WITH

A thesis submitted to the University of Manchester
for the degree of Master of Science
in the Faculty of Science and Engineering

October 1996

By
Diana Gabrielle Maynard
Department of Computer Science

Contents

List of Tables

Abstract

Prepositional attachment is a common source of structural ambiguity in NLP. The problem is often caused by the preposition being semantically ambiguous. This thesis examines the different meanings of the preposition *with* and the ways in which it is used in natural language. It investigates how such information can be used to help resolve the ambiguity, and to what extent the verb, direct object and prepositional object can help disambiguate the preposition. The results show that we can correctly predict the meaning of the preposition over a third of the time using verbal information alone, and this is improved with the addition of nominal information. The inclusion of deeper forms of information such as pragmatics and contextual clues, as used in human processing, could possibly enhance this still further.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

Copyright

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the head of Department of Computer Science.

Acknowledgements

I would like to thank my supervisors, Mary Wood and David Brée, for all their support and enthusiasm, and for the different viewpoints they provided.

Chapter 1

Introduction

The correct semantic analysis of sentences is an important prerequisite of almost any NLP application. It is unlikely that machines will ever be able to truly "understand" language in the same way that humans can, because however sophisticated they may be, machines can only follow instructions and have no innate language skills.

The task of the computational linguist is therefore to design rules and algorithms which will enable the system to produce correct interpretations of language. These do not necessarily have to emulate human processing mechanisms, but simply need to produce the same results. However, the more comprehensive our understanding of these human mechanisms, the greater our chances of success.

One of the main problems facing NLP systems is that of correctly interpreting ambiguous sentences. There are various types of ambiguities, but the one we are concerned with here is that of syntactic ambiguity, and in particular the problem of prepositional attachment.

1.1 Prepositional Attachment

The analysis of a sentence involves dividing it into syntactic constituents. The problem of prepositional attachment is thus concerned with attaching the PP to the correct constituent. PPs can be attached to verb phrases, noun phrases, adjectives, or whole sentences, as the following examples show.

VP-attachment: *The children swam <u>in the sea</u>.*

NP-attachment: *The man <u>with the beard</u> looks like a convict.*

Adj-attachment: *Blue <u>with cold</u>, the swimmers climbed out of the river.*

Sentence attachment: *<u>With so many things on his mind</u>, John couldn't sleep.*

1.1.1 VP- or NP-attachment?

The most common source of prepositional attachment ambiguity involves the possibility of the PP being attached to either the VP or the preceding NP.

John killed the man with a sledgehammer.

We would normally analyse this sentence as having the PP attached to the VP, i.e. where *with a sledgehammer* is attached to *killed.* But an alternative interpretation is also possible, where it attaches to the NP *the man.* The following sentence, however, would only have one possible interpretation for us.

John killed the man with a limp.

On syntactic grounds, this sentence is also ambiguous, but semantically, the VP-attachment interpretation does not make sense, so only the NP-attachment interpretation is valid.

1.1.2 Semantic Ambiguity of Prepositions

The reason why a PP can attach to a VP or an NP lies in the fact that many prepositions are semantically ambiguous. The preposition may thus be used in a different way depending on the constituent to which it attaches. In order to resolve the attachment problem, we therefore need to understand the meaning of the preposition.

John killed the man with a sledgehammer.

In the VP-attachment reading of this sentence, *with a sledgehammer* has an instrumental meaning, i.e. the sledgehammer is an instrument being used to carry out the killing. In the NP-attachment reading, it has a possessive meaning, i.e. the man being killed possesses a sledgehammer.

I opened the door in my pyjamas.

Similarly, in this sentence there is a difference in meaning depending on the attachment. The VP-attachment reading interprets *in my pyjamas* as a manner adverbial, i.e. how I opened the door, whilst the NP-attachment reading interprets it as a locative, i.e. where the door was.

1.1.3 Semantics of Other Constituents

However, it is not only the meaning of the preposition that plays a role in attachment ambiguity. The semantics of other constituents also determines whether they are a possible site for attachment.

I opened the door to the postman with my pyjamas on.

I opened the door to the bathroom with my pyjamas on.

In the first sentence, both VP-attachment and NP-attachment readings are possible. VP-attachment, where I am wearing my pyjamas whilst opening the door to the postman, is the most likely interpretation. NP-attachment, where the postman is wearing my pyjamas, is an unlikely but possible interpretation. In the second sentence, VP-attachment again makes perfect sense (I was wearing my pyjamas whilst opening the door to the bathroom). But NP-attachment makes no sense at all (the bathroom is wearing my pyjamas when I open the door to it). Similarly, the semantics of the verb can determine whether it is a possible attachment site.

John killed the man with a sledgehammer.

John saw the man with a sledgehammer.

In the first sentence, both VP- and NP-attachment is possible, as we have already discussed. The second sentence, however, precludes the possibility of VP-attachment, because a sledgehammer cannot be used as an instrument for seeing.

1.2 Aims

The main purpose of this thesis is to help solve the problem of prepositional attachment ambiguity by investigating one specific preposition, *with*. This preposition was chosen because it has the potential to cause many problems, since semantically it is highly ambiguous. We examine the different ways in which *with* can be used, and look at the methods that can be used to help determine the correct attachment point. In particular, we focus on the possibility that correct predictions can be made given the verb and its objects of the preposition. Experiments such as [BR94] and [WAB$^+$91] seem to suggest that this may be sufficient information to produce good results, but psycholinguistic research implies that other information such as real-world knowledge and contextual clues play a part in human sentence processing. We predict that if this type of information can be incorporated into automatic systems, better results could be obtained. This, however, would in itself be a non-trivial task, and would be most easily undertaken in restricted domains. We also predict that, because of their very nature, domain-specific texts would in any case contain fewer ambiguities and be easier to disambiguate.

1.3 Organisation

The first section, comprising the next three chapters, reviews some of the previous research that has been done in this area. Firstly we examine some of the psycholinguistic studies which investigate the sentence processing mechanisms used by humans, and the ways in which we disambiguate sentences. Then we consider the main strategies adopted for machine disambiguation, and to what extent the mechanisms used by humans can be incorporated into automatic systems. Thirdly we review some of the relevant linguistic research that has been carried out on prepositions, and in particular on *with*.

The following three chapters examine the meaning and uses of the preposition *with*. Chapter 5 develops a set of semantic categories for it, whilst in Chapter 6 we turn our attention to a large body of text, the Brown corpus, and investigate how *with* is actually used. We study the co-occurrence of certain verbs and nouns with this preposition, and see what conclusions we can draw about its behaviour. We also examine whether it is possible to predict the correct attachment site for the PP, given the verb and its object, or whether other factors such as contextual information are necessary. In Chapter 7, we focus on specific subdomains within the corpus, to see if a restricted vocabulary aids the process of disambiguation, and finally Chapter 8 outlines our conclusions from this research and highlights possible future work to be carried out.

Chapter 2

Human Sentence Processing

One of the major problems facing NLP systems is that an enormous amount of knowledge is required for comprehension. Even in restricted domain texts, it is to be expected that parsers will either encounter problematic sequences or produce incorrect analyses. This may be due to the system's incomplete linguistic knowledge or to its lack of world knowledge, or it may be due to the way in which the text is written. Although there is a large dividing line between the comprehension capabilities of an NLP system and those of a human, many texts will contain sentences for which even humans find it difficult or even impossible to decide on the correct interpretation, due to ambiguities.

It is important to distinguish between syntactic ambiguity at three different levels. At the most basic level, a naïve syntactic parser will identify, for example, any structure of the form V + NP + Prep + NP as being ambiguous, not knowing whether the preposition attaches to the preceding NP or to the V. Since most prepositions attach to the NP, it may always make the choice to attach the preposition to the NP, but in the case of *with*, this would be a bad choice since *with* is an exception, attaching more commonly to the verb [Bri93]. At the next level, a more comprehensive parser will employ some more sophisticated means of disambiguation, such as the encoding of semantic information, e.g. case frames for verbs, or the use of co-occurrence and frequency based techniques, which can improve a parser's performance by making use of lexical preferences [CM93].

At the uppermost level, syntactic ambiguity can be viewed in human terms. The fact that relatively few sentences are genuinely ambiguous to humans, in terms of structure, should be able to aid us in the development and use of disambiguation techniques. There are a number of ways in which humans can disambiguate sentences, such as using syntactic rules and making use of semantic, pragmatic or discourse information. In order to understand how human processing works, in addition to understanding the mechanisms involved, it is also important to identify what type of information is used in each particular situation.

2.1 Syntactic Information

One method of determining the correct attachment points for a prepositional phrase is to use some kind of syntactic rule-based system. Various structural rules have been proposed as the basis for decisions about attachment and other local syntactic ambiguities, such as Late Closure [Fra78], Minimal Attachment [Fra78] and Verb Dominance [Kim73].

2.1.1 Late Closure

The strategy of Late Closure predicts that, wherever possible, an item will be attached to the constituent currently being processed [Fra85], in other words that constituents will be closed as late as possible, and that attachment will take place at the lowest level. It is also known as Right Association [Kim73], since the item being processed is attached as far to the right in the parse tree as possible. For example, in the sentence *The secretary typed the letter that the manager had dictated with care*, the initial analysis will attach the PP *with care* to the last constituent being processed, i.e. to the verb *dictated*, rather than to *typed*. Other information may later be used to discard this analysis and select the alternative interpretation.

2.1.2 Minimal Attachment

Frazier's Minimal Attachment rule [Fra78] claims that we prefer the interpretation where a constituent is attached at the highest level, so that there is a minimum number of nodes. This means that where there is a choice between attachment of the PP to the VP or to the preceding NP, we choose VP-attachment. For example in the phrase *hit the man with a hammer*, we prefer the instrumental interpretation of the PP *with the hammer*, where the hammer is used to hit the man, rather than the possessive interpretation where the man being hit has a hammer. It is only if the minimal attachment interpretation is semantically incongruous that we then consider the alternative one. So in the phrase *hit the man with a beard*, we would first consider the interpretation of using the beard to hit the man, but after realising that this is an unlikely interpretation, since beards are not usually used for hitting people, we would come to the alternative conclusion that the beard is an attribute of the man.

The Minimal Attachment hypothesis thus assumes that disambiguation is initially syntactic, but can be overridden by semantics. Because it predicts the simplest analysis with the least nodes, and only constructs one representation at a time, it is very efficient in terms of processing time [FC86]. If the initial syntactic analysis turns out to be wrong, a new analysis is created, so it is predicted that sentences involving non-minimal attachment will take longer to process, because subjects will be garden pathed. This has been supported by a number of experiments, e.g. [FR82], [RCF83].

Rayner et al [RCF83] also conclude that the initial choice of attachment is syntactic rather than semantic, and argue that the two processes are completely distinct. They performed a series of experiments using pairs of sentences such as the following.

The spy saw the cop with the binoculars, but the cop didn't see him.

The spy saw the cop with a revolver, but the cop didn't see him.

In the first sentence, the most plausible interpretation is the minimal (VP-) attachment reading, whilst in the second it is the non-minimal (NP-) attachment reading. As would be expected, they found that pragmatics influences the decision making, so that the subjects eventually choose the correct interpretation, whereby the first PP is an instrumental and the second PP is a possessive. However, with the second sentence in which the non-minimal attachment reading is the most plausible one, the subjects were garden pathed. This indicates that they initially chose the minimal attachment reading, which was then overridden by pragmatic information to give the correct interpretation. If they had not been garden pathed, this would have indicated that pragmatic rather than syntactic information had been used initially to determine the correct attachment.

2.1.3 Verb Dominance

For sentences involving PP-attachment, Holmes [Hol84] explains the choice of analysis in terms of the structurally-based principle of Verb Dominance [Kim73]. This states that if a particular case or argument can follow a verb, then if there is a constituent with the potential to fill that role, it is assumed that the constituent does indeed fill that role and that it is thus attached to the verb. For example, in the sentence *The Australians are discussing the trade agreement with the Japanese*, there are two possible interpretations. The first is where the PP is attached to the VP, and has the meaning that the Australians are having a discussion with the Japanese about the trade agreement. The alternative is where the PP is attached to the NP, with the meaning that a trade agreement with the Japanese is being discussed amongst the Australians. Holmes found that the preferred interpretation for sentences such as these was VP-attachment rather than NP-attachment, which is in accordance with the Verb Dominance Principle. The verb *discuss* has a potential argument slot to be filled by a PP (*discuss something <u>with somebody</u>*). Since there is a PP capable of filling this role (*with the Japanese*), it is assumed to fill it and is therefore assumed to attach to the verb, even though there is absolutely no obligation for the PP or for any other constituent to do so.

Ford, Bresnan and Kaplan [FBK81] also agree that it is the verb which is significant in cases of structural ambiguity. They tested this theory using examples in which only the verbs differed.

