A Data-Driven Method for Classifying Connective Phrases

Alistair Knott
Human Communication Research Centre
2 Buccleugh Place
Edinburgh EH8 9LW
A.Knott@ed.ac.uk *

Chris Mellish
Department of Artificial Intelligence
80 South Bridge
Edinburgh EH1 1HN
C.Mellish@ed.ac.uk

April 15, 1996

Running Head: A Data-Driven Method for Classifying Connective Phrases

Correspondence should be sent to the first author.

*This work was completed while the first author was studying at the Department of Artificial Intelligence. The first author would like to acknowledge the Engineering and Physical Sciences Research Council for Studentship 91308508 and grant GR/K53321. We are also grateful to the two anonymous reviewers for their comments.

Abstract

This paper describes a three-stage methodology for investigating the semantics and pragmatics of sentence and clause connective phrases. The first step in the methodology is to assemble a large corpus of connectives. The second step is to organise this corpus into a hierarchical taxonomy of synonyms and hyponyms, using a pre-theoretical substitution test. The final step is to impose a theoretical interpretation on the taxonomy. The taxonomy lends itself to an analysis of intersentential/interclausal relations in terms of a number of orthogonal binary-valued features; connectives are then seen as signalling the values of one or more of these features.

1 Introduction

Researchers interested in the semantics of sentence and clause connectives are faced with a trade-off between the accuracy of the theory they develop and its scope. By accuracy, we mean how precisely the theory describes the conditions under which each connective can be used; by scope, we mean how many different connectives it describes.

The trade-off can be seen in the fact that analyses within formal semantics have tended to concentrate on particular classes of connective, rather than on connectives in general. For instance, connectives signalling temporal relationships (such as *before*, *after*, and *when*) have been studied by Moens and Steedman (1987), Lascarides and Oberlander (1993), Oversteegen (1993), Glasbey (1995); those signalling concessive/adversative relations (such as *but*, *although* etc) by Spooren (1989), Von Klopp (1993), Grote *et al.* (1995); those signalling argumentative relations (such as *because* and *since*) by Cohen (1984), Elhadad and McKeown (1990); those signalling purposive relations (*to*, *by*) by Di Eugenio (1992), Delin *et al.* (1994). As yet, there are no studies within the formal tradition which aim to account for all of these at once.

At the other end of the spectrum, theories which aim to describe the whole range of connectives tend to provide less formal analyses. Prominent amongst such theories are the systemic accounts of Halliday and Hasan (1976) and Martin (1992). The emphasis here is on producing a *classification* of connectives; and on elucidating the semantics of individual phrases in terms of the different dimensions along which they can vary. For instance, Martin makes use of general distinctions between 'paratactic' and 'hypotactic' phrases, and between 'internal' and 'external' phrases, which both serve to classify connectives across the board. However, even in these theories there is a limit to the extent to which connectives are cross-classified. It is a feature of systemic networks that the decisions made early on about a phrase can determine the set of questions that are subsequently asked about it; thus it is possible to provide altogether different classification systems for different groups of phrases. In Martin's account, for instance, four quite different systems are proposed for the handling of ADDITIVE, COMPARATIVE, TEMPORAL and CONSEQUENTIAL phrases.

In short, the conflicting demands of accuracy and scope are often reconciled by the introduction of **intermediate-level categories** of connectives. In global systemic accounts, the

categorisation is imposed by the initial choices made in a network, which can determine the kinds of choices subsequently made. In formal semantics, the categorisation is implicit in the adoption of restricted sets of connectives as the object of study.

Intermediate-level categories are useful, in that they allow researchers a certain amount of modularity in developing an account of connectives. It is hard to develop theoretical distinctions like PARATACTIC and HYPOTACTIC which can apply right across the set of connectives, because the set is very large and diverse. And it is equally hard to test any such distinctions in a systematic way. Naturally, these problems are much reduced if subsets of connectives can be considered in isolation.

But it is less clear whether intermediate categories are justified *as theoretical constructs*. Many connectives are hard to classify, or seem to be classifiable in more than one category. Consider some of Martin's high level categories, for instance. *When* is classified as a TEMPORAL phrase, but it can also be used to describe causal relations, which are classified as CONSEQUENTIAL:

(1) *When* Sue heard the news, she was shocked.

And *while* is classified as both TEMPORAL and COMPARATIVE:

(2) Bill had a cup of tea *while* the food was cooking.

(3) Bill had tea, *while* Jill had coffee.

In any of the proposed classifications, examples can be found of this kind of overlap between intermediate categories. In such circumstances, categories cannot be said to stand in simple relationships to one another—for instance, it is no longer clear what it means to distinguish between two phrases by assigning them to different categories.

Of course, we might find that intermediate-level categories do have a useful explanatory role in a theory of connectives. It might be, for instance, that there is no alternative but to use different theoretical mechanisms to account for different groups of connectives. But we certainly don't know this *a priori*. Naturally, intermediate categories can provide us with a reasonable high-level picture about the different types of connectives. But their adoption still

seems partly a matter of expedience, motivated by the combinatorial problems inherent in comparing every connective directly with every other.

In this paper an alternative methodology for classifying connectives will be presented, in which no intermediate categories are posited *a priori*. The methodology is incremental, proceeding in three distinct steps: first, the assembly of a large corpus of connective phrases; next, the organisation of this corpus into a taxonomy of synonyms and hyponyms by means of a pre-theoretical substitution test; and finally a theoretical interpretation of the phrases in the taxonomy as signallers of features of discourse relations. The taxonomy is intended as a half-way house for theorists; it provides a way of structuring the explosion of pre-theoretical data that ensues when assumptions about intermediate categories are dropped; at the same time it can be used systematically to motivate elements of a wide-ranging theory.

An overview of the methodology is given in Section 2, and the results of the first two steps (which relate to the construction of the taxonomy) are outlined. The rest of the paper concentrates on the final step, in which the taxonomy is given a theoretical interpretation. In Section 3, a number of extracts from the taxonomy are considered in detail, and used to motivate a succession of independent features. The preliminary conclusions arrived at in this section are summarised in Section 4, and some suggestions are made about how the ideas developed so far might be extended in future work to other parts of the taxonomy.

2 The Substitution Methodology

The main idea behind the methodology is to delay the onset of 'theorising' until a large body of pre-theoretical data has been gathered and organised. The first two steps of the methodology are nothing more than data collection; it is only in the final step that the data is interpreted from a theoretical point of view.

2.1 Assembling a Corpus of Cue Phrases

The first stage is to gather a corpus of phrases to study. The connectives are selected by means of a pre-theoretical **test for cue phrases**, which is designed to select from naturally-occuring monologue any phrase whose function is to link one unit of text to another. For details of the

test, see Knott and Dale (1994), Knott (1996). Very briefly, the idea is to select any phrase in a naturally-occurring text, isolate it together with its host clause, and ask whether the result requires further context in order to be interpreted.[1] For instance, Text 4 requires no additional context to be understood; while Text 5 only makes sense as the follow-up to some previous statement.

(4) Bill is six feet tall.

(5) *But* Bill is six feet tall.

If the clause cannot be interpreted without further context, but can be thus interpreted if the selected phrase is removed, then the phrase is termed a **cue phrase**.[2]

The test is functional in inspiration, and identifies phrases of a number of syntactic types, including prepositional phrases, sentential adverbs, co-ordinators, and subordinators. (Note in particular that the test does not distinguish between sentence and clause connectives.) In all, over 200 pages of text were analysed using the test, and over 200 different cue phrases were identified.

2.2 Classifying the Corpus of Cue Phrases using the Substitution Test

The next step in the methodology is to impose structure on the corpus of cue phrases. A second pre-theoretical test is used to this end, this time to determine whether one phrase can be used in the place of another. Again, it can be described very briefly: the tester chooses a context where one cue phrase X naturally occurs, and decides whether (s)he, as a writer, would be prepared to replace it with another cue phrase Y. (S)he might imagine that an earlier occurence of X has been discovered in the text, and that 'elegant variation' requires the current occurrence to be changed. In other words, the writer does not want to change

[1] Note that the test makes use of theoretical concepts from *syntax* ('phrase' and 'host clause'). It is only pre-theoretical in that it does not presuppose a particular theory of discourse relations, and then go on to define cue phrases as signallers of these relations.)

[2] This latter requirement means that a small number of relevant phrases (notably *to* and *by*) do not pass the test as it stands. However, as the test is intended to be simple and pre-theoretical, we did not want to alter it specifically to include these phrases. Our intention is to begin by developing a semantic analysis of the phrases which pass the test, and then to consider whether it extends to other phrases such as these.

what (s)he wants to say; only how it is said. The key idea is to have the tester determine whether two phrases have 'the same meaning' in the context chosen, without having to invoke an explicit theoretical conception of 'meaning'.

To give an example: in the context in Example 6, the phrase *whereas* can be substituted by *and*, but not by *because*:

(6) Bill and Bob are quite different. Bill is a born optimist, { *and* / ✓ *whereas* / # *because* } Bob is perpetually gloomy.

In this notation, the top phrase in the braces is the original phrase, and the other phrases are candidates for substitution. A tick denotes substitutability in the context given; a hash sign denotes that substitution is not possible. (Note that the hash sign does not necessarily signal *ungrammaticality*; only a lack of substitutability for the original phrase.) Again, for a full description of the test, see Knott and Dale (1994), Knott (1996).

If we generalise over contexts, there are four possible substitutability relationships between two phrases X and Y:

- X and Y are **synonymous** if in any context where one can be used, the other can also be used. For instance, the phrases *to start with* and *to begin with* can be classed as synonymous.

- X and Y are **exclusive** if they can never be substituted for one another in any context. For instance, *to start with* and *alternatively* are exclusive.

- X is a **hypernym of** Y—and Y is a **hyponym of** X—if whenever Y can be used, so can X; but there are some contexts where X can be used and Y cannot. For instance, *and* is a hypernym of *whereas*: *whereas* can always be substituted by *and*, but there are some contexts where *and* cannot be substituted by *whereas*.

- X and Y are **contingently substitutable** if there are some contexts where they can be substituted, other contexts where X can be used and not Y, and still other contexts where Y can be used and not X. *And* and *but* are contingently substitutable, as shown

in the following examples:

(7) I'm very tired, { *and* / # *but* } I don't want to be disturbed.

(8) Don't be too harsh on Bob. He arrived late, { *but* / # *and* } he's usually very punctual.

(9) Bill's a liar. He said he can run a mile in three minutes, { *and* / ✓ *but* } that's impossible.

In what follows these relationships will be represented graphically, as depicted in the legend in Figure 1. Note that the relationships of exclusivity and contingent substitutability both apply

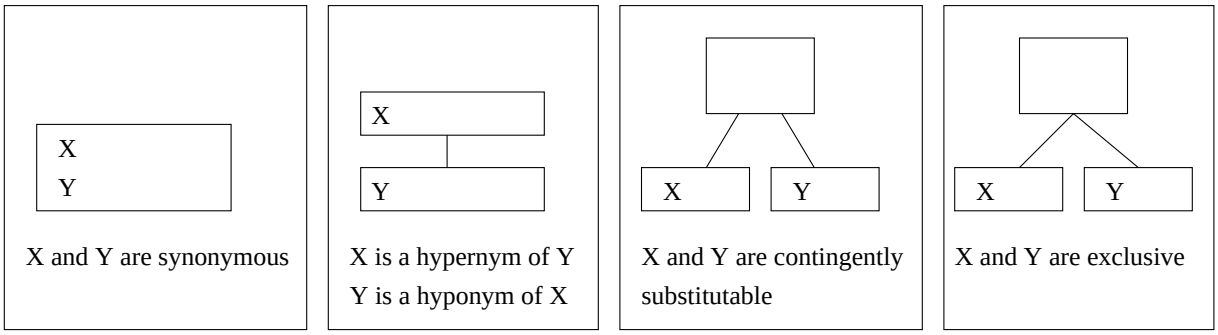

Figure 1: Diagrammatic Representation of Substitutability Relationships

between 'sister' nodes. (The difference between them relates to whether or not the arcs from the sister nodes touch on the mother node.) The notation thus allows for the representation of exclusive or contingently substitutable phrases which have hypernyms in common.

