchlin, J. Moeschler, C. Rubattel and M. Schelling (1985) *L'articulation du Discours en Français Contemporain.* Bern: Peter Lang.

Russell, B. (1967) On denoting. In I. M. Copi and J. A. Gould, eds., *Contemporary Readings in Logical Theory*, pp. 93–105. London: Macmillan. Reprinted from Mind, Vol. 14, 1905..

Sadock, J. M. (1977) Modus brevis: The truncated argument. In W. A. Beach, S. E. Fox and S. Philosoph, eds., *Papers from the Thirteenth Regional Meeting of the Chicago Linguistics Society*, pp. 545–554, Chicago.

Sadock, J. M. (1981) Almost. In P. Cole, ed., *Radical Pragmatics*, pp. 257–71. New York: Academic Press.

Schank, R. C. and R. Abelson (1977) *Scripts, Plans, Goals and Understanding.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Searle, J. (1969) *Speech Acts: An Essay in the Philosophy of Language.* Cambridge: Cambridge University Press.

Seuren, P. A. M. (1991) The ambiguity of negation. Paper presented at the DANDI workshop on natural language negation.

Seuren, P. A. M. (1985) *Discourse Semantics.* Oxford: Basil Blackwell.

Shannon, C. E. and W. Weaver (1949) *The Mathematical Theory of Communication.* Urbana, Ill.: University of Illinois Press.

Sperber, D. and D. Wilson (1986) *Relevance: Communication and Cognition.* Oxford: Basil Blackwell.

Zwarts, F. (1991) The syntax and semantics of negative polarity. Unpublished ms., June 1991.