The women discussed the dogs on the beach.

The women kept the dogs on the beach.

In the first example, NP-attachment was preferred, whilst in the second example, VP-attachment was preferred. However, there were no time restrictions in their experiments, meaning that their results reflected the final rather than the initial interpretation chosen by the subjects.

2.2 Semantic Information

2.2.1 Case Frames

Semantic information is useful in dealing with sentences where there is ambiguity in the actual meaning of the preposition. For example, in the sentence *I saw the man with the beard*, there is no ambiguity because a beard is not an instrument for seeing, thus ruling out the interpretation whereby *with* is used as an instrumental, denoting that the beard was used to see the man. The sentence *I saw* *the man with the telescope* is, however, still ambiguous because *with* can be interpreted either as an instrumental (the telescope is used in order to see the man), or as an attributive (the telescope belongs to the man). One principle governing our use of semantic information in this way develops the idea of case frames for verbs [Fod78] [FBK81], suggesting that we may use some kind of frequency- or probability-based mechanism to judge the likelihood of attachment of two constituents or to enable us to make alternative analyses [RCF83]. It is not clear though exactly what form the case frame information would take. The possibilities range from strict subcategorisation information [Cho65] to information about selectional restrictions, thematic roles, obligatoriness, etc. [CFJ+84].

2.2.2 Contextual Clues

We might also be able to make use of pragmatic information or real-world knowledge to aid us with disambiguation. For example, in the phrase *the box on the table with a broken leg*, we would attach the PP *with a broken leg* to the NP *the table* rather than to the NP *the box*, since we know that boxes do not have legs, but tables do. With the phrase *the doll on the table with a broken leg*, we would have more difficulty in deciding where to attach the PP, since both dolls and tables can have broken legs. However, given the sentence *The doll looked very miserable on the table with a broken leg*, we would be much more likely to interpet this as meaning that it was the doll and not the table that had a broken leg. Syntactic-based hypotheses such as Late Closure would predict the alternative interpretation, but our knowledge of the world would probably lead us to understand that it was because the doll had a broken leg that it looked miserable, even though we are aware that dolls do not actually have feelings.

Similarly, with a sentence such as *Mary killed the burglar with a shotgun*, we would imagine that it was Mary who had the shotgun. But if the sentence continued *...but left the one who was unarmed*, we would understand that there were two burglars, one with a shotgun and one without, and that it was the one with a shotgun that Mary killed.

Crain [Cra80] proposes that context can also be used to help determine attachment in the case of garden path sentences. His experiments used pairs of sentences such as the following:

> *The psychologist told the woman that he was having trouble with her husband.*
> *The psychologist told the woman that he was having trouble with to visit him again.*

Each of the sentences was presented in one of two contexts, either where there were two woman in the room with the psychologist, or where there was only one. Given the context where there were two women in the room, we might expect that there would be a garden path effect in the first sentence, but not in the second, whilst with the alternative context we might expect the opposite effect.

Crain and Steedman [CS85] proposed a mechanism for using context to help make attachment decisions, which was later reformulated by Altman and Steedman [AS88] as the Principle of Referential Support:

> "An NP analysis which is referentially supported will be favoured over one that is not."

This means that if the context makes one interpretation more likely than the other, this interpretation will be favoured. For example, the NP *the woman* is referentially supported if there is a single woman in the universe of discourse. If the context contains more than one woman, then it will not be referentially supported. On the other hand, the NP *the woman with a blue hat* is only referentially supported if there is more than one woman in the universe of discourse and one of these women is distinguished from the rest by virtue of her having a blue hat.

Altman and Steedman [AS88] performed similar experiments to Crain's, using pairs of sentences supported by one of two contexts. In the VP-supporting context, only one possible referent for the object of the verb was introduced, whilst in the NP-supporting context, two possible referents were introduced. The idea is that where there is only one possible referent in the context, VP attachment is supported, because the object of the verb does not need to be further restricted. Where there is more than one referent, NP attachment is supported, because it is necessary to identify which of the two possible candidates is being referred to, and hence we expect a further restriction on the NP. An example of the contexts and target sentences is given below.

NP-supporting context:

> *A burglar broke into a bank carrying some dynamite.*
> *He planned to blow open a safe.*
> *Once inside he saw that there was a safe with a new lock and a safe with an old lock.*

VP-supporting context:

> *A burglar broke into a bank carrying some dynamite.*
> *He planned to blow open a safe.*
> *Once inside he saw that there was a safe with a new lock and a strongbox with an old lock.*

NP-attached target:

> *The burglar blew open the safe with the new lock and made off with the loot.*

VP-attached target:

The burglar blew open the safe with the dynamite and made off with the loot.

Following the Minimal Attachment hypothesis, referential context should have an effect on NP-attached but not VP-attached sentences. However, the results showed that both NP- and VP-attached sentences were affected by context, and the VP-attached sentences even produced longer response times than the NP-attached ones. Altmann and Steedman offer the explanation that the difference in response time is caused by the use of different processes for establishing plausibility in each case. They claim that the subjects use their general world knowledge in the case of VP-attached sentences, and their discourse model in the case of the NP-attached sentences. Since the discourse model is relatively small compared to the knowledge base, it is unsurprising therefore that the process involving discourse knowledge is faster.

2.3 Processing Mechanisms

As a result of their experiments [RCF83] [FC86], both Rayner et al and Fereira and Clifton argue in favour of an independent syntactic module, which can process information about phrase structure and subcategorisation frames, but which does not involve the processing of semantic or pragmatic information. This is supported by the Minimal Attachment hypothesis in the case of PP-attachment, where subjects initially prefer to attach the PP to the VP rather than to the NP. The question therefore arises as to how and where non-syntactic information is processed. Rayner et al's Thematic Selection hypothesis [RCF83] provides a possible solution. This proposes that sentence processing is influenced by pragmatics only after initial syntactic analysis has taken place. The syntactic analysis identifies certain thematic structures associated with the head of the phrase, and the plausibility of these structures is then judged by the thematic selection process, which makes use of world knowledge and discourse information. If the analysis is deemed to be pragmatically implausible, the sentence must be reanalysed and the selection procedure repeated. It is not entirely clear though how this non-syntactic information is used. Fereira and Clifton [FC86] suggest either that it is represented by means of a mental model [JL83], or that there is some kind of discourse processor which interacts with the thematic processor and suggests plausible verb arguments.

The alternative viewpoint is that processing occurs interactively, with all types of information being available at any stage in the processing. If this is the case, then one would expect the initial analysis to be the most plausible one [FC86], and that subjects would not be garden pathed in sentences where the most plausible interpretation corresponds to the non-minimal attachment reading. Altman and Steedman [AS88] reject the principle of Minimal Attachment in favour of the Principle of Referential Support, which they find to be consistent with the interactive hypothesis.

Whatever the exact form of mechanism used, it is clear that the main components are similar. There must be some form of syntactic component and some form of semantic component, which will probably incorporate pragmatic and information, and possibly contextual information. In the next chapter, we examine the mechanisms incorporated into automatic processing systems, and see how they compare to human ones.

Chapter 3

Machine Disambiguation

It might seem that emulation of the human processing mechanism would be the most efficient solution for machine disambiguation, considering the efficiency of the human mechanism. This would be likely to comprise a syntactic-based module for initial analysis, followed by either an independent or an interactive semantic processor. However, we must bear in mind that although rule-based approaches can easily operate automatically, the same cannot be said of processes that involve making judgements based on deep forms of knowledge, such as pragmatics and contextual clues. Machine disambiguation does not generally make much use of this type of information, since it is hard to identify what is relevant, and to find a suitable way of encoding the information. However, although some cases of ambiguity will require this kind of knowledge, recent work has shown that more superficial forms of knowledge can be used successfully. Various researchers have attempted to resolve prepositional phrase attachment ambiguities by means of statistical approaches involving semantic word classes, e.g. [WAB+91], [BPV91], [Bri93], [HR93]. These probabilistic models produce the interpretation with the greatest likelihood of occurrence.

3.1 Rule-Based Approaches

Brill and Resnik [BR94] describe a rule-based approach to disambiguation of prepositional phrase attachment, which uses information automatically extracted from a large corpus. Their approach relies on a rule-based approach which attempts to predict the correct attachment of the preposition, with the aid of contextual information. The algorithm they use is a transformation-based error-driven learning mechanism which they tested on structures of the form V-NP-PP. In each case, attachment of the preposition could be either to VP or to the NP. Each noun in the text was annotated with information about its semantic class, and on the basis of this, rules could then be applied. For example, prepositional phrases denoting time are more likely to attach to a verb than to a noun, so if the prepositional object belongs to a semantic class denoting time, the PP will be attached to the VP rather than to the NP.

Hirst's Semantic Enquiry Desk [Hir87] applies one of a set of decision algorithms to determine attachment of PPs, depending on the nature of the ambiguity. In the standard attachment problem, where there is a choice between attachment to the VP or the preceding NP, his algorithm is based on the principles of presupposition, plausibility, and verb expectations. The algorithm looks first at referential success, a technique first used by Winograd [Win70] and later termed the Principle of Referential Success by Crain and Steedman [CS85]. This states that

> "if there is a reading which succeeds in referring to an entity already established in the hearer's mental model of the domain of discourse, then it is favoured over one that does not".

Essentially this means that if the attachment of a PP to an NP describes something that we know exists, then the attachment is probably correct. If, for example, we have the sentence *The policeman killed the man with the chainsaw*, and we know that there is a man holding a chainsaw, then it is very likely that *with the chainsaw* attaches to *the man* and not to *kill*.

If referential success fails to produce an attachment, plausibility is then examined. Returning to our sentence *She stabbed the man with a beard*, it is not very plausible that a beard was used to stab the man, since beards are not normally stabbing instruments. If we do not know of the existence of a bearded man, the referential success principle is of no help, but it is more plausible that *beard* attaches to *the man* than to *stabbed*. The Semantic Enquiry Desk uses the Frail frame language to make decisions about plausibility, but most other systems will be able to make judgements using whatever form of semantic knowledge base they incorporate. If attachment to the VP is implausible, attachment to the NP is examined for plausibility. If this too is implausible, then the verb expectations are considered. The system needs to incorporate information about what constituents are expected to follow each verb, for example in the form of case frames or slot restrictions. Some verbs demand or prohibit the presence of certain cases, although in English this is only true of a very limited number of verbs, such as *break*, which requires the presence of at least one of the Agent, Patient or Instrument cases. The system simply needs some kind of data about verb preferences, e.g. [CFJ+84]. If the verb expects some case that could be filled by the PP, then the PP is attached to the verb, otherwise it is attached to the NP. If this too fails, then the final preference is for VP-attachment.

3.2 Statistical Approaches

3.2.1 Lexical Preferences

Hindle and Rooth [HR93] propose a solution to structural ambiguity using lexical preferences rather than syntactic structures. Whittemore et al [WFB90] also favoured this approach, but were reduced to judging the preferences manually, since they could find no means of uncovering the necessary information about lexical preferences automatically. Hindle and Rooth's solution to this problem was to use a corpus-based approach, involving the co-occurrence of verbs and nouns with prepositions. For example, if the preposition *with* frequently follows the sequence *hit* + NP, this would indicate that there is a lexical association between the verb *hit* and the preposition *with*. Using the typical association rates between the noun or verb and the preposition, in the form of a co-occurrence table, they estimate whether noun or verb attachment is more likely in each case of ambiguity. However, they point out that one of the drawbacks of this kind of corpus-based approach is that it is difficult to estimate probabilities accurately when the frequency of occurrence is small.

Hindle and Rooth's experiments were thus limited to judging attachment preferences based on a verb-noun-preposition triple, in a domain-independent environment. They did not take into account the object of the preposition, which may however be significant in deciding prepositional attachment.

She stabbed the man with a beard.

She stabbed the man with a carving knife.

The first example is unambiguous to a human, since we know that beards are not normally stabbing instruments, and thus the prepositional phrase is assumed to attach to the noun and not to the verb. The second example could still be ambiguous, but the most plausible interpretation would be that the carving knife was used to stab the man, since a carving knife is a typical stabbing instrument. An interpretation whereby the prepositional phrase is attached to the noun is still feasible: for example, she could have stabbed the man in self-defence because he was holding a carving knife in his hand and threatening her with it. However, it can be seen that the preceding noun is also important for deciding attachment.

She stabbed her husband with a carving knife.

We would not consider this sentence ambiguous, because of the implausibility of a constituent *her husband with a carving knife.* So we would analyse the sentence as a case of VP-attachment, where the carving knife is the instrument used for stabbing.

She stabbed her husband with a beard.