Using this representation, we can build up diagrams containing any number of cue phrases, and depicting the substitution relationship which holds between each pair of phrases. These composite diagrams make use of a notion of inheritance; any relationship which is true of one phrase is true of all its hyponyms. For instance, in Figure 2 (i), B is a hyponym of A and exclusive with C. D is a hyponym of B, so by inheritance, D is a hyponym of A and exclusive with C. There is one case in which inheritance does not occur, however; namely, where the inherited relationship is of contingent substitutability. Consider Figure 2 (ii): here, Q is represented as contingently substitutable with R, and thus by inheritance, S should also

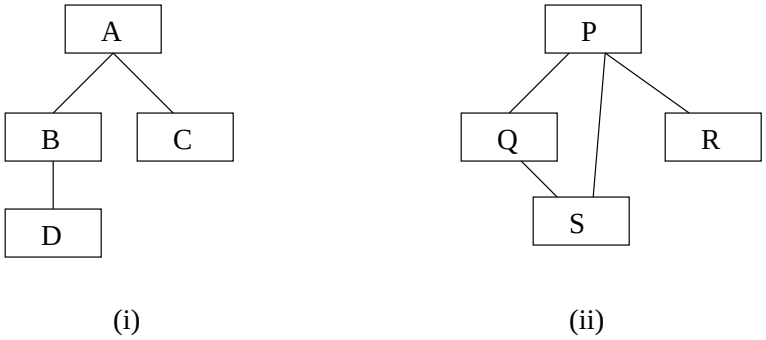

Figure 2: Inheritance in Substitutability Diagrams

be contingently substitutable with R. However, S and R are also represented as exclusive; in a case like this, the relationship of contingent substitutability is overridden.

A large taxonomy of cue phrases has been built up using this notation, containing around 150 phrases from the corpus, and documenting the relationship between each pair of phrases.[3] An extract from the taxonomy is given in Figure 3. (Note that the empty boxes below the mother node are only included to make the diagram easier to read—they reduce the number of arcs that need to be drawn. They have no theoretical significance, and the diagram could always be redrawn without them.)

The complete taxonomy, given in Knott (1996), is too large to be presented here. However, some of its characteristics can be noted. In particular, we should note that the taxonomy does not divide neatly into large exclusive subgroups of phrases. For any candidate grouping, many phrases can be found which fit into more than one group. No strong macrostructure can be detected in the taxonomy; in fact, most of the variation between cue phrases is represented at a relatively low level, close to the leaf nodes. In other words, the taxonomy provides scant evidence for the kind of intermediate-level categories discussed in Section 1.

[3]The taxonomy *specifies* a relationship for each pair of phrases, but naturally we have not examined all of the 11,000 or so possible pairings individually. The taxonomy should properly be considered as being 'under development', and as setting out *hypotheses* about substitution relationships which are subject to revision as new data is considered. Having said this, it has been extensively tested, has undergone a number of revisions, and is currently in our experience quite stable faced with new data.

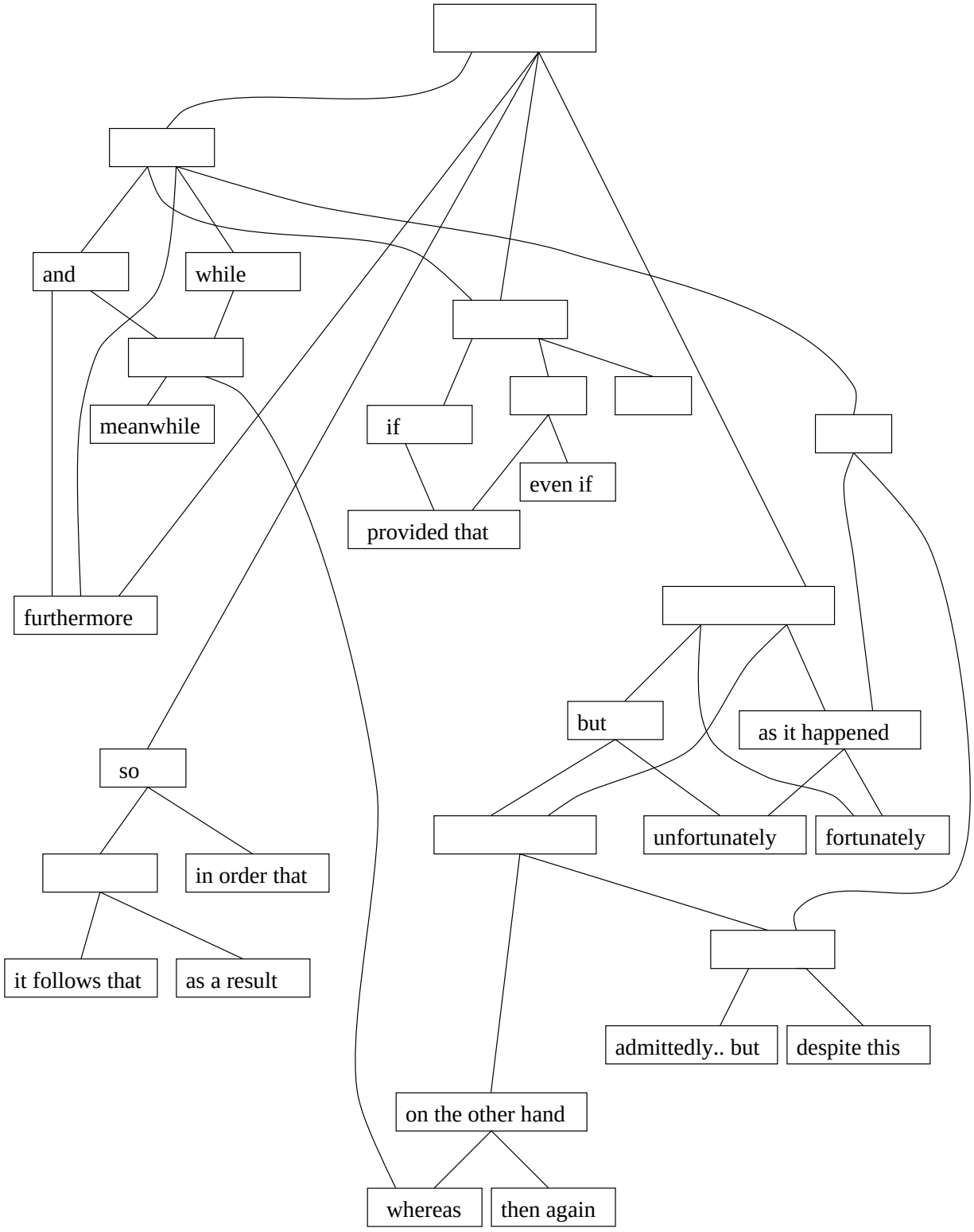

Figure 3: An Extract from the Taxonomy

2.3 A Theoretical Interpretation of the Taxonomy

The taxonomy does not in itself embody a *theory* about the semantics of cue phrases. In any such theory, an independent representation of the **relations** signalled by cue phrases must be provided, so that predictions can be made about which relations are signalled by which phrases. No such independent level of representation is present in the taxonomy.

However, the taxonomy does lend itself to a particular conception of intersentential/interclausal relations. Its hierarchical structure suggests that relations are composite constructs: a phrase high up in the taxonomy provides information about some characteristics of a relation while leaving others unspecified; and those beneath it add information about additional characteristics. And the lack of a strong macrostructure in the taxonomy suggests that there is no single characteristic which dominates the classification. A picture emerges in which relations are defined in terms of a set of **independent features**, and where cue phrases provide information about subsets of features rather than about whole relations. This model suggests a natural conception of substitution relationships between phrases:

- If X is **synonymous** with Y, then they signal the same values of the same features.
- If X is **exclusive** with Y, then they signal *different values* of at least one feature.
- If X is a **hyponym** of Y (and Y is a **hypernym** of X), then X signals all the features that Y signals, and *some other* feature(s) in addition, for which Y is undefined.
- If X and Y are **contingently substitutable**, then X and Y signal some of the same features, but in addition X is defined for a feature for which Y is undefined, and Y is defined for a feature for which X is undefined.

This feature-theoretic interpretation of substitution relationships turns the taxonomy into a useful platform for theory development. Each substitution relationship in the taxonomy can now be seen as evidence about one or more features of relations. By looking systematically at a wide range of different relationships, we should be able to build up a picture of the different features needed for a thoroughgoing cross-classification of cue phrases.

2.4 A Theoretical Framework for Motivating Features

Before we turn to the motivation of individual features, it is important to have in mind a general framework for relation definitions. Our background is in the field of text generation; so the framework we propose is based on a conception of relations as planning operators in a text generation system.

The framework has its origins in the plan-based approach to speech acts developed by Cohen and Perrault (1979), Perrault and Allen (1980) and others. Cohen *et al* developed a scheme for representing a speaker's utterances in terms of their intended effects on a hearer, and used the scheme to formalise parts of Searle's (1969) classification of speech acts. This intention-based conception of speaker utterances was extended by Mann and Thompson (1988) to account for the **rhetorical relations** between adjacent speaker utterances, in the context of a theory of discourse coherence. Relations are defined in terms of a speaker/writer's intended effect in *juxtaposing* two utterances within an extended segment of discourse. (The effect can be to create a CONTRAST for the hearer/reader, to convey a TEMPORAL SEQUENCE of events, to provide EVIDENCE for a statement, and so on.) More recently, relations have been adapted for use in automatic text generation systems (Hovy (1988, 1993), Moore and Paris (1989, 1993)). Their intentional characterisation makes them suitable for implementation as **planning operators**, defined in terms of their **preconditions** (the conditions which must hold in order for a relation to be used) and **postconditions** (the communicative effects which are achieved when the relation is used). We assume this plan-based conception of relations in the theoretical interpretation of the taxonomy.

3 Motivating Individual Features from the Taxonomy

This section describes a preliminary application of the method set out in Section 2.3. A number of small extracts from the taxonomy will be considered in turn, and used to justify the introduction of a number of different independent features. The features will be expressed in terms of the preconditions necessary for the use of different cue phrases, and the communicative effects achieved by using them.

It should be stressed that the aim of this paper is not to come up with a complete and watertight set of cue phrase definitions. Producing a complete feature-theoretic account of the taxonomy is a huge task, and well beyond the scope of the paper. The present aim is rather twofold:

- to demonstrate the utility of the substitution methodology, by noting some of the interesting theoretical constructs which emerge when it is pursued;
- to motivate a core set of features needed to describe the taxonomy, to serve as the basis for further investigations.

A great deal has already been written about the semantics and pragmatics of cue phrases, and it should be no surprise that many of the suggestions about features make reference to existing work. The main novel element in what follows is that each new theoretical construct is justified in exactly the same way; by examining a portion of the taxonomy of cue phrases, and noting a demand for features to represent the patterns of substitutability it contains.

3.1 SEMANTIC and PRAGMATIC Relations

The first extract from the taxonomy to be considered is given in Figure 4.[4] Motivating examples are given in Texts 10 and 11:

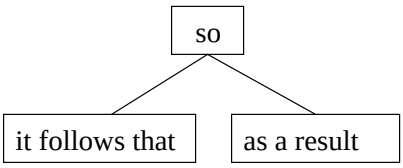

Figure 4: SEMANTIC and PRAGMATIC Phrases

(10) The footprints are deep and well-defined. { *It follows that* / ✓ *So* / # *As a result,* } the thief was a heavy man.

[4]The extracts from the taxonomy given in this section will not be labelled with the features they serve to motivate. A labelled version of the taxonomy is given in Section 4; the reader might find it useful to refer forward to this section to see how feature values are assigned to cue phrases.

(11) I had a puncture on the M25 on my way back from work. { *As a result,* / ✓ *So* / # *It follows that* } I missed most of the first half.

As a result seems strange in the context of Example 10, because it suggests that the thief's heaviness is *caused* by the footprints being deep. Conversely, *it follows that* is odd in Example 11, because it suggests that the writer is *deducing* the fact that she missed most of the first half—while in fact she is reporting from her own experience. Note that *so* is acceptable in both cases.

Both examples are commonly analysed as involving a causal/inferential relation of some kind, the difference between them being to do with what this relation holds between. Traditionally, the relation in examples like 11 is taken to be between the events in the world described by the two clauses: the puncture causes the missed first half. In examples like 10, the relation is taken to involve linguistic events themselves, not just the events they represent. For Martin (1992), Sanders *et al.* (1992) and others, the writer's *statement* that the thief was heavy is caused by her belief that the footprints are deep. Commentators have used a variety of terms to represent these distinctions. Martin (1992) and Halliday and Hasan (1976) use the labels EXTERNAL and INTERNAL to refer to examples like 11 and 10 respectively; Redeker (1990) uses the labels IDEATIONAL and PRAGMATIC; van Dijk (1979) and Sanders *et al.* (1992) talk of SEMANTIC and PRAGMATIC relations.