We would normally consider this sentence extremely weird. Two interpretations are possible, although neither is very likely. If the prepositional phrase is attached to the noun, we would have to assume that the woman had more than one husband, and that it was the husband who had a beard, rather than the husband without a beard, whom she stabbed. If the prepositional phrase is attached to the verb, we must assume that the beard was used as the stabbing instrument, and that it was therefore either a very unusual beard in the normal sense of the word, or that the word beard can also refer to some type of sharp pointed instrument. The fact that the object of the preposition was ignored did not appear to disrupt Hindle and Rooth's experiment, however, since the predictions of the human judges were, surprisingly, relatively accurate.

3.2.2 Case Frames

Weischedel et al [WAB$^+$91] performed a series of experiments to test the effectiveness of probabilistic models in learning case frame information about unknown verbs from a corpus of text, using domain-specific knowledge. They measured this effectiveness by testing how well those case frames can predict prepositional attachment correctly. They found that using semantic probabilities alone produced the poorest performance, using syntactic constraints (closest attachment) was more effective, and that a combination of the two brought about a significant increase in performance. However, they do point out that the reduction in error rate may not be definitive, due to factors such as the size of the corpus used and the incompleteness of the semantic information. Weischedel et al's experiment is slightly more realistic than Hindle and Rooth's, in that they considered all possible attachment points for the prepositional phrases, which could be three or more, whereas Hindle and Rooth considered only two possible sites in the limited context VP - NP - PP. Weischedel et al also included the object of the preposition in their model, whilst Hindle and Rooth ignored it.

3.3 Effectiveness of Systems

Rule-based systems can be quite effective, but it is difficult for them to cope with exceptions or irregularities, which means that there will almost always be a certain tradeoff between precision and recall. The Semantic Enquiry Desk does not always produce the correct analysis for ambiguous sentences, in spite of its fairly comprehensive algorithm. Hirst [Hir87] suggests that the main reason for this stems from the fact that people have a far greater store of world knowledge and a better ability to use it. A successful probability-based system needs a large training corpus on which to base its frequency tables, as Hindle and Rooth [HR93] and Weischedel et al [WMS+93] discovered. The problem is that there may be examples that occur very rarely in English, however large the corpus, which can lead to erroneous statistics.

As we have seen, rule-based and stochastic approaches are currently the most popular methods of machine disambiguation. Although they do not use the same techniques as humans, they have a reasonably high success rate. However, use of contextual information could be invaluable in improving this further. There is still debate among psychologists as to the exact role that context plays in human sentence processing, but there is no doubt that it is a significant factor.

Deep forms of knowledge such as pragmatics are also essential to human understanding. Although it is difficult to envisage how the enormous volumes of information needed would be effectively stored and retrieved, there are possibilities for this type of approach within a restricted domain, as machine translation systems such as METEO [Cha76] have shown.

Chapter 4

Linguistic Studies of *With*

This chapter considers some of the linguistic studies that have been carried out on prepositions, and in particular on *with*. Most early work concentrated on the syntactic environment in which prepositions occur, whilst in recent years more attention has been paid to semantics. We first consider how prepositions can have different meanings according to the context, and see that the choice of preposition is often determined not by its meaning but by the constituent to which it attaches. For example, some verbs require specific prepositions to follow them. Focusing on the preposition *with*, we then examine some of the ways in which the different uses of this preposition have been categorised, according to syntactic and semantic distinctions.

4.1 Meanings of Prepositions

Most common prepositions have meanings which can easily be understood, because they relate two entities in a cleary definable way. For example, we can say

The blue box is <u>on</u> the yellow box.

The chair is <u>behind</u> the table.

We had coffee after the meal.

However, prepositions can also be used in other ways where the meaning of the preposition itself is not so clearly definable, e.g.

Bill stood outside the cinema for an hour.

What's the matter with you?

The chairman presented the winner with his award.

It is particularly hard to specify precise meanings for the preposition *with*, because it can be used in such a wide variety of contexts with different, unrelated meanings. The following examples illustrate just a few of the different ways in which the same sentence can be modified by a PP containing *with*.

Comitative-Agentive: *John drove the car with his driving instructor.*

Manner: *John drove the car with no concern for other drivers.*

Attributive: *John drove the car with three wheels.*

Circumstance: *John drove the car with all its faults.*

Instrument: *John drove the car with his feet.*

With can even be used to denote completely contrasting meanings.

Support: *Mary agreed with her sister.*

Opposition: *Jim fought with his brother all the time.*

It is perhaps more appropriate to say that the meaning of *with* is dependent on the context and on the words that it relates [Fri40][Hir87]. Charles Fries [Fri40] studied the occurrence and distribution of prepositions in English, and found that there were 9 common prepositions which accounted for 92.6% of all occurrences of prepositions in Standard English: *at, by , for, from, in, of, on, to* and *with.* He claimed that it is precisely because these prepositions have so many different meanings that they occur so frequently, the average number of senses recorded in the Oxford English Dictionary for each preposition being 36.5. The Oxford English Dictionary (OED) lists 40 uses of *with* as a preposition, although many of these are now obsolete, whilst the Concise Oxford Dictionary (COD) lists 14 uses, all of which are current.

4.2 What determines the choice of preposition?

If prepositions do not always have distinctive meanings, then something else must determine not only the meaning but also the choice of preposition in a given context. Barkman [Bar66] studied the syntactic distribution of English prepositions, in an attempt to establish an unambiguous class of prepositions and to define the nature of a preposition according to its syntactic behaviour. He found that, with the exception of *less* and *unless*, which he assigned to a special category, all the items under consideration as prepositions could occur in the following environments.

P + N : *With a democratic administration, ...*

V + P + N : *... voted with Roberts, ...*

N + P + N : *... dispute with the county school board.*

Adj + P + N : *... complete with ivory tusks, ...*

Adv + P + N : *... effectively with two problems.*

and/or + P + N : *... and with Joseph T. Sharkey, ...*

Given that syntax is instrumental in distinguishing the class of prepositions as a whole, it is unlikely to be a real factor in determining the choice of preposition, since many prepositions can occur in the same types of environment.

Charles Kreidler [Kre66] looked instead at the semantic environment as a basis for distinguishing prepositions. He examined 200 verbs in the context V (+ P) + N + P + N, and arranged them into 24 groups according to the prepositions that could occur in this context. Five of these groups contained the preposition *with*:

V with somebody about something, e.g. *agree, quarrel*

V with somebody for something, e.g. *bargain, compete*

V with somebody in something, e.g. *alternate, cooperate*

V somebody with something, e.g. *assist, provide*

V something with somebody, e.g. *discuss, share*

Although the same preposition can occur in different syntactic environments, this may be limited by the choice of verb.

John assisted Bill with his homework.

**John assisted with Bill in his homework.*

**John assisted his homework with Bill.*

Kreidler's categorisation aims to show how in certain cases the preceding verb can determine the choice of preposition in a given syntactic context.

The lecturer agreed with the student that she was right.

**The lecturer agreed to the student that she was right.*

**The lecturer conceded with the student that she was right.*

The lecturer conceded to the student that she was right.

The choice of preposition may also be determined by a preceding noun or adjective.

John's concern for his sister.

**John's concern with his sister.*

**John was pleased for his results.*

John was pleased with his results.

The verb, noun or adjective does not always uniquely determine the choice of preposition, as can be seen in Kreidler's classification where some verbs fall into several categories. Even in exactly the same context, several prepositions may be equally valid.

John was pleased with his results.
John was pleased at his results.
John was pleased about his results.
John was pleased by his results.

4.3 Syntactic Distinctions

Although we have seen that it is largely semantics that governs the choice of preposition, it is possible that syntax does play a role in determining the meaning of a particular preposition. According to Fries [Fri40], there are 5 main types of relation occurring between a PP and the entity it relates to:

– The PP modifies the preceding NP, e.g.

Sally hates men with beards.

– The PP modifies the preceding V, e.g.

She attacked her husband with a pitchfork.

– The PP modifies the preceding Adj, e.g.

The swimmers were blue with cold.

– The PP modifies the subject of the verb *be*, e.g.

The suspect was with his wife on the night of the murder.

– The PP is some kind of object of the verb, e.g.

The teacher supplied the class with paper.

A distinction can also be made in terms of the syntax of the PP itself.

– P + NP, e.g. *My father built the house with the yellow roof.*

– P + pronoun, e.g. *Lucy made her exam paper into a plane, and I did the same with mine.*

– the PP is a clause, e.g. *With the rain falling so heavily, he knew he would get wet.*

It can be seen, however, that semantic and syntactic distinctions are not totally unrelated. For example, the meaning of *with* tends to differ according to the constituent it is modifying, although there may also be variations of meaning within the same class of relation.

4.4 Semantic Distinctions

In his scheme for the tagging of the Susanne Corpus [Sam95], Sampson distinguishes between 10 categories of adjunct on semantic grounds, noting that there is a distinct correlation between semantic and syntactic behaviour. The categories are based on a regrouping of the subcategories found in [Q+85], and are as follows:

Manner / Degree: *Ordinary Carey Williams, armed with a pistol, stood by at the polls to insure order.*

Modality: *Hartsfield has been mayor of Atlanta, with exception of one brief interlude, since 1937.*

Contingency: *With all his musical activities, did he have the time ...?*

Respect: *She helped him a little with his book.*

Comitative: *All Dallas members voted with Roberts, except Rep. Bill*

Jones, who was absent.

Absolute: *Congress starts another week tomorrow, with sharply contrasting forecasts for the two chambers.*

Place: *He lay on his bed.*

Direction: *She walked to the bus stop.*

Benefactive: *I have found a place for the magnolia tree.*

Time: *She drove to Chicago on Sunday.*

The categories are designed to encompass all types of adjunct, and thus are fairly broad in their coverage. Since they do not refer specifically to *with*, but to prepositions in general, the categorisation may not be particularly suited to this preposition, which can belong to the first 6 categories. For example, Sampson does not propose an instrumental category, but instead groups all forms of instrumental under the manner category. This is quite justifiable on the grounds that the distinction between the two is sometimes unclear, but it is a somewhat unusual decision.

Sampson also argues that *with* and *of* should be distinguished from other prepositions because of their special properties. He claims that *with* is different in that it forms part of a particular syntactic construction which he names the "W-clause". This is a non-finite clause introduced by *with*, *without*, or *what with*, and has two uses. Firstly it can act as a post-nominal modifier, with a possessive meaning.

The dead body with blood all over it was the most horrible thing he had ever seen.

Alternatively, it can act as a sentence adverbial, attaching to the whole main clause, where *with XY* has the interpretation *since X is Y*.

With the door wide open, there was a draught.

With his helmet on, he felt safe.

Since Sampson's categorisation scheme for adjuncts is semantically based, it might be considered strange that he makes this syntax-based distinction. However, there is a difference in meaning between *with* used in a prepositional phrase and *with* used in a W-clause. The clausal use generally relates to a temporary and clearly delimited situation, whilst the prepositional use relates to a more permanent attribute or characteristic.

the book with dust all over it

the book with the red cover

**the book with dust*

**the book with the red cover all over it*

However, it is debatable as to whether this special use of *with* is enough to merit it being placed in a separate category.

Most classification schemes seem to distinguish between the two different meanings of *with* within the W-clause, but do not distinguish the W-clause usage as being separate from the prepositional usage. The usage as a postnominal modifier, signifying *having*, is grouped in the same category as the ordinary possessive use of *with*, although both syntactically and semantically they behave slightly differently, as demonstrated above. The usage of *with* as a sentence adverbial is placed in a separate category, and is given interpretations such as "accompanying circumstances" [Swa80], or "taking into consideration" [Ale90].

Classification schemes aimed specifically at *with* tend to be more consistent in type, varying only in the broadness of their coverage. Most take into account only the more common meanings, and either ignore more specific meanings or include them in other groups. For example, Comitative is not usually subdivided into separate types, nor are there special categories for meanings such as Responsibility or Cause.

In their Comprehensive Grammar of the English Language, Quirke et al. [Q^+85] distinguish 8 meanings of *with.*

Manner: *We were received with the utmost courtesy.*

Means & Instrument: *Someone had broken the window with a stone.*

Accompaniment: *Jack, with several of his friends, was drinking till after 2 in the morning.*

Support & Opposition: *Remember that every one of us is with you.*

Having: *an industrialist with a house on the Costa Brava*

Ingredient: *This cake is made with lots of eggs.*

Space: *I left the keys with my wallet.*

Pervasive: *The garden was buzzing with bees.*

In his study of the Instrumental Case, Nilsen [Nil73] proposes 6 basic semantic types, but does allow for subcategories within Comitative and Instrumental.

Manner: *She sang with unexpected enthusiasm.*

Instrumental: *He hit the ball with a bat.*

Comitative: *He went with John.*

Possessive: *He saw a monster with six heads.*

Objective: *He stocked the stream with fish.*

Locative: *He remained with John.*

Lambert [Lam69] also proposes 6 types, but replaces Objective with Material, whilst many English grammars, e.g. [Swa80], [TM86], [Ale90], propose similar but nameless categories.

The linguistic studies discussed have focused either on the class of prepositions as a whole, or on a particular aspect of prepositions. This means that the observations made have been fairly general. Although some of them have devoted particular attention to *with*, there have been few specific studies devoted to it.

Since it is one of the most common prepositions with a wide variety of meanings, and is so often at the root of attachment ambiguities, we feel that it is worthy of closer investigation. The next chapter will look in more detail at the semantics of this preposition and propose a new classification scheme accordingly.

Chapter 5

Semantic Analysis of *With*

We claimed earlier that the meaning of certain prepositions such as *with* is dependent on the context and on the words related by the preposition, rather than on any intrinsic properties of the preposition itself. In order to study this further, we distinguished between the different ways in which *with* can be used, in terms of a set of categories. Although categories for *with* have been defined before, e.g. [Sam95], [Q^+85], [Nil73], we propose a more comprehensive set, based on semantic criteria.

5.1 Semantic Categories

It is difficult to devise a set of categories that will allow us to differentiate between the meanings of the word and its usages, but which makes useful distinctions and is manageable in size. We need a balance between the OED's 40 meanings, which are perhaps somewhat arbitrary in their distinctions, and the 5 or 6 distinctions made in English Grammars such as [Ale90], [Swa80], [TM86], which either cite only the most common uses, or else group very different ones in all-encompassing categories. Our final set contains 18 categories.

5.1.1 Instrument

The instrumental use is probably the one that is most often associated with *with*, although other prepositions such as *by* may also be instrumentals. The preposition can usually be substituted by alternatives such as *using* or *by means of*. The following three sentences are all equivalent.

He killed his wicked stepmother with an axe.

He killed his wicked stepmother using an axe.

He killed his wicked stepmother by means of an axe.

The object of the preposition is typically some kind of tool or instrument used to perform the action, as in the previous example where the axe is used to carry out the killing.

It is a debatable issue whether various other types of object should be classed as instrumentals, or whether they should be accorded a separate category. These include nouns denoting means, body parts, natural forces, materials, etc.

Means

Noun phrases denoting means are usually classified along with instrumentals. Whereas a tool such as an axe is a physical instrument, noun phrases denoting means are figurative instruments which are usually, but not necessarily, abstract.

Angela lost weight with a low-fat diet.

In the above sentence, *with* could be replaced by *by means of* but not by *using*, which generally requires a physical tool.

Natural forces

Natural forces such as wind, fire, earthquakes, floods and storms are sometimes considered to be distinct from instruments. Semantically they form a subclass of their own, and do not always behave in the same way as ordinary instrumentals. It is often said that natural forces are not really instrumentals but agents, since they are the cause of the action. They also occur much more frequently as the object of *by* than as the object of *with*.

John's house was destroyed by an earthquake.

**John's house was destroyed with an earthquake.*

However, there are circumstances in which they can occur with *with*.

John was baptised with fire.

There is also a special sense in which a natural force can occur as part of a prepositional phrase in an active sentence, where it is not the agent.

God destroyed the village with an earthquake.

The gods vented their anger on the mariners with a terrible storm.

In the circumstances where they do occur as part of a prepositional phrase following *with*, there is no real difference between their behaviour and that of other instruments, and in fact they could be seen as the instruments of God.

Body parts

Body parts are another semantic group which are sometimes regarded as distinct from instruments. However, sometimes a body part can be used in an identical way to an instrument.

The burglar smashed the window with his fist.

The burglar smashed the window with a hammer.

The burglar used his fist to smash the window.

The burglar used a hammer to smash the window.

Sometimes it is less obvious whether the PP is really a manner adverbial or an instrumental.

The boy kicked the ball with his left foot.

It can again be paraphrased in the same way as instrumentals:

The boy used his left foot to kick the ball.

The boy kicked the ball by means of his left foot.

But the PP *with his left foot* could equally be said to represent the way in which the boy kicked the ball. In fact, it could be said that all instrumentals are a subset of Manner, since they all refer to the way in which the action is performed. The difference between the two lies in the use of another object to perform the action, so we should really classify the above example as an instrumental rather than as manner.

5.1.2 Ingredient

Ingredients and materials are sometimes classified along with instrumentals, but we have chosen to place them in a separate category. Although the object of the preposition is used in performing the action, it differs from an instrument in that it becomes a part of the final product, which is the object of the verb.

You always make a cake with eggs.

He built his house with bricks.

In the first sentence, the eggs are an ingredient used in making the cake, and are part of the final product. In the second sentence, bricks are the material out of which the house is built, and again are part of the final product. In semantic terms, we cannot say that the eggs are a literal or even figurative instrument used to make the cake, unlike a food processor or an egg whisk.

You always make a cake with a food processor.

In this sentence the food processor is an instrument used to assist the cake-making process, and is not a part of the final product.

With ingredients and materials, the preposition *with* can often be replaced by *using* or *out of*.

You always make a cake using eggs.

He built the house out of bricks.

The verb to which the PP attaches is generally a verb of creation, such as one connected with cookery or building. The object of *with* is almost always a physical object such as a foodstuff or material.

It is also possible that there is legitimate ambiguity between NP- and VP-attachment with this type of PP. The following sentence would normally be interpreted as a case of VP-attachment, as with the previous examples.

Peter made the cake with raisins.

But in the next example, readers would probably be garden pathed, interpreting the first PP as VP-attachment, and later changing it to NP-attachment after reaching the second PP.

Peter made the cake with raisins, and his brother made the one with cherries.

The reason for this is that the second PP can only be a case of NP-attachment, since *the one* cannot be a constituent in its own right, and must therefore form part of the larger constituent *the one with cherries*. If this PP is interpreted as attaching to the NP, then by analogy the preceding one will also be interpreted as attaching to the NP.

5.1.3 Pervasive

Similar to the Ingredient category is what Quirke calls Pervasive [Q^+85] and the COD calls Addition. It occurs with verbs of covering, filling, decorating etc., where the object of the preposition is added in some way to the object of the verb, resulting in some sort of pervasion.

> *The women decorated the church with flowers.*
> *Jack filled the bucket with water.*
> *The field was covered with snow.*

Unlike with ingredients and materials, the object of the preposition does not become a part of the final product, in fact there is no new final product as such created. The objet of the verb is an entity which existed before the action is performed, unlike with ingredients where the object of the verb is an entity created by the action of the verb.

> *Jack made the jelly with vodka.*
> *Jack filled the bucket with vodka.*
> *Jack filled the bucket with the hose.*

These three examples illustrate the difference between the three categories Ingredient, Pervasive and Instrument. In the first sentence, vodka is an ingredient because it becomes a part of the jelly. In the second sentence, vodka is pervasive, because it does not become a part of the bucket, but it is contained in it once the action has been carried out. In the third sentence, the hose is involved in the filling, but does not become a part of the bucket, nor is the bucket full of hose after the action (except in the slightly strange interpretation whereby the hose is coiled up in the bucket, rather than being used to fill the bucket with a liquid.)

5.1.4 Possession

The category Possession is another very frequent use of *with*, but this time the PP is attached to the preceding NP rather than to the VP. Since the most common source of attachment ambiguity is between VP- and NP-attachment, it is important that we can correctly identify the 2 cases clearly. The object of *with* is some kind of attribute or possession of the NP to which it attaches. It can indicate physical characteristics, hairstyles, clothes, objects carried, etc.

> *people with big noses*
>
> *the girl with the pink trainers*
>
> *the man with the orange briefcase*
>
> *the student with a hangover*

The PP acts as a postnominal modifier and has the function of a restrictive relative clause containing the verb *have*, used to identify the entity it modifies.

> *Sally hates men with beards.*
>
> *Sally hates men who have beards.*

There is some dispute over whether this category should include the possessive meaning of the W-clause described in Chapter 2. Although there is a semantic distinction between the two, we do not consider the difference in meaning significant enough to merit establishing a separate category.

5.1.5 Interaction

This category also involves NP-attachment rather than VP-attachment, and denotes some kind of interaction between two entities. The direct object is usually a noun such as *meeting*, *contact*, *talk*, etc.

The President held a meeting with his advisers to discuss security issues.

Do you still have any contact with your sister?

5.1.6 Comitative-Agentive

The Comitative categories denote an association with other people or things in the sense of *together with* or *along with*. We have subdivided the category into 4, based on Nilsen's categorisation [Nil73]: Comitative-Agentive, Comitative-Objective, Comitative-Spatial and Comitative-Temporal.

With Comitative-Agentive PPs, the object of the verb is accompanied by a person or animate being in performing the action.

Sally went to the party with her new boyfriend.

He was pleased to be with friends again.

The sentence can usually be rephrased by substituting a conjunctive phrase for the agent.

Sally and her new boyfriend went to the party.

5.1.7 Comitative-Objective

This is similar to Comitative-Agentive but involves an inanimate object instead of an animate one. It represents two objects that are in some way associated with each other, or an object associated with a person.

The perfume comes with the dress.

John came to the party with a bottle of wine.

5.1.8 Comitative-Spatial

With Comitative-Spatial, the two entities, which may be animate or inanimate, are associated by their spatial location.

I saw David with Mary in the supermarket.

I always keep my cheque book with my wallet.

Sometimes a distinction is made between entities that are related by a specific location (Comitative-Agentive) and ones that are related by their relative location (Proximity) [Nil73]. In the first example, David and Mary are together in a specific location, in the supermarket. In the second example, the cheque book and the wallet are always together but are not restricted to a particular location. We do not find any particular advantage to be gained by making this distinction, however, since the two types are so closely related.

5.1.9 Comitative-Temporal

With Comitative-Temporal, the two entities are related by time. The preposition has the meaning *at the same time as*.

Werewolves come out with a full moon.

The temperature increases with the pressure.

5.1.10 Manner

The Manner category is a fairly straightforward one, the only problem being the close dividing line between it and the Instrument category, as we have already seen. The PP is like an adverb of manner, describing the way in which the action is performed.

The mountaineer reached the summit with ease.

Pat listened to the music with tears in her eyes.

Angela unpacked her new computer with care.

The PP can often be replaced by an adverb, where one exists.

Angela unpacked her new computer carefully.

The object of the preposition is usually, but not always, an abstract noun.

5.1.11 Opposition

The Opposition category has the meaning *against*, occurring with verbs such as fighting and quarrelling, and playing against an opponent.

He fought with his brother at every possible opportunity.

Ann played chess with Simon.

It can also occur with nouns of opposition, such as game, fight, quarrel, etc.

Norman had constant disagreements with his supervisor.

5.1.12 Support

This has the meaning of being on someone's side, occurring with verbs such as *agree* and *concur*.

He sided with his party on the issue of abortion, even though he did not believe they were right.

He agreed with the Minister that something ought to be done immediately.

It can also occur with nouns of support, such as agreement.

He made an agreement with his brother that they would never fight over a woman.

5.1.13 Company

The object of the PP here is some sort of company, association or team to which a person belongs.

His ambition is to work with Rolls Royce.

Mark has been with the Greensox for four years.

Attachment may be to the VP, as in the examples above, or to the NP, as below.

John, the lawyer with Easton and Mita, is the best baseball player I know.

5.1.14 Objective

The Objective category [Nil73] is where the object of the PP is some kind of object of the verb. A handful of verbs are optionally ditransitive, such as *provide* and *supply*. Wherre they occur with two objects, the second of these occurs as part of a PP headed by *with*.

Newspaper samples provide us with amusing examples.

The recipient can be omitted, in which case the prepositional object becomes the direct object and a PP is no longer needed.

Newspaper samples provide amusing examples.

The recipient can also sometimes occur inside another PP following the object.

Newspaper samples provide interesting examples for us.

The object, however, cannot be omitted.

**Newspaper samples provide us.*

The construction has some similarity with the one occurring with mandatorily ditransitive verbs such as *give*.

John gave a present to his mother.

John gave his mother a present.

John gave a present.

**John gave his mother.*

Prepositional and phrasal verbs such as *put up with*, *go along with* and *start with* also have Objective PPs attached to them.

Their relationship started with friendship and soon progressed to romance.

Fred found he could no longer put up with his wife's snoring.

Nilsen included what we have classed as Pervasive within his Objective category [Nil73].

He stocked the stream with fish.

Although this may appear similar to our examples of optionally ditransitive verbs, there is a difference both in meaning and in the syntactic behaviour of verbs like this, which is why we prefer to classify them separately.

5.1.15 Responsibility

This category is similar to the Comitative one, except that the object of *with* is in some way responsible for the object of the verb, which may be animate or inanimate.