A useful modification of the notion of PRAGMATIC relations is introduced by Sweetser (1990). For her, the relation in examples like 10 primarily describes the cause of the writer's *conclusion* that the thief must have been heavy, and only indirectly describes the cause of her *statement* to this effect. The important relation in the example is the logical one, between two of the writer's beliefs. Sweetser calls relations involving the writer's beliefs EPISTEMIC, and defines a further category of SPEECH ACT relations which make reference to actual writer utterances. Prototypical of SPEECH ACT relations are examples like the following:

(12) What are you doing tonight, because there's a good movie on.

Here it is certainly appropriate to analyse the relation as describing the cause of the writer's utterance *What are you doing tonight?*.

Sweetser's definition of EPISTEMIC relations is an improvement on the previous definitions. However, it still leaves something to be desired. Consider again Sweetser's EPISTEMIC analysis of Example 10, as a statement about the writer's conclusions and how they were reached. The text, according to the new analysis, is still fundamentally descriptive; instead of describing the external world, it now contains a description of the writer's own thought processes. What is missing is an account of how an argumentative text like this one achieves a rhetorical *effect* on the reader—how it *persuades* the reader that the thief was heavy, where a simple statement like *The thief was heavy* might not have sufficed. Of course, in offering the reader a trace of the writer's reasoning, the text suggests how the reader might come to the same conclusion. But while it is vital for the writer's purposes that the reader take this last step, it is not represented in Sweetser's analysis.

Thinking about utterances in terms of their intended effects on the reader suggests an alternative definition for EPISTEMIC relations. The new definition expresses the intended effect of a text containing two related utterances *as a relation between the intended effect of each individual utterance.* The suggestion is, for instance, that in a text containing an EPISTEMIC *so*, the writer's intended effect is not that the reader *believe* a statement about the causes of the writer's beliefs, but rather that a causal relation *actually does* hold, in the real world, between the intended effects of the two related utterances—in other words, between two reader beliefs.

Consider how this definition works in the case of Example 10. The text is presented again below, and the intended effects of its two clauses are shown in italics:

(13)

The footprints are deep.		*So*	the thief was a heavy man.
R believes the footprints are deep.		*So*	*R believes the thief was a heavy man.*

On this interpretation, the intended effect of the text is that the reader's belief that John is sick *causes* the reader to believe that John has stayed at home, where otherwise the reader would not have been inclined to believe this latter statement. It is because of *this* cause that the relation has persuasive force.

Note that the proposed new definition of EPISTEMIC relations actually extends to some relations Sweetser considers as SPEECH ACT. For Sweetser, "if an utterance is imperative

(...) in form, then it cannot reasonably be causally conjoined to another utterance except at the speech act level" (p78). Thus Sweetser would interpret an example like 14 as SPEECH ACT:

(14) Hurry up, because we haven't much time!

But here again, there are advantages in expressing the relation in terms of intended effects. The writer's main intention in such a case is not to inform the reader about the cause of her utterance; but rather the realisation that they haven't much time should motivate the reader to hurry up. Using the new definition, this is just what is expressed. The intended effect of the imperative *Hurry up* is that the reader hurry up; the intended effect of the statement *we haven't much time* is that the reader believe they haven't much time; and the intended effect of the whole utterance is that this belief causes the reader to hurry up.

Because the new definition encompasses examples such as this one, we have decided to revert to the label PRAGMATIC to refer to the relations it describes; and consequently to return to the label SEMANTIC for what Sweetser calls CONTENT relations. The definition for a feature with alternative values SEMANTIC and PRAGMATIC relations can now be given. After Sanders *et al*, we can call this feature SOURCE OF COHERENCE:[5]

SOURCE OF COHERENCE

SEMANTIC: the intended effect of the text containing the relation is that the reader believes some relation holds between two propositions A and C. A and C are the propositional contents of the two related text spans S_A and S_C.

PRAGMATIC: the intended effect of the text containing the relation is that some relation actually holds between two propositions A and C. A and C are the intended effects of the two related text spans S_A and S_C.

[5]It should be noted at this point that, as new features are motivated, changes are sometimes required to the definitions of features already introduced. In consequence, the feature definitions presented one by one in this chapter should not be regarded as final, but only as sufficient to account for the data so far introduced.

Note that the new notion of PRAGMATIC by no means covers all of Sweetser's SPEECH ACT relations. (For example, Text 12 is still much better analysed as describing the causes of the writer's speech act.) However, the class of SPEECH ACT relations has not yet been motivated from the taxonomy, as it is hard to find cue phrases which are specific to this class.

3.2 POSITIVE and NEGATIVE POLARITY Relations

A second portion of the taxonomy is given in Figure 5. Some motivating examples are provided in Texts 15 and 16.

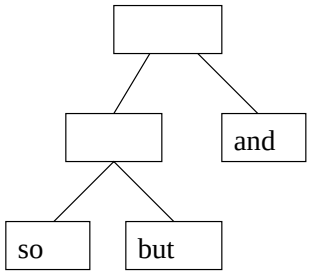

Figure 5: POSITIVE and NEGATIVE POLARITY Phrases

(15) Jim had just washed his car, { *so* / ✓ *and* / # *but* } he wasn't keen on lending it to us.

(16) It was odd. Bob shouted very loudly, { *but* / ✓ *and* / # *so* } nobody heard him.

The fact that the phrases *but* and *so* can never be substituted for one another is clear from consideration of examples like these. But the examples also show that the phrase *and* is contingently substitutable both for *but* and (in other contexts, of course) for *so*. In feature-theoretic terms, we can conclude that *but* and *so* are defined for different values of some feature; and that *and*, being contingently substitutable for both, is undefined for this feature. It remains now to decide what the feature is.[6]

[6]Note that in order to motivate a feature which distinguishes between *so* and *but*, we are not obliged to find a common hypernym of the two phrases, as we did in the previous section. The fact that *and* can sometimes be substituted for both *but* and *so* is sufficient to show that it *cannot* be defined for any feature which takes

Many different suggestions have been made as to the similarities and differences between phrases like *but* and *so*. It is uncontroversial (as far as it goes) that *A, so C* signals some kind of implication or cause, with *A* as the antecedent/cause and *C* as the consequent/result. And it is likewise uncontroversial to say that *A, but C* signals (or at least can signal) a violation of the type of relation signalled by *so*. To illustrate with reference to the above examples: in Text 15, *so* signals that it follows from the fact that Jim had just washed his car that he was unwilling to lend it to us. In Text 16, *but* signals that it *normally* follows from the fact that Bob shouts loudly that people hear him, but in this case no-one does. Both phrases can thus be thought of as having a consequential component: for *so*, the consequence relation is specified as succeeding; while for *but*, an expected consequence is not forthcoming. The phrase *and* simply does not specify whether or not the consequence relation succeeds—indeed, it does not even specify that there *is* an expected consequence relation—the information is left to be inferred by the reader.

The important question is how to express the above ideas more precisely. It has been common to begin formalising the difference between relations signalled by *so* and those signalled by *but* by making reference to a 'statement of implication' $P \rightarrow Q$ which underlies both types of relation. The difference between the two relations is then expressed in terms of the relationship between P and Q and the propositions in the related spans of text. For *so*, P relates to the proposition in the first span and Q to that in the second span. For *but*, P relates to the proposition in the first span and Q to the *negation* of that in the second span. This story is roughly that given by Longacre (1983) in distinguishing between 'consequence' and 'frustrated consequence' relations. Sanders *et al.* (1992) give a similar story to distinguish between POSITIVE and NEGATIVE POLARITY relations; we will use these latter terms in what follows.

Central to the distinction between *so* and *but* is the notion that causal or consequential rules can be **defeated**. A number of recent accounts of concessive relations have employed the notion of defeasible rules; in particular Oversteegen (1995) and Grote *et al.* (1995). Defeasible

alternative values for these two phrases. If it were, there would have to be an exclusive relationship between it and one or other of the phrases.

rules provide a useful method for representing the kind of common-sense generalisations which people rely on in order to make up for their partial knowledge of the world. They are becoming increasingly popular in computational linguistics, as a tool for modelling the influence of the reader's world knowledge on the resolution of ambiguities. For instance, Lascarides and Asher (1991), Lascarides *et al.* (1992) use a system of defeasible rules to develop a framework for deciding which coherence relation is present at a particular point in a text when this is not signalled explicitly. Hobbs *et al.* (1993) use defeasible rules to model a range of processes in text interpretation, including the resolution of anaphora, lexical ambiguities and compound nominals. The use of defeasible rules in the present context is somewhat different, however. They are not being proposed as a way of *deciding* about the interpretation of some part of a text, but as a part of the interpretation *itself*—the defeasible rules used by a reader and writer to model the world are actually implicit in the semantics of phrases like *but*. As an initial model, then, we can propose that the phrases *so* and *but* are each associated with a defeasible rule, which in the case of *so* succeeds and in the case of *but* is defeated.

A number of questions still remain, however. Most importantly, what is the communicative status of the defeasible rule? Is it something which the reader must already know as a precondition to understanding the text, or is it something which the reader is told in the text? Along with Oversteegen (1995), we suggest that the existence of the defeasible rule should be thought of as a precondition. It is problematic to suggest that the rule itself is part of the information communicated to the reader by the writer. For one thing, the reader is only given one instance of the rule—it would then be necessary to abstract away from this to the rule itself; a process which is very underconstrained. Moreover, it is questionable whether causal or inferential *rules* constitute the kind of information that a reader will accept 'on authority' from a writer in any circumstance. Consider again the statement in Text 15: *Jim had just washed his car, so he wasn't keen on lending it to us.* This may certainly provide new information, but it is implausible to suggest that the writer is *informing* the reader that 'if a person with temperament T has just washed his car, he normally doesn't like to lend it to others'. It is more plausible to suggest that a rule along these lines is already known by the reader, and what is being communicated is the fact that the rule succeeds in this instance.

Note that the information conveyed by such a statement might be more than the bare assertion that 'there is nothing unusual about the situation being described'. Knowing that the rule in question is *triggered* would allow the reader to infer that Jim is of type T, for instance, if this was not already known. But it is much easier to imagine the reader adding such facts to his database than whole causal rules.

To sum up: we can hypothesise a feature called POLARITY, with alternative values NEGATIVE and POSITIVE. It is assumed that each relation presupposes the presence of a defeasible rule $P \rightarrow Q$.[7] The relationship between P and Q and the propositions A and C (defined in the SOURCE OF COHERENCE feature) is determined by the different values of the POLARITY feature, as follows:

POLARITY

POSITIVE: $A = P$; $C = Q$. The rule is specified to succeed.

NEGATIVE: $A = P$; C is inconsistent with Q. The rule is specified to fail.

3.2.1 Conditional NEGATIVE and POSITIVE POLARITY

Another portion of the taxonomy which can be used to motivate the POLARITY parameter is given in Figure 6. Motivating examples are given in Texts 17 and 18:

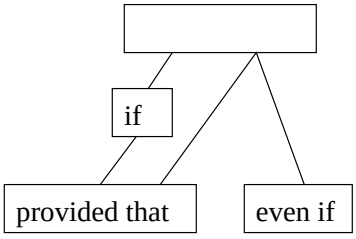

Figure 6: Conditional POSITIVE and NEGATIVE POLARITY Phrases

[7] We will not at this point buy into any particular formalism for representing defeasible rules. However, some of the requirements for the formalism eventually to be adopted will emerge from the discussion in following sections.

(17) You can sit in the front seat, { *provided that* / ✓ *if* / # *even if* } you put your seatbelt on.

(18) I wouldn't vote for Major { *even if* / ✓ *if* / # *provided that* } you gave me a thousand pounds.

The phrases in Figure 6 can be compared to those in Figure 5, a crucial difference being the fact that the latter group of phrases relate hypothetical eventualities while the former phrases relate actual ones. Motivation for a feature representing this dimension of variation will be provided below, in Section 3.8; for now the important thing to note is the variation in polarity exhibited by the phrases. For each phrase, an underlying defeasible rule $P \rightarrow Q$ can be identified. For *C, provided that A*, *A* and *C* map onto P and Q respectively and the rule is represented as succeeding. For *C, even if A*, *A* and *C* map onto P and $\neg Q$ respectively, and the rule is represented as failing. Thus in Example 17, the rule that putting a seatbelt on causes being allowed to sit in the front seat is asserted to succeed, while in Example 18 the rule that giving people lots of money causes them to vote against their will is asserted to fail in the case of the writer.