I leave the baby with my mother when I go shopping.

He left his thesis with his supervisor.

With could be replaced by a phrase such as *in the care of*.

> *I leave the baby in my mother's care when I go shopping.*

Unlike with the Comitative category, the two entities here are not performing the action together, nor are they necessarily linked to a relative or absolute location.

5.1.16 Cause

The Cause category denotes a situation where the object of the preposition is in some way the cause or reason triggering a state or action, and is almost always abstract. The PP may be attached to an adjective following the verb *be*, or to a VP.

> *The young girl shivered with cold.*
> *His cheeks were red with anger.*
> *The room echoed with the roar of the tyrant.*

5.1.17 Situation

A PP belonging to this category refers to a particular case or situation, *with* having the interpretation *in the case of*.

> *With her husband, she knew where she was, but other men were a different kettle of fish.*
> *You have to study hard with a law degree.*
> *With Action Aid, you know exactly where your money goes.*

It is not always clear whether the PP attaches to the VP or to the whole sentence, but it cannot attach to an NP.

5.1.18 Circumstances

This category is similar to Situation, except that it refers to a general situation or circumstance in which the action takes place, and always attaches to the whole sentence. It can be interpreted by *in spite of*, *because of* or *taking into consideration.*

With all his hard work, he nevertheless failed his degree.
With our limited budget, a holiday in France would be too expensive for us.
She should be able to find a job easily, with all her qualifications.

5.2 Semantic Features of the Nouns

For some of these categories, the direct object and/or prepositional object have typical semantic features, as we have discussed. These can help us identify the correct category of the preposition. For example, most abstract prepositional objects denote Manner. Table 5.1 depicts typical features associated with each category. It must be stressed that these features are not obligatory, and that where a feature exists for a direct object, this does not imply that the direct object itself is obligatory.

5.3 Frequency of Occurrence

As we have already mentioned, some of the categories occur much more frequently than others. The most commonly found ones are Manner, Instrument, Objective, and the Comitative group. Some, such as Company, Ingredient and Pervasive, occur extremely rarely or only with a small minority of verbs or nouns. For example, Pervasive PPs would never be found with verbs such as *go* or *have*,

Category	Attachment	Direct Object	Prepositional Object
Possession	NP	concrete	concrete
Interaction	NP	abstract	animate
Company	NP/VP	job title	company name
Support	NP/VP	any	animate
Opposition	NP/VP	any	animate
Com-Agentive	VP	any	animate
Com-Objective	VP	any	inanimate
Com-Spatial	VP	any	concrete
Com-Temporal	VP	any	any
Manner	VP	any	abstract
Instrument	VP	concrete	inanimate
Pervasive	VP	concrete	concrete
Ingredient	VP	concrete	concrete
Objective	VP	any	any
Responsibility	VP	concrete	animate
Cause	VP/Adj	–	abstract
Situation	VP/S	any	any
Circumstance	VP/S	any	any

Table 5.1: Semantic Features of Nouns

but only with verbs such as *fill*, *decorate* and *arm*. In the next chapter we shall investigate exactly which categories occur with the most common verbs in the corpus, and with what frequency.

Chapter 6

Analysis of the Brown Corpus

Most studies of prepositional attachment, e.g. [BR94] have concentrated solely on the sequence V–NP–PP. It was decided not to limit our study in this way, but to examine the sentences in their entirety. There were two reasons for this. Firstly, this is not the only sequence in which ambiguity can occur. The sequence V–NP–PP–PP is also a common cause of attachment ambiguity, where the second preposition can attach to either the object of the verb or to the object of the first preposition.

Put the pen in the box with the books.

In this example, one interpretation is that the pen which is located in the box must be put next to the books, i.e. attachment of *with the books* is to the verb. The alternative interpretation is there is a box containing books, into which the pen must be put, i.e. attachment of *with the books* is to *the box.*

There may also be other items such as adverbs inserted between the main constituents. All too frequently, sentences containing these are ignored, because extraction of the relevant sequence is carried out automatically.

The second reason is that the purpose was not only to examine cases of potential ambiguity, but to investigate the co-occurrence of specific verbs and nouns with different uses of the preposition. For example, it has been claimed that certain verbs preclude the attachment of certain prepositions, e.g. that *know* cannot flag any case of *with* [Hir87]. If this were true, then in any sentence containing the sequence V–NP–PP, where the verb is some form of *know* and the preposition is *with*, the preposition must always attach to the NP rather than to the VP. For example, the sentence *I knew the man with John* must be a case of NP-attachment. However, cases were found in the corpus where *with* could be attached to *know*.

> *I knew then, knew with a heart-stopping shock.*
> *And he knew in that moment, with a cold sinking of despair, a dying of old hopes, that Mae had spread some kind of word there among the neighbors.*

6.1 The Verb

Because so many different verbs can occur with *with*, we decided to examine the 50 most common verbs in the corpus, and see which uses of *with* occur with each verb. This should give us an idea of the extent to which the verb limits the possibilities of occurrence of different uses. Table 6.1 depicts the frequency of occurrence of each combination of verb and use.

6.1.1 Uses of *with*

The ways in which *with* was used differed widely according to the verb to which it was attached. Attachment to the NP is theoretically possible for all transitive verbs, because the actual verb is irrelevant. Use of the PP as a manner adverbial was also found to be possible for almost all verbs, although examples of this did not always appear in the corpus. Intuitively, this makes sense because anything

Verb																		
come					1	11	14		6		12	5					1	
go						29	1		7		5					1	2	
mean																		
run						3	2		1		2							
seem																	1	
think											2							
work			2			11				3	4	1						2
give									1		1	1						
put								2			2	2					2	
ask											1							
be	73	5	2		1		6	21				4			2	26	4	12
become											2						1	
begin	2		1						3	30	4							
bring	2					8	2		2		2							
call											2							
do	1	2	1			1				4	13	11	1				4	4
feel											4							
fill											2			74			1	
find	3																	
follow							1				4	3						
get						4												
have	14	27			1	12		3		1							2	
hear											2						1	
help	2									4								
hold	4			1		6					3	3						2
include	3							1										
keep								1									1	
know						1					3							
leave						2	4	4		12	4				4			
like	2																	
look											14	2					1	
make	2	2				6				12	14	20	5					
meet							1			15	2	3						
move						3					5	1						
need																	1	
pay												4						
play	1		2		2	12				7	9							
provide	1	1							2	36								
say											27	1						
see	9			1				2			5	3					1	
set								1				2		2				
show	1						1		1									
start		1					2			9	5	2						
take	4	2	1			13		4		2	3	2						
tell											5							
try											1	1						
turn						1	1				2					1		
use	1										6						1	3
want		1									1							
write	2										12	1					1	

Table 6.1: Distribution of Categories of *With* with 50 Verbs in the Brown Corpus

that can be done can be done in a certain manner.

With five of the verbs: *find, need, mean, seem* and *like*, there were no examples of VP-attachment at all. With the first four of these at least, VP-attachment is possible.

> *The children found the treasure with a metal detector.*
>
> *He needed money with great urgency.*
>
> *He meant what he said with all his heart.*
>
> *The days seemed longer with every passing week.*

For some verbs, such as *become*, there were no examples of VP-attachment except where the PP was a manner adverbial, e.g.

> *His rhythm became steadier, with the rigid monotony of an unskilled reader.*

With other verbs, such as *leave*, examples of VP-attachment exhibited a wide variety of uses of *with*.

> Manner – *He left with all the joyous spirit of a child going on holiday...*
>
> Comitative-Agentive – *Discovery that the gloves frequently left with the customers made the wise peddler of spiced sausage-meat come upon a compromise.*
>
> Comitative-Objective – *Richard left America with his clothes, his biwa and his image of Acala ...*
>
> Comitative-Spatial – *No need leave a note with it, either ....*
>
> Responsibility – *...and I can't leave the baby with Grandma because she isn't strong enough and the baby's too young to be put in a nursery.*
>
> Objective – *The departure of the Giants and the Dodgers to California left New York with only the Yankees.*

It is evident that the examples found in the corpus are not a definitive guide, since they can only provide positive evidence of what can occur, rather than negative information about what cannot occur. We cannot make the assumption that because something does not occur in the corpus, it can never occur, although we can make assumptions about the probability of something occurring. So it is impossible to make judgements about what cannot be flagged by a particular verb, but only about what can.

6.1.2 Predicting the Category of *With* from the Verb

Although certain verbs are likely to have fewer meanings of *with* attached to them, for the majority of verbs this does not help us to predict the correct one. Verbs such as *need* are unlikely to be an attachment site for *with*, but for most verbs, we need more information than the verb alone. For example, we might predict that the verb *pay* will be followed by an instrumental, but it is also found in the corpus followed by a possessive PP.

> *If an expensive and specialized piece of machinery is needed – such as a spray rig, a combine, or a binder – it is better to pay someone with a machine to do the work.*

6.1.3 Semantic Classes of Verbs

On dividing the verbs into semantic classes, it becomes apparent that the more abstract verbs, such as verbs of emotion and sensation, flag fewer meanings of *with* than physical verbs such as verbs of transfer and motion. There also appears to be a connection between the number of meanings of the verb and the number of meanings of *with* occurring with it. Using the COD, we compared the number of meanings of each and found a positive correlation (see Table 6.2).

Meanings of With	Av. Meanings of Verb
0	4.5
1	6.8
2	14.1
3	15.2
4	22.5
5	21.6

Table 6.2: Comparison of Meanings of the Verb and of *With*

Semantic Class of Verb	Av. Meanings of With	Av. Meanings of Verb
Emotion	1.2	6.2
Communication	1.5	11.4
Cause	3.0	14.0
Transfer	3.1	14.7
Location	3.6	17.4
Motion	3.7	19.9

Table 6.3: Comparison of Semantic Class and Meanings

However, it may not be intrinsically the number of meanings of a verb that is the deciding factor, but rather the semantics of the verb. The semantic class of the verb also appears to be related to the number of meanings it has. We divided the verbs into 6 basic semantic classes, and compared the average number of meanings of verbs in each class with the average number of uses of *with* occurring with verbs in that class. The results show that the more abstract verbs had fewer meanings and fewer uses of *with* than the more concrete verbs (see Table 6.3).

Although this is only very small-scale, it does show that there is at least some kind of correlation. The implications of this are not only that *with* is used in more ways with some types of verbs than others, but that if verbs are used in a very specific way, this may limit the number of uses of *with* occurring with it, which may in turn make it easier to disambiguate.

It appears to be much harder to predict which uses of *with* will appear with verbs from a particular semantic group, or vice versa. The most common uses of *with*, Manner and Instrumental, occur with all semantic groups, while the less frequent uses, such as Responsibility and Support, do not seem to be linked to any specific group or groups, but to individual verbs, on account of their very specific meanings. Because most of the verbs can occur in a variety of contexts and with different meanings, it is difficult to draw any specific conclusions.

6.1.4 Probability of Occurrence

So far, we have only considered the possibilities of each use of *with* occurring with each verb, which does not help us much to formulate any predictions. If we turn our attention instead to the relative frequencies of occurrence, a different picture emerges. Although many combinations are theoretically possible, in practice only a few are probable.

Returning to Table 6.1, it immediately becomes clear that certain verbs are highly strong predictors of attachment. For example, the verb *fill* occurs 74 times with the Pervasive category, and only 3 times with other categories. Similarly, *provide* occurs 36 times with Objective and 4 times with other categories, and *say* occurs 27 times out of 28 with Manner. The total number of instances of *with* found in combination with the 50 verbs was 955. Out of these, it was found possible to predict the correct use of *with* over a third of the time, without the need for any other information.

However, other verbs such as *come*, *take* and *make* seem to have a more random distribution of categories, which means that it is not so easy to predict the use of *with* occurring with them. Many of the verbs also have very few occurrences in total, which seems surprising in view of the fact that they are some of the most frequently occurring verbs in the corpus. The corpus as a whole contains 6255 occurrences of *with*, which means that less than a sixth of the total occurrences occurred with the 50 most frequent verbs.

6.2 The Direct Object

We might expect that knowledge about the direct object of the verb will help us with the attachment problem. If the PP attaches to the NP, then presumably it is some feature of the NP rather than the V which enables this to occur. Because PP-attachment acts like a restrictive relative, this means that any NP which has a unique referent has no need to be further restricted. So proper nouns and pronouns will not normally have a PP attached to them.

> **me with a beard*
>
> **John with a hammer*

There are a few exceptions to this, since proper nouns do not always have a unique referent. For example, we may want to distinguish between two different Johns, and therefore refer to them as something like *John with a beard* or *John with a PhD*.