3.2.2 Semantic and Pragmatic Negative Polarity Relations

The distinction between POSITIVE and NEGATIVE POLARITY also cuts across the SEMANTIC/PRAGMATIC distinction. Consider Figure 7, for which motivating examples are given below:

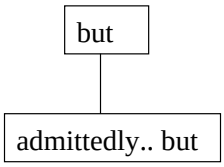

Figure 7: Semantic and Pragmatic Negative Polarity Phrases

(19) United have some key players injured; { *admittedly. . . but* / ✓ *but* } they're still bound to win.

(20) Mary was behaving oddly. She ordered a pizza, { *but* / # *admittedly... but* } she didn't eat any of it.

(The construction *admittedly... but* in these examples is to be read as distributed between the two clauses in the relation. The first example should thus read '*Admittedly,* United have some key players; *but...*', and the second example should read '*Admittedly,* she ordered a pizza, *but...*'.)

The point is that *Admittedly... but* signals the breaking of a defeasible rule just as *but* does; yet it has a specifically argumentative flavour. In Text 19, *admittedly* introduces a proposition which suggests one conclusion, and the negation of that conclusion is then asserted. *But* on its own can also be used in the absence of any argument, as in Text 20: here, the writer is simply informing the reader about an unusual state of affairs, and *admittedly* is quite out of place.

The SEMANTIC/PRAGMATIC distinction is useful in capturing the difference between these NEGATIVE POLARITY phrases. *Admittedly... but* can be defined as signalling the value PRAGMATIC, and *but* can be thought of as undefined for the feature. The difference between SEMANTIC and PRAGMATIC NEGATIVE POLARITY relations can be thought of as follows. In the SEMANTIC case, the writer's aim is to inform the reader that some general rule in the reader's model of the world is defeated in the situation being described. Thus in the above example, the defeated rule is that people who order food generally eat it. In the PRAGMATIC case, the writer's aim is that some general rule in the world itself *actually* fails in the present instance. This rule holds between two reader beliefs—the intended effects of the first and second clauses taken individually. In the above example, the intended effect of the first span is that the reader believe that United has some key players injured; that of the second span is that the reader believe that United will win. Normally, if the reader believes a team has several players injured, he will believe they will lose; but in this case, the writer's intention is that this conclusion is not drawn.

3.3 UNILATERAL and BILATERAL Relations

The next portion of taxonomy to be considered, given in Figure 8, also involves NEGATIVE POLARITY phrases. Motivating examples are given in Texts 21 and 22:

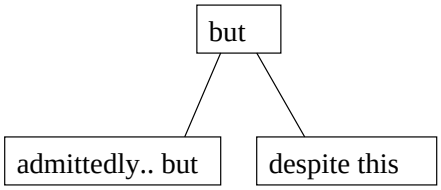

Figure 8: UNILATERAL and BILATERAL Phrases

(21) Bill lost the 400m last year. { *Admittedly... but* / ✓ *But* / ✓ *Despite this,* } He should win it this year.

(22) Bill should win the 400m. He lost last year; { *Admittedly... but* / ✓ *but* / # *despite this,* } they're running at altitude this time.

The relation in each of these examples can be signalled by *admittedly... but*, and can thus be considered as PRAGMATIC NEGATIVE POLARITY. However, the phrase *despite this* is only acceptable as a substitute in Text 21. In Text 22 it appears odd; it suggests that Bill's losing last year would normally lead to them not running at altitude this time—an implausible assumption.

The difference between the two examples appears to be to do with the status of the second span in the relation (the one introduced by *but*). In each case, the first span presents a premise P_1 which suggests a conclusion C. In Text 21, the second span presents the negation of this conclusion, $\neg C$. In Text 22, the second span presents another premise P_2, which is more telling than P_1, and suggests an alternative conclusion. This dimension of variation suggests another parameter, which we can call PATTERN OF INSTANTIATION.

To allow for multiple premises in the rule underlying the relation, we need to review the definition of POLARITY presented in the previous section. The assumption must now be that each relation presupposes a rule of the form $P_1 \wedge \ldots \wedge P_n \rightarrow Q$. We need to map elements

from this rule onto the variables A and C, defined in the feature SOURCE OF COHERENCE. Let us assume that A is always on the left-hand side of the rule.[8] The POLARITY feature presented before assumed that C always mapped onto the conclusion of the rule (Q); but we must now abstract away from this idea. We must define a new variable—call it C'—whose relationship to C is determined by the value of the POLARITY feature. For symmetry, we will also introduce a variable A', which always equates directly with A.[9] The POLARITY feature can now be thought of as specifying a function from A and C to A' and C'. Its revised definition now looks like this:

POLARITY (2nd definition)

POSITIVE: $A' = A$; $C' = C$. The rule is specified to succeed.

NEGATIVE: $A' = A$; C' is inconsistent with C. The rule is specified to fail.

The mapping between A' and C' and the rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$ is now given by the new feature:

PATTERN OF INSTANTIATION

UNILATERAL: C' is on the same side of the rule as A' (but not the same as A).

BILATERAL: C' is on the opposite side of the rule to A'.

The phrase *despite this* can now be thought of as defined as BILATERAL, while both *but* and *admittedly... but* are undefined for the feature. Thus in Example 21, for instance, A' is the proposition 'Bill lost the 400m last year' and C' is the proposition 'It is *not* the case that Bill should win this year'. These two propositions can be thought of as premise and

[8] This assumption itself changes in Section 3.5, for reasons developed in that section.

[9] This is another assumption which will be re-examined as further portions of the taxonomy are considered; see Section 3.6 for details.

conclusion of a defeasible rule. The relation can thus be thought of as BILATERAL, (and thus *despite this* is appropriate). Note that since the relation is also NEGATIVE POLARITY, the point is that the rule is defeated.

Several examples of phrases defined as UNILATERAL will be given in the following sections.

3.4 CAUSAL and INDUCTIVE Relations

Consider next the extract of the taxonomy in Figure 9, again featuring NEGATIVE POLARITY phrases. Motivating examples are given below:

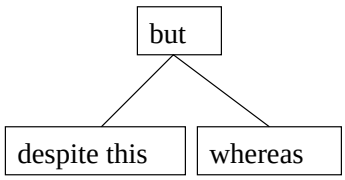

Figure 9: CAUSAL and INDUCTIVE Phrases

(23) Bill and Jill are like chalk and cheese. Bill lives for his books; { *whereas* / ✓ *but* / # *despite this,* } Jill is only interested in Tae Kwan Do.

(24) Bob was out of training; { *despite this* / ✓ *but* / # *whereas* } he completed the marathon in record time.

These phrases draw attention to another dimension of variation in the phrase *but*. Text 24 is of a kind we have already seen, where a causal or inferential rule is defeated: the rule in this case is that people who are out of training do not normally break records. But no such rule appears to underlie Text 23: knowing that Bill lives for his books gives no grounds for thinking that Jill is not just interested in Tae Kwan Do.

Whereas intuitively signals a contrast between two propositions. On a conventional analysis (see e.g. Spooren (1989)), we are able to say S_1, *whereas* S_2 if S_1 and S_2 allow the inference of $p(A)$ and $\neg p(B)$ respectively, for some predicate p and two entities A and B.[10]

[10] In Text 23, for instance, we can infer from *Bill lives for his books* that *Bill is not only interested in Tae Kwan Do*, which is an explicit negation of the predicate in the second span of the relation.

The question posed by the extract given in Figure 9 is: what does this have in common with the violated expectation analysis required for *despite this*? The feature-theoretic interpretation of the diagram requires us to find some feature or features of *whereas* that are shared by *despite this*; *but* is defined for this common component, and undefined for the feature(s) which distinguish the two phrases.

A point to note about $p(A)$ *whereas* $\neg p(B)$ is that A and B are required to belong in some sense to the same category of entities. Constrasts are not made between objects which have nothing in common at all. Thus Bill and Jill might be brother and sister, or friends, or two candidates for some job that needs doing. Put another way: relations signalled by *whereas* highlight an inability to *generalise* over the objects in a given class as regards some property p. This idea prompts the suggestion that a different type of rule underlies such relations; namely **inductive rules**.

An inductive rule is of the following general form: if property p is true of a certain finite number of elements from a particular class, then it follows that p is true of *all* the elements in that class. Clearly, as rules of inference, such statements are not sound. But inductive rules are nonetheless a mainstay of human reasoning: in the end, all our generalisations about the world are arrived at by inducing from particular instances. They can in fact be considered as another kind of *defeasible* rule, albeit quite different from those which we have so far been considering. Both kinds of rules are to be relied on when no information to the contrary is found, but are overruled if contrary information is forthcoming.

A proposition like $p(A)$ can thus be thought of as forming part of the left hand side of an inductive rule. If enough other objects in the same class as A have property p, then the inductive generalisation that $\forall X(X \in C \rightarrow p(X))$ will be triggered. Even the single instance of $p(A)$, in the absence of any other information, is presumably enough to trigger some wery weak hypotheses about other similar objects. However, the inductive rule is defeated as soon as $\neg p(B)$ is presented.

In summary, what *whereas* and *despite this* seem to have in common is that they both presuppose a defeasible rule of some kind, and both signal its defeat. In one case the defeasible rule is causal and in the other case, inductive; *but* can then be seen as undefined with respect

to the type of rule which is defeated. A new feature RULE TYPE is now motivated:

RULE TYPE

CAUSAL: the defeasible rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$ is a causal rule.

INDUCTIVE: the defeasible rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$ is an inductive rule.

As might have been noticed, *whereas* and *despite this* differ not only as regards the feature RULE TYPE, but also as regards the feature PATTERN OF INSTANTIATION. *Despite this* is defined as BILATERAL, as we have already seen in Section 3.3. *Whereas* must be defined as UNILATERAL: its two spans present two propositions from which generalisations can be drawn, and the generalisation itself (or rather the lack of it) remains implicit. Consider Example 23in more detail. A and C can be identified as 'Bill lives for his books' and 'Jill is only interested in Tae Kwan Do' respectively. The relation is NEGATIVE POLARITY, so while A' is identical to A, C' is inconsistent with C. *Whereas* is defined as UNILATERAL; so A' and C' are both on the left-hand side of some defeasible rule. The relation is INDUCTIVE, so each item on the left-hand side of the rule will be attributing the same predicate to a different member of a given class. In this case, then, C' must be defined as 'Jill lives for her books', or something to that effect. As this is inconsistent with the value of C, the inductive rule fails.

3.4.1 SEMANTIC and PRAGMATIC INDUCTIVE Phrases

The distinction between CAUSAL and INDUCTIVE relations cuts across that between SEMANTIC and PRAGMATIC relations. Consider the extract in Figure 10, motivated by the examples below:

(25) Bill and Jill are like chalk and cheese. Bill lives for his books; { *whereas* / ✓ *on the other hand,* / ✓ *but* / # *then again,* / # *despite this,* } Jill is only interested in Tae Kwan Do.

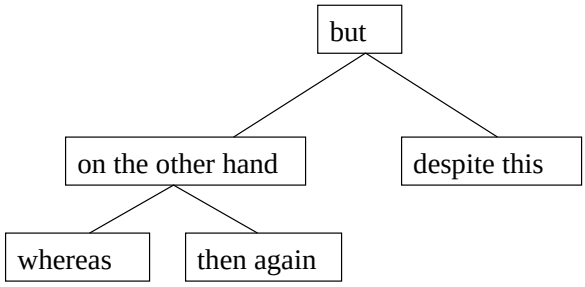

Figure 10: SEMANTIC and PRAGMATIC INDUCTIVE Phrases

(26) I don't know where we should eat tonight. The Kalpna is great value for money; { *then again,* / ✓ *on the other hand,* / ✓ *but* / # *whereas* / # *despite this* } Sid isn't crazy about Indian food.

Example 25 is the kind of contrast we have already considered, between objects in the world with inconsistent predicates. Example 26 is also a contrast of sorts, but here it seems to involve two arguments which push in opposite directions. The fact that the Kalpna is good value for money suggests that we should eat there. But the fact that Sid doesn't like Indian food suggests an alternative plan.

The interesting relationship in this diagram is between *whereas*, *then again* and *on the other hand*. *Then again* seems quite wrong in the first example, and *whereas* is out of place in the second one. But note that *on the other hand* is acceptable in both cases. It will be suggested in the remainder of this section that all three phrases are signallers of NEGATIVE POLARITY INDUCTIVE relations; that the exclusivity between *whereas* and *then again* is due to the SEMANTIC/PRAGMATIC distinction; and that *on the other hand* is undefined for this latter feature and hence substitutable for both phrases.

The contrast signalled by *whereas* relates to the propositional content of the related spans, and the writer's objective in presenting it is to make the reader aware of the generalisation which fails. It can thus be thought of as a SEMANTIC contrast. But clearly, the contrast signalled by *then again* does not have to relate to the propositional content of the spans. No inconsistent predicates are present in the contents of the two spans in Example 26, for instance. In such cases, different kinds of objects are apparently being compared. A PRAG-

MATIC analysis of these cases will be suggested here: in this analysis, the objects are reader beliefs (rather than objects in the world), the predicates about the objects concern the different conclusions supported by different beliefs, and the classes into which beliefs fall concern the conclusions to which they are relevant.

Consider what happens in the process of 'reaching a conclusion by examining premises'. There must first be an attempt to delineate those premises which will be relevant; an exhaustive search will not be feasible in any system with a reasonably sized set of facts and rules. Even the set of relevant propositions is likely to be too large to be exhaustively searched, and further heuristics will need to be used to consider these selectively. It is reasonable to suppose that inductive principles play a part in these heuristics: if we consider a certain number of relevant premises, and each one supports the same conclusion, there will come a point at which we decide that the conclusion is true, and stop looking for additional premises. The notion of 'reaching a conclusion' can be thought of as the moment when an inductive generalisation is made, and we assume that *all* premises relevant to the conclusion in fact support the conclusion.

Now consider what happens with *then again*. Here two relevant premises are presented which support opposite conclusions. This has the effect of blocking the generalisation, making the reader unable to reach a decision. In Lascarides and Asher's (1991) terminology, the reader has encountered a 'Nixon Diamond', where two defeasible rules are triggered, and neither takes precedence. Note that as this is a PRAGMATIC relation, the effect is not simply that the reader *realises* that inconsistent premises have been presented, but that a Nixon Diamond *actually happens* in the reader's theorem proving system. However, we must also note that the system does not freeze up altogether as a result of this impasse: it is only as regards one particular conclusion that no decision can be reached. Inductive generalisations can still be made to reach other conclusions.

3.4.2 POSITIVE and NEGATIVE POLARITY INDUCTIVE Phrases

Finally, consider the diagram in Figure 11, motivated by Texts 27 and 28.

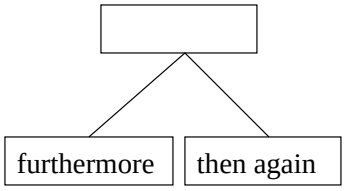

Figure 11: POSITIVE and NEGATIVE PRAGMATIC INDUCTIVE Phrases

(27) United are bound to win. They have a great team; { *furthermore,* / # *then again,* } they're playing at home.

(28) I don't know whether United will win. They have a great team; { *then again,* / # *furthermore,* } they're playing away from home.

If *then again* signals the defeat of a PRAGMATIC INDUCTIVE rule, then *furthermore* can be regarded as signalling the success of such a rule. In Example 27, a conclusion is reached: that United will win. Two relevant premises needed to be considered in order for the inductive rule to fire in this instance.

In other cases, more than two premises need to be advanced. Lists of premises signalled by phrases like *furthermore* can in principle be of any length. In the present model, these lists are analysed as nested applications of a binary relation, as in Figure 12 (i). (Premises are

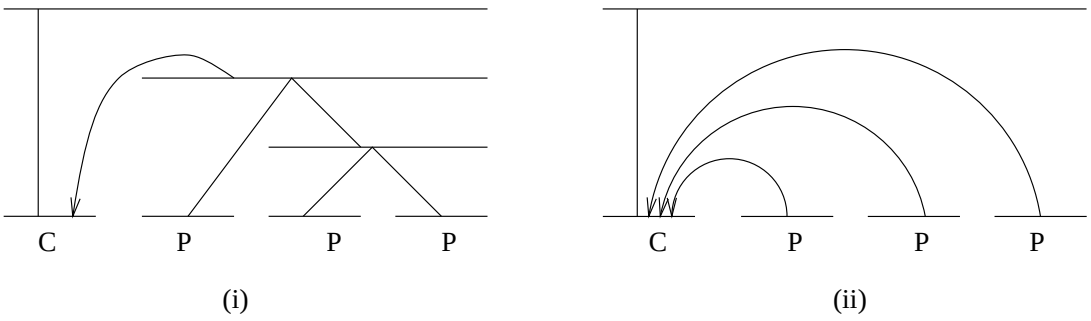

Figure 12: Alternative Structural Analyses of *Furthermore*

marked with a P; the conclusion with a C.) As regards the relations between premises, the topmost relation links the first premise with a complex span consisting of another relation between two further premises. It is easy to see how this pattern could be extended.

Figure 12 (ii) shows an alternative structural analysis for a text containing a sequence of premises; the kind that is given in Rhetorical Structure Theory (Mann and Thompson, 1988). In this analysis the premises are not represented in relation to each other, but in relation to the conclusion they support. The conclusion span is thus related to several adjacent text spans. However, a disadvantage of this approach is that it can only be used if the conclusion is represented explicitly in the text; and this is far from always the case. Maier and Hovy (1991) counter this problem by adding a separate level of 'textual' relations to the diagram in Figure 12 (ii), which link the adjacent premises. But in the present system, two levels of relations are not necessary: the relation defined by POSITIVE PRAGMATIC INDUCTIVE features is sufficiently abstract to capture both the relationship between two premises and that between the premises and the conclusion.[11]

3.5 CAUSE and RESULT-DRIVEN Relations

A further extract from the NEGATIVE POLARITY portion of the taxonomy is given in Figure 13. Motivating examples appear below.

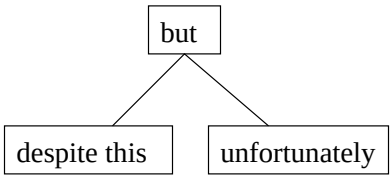

Figure 13: CAUSE-DRIVEN and RESULT-DRIVEN Phrases

(29) Bill took the lid off the pot. { *Unfortunately,* / ✓ *But* / # *Despite this,* } there was nothing inside.

[11] A question remains as to the nature of the relation which links the set of premises to the conclusion: it could either be PRAGMATIC POSITIVE CAUSAL or PRAGMATIC POSITIVE INDUCTIVE. We favour the former suggestion; note, for instance that *it follows that* can be used to introduce a conclusion supported by several premises.

(30) Sue had been up all night; { *despite this,* / ✓ *but* / # *unfortunately,* } she was looking fresh as a daisy.

The text in Example 29 can be thought of as presenting an unsatisfied desire. Let us say Bill wants to eat something; one way of achieving this goal would be if the pot contained food and its lid were removed. He lifts the lid, but there is nothing inside. *But* is certainly appropriate to describe the circumstance which blocks fulfilment of the goal; as is *unfortunately* (understood as 'unfortunately *for Bill*'). But *despite this* is quite wrong—it suggests that lifting the lid is expected to cause or entail that the pot is not empty. In Example 30, a text with a more conventional violated expectation is presented. Here *but* and *despite this* are acceptable, but *unfortunately* is inadmissible as a substitute.

These two different uses of *but* have been noted by others;[12] the following two examples of the unsatisfied-desire type are respectively from Longacre (1983) and Spooren (1989):

(31) I intended to go, but we had visitors that night.

(32) I went to the church, but the vicar was not there.

Both of these commentators attempt an explanation of such texts by proposing more abstract ways in which expectations are violated. For Longacre,

> something is presupposed here like the Newtonian assumption (inertia) that a body in motion in a given direction will keep moving in that direction unless some force deflects or stops it...

For Spooren, the expectation arises as a result of implicatures that follow from the statement of intention:

> part of our world knowledge is that going to church probably means that the vicar is in the church.

Neither of these explanations is very convincing. It is far from the case that intentions are normally achieved. It is certainly possible to set up contexts where an intention has no chance

[12] However, no-one to ourknowledge has suggested using *unfortunately* as a diagnostic for the unsatisfied-desire use.

of being achieved and yet *but* is still appropriate. Imagine Jim is in a prison cell from which he cannot escape. We could still say that

(33) Jim looked around for food, but there was none to be found.

No amount of wanting or looking for food is going to satisfy Jim's intention, so the 'inertia' explanation is ruled out. His looking for food does not probably mean that there is food to be found, so the introduction of implicatures is similarly inadmissible. Yet the phrase *but* does not seem out of place.

A preferable explanation—and one that is motivated by the pattern of substitutability in Figure 14—is that *but* is undefined for a further feature, for which *despite this* and *unfortunately* signal different values. The values of the feature relate to the manner in which the rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$ is used: are we *predicting* Q from our knowledge of $P_1 \wedge \ldots \wedge P_n$, or are we seeking to *achieve* Q, and thus investigating whether $P_1 \wedge \ldots \wedge P_n$ are true or themselves achieveable? We can call this feature ANCHOR, to reflect whether the 'certainty' relates to the knowledge of the premises, or the desirability of the conclusion. In each case, the 'certain' proposition will be identified as A (for **anchor**), and the other proposition as C (for **counterpart**). The feature can be defined as follows:

ANCHOR

CAUSE-DRIVEN: $A \in P_1 \ldots P_n$; $P_1 \wedge \ldots \wedge P_n$ is true.

RESULT-DRIVEN: A corresponds to Q; and A is desired by the protagonist.

To take an example, consider again Text 31: *I intended to go, but we had visitors that night.* This is a RESULT-DRIVEN relation: A, the first clause, which presents the writer's intention to go, corresponds to the right-hand side of the rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$. C is the second clause, presenting the fact which prevents the intention being achieved. The relation is BILATERAL, since C relates to a fact on the left-hand side of the rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$. It is NEGATIVE POLARITY, since the relevant fact in the rule (C') is inconsistent with C.

In one respect, the above definition of the ANCHOR feature is slightly fudged. Consider the original example of a RESULT-DRIVEN relation: *Bill took the lid off the pot; but there was nothing inside.* The first clause must still be considered A, but note that it does not *itself* present the intention. Rather, it presents an action performed to achieve the intention. There will be more to say about such cases in Section 3.7, which deals with presupposition. For the moment, note that the intention behind the action in the current example can be expressed in a subordinate clause:

(34) Bill took the lid off the pot *to* get some food; but there was nothing inside.

Note that the revised definitions of POSITIVE and NEGATIVE POLARITY presented in Section 3.3 are still serviceable for both CAUSE-DRIVEN and RESULT-DRIVEN relations. In POSITIVE POLARITY relations, nothing is negated, and so no problems arise. And in NEGATIVE POLARITY relations, it is always C which is negated. For a CAUSE-DRIVEN NEGATIVE POLARITY relation, the anchor A is on the left-hand side of the rule, and C is the negation of the expected conclusion. For a RESULT-DRIVEN NEGATIVE POLARITY relation, the right-hand side of the rule is desired by the protagonist, and C is the negation of one of the conditions necessary for this desire to be brought about.

It should also be noted that the definition of PATTERN OF INSTANTIATION in Section 3.3 does not need amendment. This definition determines whether C is on the same side of the rule as A (UNILATERAL), or whether they are on opposite sides (BILATERAL). In combination with the two values of the ANCHOR feature, four possible patterns of instantiation can now be expressed: A and C can both be on the left of the rule; or they can both be on the right; or A can be on the left and C on the right; or C can be on the left and A on the right.

Figure 14 provides some additions to the diagram in Figure 13. Again, motivating examples are provided:

(35) Bill took the lid off the pot. { *Unfortunately,* / ✓ *But* / ✓ *As it happened,* / # *Fortunately,* / # *Despite this,* / # *As a result,* } there was nothing inside.