6.2.1 Semantic Features of the Direct Object

On examining the data in the corpus, we did not find any semantic features on the direct object that consistently helped identify the correct attachment of *with*. Taking our list of verbs, we divided the direct objects occurring with them into the following 6 semantic groups: proper nouns, animate nouns, pronouns, places, abstract nouns and concrete objects. We then looked at the verbs and the uses of *with* occurring with each group. We found that there appeared to be no consistency in the combination of verb, object and use of *with* that would enable us to make any kind of prediction, apart from the fact that NP-attachment did not occur with proper nouns or pronouns. The data used was, admittedly, fairly small, but it seems unlikely from what we found that an increase in the size of the data would significantly alter the result.

The reason for the negative results could be described as follows. The combination of the verb and the direct object does not seem to be sufficient to predict attachment. Assuming that the direct object is capable of permitting attachment, let us consider the possibilities. If the PP attaches to the verb, then the direct object should have little relevance. If the PP attaches to the NP, then it may be the combination of the features of the direct object and the prepositional object which is important, rather than just the features of the direct object. Alternatively, it may be that the semantics of the direct object are irrelevant altogether, and that the combination of the verb and the prepositional object provides the necessary information. The most likely possibility, however, is that a combination of all three constituents is needed, since this would provide optimum information.

6.3 The Prepositional Object

Let us take the hypothesis that it is the prepositional object which can help us predict attachment. For example, in the following three sentences, we would predict the first PP to be an instrumental, the second to be a possessive, and the third to be a manner adverbial, basing our distinction solely on the semantic features of the prepositional object.

Lucy read the book with a magnifying glass.

Lucy read the book with the small print.

Lucy read the book with a smile.

6.3.1 Semantic Features of the Prepositional Object

From studying the nouns that occur as objects of *with* in the corpus, we can make some basic assumptions. Abstract nouns such as those depicting sound, emotion, sensation, etc. are likely to be associated with prepositional phrases of manner.

They moved towards the skiffs with shocking eagerness, pushing and shoving.

Pronouns and animate nouns are likely to be associated with Comitative-Agentive or Comitative-Spatial PPs.

Glendora dropped a chicken and a flurry of feathers, and went with her through the drizzle, to the storehouse.
He moved with his parents to West Boxford when he was sixteen years old.

Concrete nouns tend to be associated with the instrumental and possession cases.

This tool can also be made with a lathe.
When I drove a car with tail fins, I had plenty status at the wind-and-water oases.

However, this alone is not a very useful guide. For a start, it does not help us distinguish between instrument (VP-attachment) and possession (NP-attachment). Secondly, these assumptions only indicate general trends. For example, an abstract noun could still indicate possession.

From the beginning, it took a man with extraordinary qualities to succeed in Mexico.

Thirdly, it is clear that sometimes we need to know the verb as well as the prepositional object.

John shocked his wife with a bright yellow hat.
John saw his wife with a bright yellow hat.

In the first sentence, the PP is an instrumental - it was by means of the hat that John shocked his wife. In the second sentence, the PP is possessive - John's wife was wearing a bright yellow hat. We can distinguish between the two because a hat could be an instrument for shocking someone, but it cannot be an instrument for seeing someone (except in some very obscure sense). So in this case it is the combination of the verb and the prepositional object that enables us to make the distinction.

6.3.2 Predictions Using the Verb and Prepositional Object

In order to see whether this combination of information really is sufficient to predict attachment, we examined examples from the corpus.

Let us look first of all at some of the prepositional objects found in combination with the verb *see* in the corpus. Since *eye* is a concrete noun and it is a possible instrument used for seeing things, we might assume that it will be an instrumental.

> *I can't tell you now the distance in lightyears or what star this was, though you can see it with the naked eye at night in this hemisphere.*

Although there are none in the corpus, it would, however, be possible to find sentences where *eye* indicated possession.

> *He became increasingly nervous as he saw the man with the glass eye staring at him.*

Perspective is an abstract noun and cannot be an instrument for seeing, therefore we would expect it to indicate Manner, as indeed it does in the corpus.

> *Living in San Francisco I saw them seldom enough to see them with a perspective which was not distorted by exasperation or fatigue.*

Beard is a concrete noun, but it is not a possible instrument for seeing, and therefore we would probably assume that it indicates Possession.

> *He saw a pint-sized man with a graying spade beard and an unusually large head.*

It is extremely unlikely that a sentence would exist where *beard* in combination with *see* indicated Manner or Instrument.

Ax is a concrete noun which, when associated with other verbs, often indicates the Instrument category. It cannot, however, be an instrument used to see things, and it is unlikely to indicate Manner, so we would again presume it to indicate Possession.

> *Mrs. Borden would not have been alarmed if she saw Morse with an ax or hatchet in his hand.*

Since *crew* is an animate collective noun, we might expect it to be associated with the Comitative-Agentive or Comitative-Spatial category. In the example found in the corpus, it indicates the latter.

> *Over his shoulder he could see Prevot with the machine gun crew.*

It could also indicate Possession or Company, although there were no examples of this in the corpus.

> *Over his shoulder he could see the corporal with the machine gun crew giving instructions.*

The examples above illustrate correct predictions based on the semantic features of the verb and the noun. However, it is rare that we can guarantee a correct prediction. There are many instances in the corpus where the verb and the noun alone do not provide sufficient information to make the correct analysis. Some verbs are more prone to this than others: for example, *discuss* and *replace*, which can be used either ditransitively or transitively. If we have the sequence *replace X with Y*, Y may be either attached to *replace* or to X, as in the following two examples show.

> *We'll have to replace the cup with the broken handle.*
>
> *...plans were being made to replace the wooden dome of Bullfinch with the present much larger structure.*

In the first sentence, *the cup with the broken handle* forms one constituent, and it is this whole constituent that is to be replaced. In the second sentence, *the present much larger structure* is the proposed replacement for *the wooden dome*, and is therefore attached to the VP.

In the corpus, we found many examples of incorrect predictions, such as the following. As an abstract noun, we might expect *questioning* to indicate Manner when combined with the verb *hold*. But in the corpus it indicated Possession.

> *"Well", Heiser ventured, "why don't we hold an investigation with questioning and –".*

Program is not a noun which we would normally associate with the Instrument category, when in combination with the verb *make*. But in one example in the corpus where it might have been expected to indicate Possession, it indicated Instrument.

> *It seems that Khrushchev himself took a very special pride in having made a world-shaking contribution to Marxist doctrine with his draft program ...*

Gun would often be representative of an instrumental PP, but when associated with *meet*, this is unlikely. The next most obvious choice would therefore be Possession.

> *The important thing was that Burton had gone somewhere to meet a blackmailer with a gun in his pocket.*

In fact, this sentence is ambiguous since Possession is a possible interpretation (where it is the blackmailer who has the gun). However, the surrounding context tells us that Comitative-Objective is a more probable interpretation.

One of the reasons why information about the verb and prepositional object is not always sufficient to guarantee correct prediction is that most features apply to more than one category of *with*, as we saw in Table 5.1. So although an abstract prepositional object is likely to indicate Manner, it can also indicate Cause, Objective, Situation, Circumstance or Possession on occasion. Similarly, animate prepositional objects may indicate not only Comitative-Agentive, the most likely category, but also Support, Opposition, Interaction, Pervasive and Responsibility, to name but a few.

6.4 Combining all Three Constituents

If the verb and prepositional object do not provide sufficient information to make a correct prediction, it is possible that adding information about the direct object may improve our success rate. The problem here is similar to our previous one, however. Although the combination of features will apply to fewer categories, there are still no combinations of features which are guaranteed to be unique to a single category.

We examined some combinations of features likely to apply to a particular category, as shown in Table 5.1. The combination of an abstract direct object and an animate prepositional object seems likely to point to the category Interaction.

> *President Eisenhower held an 8:30am meeting with top military and foreign-policy advisers.*

But there were no verbs in our list for which Interaction was the only category having this combination of features. Other categories were Objective, Manner, Instrument, Circumstance, Situation, Support and Opposition. For example, both Interaction (as in the example above) and Situation (as in the example below) occur with the verb *hold* and the features abstract and animate for the two nouns respectively.

> *Rhythmic interest, melodic beauty and the expansiveness of the writing are all qualities which hold one's attention with the Gilels playing.*

One solution to the problem would be to make the semantic features of the nouns more explicit. However, this is not possible here due to the size of the corpus, because there would not be enough data to test this effectively. Indeed, this is the case with categories such as Company, which occur very infrequently. Theoretically it is the only category likely to have the feature "job title" for the direct object and "company name" for the prepositional object. But because these features occur so rarely in the corpus, we are unable to verify this.

6.5 Probability of Occurrence

Again we have been focusing on possibility rather than probability. There are very few hard and fast rules about which semantic features on the noun or verb can license or prohibit the attachment of different uses of *with*. So we can very rarely guarantee that a certain type of attachment will be correctly predicted. However, we can make sensible judgements about the most likely form of attachment based on what occurs most frequently. In terms of probability, if we have an outcome which is better than chance, then we have a degree of success.

In Table 6.1, we saw that some verbs are reasonably reliable predictors of attachment, whilst others have a more random distribution of possibilities. In the case of the latter, we need further information to distinguish between the uses of *with*, so we looked at the direct and prepositional objects occurring with these verbs.

6.5.1 Predicting Attachment from the Nouns

As before, we divided the direct objects into broad semantic classes, since each noun does not occur frequently enough to form any kind of judgement. We then looked to see whether, for each verb, we could make any kind of comparison between the use of *with* and the class of noun, again based on probability rather than possibility. We then did exactly the same for the prepositional objects, and compared first the use of *with* and the prepositional object on its own, and then the use of *with* and the combination of the two objects.

One of the problems which cropped up again was the insufficiency of data to make accurate predictions. However, for some verbs there were definite patterns in the semantic classes of the nouns and the type of attachment. With the verb *have*, most cases of Possession entailed both objects being concrete.

We now have certain squares with three corners on C.

For Interaction and Opposition the direct object was always abstract and the prepositional object was always animate or concrete.

Sam is having his troubles with Layton Hanover...

With Comitative-Agentive and Comitative-Spatial, the prepositional object was always animate.

I have it with me, right here.

Table 6.4 shows the frequencies of the semantic features associated with the direct and prepositional objects of *have*. Note that it does not depict how the features of the two objects are combined, as this appears to be somewhat more complicated.

Use	Direct Object		Prep. Object	
Possession	concrete	9	concrete	7
	abstract	5	abstract	3
	animate	2	animate	2
Com-Agentive	meal	7	animate	12
	abstract	3		
	—	2		

Table 6.4: Semantic Features of the Objects of *Have*

The patterns did vary somewhat according to the verb. With the verb *be*, cases of Possession were more likely to have an animate direct object and a concrete prepositional object.

Marcus Jastrow was a friendly, round-faced fellow with a little mustache.

It was also commonly found with an abstract prepositional object.

This was the big man with the proprietary air.

Table 6.5 shows the frequencies of the semantic features of the direct and prepositional objects associated with the verb *be*.

With the verb *come*, we can use the prepositional object to distinguish between the 3 most common uses of *with*: Manner, Comitative-Agentive and Comitative-Objective. Manner occurs mainly with abstract prepositional objects, Comitative-Agentive with animate ones, and Comitative-Objective with concrete ones. However, we do not have sufficient information to enable us to distinguish between

Use	Direct Object		Prep. Object	
Possession	animate	41	concrete	45
	concrete	27	abstract	30
	abstract	7	animate	4
Interaction	abstract	5	animate	4
			concrete	1
Com-Spatial	—	20	animate	20
	animate	1	concrete	1
Cause	—	26	abstract	21
			concrete	4
			animate	1

Table 6.5: Semantic Features of the Objects of *Be*

Use	Prep. Object	
Com-Agentive	animate	11
Com-Objective	concrete	11
	abstract	5
Com-Temporal	abstract	4
	concrete	1
	animate	1
Manner	abstract	8
	concrete	1
Instrument	abstract	4
	concrete	1

Table 6.6: Semantic Features of the Prepositional Object of *Come*

Manner and Instrument, since the prepositional object is most likely to be abstract in both cases. Table 6.6 shows the frequencies of the semantic features associated with the prepositional object of *come*.

With the verb *leave*, we can distinguish between the categories Objective, Comitative-Objective, and Responsibility using the semantic features of the prepositional object, as Table 6.7 shows. All Comitative-Objective prepositional objects were concrete, most Objective ones were abstract, and all Responsibility ones were animate.

With the verb *do*, we can use the features of the prepositional object to distinguish between the two most frequently occurring uses, Manner and Instrument.