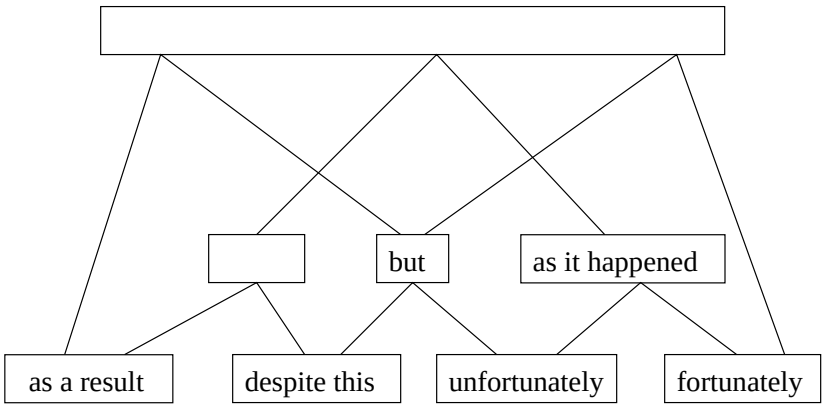

Figure 14: Additional CAUSE-DRIVEN and RESULT-DRIVEN Phrases

(36) Bill took the lid off the pot. { *Fortunately,* / ✓ *As it happened,* / # *Unfortunately,* / # *But* / # *Despite this,* / # *As a result,* } there was something inside.

Note that the exclusivity of *fortunately* and *unfortunately* depends crucially on the fact that the two phrases set up different implicatures about what Bill wants. Both of the above texts presuppose that Bill's plan requires there to be something in the pot. It might also have been that the plan required the pot to be empty—in which case the appropriateness of *fortunately* and *unfortunately* would be reversed. The two phrases cannot be substituted for one another, as to do so requires a change in the assumptions about what Bill's plan involves.

The most significant fact about the diagram in Figure 14 is that *as it happened* is substitutable for both *fortunately* and *unfortunately*, while these two latter phrases are exclusive. All three phrases are exclusive with CAUSE-DRIVEN phrases like *as a result* and *despite this*, and it is plausible that they should all be labelled as RESULT-DRIVEN. The variation within the phrases can then be traced to the POLARITY feature. *Fortunately* has POSITIVE POLARITY: the intention behind the first clause is achieved if the second clause is true. *Unfortunately* has NEGATIVE POLARITY: the intention behind the first clause is achieved if the *negation* of the second clause is true. And *as it happened* is undefined with respect to the POLARITY feature.

3.5.1 Conditional Variants of CAUSE-DRIVEN and RESULT-DRIVEN Phrases

Evidence for the distinction between *cause-driven* and *result-driven* relations also shows up in the conditional phrases in the taxonomy. Consider the two examples below:

(37) If you feel like cooking tonight, make something spicy.

(38) If you feel like cooking, there's a kitchen on the first floor.

Both of these examples are of POSITIVE POLARITY PRAGMATIC relations. In Example 37, the rule which the writer wants to succeed is that people who feel like cooking and who have a certain disposition make something spicy. (The writer's aim is thus that the reader takes on this disposition.) But we cannot envisage a similar rule for Example 38. The kitchen is on the first floor whether or not the reader feels like cooking. The point is rather that it is only *relevant* for the reader to know where the kitchen is if he feels like cooking. Treating the *if* in this text as RESULT-DRIVEN thus provides a better analysis. We assume there is a goal underlying the first clause; namely that the reader cooks. This will happen if the reader feels like cooking, *and* knows where the kitchen is.[13]

A similar story can be told for NEGATIVE POLARITY relations.

(39) Even if you manage to break out of the prison, you'll never make it home.

(40) You'll never make it home. Even if you manage to break out of the prison, the jungle round here is impenetrable.

[13] An alternative analysis for this kind of text is given by Sweetser (1990). She considers such a text to be an example of a SPEECH ACT conditional, to be read as 'if you feel like cooking, then (let us consider that) I *inform* you that there's a kitchen on the first floor'. According to this analysis, the speech act of informing is only to be understood as having occurred if the information about the kitchen is considered relevant. However, it is odd to suggest that the speech act simply disappears if its content is not relevant. The information about the kitchen is conveyed to the hearer under any circumstances, even if it is not relevant; and it is hard to see what there is to an informative speech act beyond the deliberate conveying of information. The problem is exacerbated if the analysis is extended to cover examples such as *Whenever you feel like cooking, there's a kitchen on the first floor*. If we interpret this as a SPEECH ACT conditional, we must envisage a whole series of informative speech acts, one for each time the hearer feels like cooking. At this point, we are clearly stretching the notion of a 'speech act' beyond its normal use.

In Example 39, the presupposed rule is that if one breaks out of prison, one can normally get home. The rule is defeated in the present case. In Example 40, however, there is no rule stating that breaking out of prison normally entails the jungle not being impenetrable. Rather, we must assume a goal behind the breaking out of prison, and a rule stating that the goal will be achieved if the outbreak occurs and the jungle is *not* impenetrable.

The distinction between CAUSE-DRIVEN and RESULT-DRIVEN conditional relations does not emerge systematically from the taxonomy. There are not always phrases which differ only with regard to this feature (although see Section 3.8 for a discussion of the phrase *in case*). The only systematic evidence for the feature in hypothetical relations is thus that the phrases *if* and *even if* can take either of its values. However, indirect evidence can be obtained by converting 'hypothetical' texts to 'actual' ones, and observing which cue phrases are now appropriate. The following two texts are 'actual' versions of Texts 40 and 38; the patterns of substitutability for the phrases *fortunately*, *unfortunately*, *despite this* and *so* are what we would expect for CAUSE-DRIVEN and RESULT-DRIVEN relations.

(41) Bill managed to escape from prison. { *Unfortunately,* / ✓ *but* / # *despite this,* } the jungle was impenetrable.

(42) Bill felt like cooking. { *Fortunately,* / # *So* } there was a kitchen on the first floor.

3.6 ANCHOR-BASED and COUNTERPART-BASED Relations

Another portion of the taxonomy is given in Figure 15. The motivating examples are as follows:

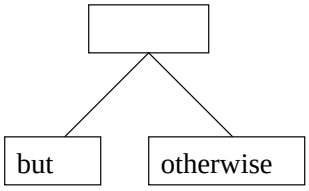

Figure 15: ANCHOR-BASED and COUNTERPART-BASED Phrases

(43) Bob put his hands up, { *otherwise* / # *but* } Jill would have shot him.

(44) Bob kept his hands by his sides, { *but* / # *otherwise* } Jill didn't shoot him.

It seems that both of these texts involve a rule along the following lines:

If Bob doesn't put his hands up, Jill will shoot him.

For both texts, the anchor relates to the left-hand side of this rule, and the counterpart to the right-hand side. The texts are similar, in that the conclusion of the rule is avoided in each case. However, the reason for this is different in the two cases. In Example 44, the premise of the rule is true, but the rule is defeated: some stronger conflicting rule must therefore be supposed to have taken precedence. In Example 43, the premise of the rule does not even occur: the protagonist takes action to avoid a conclusion which is inconsistent with his goals.

In order to link the propositions related by the cue phrases onto the premise and conclusion of the relevant rule, another dimension of variation must be introduced. At present, in NEGATIVE POLARITY relations, it is always the *counterpart* span which needs to be negated to map back onto the rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$. This is still the case for Example 44; the counterpart is the second span, which is the negation of the expected conclusion. But for Example 43, it is rather the anchor (*Bill put his hands up*) which must be negated to map onto the rule, and the counterpart (*Jill would have shot him*) which maps onto the rule without being negated.

Examples such as this suggest the need for a new feature, which we can call FOCUS OF POLARITY, to specify whether the POLARITY transformation (if there is one) operates on the **anchor** or on the **counterpart**.

The definition of FOCUS OF POLARITY will require another modification to the definition of the POLARITY feature. In the current definition in Section 3.3, it is always the counterpart of the rule (C) which is a candidate for negation: the anchor A is never negated. We now need to express the definition so as to allow the candidate for negation to be determined by FOCUS OF POLARITY. To this end, we must introduce some new variables: the **focus of polarity** (F), which is the candidate for negation, and the **invariant** (I), which is never negated. The

variables F' and I' will be used to represent F and I after the polarity transformation has taken place. The new definition for POLARITY now looks like this:

POLARITY (3nd definition)

POSITIVE: $F = F'$; $I = I'$;

NEGATIVE: $F = \neg F'$; $I = I'$.

The FOCUS OF POLARITY feature now identifies F and F' with A and A' (and I and I' with C and C') or F and F' with C and C' (and I and I' with A and A'.

FOCUS OF POLARITY

ANCHOR-BASED: $F = A$; $F' = A'$; $I = C$; $I' = C'$.

COUNTERPART-BASED: $F = C$; $F' = C'$; $I = A$; $I' = A'$.

Consider how these new definitions work with the phrases *but* and *otherwise*. For *Bob kept his hands by his sides, but Jill didn't shoot him*, the first clause is A and the second clause C. It is BILATERAL CAUSE-DRIVEN, so A' is part of the left-hand side of $P_1 \wedge \ldots \wedge P_n \rightarrow Q$, and C' is Q. The relation is COUNTERPART-BASED, so F is C and F' is C'. It is NEGATIVE POLARITY, so F is $\neg F'$; which means that C' is $\neg C$. A is the invariant, and hence maps straight onto A'. The effect is a violated expectation.

For *Bob put his hands up; otherwise Jill would have shot him*, the first clause is again A and the second clause C. The relation is BILATERAL CAUSE-DRIVEN, so A' is part of the left-hand side of $P_1 \wedge \ldots \wedge P_n \rightarrow Q$, and C' is Q. This time the relation is ANCHOR-BASED, so F is A and F' is A'. It is NEGATIVE POLARITY, so F is $\neg F'$; which means that A' is $\neg A$. This means that the rule does not trigger, and the right-hand side of the rule (C) does not occur.[14]

[14] In fact, the story is likely to be more complicated than this. The reason why C does not occur is *because it is not desired by Bob*, and he takes action to avoid it. There thus seems likely to be a result-driven component to the relation which has not so far been captured.

3.6.1 SEMANTIC and PRAGMATIC ANCHOR-BASED Phrases

As it is expressed in terms of A and C, the definition of FOCUS OF POLARITY serves equally well for SEMANTIC relations and for PRAGMATIC ones. Consider the following two cases:

(45) Bob put his hands up, *otherwise* Jill would have shot him.

(46) Put your hands up, *otherwise* I'll shoot you.

Example 45 requires that the reader believe that a protagonist in the world being described (Bill) does not want some eventuality (being shot) to occur. It is thus a SEMANTIC relation, holding between the propositional contents of the related spans. However, for Example 46 it is a precondition that someone in the real world (namely the reader) does not want to be shot. The intended effect of the relation is that the reader actually put his hands up in order to avoid this eventuality. (Note that it is not specified whether the eventuality is *in fact* avoided; or even whether it is true that the writer would have shot the reader if he had not obeyed her instruction.)

3.6.2 Distinctions Amongst ANCHOR-BASED Relations

Otherwise is only one of a number of ANCHOR-BASED phrases. There are many others that seem to fall into this category: for example, *or*, *unless*, *until*, and *before* (see Section 3.7.2 for a discussion about this latter phrase). There is not room to talk about them all here, but the dimensions of variation between them look likely to correspond to those identified by other features. Deciding whether this is indeed the case is a matter for further research.

3.7 PRESUPPOSED and NON-PRESUPPOSED Relations

Another informative extract from the taxonomy is given in Figure 16. Motivating examples are given in Texts 47 and 48:

(47) I haven't always been unfit. I played a lot of rugby { *while* / ✓ *when* / # *meanwhile* } I was at college.

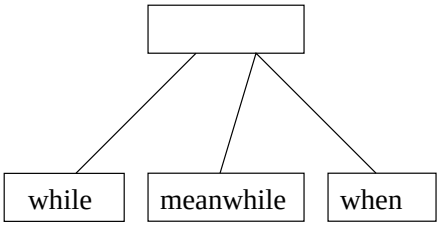

Figure 16: PRESUPPOSED and NON-PRESUPPOSED Phrases

(48) They set about preparing the meal. Bill marinated the meat; { *meanwhile,* / ✓ *while* / # *when* } Bob lit the barbecue.

All three of the phrases in the diagram convey information about temporal simultaneity, among other things. However, there is a syntactic difference between *when*, which is a temporal subordinator, and *meanwhile*, which is a sentential adverb. *While* is appropriate as a substitute for either phrase, as the examples show.