Use	Direct Object		Prep. Object	
Manner	place	2	abstract	3
	—	2	concrete	1
Com-Objective	place	2	concrete	4
	concrete	1		
	—	1		
Com-Spatial	animate	3	animate	2
	concrete	1	concrete	2
Objective	animate	4	abstract	6
	—	4	concrete	2
	place	1	animate	1
Responsibility	animate	2	animate	4
	concrete	1		
	—	1		

Table 6.7: Semantic Features of the Objects of *Leave*

Use	Direct Object		Prep. Object	
Manner	—	7	abstract	10
	concrete	4	concrete	1
	abstract	2	animate	1
Instrument	abstract	9	concrete	8
	—	2	abstract	3

Table 6.8: Semantic Features of the Objects of *Do*

Manner occurs mainly with abstract prepositional objects, whilst Instrumental occurs mainly with concrete ones. Table 6.8 shows the frequencies of the semantic features of these prepositional objects.

6.6 Other Indicators of Attachment

Our analysis of the Brown corpus has shown that it is difficult to make absolute judgements about the type of attachment. Using the verb, the direct object or the prepositional object on their own often does not provide enough information to identify the correct attachment. A combination of two of these constituents provides a more reliable guide, but even the combination of all three does not guarantee success. But what we can do is to make predictions about what is most likely to occur, based on what usually occurs. For some verbs this is easier than others, but from just a few rules we can get over a third of the predictions right using just the verb alone, and this appears to increase considerably once we add information about the objects. If we want to be more accurate than this, we need still further information. This might be obtained by looking at a wider structure than just the verb and the noun, possibly beyond the sentence itself, or by using more detailed information about semantic or even pragmatic features. In the next chapter, we examine whether using information about the subject area and text type can improve our chances of success, by focusing on restricted subdomains.

Chapter 7

Comparison of Subdomains

7.1 Ambiguity and Corpus Type

The Brown corpus was originally compiled with the specific aim of including a variety of texts and styles. The subject matter covers politics, sport, religion, the arts, specialist skills and general fiction, whilst the discourse type includes reviews, news bulletins, manual instructions, scientific writing and narrative writing. A study of the corpus should therefore provide linguistic data which is representative of the English language as a whole. However, in order to make specific observations, it may be more useful to examine specific types of text and subject areas individually.

It might be predicted that both the subject matter and the discourse type would affect the level of ambiguity and the way in which *with* is used. For example, texts providing instructions to a user demand clear, precise information, and might therefore be less likely to contain ambiguities than general texts, since it is important that there are no misunderstandings. Texts belonging to a particular subdomain might also be less ambiguous than texts of a general nature, because we are able to make use of a restricted knowledge base and context to aid us with processing.

7.2 The UNIX Corpus

The on-line UNIX manual was chosen as an example of a technical text set in a restricted subdomain. We studied the way in which *with* is used in this corpus, and how examples of ambiguity differ from those found in the Brown corpus. The first observation to be made is that the text has various unusual characteristics. The sentence structure is generally short and simple, as is consistent with giving instructions, and verbs are almost always found in either the present tense or the imperative. The most striking feature is that punctuation marks and symbols often form part of the text itself.

> *fix need not begin with a '.' to be recognised.*
>
> *It can be used later by surrounding it with* { }.

Also, many words are used in unusual ways, particularly those that are UNIX commands, for example *make*, *bash*, *cat*, and *head*.

> *If the command has a non-zero exit status, make halts with an error.*

In this example, the usual interpretation would probably be where *make* is an imperative verb and *halts* is the direct object. But the correct interpretation in this case is where *make* is the subject and *halts* is the verb.

Verbs such as *run* may also be used transitively where normally this would be impossible or unlikely, and have a different meaning from their usual one.

> *View runs vi with the readonly flag set.*

7.2.1 Ambiguity in the UNIX Corpus

From our investigation, it appears that it is the subject matter rather than the text type which is relevant to ambiguity, and that factors such as sentence complexity and style of writing do not seem to have much significance. Somewhat surprisingly, technical texts actually appear to contain more ambiguities than general ones. The rationale behind this is that people may have difficulty comprehending texts such as cookery recipes, gardening tips, computer manuals, etc. if they are not completely familiar with the vocabulary. Thus it may be a lack of specialist knowledge that turns potential ambiguities into real ones. Presumably, this would be less of a problem for a machine, because it could be provided with the specific knowledge relevant to the subdomain.

Another factor is that instruction manuals and the like are designed to be used in conjunction with some form of visual aid. A UNIX manual is not designed for bedtime reading, but rather to be consulted whilst using one's computer. Similarly, a car manual is consulted in the presence of the car and its parts. For example, if we have a set of instructions *Attach the plimble to the fromp with the sprigget*, the sentence will be ambiguous unless we know what a fromp and a sprigget are. If we know that a *sprigget* is some kind of tool, then we can be fairly certain that attachment is to the VP. If, on the other hand, we know that there are different types of *fromps*, some of which have *spriggets* and some of which do not, then we can assume that attachment is to the NP. Or supposing that a *sprigget* could be either a tool or a part of a *fromp*, depending on the circumstance, then we might be able to make use of visual aid instead. By looking at the various parts, we should be able to see if one of the *fromps* has a *sprigget*, or if it is likely that we will need to use a *sprigget* to attach the *plimble*.

The difference between the use of words in the two corpora may seem trivial, but if we know in advance how words are likely to be used, this may be beneficial in understanding attachment. For example, it is unlikely that the verb *run* would occur in a general text with an instrumental PP attached to it.

> **The boy ran with his legs.*

But it would be quite possible in a text concerning computers to find an instrumental PP attached to *run*.

> *Run this utility program with SunView.*

For someone who is not a UNIX expert, many of the sentences in the UNIX corpus with a PP containing *with* could be construed as ambiguous. This mainly occurs because it is not clear whether the object of *with* is an instrument or is an attribute of the preceding NP. The following two examples are instances of the classic attachment problem.

> *In the example above, we are creating a bar-file with the c key letter.*

If we do not know what a *bar-file* or a *c key letter* are, we may not know whether a bar-file can contain a c key letter, or whether it is possible to use a c key letter to create a bar-file.

> *Individual users may set up a similar private equivalence list with the file .rhosts in their home directories.*

Similarly, in the above example we may not know whether the file .rhosts is used to set up the equivalence list, or whether it is an equivalence list with the file .rhosts that is being set up.

Sometimes the surrounding context can help us identify the correct attachment. The sequence *choosing the item with an arrow* is ambiguous, however familiar one may be with computer terminology. By looking at the whole sentence, we may be able to disambiguate correctly.

> *Choosing the item with an arrow (the "pull-right item") invokes the first item on the submenu.*

Both semantically and syntactically, use of the term *the pull-right item* suggests that *the item with an arrow* forms one constituent. Visual clues could also play a part here. Looking at the screen , we would be able to see if there is an item which has an arrow, or whether we would need to use an arrow to choose the item.

There are also cases where it is unclear which verb is the source of PP-attachment.

> *In the example above, we are ... telling bar it is reading from a pipe with the B option.*

Here it is not obvious whether the PP *with the B option* attaches to *reading* or to *telling*, i.e. whether bar is using the B option to read from the pipe, or whether we are using the B option to tell bar it is reading from the pipe. In fact a third interpretation is possible, whereby the PP attaches to the preceding NP *the pipe*, but this is more unlikely since pipes do not have options. Even if we were not familiar with computers, we would probably assume that an option is used to perform an action, hence it would be more likely to be an instrumental attaching to a VP than an attributive attaching to an NP.

7.2.2 Uses of *With*

We selected those verbs out of our original set of 50 which occurred in the UNIX corpus with *with*, and noted the uses of *with* which occurred with each verb. Table 7.1 depicts a comparison between the uses of *with* as found in the Brown and the UNIX corpora. Since the UNIX corpus is much smaller than the Brown corpus, it is unsurprising that there are more occurrences in the Brown corpus than in the UNIX corpus (although this is also due to the restricted vocabulary used in the specialised subdomain). It is interesting to note the uses that occur in the UNIX corpus but not in the Brown corpus, such as Instrument occurring with *run* and *include*.

Corpus																														
Verb																														
Pos																														
Int																														
C-Ag																														
C-Ob																														
C-Sp																														
C-Te																														
Obj																														
Comp																														
Supp																														
Opp																														
Man																														
Inst																														
Ing																														
Perv																														
Resp																														
Cau																														
Circ																														
Sit																														

Table 7.1: Comparison of the Brown and Unix Corpora

The main differences between the two corpora lie in the Instrument, Manner, Possession and Interaction categories. Out of these four categories, Instrument is the only one to have examples from the UNIX but not the Brown corpus. The other three categories have a significantly higher frequency of occurrence in the Brown corpus. With categories such as Pervasive, it is hard to make a judgement since the data in the UNIX corpus is insufficient. Although Pervasive occurs 74 times with *fill* in the Brown corpus, and only once in the UNIX corpus, this does not necessarily mean that Pervasive is unlikely to appear at all in the UNIX corpus, just that it is unlikely to occur with *fill*. It may well occur with other verbs which we have not investigated.

7.3 Subdomains of the Brown Corpus

Since it appears that *with* may be used differently in a specific subdomain, we decided to examine how it is used in 4 subdomains of the Brown corpus: sport, hobbies and skills, press reportage and adventure fiction. The three most common uses of *with* in the Brown corpus as a whole are also the most common uses in each subdomain, but the relative frequency of these uses varies according to the subdomain. Table 7.2 depicts the number of occurrences of each use of *with* as a percentage of the total occurrences of *with* in each subdomain.

The fact that the most common uses of *with* remain the same across different subdomains suggests that there is not a huge difference in usage patterns. The different distribution of these three uses could be purely due to chance, but there are underlying factors which support this in each case, as will be explained.

Use of *With*	Sport	Hobbies	Fiction	Reportage
Manner	36%	17%	25%	22%
Instrument	10%	28%	15%	11%
Objective	20%	15%	15%	37%

Table 7.2: Most frequent uses of *with* in each subdomain

7.3.1 Sports Reviews

This domain can be subdivided further into different types of sports. In general, Manner was the most common use of *with* found, closely followed by Objective.

Gymnastics

In stylistic sports such as gymnastics and bodybuilding, nost of the PPs are Manner adverbials, because it is how the action is performed that is of interest.

> *They saw completely masculine and obviously virile men performing with incredible grace.*

Golf

In sports such as golf, more instrumentals will be used, because an instrument is used to perform the action, and we want to know specifically what this instrument is (e.g. what type of golf club).

> *Player attacked with his four-wood and hit a shot that few who saw it will ever forget.*

Manner is also important because we want to know how the ball is being hit, e.g. stance, type of swing, force used, etc.

It is almost axiomatic that golfers who dominate the game of golf for any period of time attack their shots with a vehemence bordering on violence.

Baseball

In baseball, on the other hand, the circumstances surrounding the action are important. When a player steps up to bat or when a pitcher prepares to deliver the ball, it is important to know details such as how many outs there are and how many runners there are on base.

... the drive fizzled when John Edwards fouled out with men on second and third and two out.

Company is also commonly found in baseball, because details about which team someone is playing for are often mentioned.

Four years later he resigned to take a similar job with the Greenbay Packers.

7.3.2 Hobbies and Skills

Again, the subdomain of hobbies and skills can be divided into different subject areas.

Crafts

In pottery and DIY, instrumentals are frequently found because specific tools are used to perform the actions.

Dress all surfaces with a file, cleaning off all solder and drilling burrs.

Pervasive is also common because it occurs with verbs of covering, filling, decorating, etc.

Fill small hole in bowl with clay.

Cookery

Cookery recipes contain a lot of Comitative-Spatial, because the association of ingredients is important.

Serve the chili and kraut hot with the franks.

Again, Pervasive is common due to verbs such as spreading and sprinkling.

Brush frankfurter twists with about 1/2 cup melted butter.

Hunting

The subject area of hunting contains many intricate descriptions of the rifles used, and therefore Possession is the most frequent use of *with*.

The Model 99 ... is a light-weight, streamlined rim-fire auto with a tubular magazine that holds 18 Long Rifles.

7.3.3 Adventure Fiction

Since adventure fiction mainly comprises descriptions of people and events, it is unsurprising that a quarter of the PPs containing *with* belonged to the Manner category.

Clapping spurs to the bronc, he set off at a sharp canter with growing alarm.

Cause is also much more frequently found here than in many other subdomains. This is consistent with the fact that the text is largely descriptive.

His mouth was open, his neck corded with the strain of his screams.

However, because there is no specific subject area or vocabulary, we can only be very general about this subdomain.

7.3.4 Press Reportage

In this subdomain there is again no specific subject area, but the uses of *with* found seem to differ considerably from those found in the Adventure Fiction subdomain. The Objective use features strongly here, whereas in adventure fiction its use was much more limited.

> *A citation from Conservation Commissioner Salvatore A. Bontempo credits his supervision with a reduction in the number of forest fires in the state.*

Manner and Cause were also relatively infrequent here. This could be explained by the fact that the style of writing is much more concise, and details are not entered into. The bare facts are what is important, rather than the circumstances surrounding them or the way in which actions are carried out.

7.4 Disambiguation and the Subdomain

The question that arises from our study of various subdomains is whether knowledge of the subdomain helps us in any way to determine the correct meaning of *with* and its attachment. The answer is rather vague: sometimes. If the subdomain is a very specific one, such as the UNIX corpus, it may be of help, because we know that certain words may be used in a particular way, or will have meanings not normally associated with them, as with the verb *run* discussed earlier. Although this type of text may appear to be much more ambiguous, this may be resolved if we have specialist knowledge about the subdomain. Knowledge of the vocabulary of a particular sport or skill may also be beneficial in the same way.

Subdomains which are more general in subject matter but which have a specific text type, such as press reportage, tell us far less about the way in which *with* will be used. Although the frequency of each use may be different, it gives us little practical assistance in determining the correct attachment. So, although the type of text does in fact make a difference to how *with* is used, it is difficult to make use of this fact. If, on the other hand, the text contains a very specific vocabulary relative to the subdomain, we may be able to use this information to help us resolve ambiguities.

Chapter 8

Conclusion

8.1 Summary

This thesis has investigated the meanings of the preposition *with* and the ways in which it is used in a corpus of text. The main purpose behind this was to gain some insight into the problem of prepositional attachment and disambiguation.

Firstly we examined the methods used by humans to disambiguate sentences. Opinions differ on the exact form of processing mechanism used, but it is generally agreed that it must involve at least a syntactic component, a semantic component and the use of some form of contextual and/or pragmatic information.

We then focused on machine disambiguation, which has the same ultimate goal but tends to use a different means of achieving it. Most systems use mainly syntactic and some semantic information, but no deeper forms of knowledge. Both rule-based and stochastic approaches have been used, with reasonably accurate results, but it is possible that they could be further improved by incorporating techniques borrowed from human processing.

Our discussion of previous studies of prepositions showed that there appears to be no general consensus on how meanings should be categorised for the preposition *with*. The number and type of distinctions made seems to depend on the purpose for which the classification is intended.

Our own semantic analysis of *with* was based on a variety of previous classifications and from a study of its use in real language. The aim of this was to provide a basis for analysing examples of attachment in the Brown corpus. Our analysis of the corpus first involved a study of ambiguity. The discrepancy here between potential sources of ambiguity and sentences which are actually ambiguous to humans is vast. What will be ambiguous or incorrectly interpreted by a machine depends on the system's capabilities, but there is huge scope for improvement from the level of a naïve syntactic parser to the level of a human (the ultimate goal for a system). Most systems involve using information from some or all of the sequence V - NP - PP for disambiguation.

Our initial study of the verbs in the corpus showed that whilst some verbs occur with a wide range of uses of *with*, others occur with very few or even no uses, even though these verbs themselves may be very common in the corpus. This difference appears to be linked to the semantics of the verb. Verbs which display a high degree of lexical ambiguity are likely to occur with a number of different uses of *with* and are also likely to belong to a different semantic group from those which are less ambiguous and occur with fewer uses of *with*.

When we looked at the frequency of occurrence of verbs with uses of *with*, as opposed to just the possibility of occurrence, it became clear that certain verbs were very good predictors of attachment, occurring mostly with one particular use of *with*. We found that we could make correct predictions over a third of the time for the verbs we studied, using only this information. Other verbs occurred in equal proportions with a variety of uses, so we turned our attention to the direct and prepositional objects. Again, looking at possibility of occurrence, we found that a variety of types of noun could occur with each use of *with*. However, there was a strong possibility for most verbs that the semantics of the nouns would indicate a particular use of *with*. This tended to be specific for each verb, although some generalisations could be made. In a few cases, there was insufficient information to make a prediction.

We then considered whether language is used sufficiently differently in specific subdomains to make predictions any easier. It is plausible that knowledge about the text and vocabulary type might improve our chances of correct prediction. We may be able to make more accurate judgements about attachment if we know in advance how certain words are likely to be used. However, we found that this is only really useful in restricted domains, particularly where the language is very technical. Subdividing the Brown corpus into subdomains did not yield much benefit. There are two possible reasons for this. Firstly, even the subdomains are still contain fairly general language, with the exception of some of the texts concerning sport and hobbies. Secondly, the subdomains were mostly fairly small, and thus often there was insufficient data to provide clear results.

8.2 Further Work

On the whole, we can make fairly accurate predictions about the way in which *with* will be used in a particular situation, using only a few simple rules. Often, just the verb is sufficient, whilst at other times semantic information about the direct object and/or prepositional object may be needed as well. However, our study of the corpus only involved certain verbs, and there were a great many cases which were not examined. An approach which involved more verbs and possibly a larger corpus would yield further results. It is probable that this would enable more predictions to be made than our study could produce.

Our analysis of the corpus revealed better results than we had originally predicted. From studies of human processing techniques, it was envisaged that we would need much more information than just syntactic and low-level semantic knowledge to enable us to disambiguate the preposition. However, as we have discussed, there is a big distinction between possibility and probability. From the linguist's point of view, it is hard to formulate any rules which enable us to identify the correct use of the preposition. But from a statistical point of view, we can make quite good predictions. This highlights the distinction between competence and performance. Most of the uses of *with* can occur in a wide variety of situations, with different types of verb and objects. However, in reality many of these possibilities only occur very rarely, if at all, so in a statistical analysis they become virtually irrelevant.

We know that humans often need to make use of other forms of information to enable correct understanding. One such source is visual information. Unfortunately machines cannot make use of this in the same way that humans can, so other sources must be found. In restricted subdomains, specialised knowledge about the vocabulary of the field could make a difference. This is one area where humans are not so capable, because we usually only have specialised knowledge of a very limited number of fields. Another source of information might be intra-sentential context. In particular, the rest of the sentence following the PP is likely to be of most interest. Pragmatic information is harder to store and manipulate, because it is difficult to know exactly what is useful and how it could be used. This would again be easier to incorporate in specialised subdomains.

Few attempts have at present been made to design systems incorporating these kinds of deeper information, although some research has been carried out involving knowledge bases for machine translation systems, e.g. [CT87], [NRT86]. It may be that this is deemed unnecessary as statistical techniques improve in accuracy, but it is possible that it could add that vital edge.

Bibliography

[Ale90] L.G. Alexander. *English Grammar Practice.* Longman, 1990.

[AS88] G. Altmann and M. Steedman. Interaction with context during human sentence processing. *Cognition*, 30:191–238, 1988.

[Bar66] L.B. Barkman. *A Distributional Definition of English Prepositions.* PhD thesis, Georgetown University, Washington D.C., 1966.

[BPV91] R. Basili, M. Pazienza, and P. Velardi. Combining NLP and statistical techniques for lexical acquisition. In *Proceedings of the AAAI Fall Symposium on Probabilistic Approaches to Natural Language*, Cambridge, Massachusetts, 1991.

[BR94] E. Brill and P. Resnik. A rule-based approach to prepositional phrase attachment disambiguation. In *Proceedings of the Fifteenth International Conference on Computational Linguistics (COLING-1994)*, Kyoto, Japan, 1994.

[Bri93] E. Brill. *A Corpus based Approach to Language Learning.* PhD thesis, Dept. of Computer and Information Science, University of Pennsylvania, 1993.

[CFJ+84] C. Connine, F. Ferreira, C. Jones, C. Clifton, and L. Frazier. Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research*, 13(4):307–319, 1984.

[Cha76] J. Chandioux. METEO: un système opérationnel pour la traduction automatique des bulletins météorologiques destinés au grand public. *META*, 21:127–133, 1976.

[Cho65] N. Chomsky. *Aspects of the theory of syntax*. MIT Press, Cambridge, MA, 1965.

[CM93] K.W. Church and R.L. Mercer. Introduction to the special issue on computational linguistics using large corpora. *Computational Linguistics*, 19(1):1–24, 1993.

[Cra80] S. Crain. Pragmatic constraints on sentence comprehension. Unpublished PhD dissertation, University of California at Irvine, 1980.

[CS85] S. Crain and M. Steedman. On not being led up the garden path: The use of context by the psychological syntax processor. In D.R. Dowty, L.J. Kartunnen, and A. Zwicky, editors, *Natural language parsing: Psychological, computational and theoretical perspectives*, Studies in natural language processing, pages 320–258. Cambridge University Press, Cambridge, 1985.

[CT87] J.G. Carbonell and M. Tomita. Knowledge-based machine translation, the CMU approach. In S. Nirenburg, editor, *Machine Translation: theoretical and methodological issues*, Studies in Natural Language Processing, pages 68–89. Cambridge University Press, Cambridge, 1987.

[FBK81] M. Ford, J.W. Bresnan, and R.M. Kaplan. A competence-based theory of syntactic closure. In J.W. Bresnan, editor, *The mental representation of grammatical relations*, pages 727–796. MIT Press, Cambridge, MA, 1981.

[FC86] F. Ferreira and C. Clifton. The independence of syntactic processing. *Journal of Memory and Language*, 25:348–368, 1986.

[Fod78] J.D. Fodor. Parsing strategies and constraints on transformations. *Linguistic Inquiry*, 9:427–474, 1978.

[FR82] L. Frazier and K. Rayner. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14:178–210, 1982.

[Fra78] L. Frazier. *On Comprehending Sentences: Syntactic parsing strategies*. PhD thesis, University of Connecticut, Indiana University Linguistics Club, 1978.

[Fra85] L. Frazier. Syntactic complexity. In David R. Dowty, Lauri Juhani Kartunnen, and Arnold Zwicky, editors, *Natural language parsing: Psychological, computational and theoretical perspectives*, Studies in natural language processing, pages 129–189. Cambridge University Press, Cambridge, 1985.

[Fri40] C.C. Fries. *American English Grammar*. Appleton-Century-Crofts, New York, 1940.

[Hir87] G. Hirst. *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press, Cambridge, 1987.

[Hol84] V.A. Holmes. Parsing strategies and discourse context. *Journal of Psycholinguistic Research*, 13(3):237–257, 1984.

[HR93] D. Hindle and M. Rooth. Structural ambiguity and lexical relations. *Computational Linguistics*, 19(1):103–120, 1993.

[JL83] P.N. Johnson-Laird. *Mental Models.* Harvard University Press, Cambridge, MA, 1983.

[Kim73] J. Kimball. Seven principles of surface structure parsing in natural language. *Cognition*, 2:15–47, 1973.

[Kre66] C.W. Kreidler. English prepositions. *English Language Teaching*, XX(2):119–122, 1966.

[Lam69] D.M. Lambert. *The Semantic Syntax of Metaphor: A Case Grammar Analysis.* PhD thesis, University of Michigan, 1969. unpublished.

[Nil73] D.L.F Nilsen. *The Instrumental Case in English: Syntactic and Semantic Considerations.* Mouton, The Hague, 1973.

[NRT86] S. Nirenburg, V. Raskin, and A. Tucker. On knowledge-based amchine translation. In *Proceedings of COLING-86*, pages 627–632, 1986.

[Q+85] R. Quirk et al. *A Comprehensive Grammar of the English Language.* Longman, 1985.

[RCF83] K. Rayner, M. Carlson, and L. Frazier. The interaction of syntax and semantics during sentence processng: Eye movements in the analysis of semantically biased sentences. *Journal of verbal learning and verbal behaviour*, 22(3):358–374, 1983.

[Sam95] G. Sampson. *English for the Computer: The Susanne Corpus and Analytic Scheme.* Clarendon Press, Oxford, 1995.

[Swa80] M. Swan. *Practical English Usage.* Oxford University Press, Oxford, 1980.

[TM86] A.J. Thomson and A.V. Martinet. *A Practical English Grammar*. Oxford University Press, Oxford, 1986.

[WAB+91] R. Weischedel, D. Ayuso, R. Bobrow, S. Boisen, R. Ingria, and J. Palmucci. Partial parsing: a report of work in progress. In *Proceedings of the Fourth DARPA Speech and Natural Language Workshop*, 1991.

[WFB90] G. Whittemore, K. Ferrara, and H. Brunner. Empirical study of predictive powers of simple attachment schemes for post-modifier prepositional phrases. In *Proceedings, 28th Annual Meeting of the Association for Computational Linguistics*, pages 23–30, 1990.

[Win70] T.A. Winograd. *Procedures as a representation for data in a computer program for understanding natural language*. PhD thesis, Department of Mathematics, Massachusetts Institute of Technology, 1970.

[WMS+93] R. Weischedel, M. Meteer, R. Schwartz, L. Ramshaw, and J. Palmucci. Coping with ambiguity and unknown words through probabilistic models. *Computational Linguistics*, 19(2):359–382, 1993.