The clauses introduced by temporal subordinators are conventionally thought of as **presupposed** (Karttunen (1973), Keenan (1971), Lascarides and Oberlander (1993)). They describe an eventuality with respect to which the material in the main clause is temporally situated: this eventuality must either be already known to the reader, or must be **accomodated** prior to the addition of the temporal relation. The asymmetry of these sentences can be demonstrated by swapping the main and subordinate clauses. This typically results in incoherence, although the relation of 'temporal overlap' between the clauses is presumably unchanged:

(49) I haven't always been unfit. ?? I was at college *when* I played a lot of rugby.

On the other hand, no harm is done by changing the clause introduced by *meanwhile*:

(50) They set about preparing the meal. Bob lit the barbecue; *meanwhile* Bill marinated the meat.

The question of when accommodation is possible is addressed by Lascarides and Oberlander (1993). They propose two different mechanisms, which apply in different situations. If the subordinate clause functions simply as a temporal adverbial, as in Example 47, then there

is no need to find a coherence relation between it and the preceding context; the important relation ('BACKGROUND', in this case) is between the main clause and the preceding context. But in other cases, as in the following example, the subordinate clause has an important narrative function:

(51) The backbenchers were in revolt. They were pacified after Major launched a charm offensive.

Here, a coherence relation between the subordinate clause and the preceding context must be found to allow accommodation, and only after the subordinate clause has been attached is the main clause considered. The explanation of the asymmetry introduced by the subordinator thus turns on the *order of attachment* of the two clauses.

A similar explanation of temporal subordinators must be sought for the present theory. Here, however, it must emerge from the definitions of the relations marked by these phrases, rather than from an account of the algorithm used to interpret them. We therefore need to think about how to express the preconditions for a relation between two spans in terms of how these link to the immediately preceding context. We need a feature which takes one value for subordinating phrases like *when*, and another for non-subordinating phrases.

A feature with alternative values PRESUPPOSED and NON-PRESUPPOSED is used here to capture this difference. The feature introduces another defeasible rule, $X_1 \wedge \ldots \wedge X_n \rightarrow Y$, similar in structure to the one which the feature definitions have so far referred to. This rule makes a link between the span in the preceding context (which we will call *Precond*) and the anchor A. We can define the new feature as follows:

PRESUPPOSITIONALITY

PRESUPPOSED: *Precond* is part of the left-hand side of the rule $X_1 \wedge \ldots \wedge X_n \rightarrow Y$, and A is Y.

NON-PRESUPPOSED: *Precond* is A itself.

We can now give an account of what happens in Example 51. The first clause *the backbenchers were in revolt* is *Precond*. The rule $X_1 \wedge \ldots \wedge X_n \rightarrow Y$ has this clause as part of

its left-hand side, and anchor clause *Major launched a charm offensive* as its right-hand side. When *Precond* occurs, the rule is thus triggered.

A similar story can be told for other temporal subordinators. Consider the following example:

(52) Bob heated the water. When it boiled, he stirred in the sugar.

The first clause *Bob heated the water* is *Precond* here. It can be seen as triggering a causal rule whose right-hand side is *the water boiled*, which is the anchor of the relation signalled by *when*.

We can now consider what happens in NON-PRESUPPOSED relations. An example of such a relation appears in the text below.

(53) Jill was curious, so she pulled the lever. Instantly, an alarm went off.

In fact, there appear to be two overlapping relations in this example: one between the first and second clauses (signalled by *so*); the other between the second and third clauses (signalled by *instantly*). It seems quite plausible that the counterpart of the first relation actually *identifies with* the anchor of the second relation. This is what is captured in the definition of NON-PRESUPPOSED: an identity is specified to hold between A and *Precond*.

3.7.1 RESULT-DRIVEN PRESUPPOSED Relations

It was noted above, in connection with Example 49, that swapping the main and subordinate clauses in a PRESUPPOSED relation leads to incoherence. However, it is interesting that where the subordinate clause has a narrative function (rather than just acting as a temporal adverbial), coherent texts can be created by swapping the main and subordinate clause, provided that the right subordinating phrase is chosen. Consider the following variation on Example 51:

(54) The backbenchers were in revolt. Major launched a charm offensive to pacify them.

The anchor for the presuppositional relation is now 'The backbenchers are pacified', and this is what must be attached first to the preceding context. However, it is no longer the pacification

itself which is caused by the context. Rather it is the *intention* that the backbenchers be pacified. We have now set up a context where *Major launched a charm offensive* can act as the anchor for a RESULT-DRIVEN rule of the kind discussed in Section 3.5.

3.7.2 ANCHOR-BASED PRESUPPOSED Phrases

Finally, it is interesting to note that the rule $X_1 \wedge \ldots \wedge X_n \rightarrow Y$ is defeasible, just like $P_1 \wedge \ldots \wedge P_n \rightarrow Q$. Consider this example:

(55) Tidy your room, before I lose my temper.

The relation signalled by the subordinator *before* is PRESUPPOSED; in other words, the subordinate clause 'W loses her temper' is A, and the main clause 'Reader tidies his room' is C. Since it is PRESUPPOSED, there is a rule running from the preconditions of the text (i.e. the situation that is currently true) to A. In other words, A is predicted to happen as things stand. However, the relation is also RESULT-DRIVEN, NEGATIVE POLARITY, and ANCHOR-BASED. These parameters describe the relationship between A and C. They specify that there is a rule $P_1 \wedge \ldots \wedge P_n \rightarrow Q$, whose left hand side includes C, the action which the writer intends the reader to perform, and whose right-hand side is the *negation* of A. 'Not A' is a desire of the reader's: he wants it not to be the case that the writer loses her temper. We thus have two conflicting rules: one leading to A and one leading to $\neg A$. The left-hand side of the former rule is currently true. The left-hand side of the latter rule contains what is currently true *plus* the desired reader action. The latter rule is the one which is intended to fire, and thus the rule leading to A is defeated.

The above account of counterfactual *before* still needs to be worked out in detail. However, it is interesting that the features being developed here seem well-suited for handling such cases.

3.8 HYPOTHETICAL and ACTUAL Relations

A final extract from the taxonomy is given in Figure 17. It is motivated by the following examples:

(56) We had a strict upbringing. { *If* / ✓ *When* } we were naughty, we were sent to bed with no supper.

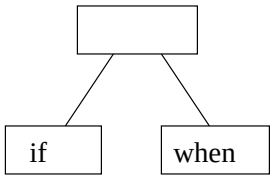

Figure 17: ACTUAL and HYPOTHETICAL Phrases

(57) You can stay up $\left\{ \begin{array}{l} \checkmark\ if \\ \#\ when \end{array} \right\}$ you don't squabble.

(58) $\left\{ \begin{array}{l} When \\ \#\ If \end{array} \right\}$ Mary gets home, ask her to call me.

All of the phrases in these examples are PRESUPPOSED. Their suitability in the different contexts seems determined by whether the presupposed span (A) is known or unknown. A feature called MODAL STATUS can thus be proposed, with alternative values ACTUAL and HYPOTHETICAL.

The MODAL STATUS feature interacts productively with a number of other features. We have already talked about 'conditional' phrases in a number of other places; for instance in connection with the POLARITY feature (*if* versus *even if*) and the ANCHOR feature (for which *if* is undefined). The question is now how to define it.

A simple idea would be to identify the contexts where the anchor A is known as ACTUAL, and those where it is not known as HYPOTHETICAL. However, there are a number of problems with this approach. Consider the case of Text 58. It cannot be that the writer actually *knows* that Mary gets home in this example, as it is an event in the future. So what is it which makes *if* and *when* different in this case? It is plausible to suggest that *when* is sanctioned by the writer's ability to *predict* Mary's return before it has happened. It would thus be preferable to define the feature in terms of knowledge (or lack of knowledge) of the *cause* of A rather than of A itself. This being the case, we can thus make use of the variable *Precond* introduced by the PRESUPPOSITIONALITY feature in Section 3.7, which for PRESUPPOSED relations such as the above, represents the cause of A. The definition of MODAL STATUS would then be as follows:

MODAL STATUS

ACTUAL: *Precond* is known by the protagonist/writer.

HYPOTHETICAL: *Precond* is not known by the protagonist/writer.

It should also be possible to talk about the HYPOTHETICAL/ACTUAL distinction for NON-PRESUPPOSED phrases. Consider the following examples:

(59) Bob piled up the boxes underneath the bananas. { *Then* / # *In that case,* } he was able to reach them easily.

(60) Bob might try piling up the boxes underneath the bananas. { *Then* / # *In that case,* } he'll be able to reach them easily.

Both *then* and *in that case* are NON-PRESUPPOSED in the above contexts, so *Precond* is identified with the anchor clause, which is the first clause in each case. *Then* seems to be undefined for MODAL STATUS, being appropriate in both contexts; but *in that case* seems to require an anchor which is unknown.

3.8.1 RESULT-DRIVEN HYPOTHETICAL Relations

A possible instance of a HYPOTHETICAL RESULT-DRIVEN phrase is *in case*. Consider the following text:

(61) Bill tidied the house, { *in case* / # *because* } his parents came home early.

This text requires an inference about one of Bill's goals; namely that he does not want his parents to come home early and find that the house is not tidy. It is not certain that his parents will come home early—hence the relation is HYPOTHETICAL—but it is sufficiently likely to make Bill tidy the house. We can therefore assume that the two related spans *Bill tidied the house* and *his parents came home early* are both part of the left-hand side of some rule whose right-hand side is a state of affairs desired by Bill, such as 'Bill's parents are not angry'. Note that the ACTUAL phrase *because* is inappropriate for signalling this relation.

Name of Feature	Possible Values	
SOURCE OF COHERENCE	SEMANTIC	PRAGMATIC
ANCHOR	CAUSE-DRIVEN	RESULT-DRIVEN
PATTERN OF INSTANTIATION	UNILATERAL	BILATERAL
FOCUS OF POLARITY	ANCHOR-BASED	COUNTERPART-BASED
POLARITY	NEGATIVE	POSITIVE
PRESUPPOSITIONALITY	PRESUPPOSED	NON-PRESUPPOSED
MODAL STATUS	HYPOTHETICAL	ACTUAL
RULE TYPE	CAUSAL	INDUCTIVE

Figure 18: The Features So Far Motivated, and Their Alternative Values

4 Summary of Features Motivated

Now that a number of features have been individually motivated, we can begin to put them together to build up the complex definitions required for cue phrases and relations. A summary of the eight features and their possible values is given in Figure 18.

Figure 19 presents the original extract from the taxonomy, labelled with those features which have so far been defined. The feature-theoretic interpretation of the taxonomy is clearly illustrated here: exclusive phrases are defined for alternative values of at least one feature; hyponyms inherit all of the feature values associated with their hypernyms and are defined for other features in addition, and so on.

The diagram in Figure 19 documents many of the substitutability relationships shown in the smaller diagrams in Section 3, as well as many relationships between phrases which appeared in different diagrams. The figure divides roughly into four exclusive groups of phrases:

- POSITIVE POLARITY CAUSAL ACTUAL relations (dominated by the phrase *so*);
- NEGATIVE POLARITY ACTUAL relations (dominated by the phrase *but*);
- HYPOTHETICAL relations (involving *if* and *even if*);

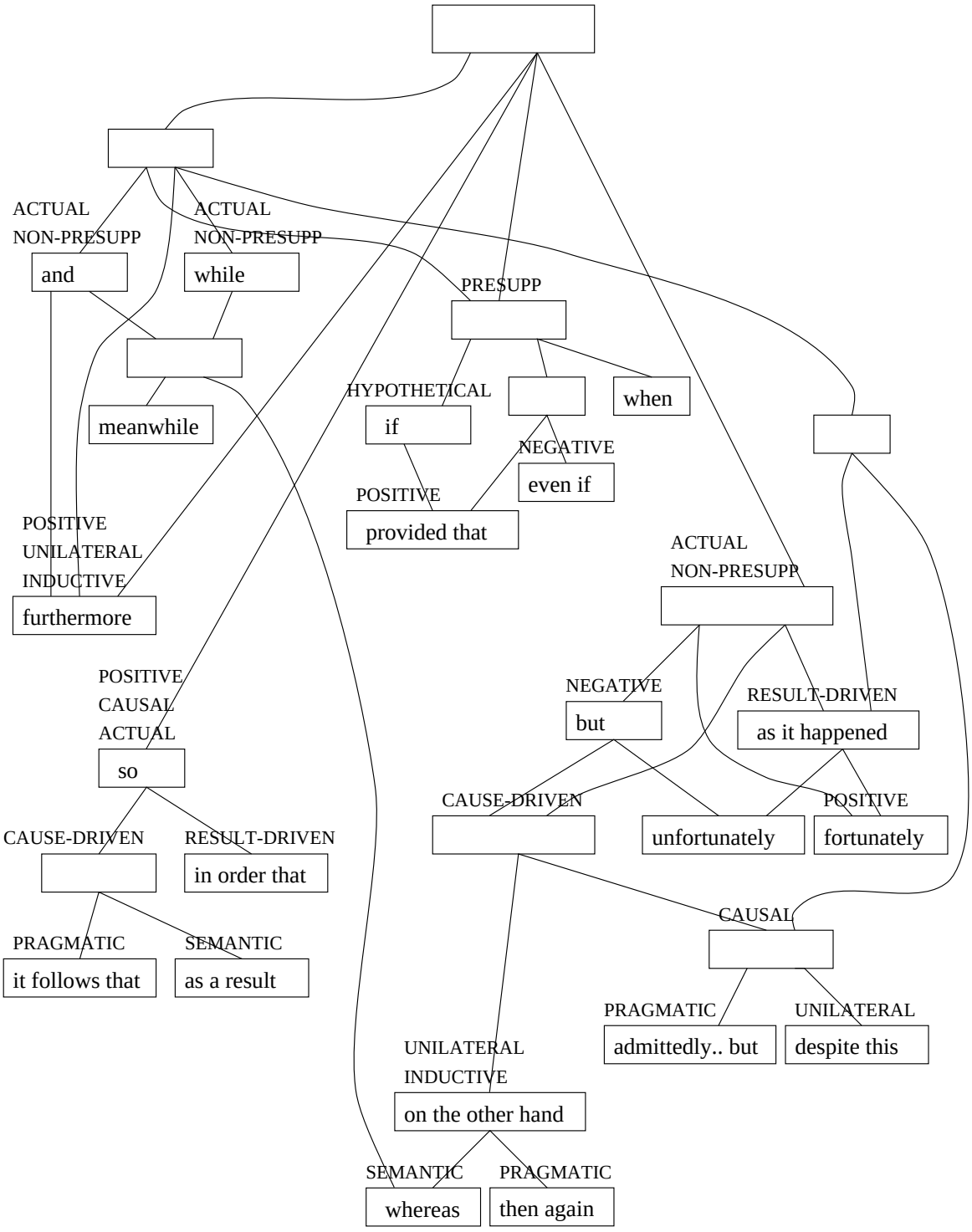

Figure 19: Summary of Motivated Features: A Labelled Extract from the Taxonomy

- POSITIVE POLARITY INDUCTIVE relations (the single phrase *furthermore*).

Much of the complexity in the diagram is due to high-level phrases such as *and* and *while*, which cut across these divisions.

The features with which the phrases are labelled are unlikely yet to be sufficient as definitions, as many additional features have still to be motivated from the taxonomy. Even in this diagram—still just a small portion of the overall taxonomy—there remain relationships which are not yet explained by the features provided. (For instance, the contingent substitutability between *while* and *whereas* remains unexplained. So does the exclusivity between *and* and *it follows that*.) However, the current set of features already give good approximate definitions in most cases, and at very least serve to indicate the flavour of the definitions which will eventually be reached.

A larger set of composite definitions is given in Figure 20, this time in tabular form. Again, the definitions are not always complete;[15] they should rather be thought of bringing together a collection of additional hypotheses to investigate in the style of Section 3. But they already begin to provide a reasonable account of the variations between the phrases.

5 Conclusion

This paper has presented a methodology for developing a theory about the semantics of sentence/clause connectives. The methodology involves the gathering of a corpus of connectives, the organising of this corpus into a hierarchical taxonomy, and the interpretation of the taxonomy in the light of a feature-based conception of interclausal/intersentential relations. The bulk of the paper was devoted to this latter task. Much remains to be done before an adequate feature-theoretic account of the taxonomy is provided; however, the features developed to date provide a promising basis for further work undertaken to this end. At very least, they should serve to justify the methodology being proposed; and in particular to justify the suggestion that a theory of connectives can be developed which does not require the postulation of intermediate-level categories.

[15] The boxes where we are not sure of the value of a particular feature have been left empty; they should be distinguished from boxes containing a dash (—), which denote that a feature is undefined.

Cue Phrase	SOURCE OF COHERENCE	ANCHOR	PATTERN OF INSTANTIATION	FOCUS OF POLARITY	POLARITY	PRESUPPOSITION-ALITY	MODAL STATUS	RULE TYPE
if S_A, S_C	—	—	—	COUNT	—	PRES	HYP	CAUS
S_A; despite this, S_C	PRAG	CAUS	BILAT	COUNT	NEG	NON	ACT	CAUS
even if S_A, S_C		—	—	COUNT	NEG	PRES	HYP	CAUS
even when S_A, S_C			—	COUNT	NEG	PRES	ACT	CAUS
S_A; otherwise S_C		RES	BILAT	ANCH	NEG	NON	ACT	CAUS
unless S_A, S_C	PRAG	CAUS	BILAT	ANCH	NEG	PRES	HYP	CAUS
until S_A, S_C	SEM	CAUS	BILAT	ANCH	NEG	PRES	ACT	CAUS
because S_A, S_C	PRAG		BILAT	COUNT	POS	PRES	ACT	CAUS
S_A, in order that S_C	PRAG	RES	BILAT		POS	PRES		CAUS
to S_A, S_C	PRAG	RES			POS	PRES		CAUS
by S_A, S_C	PRAG	RES	BILAT		POS	PRES	ACT	CAUS
S_A; instead S_C			UNIL	COUNT	NEG	NON	ACT	CAUS
S_A; rather S_C			UNIL	COUNT	NEG	NON	ACT	CAUS
S_A or S_C	—	—	—	ANCH	NEG	NON	HYP	CAUS
S_A; besides, S_C	PRAG		UNIL	COUNT	NEG	NON		CAUS
S_C, before S_A	SEM					PRES	ACT	CAUS
S_A; instantly S_C	SEM					NON	ACT	CAUS
S_A; before long S_C	SEM					NON	ACT	CAUS
S_A; suddenly S_C	SEM					NON	ACT	CAUS
S_A; not that S_C	PRAG					NON		
S_A; even then S_C					NEG		—	CAUS
S_A; then S_C	—				POS		—	
S_A; next S_C	—				POS	NON	ACT	
S_A; finally S_C	—				POS	NON	ACT	
only if S_A, S_C	—				NEG	PRES	HYP	CAUS
only when S_A, S_C	SEM				NEG	PRES	ACT	CAUS
S_A; in that case, S_C					POS	NON	HYP	CAUS
S_A, so S_C	—	—	BILAT		POS			CAUS
S_A; it follows that S_C	PRAG		BILAT		POS	NON	ACT	CAUS
S_A; as a result S_C	SEM	CAUS	BILAT	COUNT	POS	NON	ACT	CAUS
S_A; but S_C	—	—	—	COUNT	NEG	NON	ACT	—
S_A; as it happened, S_C		RES	UNIL		—	NON	ACT	CAUS
S_A; fortunately, S_C	SEM	RES	UNIL		POS	NON	ACT	CAUS
S_A; unfortunately, S_C	SEM	RES	UNIL		NEG	NON	ACT	CAUS

Figure 20: Some Further Preliminary Cue Phrase Definitions

References

Cohen, P. and Perrault, C. (1979). Elements of a plan-based theory of speech acts. *Cognitive Science*, **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*, pages 251–258.

Delin, J., Hartley, A., Paris, C., Scott, D., and Vander Linden, K. (1994). Expressing procedural relationships in multilingual instructions. In *Proceedings of the 7th International Workshop on Natural Language Generation*, pages 61–70, Kennebunkport, Maine.

Di Eugenio, B. (1992). Understanding natural language instructions: The case of purpose clauses. In *Proceedings of the 30th Conference of the Association for Computational Linguistics*, pages 120–127, Delaware.

Elhadad, M. and McKeown, K. R. (1990). Generating connectives. In *COLING-90*, pages 97–101.

Glasbey, S. (1995). 'when', discourse relations and the thematic structure of events. In *Proceedings of the Workshop on Time, Space and Movement*, Toulouse.

Grote, B., Lenke, N., and Stede, M. (1995). Ma(r)king concessions in english and german. In *Proceedings of the Fifth European Workshop on Natural Language Generation*, Leiden, the Netherlands.

Halliday, M. and Hasan, R. (1976). *Cohesion in English*. Longman.

Hobbs, J., Stickel, M., Appelt, D., and Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, **63**.

Hovy, E. (1993). Automated discourse generation using discourse structure relations. *Artificial Intelligence*, **63**, 341–385.

Hovy, E. H. (1988). Planning coherent multisentential text. In *Proceedings of the 26th ACL Conference*, pages 163–169, Buffalo.

Karttunen, L. (1973). Presuppositions of compound sentences. *Linguistic Inquiry*, **4**, 169–193.

Keenan, E. (1971). Two kinds of presupposition in natural language. In C. Fillmore and D. Langendoen, editors, *Studies in Linguistic Semantics*. Holt, New York.

Knott, A. (1996). *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. Ph.D. thesis, Department of Artificial Intelligence, University of Edinburgh.

Knott, A. and Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, **18**(1), 35–62. Also available as Technical Report RP-34, Human Communication Research Centre, University of Edinburgh, 1992.

Lascarides, A. and Asher, N. (1991). Discourse relations and defeasible knowledge. In *Proceedings of the 29th Conference of the Association for Computational Linguistics*, pages 55–63, Berkeley, CA.

Lascarides, A. and Oberlander, J. (1993). Temporal connectives in a discourse context. In *Proceedings of the 6th Conference of the European Chapter of the Association for Computational Linguistics*, pages 260–268, Utrecht, The Netherlands.

Lascarides, A., Asher, N., and Oberlander, J. (1992). Inferring discourse relations in context. In *Proceedings of the 30th Conference of the Association for Computational Linguistics*, pages 1–8, Delaware.

Longacre, R. (1983). *The Grammar of Discourse: Notional and Surface Structures*. Plenum Press, New York.

Maier, E. and Hovy, E. H. (1991). A metafunctionally motivated taxonomy for discourse structure relations. In *Proceedings of the 3rd European Workshop on Natural Language Generation*, pages 38–45, Judenstein, Austria.

Mann, W. C. and Thompson, S. A. (1988). Rhetorical structure theory: A theory of text organisation. *Text*, **8(3)**, 243–281. Also available as Tech Report RR-87-190, USC Information Sciences Institute, Marina del Rey, CA.

Martin, J. (1992). *English Text: System and Structure*. Benjamin, Amsterdam.

Moens, M. and Steedman, M. (1987). Temporal ontology in natural language. In *Proceedings of the 25th ACL Conference*, pages 1–7, Stanford, CA.

Moore, J. D. and Paris, C. L. (1989). Planning text for advisory dialogues. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pages 203–211.

Moore, J. D. and Paris, C. L. (1993). Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, **19**, 651–694.

Oversteegen, E. (1993). Dutch temporal connectives. Technical Report R1.3.2a, The DANDELION Consortium, CLS, Tilburg/Nijmegen.

Oversteegen, E. (1995). Causal and contrastive connectives. Technical Report R1.1.3b, The DANDELION Consortium, CLS, Tilburg/Nijmegen.

Perrault, C. R. and Allen, J. F. (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, **6**(3), 167–182.

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, **14**, 367–381.

Sanders, T. J. M., Spooren, W. P. M., and Noordman, L. G. M. (1992). Towards a taxonomy of coherence relations. *Discourse Processes*, **15**, 1–35.

Searle, J. R. (1969). *Speech Acts*. Cambridge: Cambridge University Press.

Spooren, W. (1989). *Some Aspects of the Form and Interpretation of Global Contrastive Coherence Relations*. Ph.D. thesis, Catholic University of Nijmegen, the Netherlands.

Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press.

van Dijk, T. A. (1979). Pragmatic connectives. *Journal of Pragmatics*, pages 447–456.

Von Klopp, A. (1993). *Negation: Implications for Theories of Natural Language*. Ph.D. thesis, Centre for Cognitive Science, University of Edinburgh.