MOSES AND THE ARK:

EXPLORING SEMANTIC ILLUSIONS

Anke Caroline Büttner

Thesis submitted in partial fulfilment of the requirements of the award of Doctor of Philosophy awarded by Oxford Brookes University

October 2002

Declaration:

Moses and the Ark: Exploring semantic illusions

I declare that this thesis is my own work, and that none of the work referred to has been submitted in application for a degree in any other institute or university.

All quotations have been distinguished by quotation marks and the sources of information acknowledged.

Anke Caroline Büttner
October 2002

To Evi and Hans-Willi who believed in me,

and to Ian who put up with me.

Abecedarium

A is for Abstract, short and sweet,
B is for Books that I had to read,
C’s the Computer that’s crashed left and right,
D is for Data amassed with all might,
E is Experiment, what else could it be?
F are the Findings I wanted to see.
G is for Groups which each study required,
H is for Help often greatly desired,
I’s Inspiration which sometimes did lack,
J is for Judgements that kept me on track,
K is for Keen, which I was at the start,
L ‘s all the Lit’rature taken apart,
M is for Meaning which all was about,
N is for Nerve – I nearly ran out!
O for “Oh nos!” which occurred now and then,
P’s for Participants needed again,
Q is for Questions both asked and designed,
R are the Results by Stats undermined
S – see above, what more can I say?
T is the Thesis – nearly finished, hurray!
U’s Understanding endeavoured to reach,
V is the Viva, last bastion to breach.
W for Worries survived by and by,
X: eXplanations made up on the fly,
Y for the Years that it took to complete,
And Z for the sounds of much needed sleep.

Table of Contents

Acknowledgements

My thanks go to:

–Roger and Val, my supervisors

–Ian, my rock of sanity and tea maker extra-ordinary

–Stephie, a role-model and invaluable proofreader and advisor

–Philippa, my (almost official) chauffeur and work-provider

–Clemens, Almut and Maike for encouraging me like only siblings can

–Matthew and Jian-Bin, two more role-models

–David, a good listener

–Russ, Ruth and Rosanna, who know what it's like

–Stephen, a great line manager

–George, office-mate and computer-rescuer

–Russell for good conversations not related to research

–Paul for providing many a meal

–all who participated in my various studies

and above all to Wakefield who helped me slay the monster Statistics.

Abstract

In response to questions like “How many animals of each kind did Moses take on the Ark?”, a large number of people will say “two”, failing to notice the substituted name. This semantic illusion occurs even though people can be shown to possess the knowledge that Noah and not Moses sailed the Ark.

The aim of this thesis is to explore semantic illusions and to examine possible mechanisms underlying them. Semantic illusions are considered to be of theoretical relevance to theories of sentence comprehension, because of what they can reveal about the mechanisms underlying ordinary processing.

A series of studies examined current theories of semantic illusions. Evidence was produced that semantic illusion sentences are materially different from sentences used in ordinary discourse. The processing requirements of semantic illusion sentences were also explored, both in terms of participant expectations of the task at hand, and in terms of processing load.

In Part One, three experiments investigated the effects of the surface structure of semantic illusion sentences upon semantic illusion rate (Chapters 3 to 6), but only a comparison of question and statements revealed any significant effects, with questions leading to more semantic illusion responses. To explore the implications of this lack of effect, a rating scale study was designed to provide an overview of how semantic illusion sentences compare to sentences used in ordinary discourse: semantic illusion type sentences were found to differ significantly from other sentences along a number of salient dimensions.

In Part Two, three further experiments related semantic illusions to problem solving and examined the processing requirements of semantic illusions. Findings indicated that semantic illusions are subject to a kind of ‘functional fixedness’, which prevents thorough processing (Chapters 9 and 10). This may in part be explained by the load that semantic illusion sentences place on working memory, as was indicated by the results of two further experiments, which investigated the role that the different components of working memory play in semantic illusion processing (Chapter 11).

Part One:

Describing semantic illusions

Chapter 1:

Semantic Illusions: An Introduction

When asked “How many animals of each kind did Moses take on the Ark?” most people will respond with something like “Two”, as if the question was perfectly unexceptionable (Erickson and Mattson, 1981). Most people know that there were two animals of each kind on the Ark – they learnt it in Sunday school, for example – just as they learnt that Noah was the man who sailed the Ark and not Moses. What is interesting about this question, which often appears as a playground game, is not just that people do not notice the substituted term, but that it appears impossible to filter out errors of this kind in a reliable fashion. The existence of such a processing glitch calls into question many of the assumptions that have been made about sentence processing and discourse comprehension. Over the course of the research dealing with this phenomenon, it came to be called the *semantic illusion*. For the purpose of this thesis, a semantic illusion will be defined as follows:

> *A semantic illusion occurs when an individual interprets a sentence containing a substituted word of similar but distinct semantic content as if no substitution had been made, under circumstances where the individual can be shown to have the correct knowledge of the idea expressed in the sentence.*

Semantic illusions are of interest to theories of sentence processing in much the same way that visual illusions are to theories of visual perception. The processes responsible for the strange perceptual experiences caused by illusion stimuli are exactly the same processes with which people perceive the ordinary world and make sense of it (e.g. Jackendoff, 1993). Visual illusions allow researchers to contemplate visual processes in situations where these do not lead to a desired outcome, but instead to a mis-understanding of the data that is available to the eye. This suggests that the brain makes an active contribution to the interpretation of input, drawing on previous experiences and knowledge about the world to make sense of what is seen. It appears that there is a reflexive (as opposed to reflective) quality to perceptual processes (e.g. Fodor, 1983, 1985, 1986), and these 'mental reflexes' can also be shown to affect language perception. Proof-reading errors (e.g. Healy, 1981) are an example of this, and so is what Jackendoff (1993) refers to as a 'mental grammar' which allows people to make accurate judgements about the acceptability of an indefinitely large range of sentences.

By analogy to visual illusions, semantic illusions provide evidence of the processes used to deal with ordinary language. Drawing on previous knowledge and experiences, the brain makes an active contribution to the interpretation of the linguistic input and people make sense of what they hear or read. In the case of semantic illusions the same strategy is used but does not lead to the correct outcome. The sentence is processed and made sense of, even though the input does not technically make sense as such. Thus, unlike many visual illusions, the outcome of the glitch in processing that leads to a semantic illusion does not enter the reader/listener's consciousness. Instead, the substituted word is not detected as inappropriate in its context, but appears to fit in adequately.

This apparent failure to notice a word in the middle of a sentence, even though it appears to be processed in some superficial way, calls into question many assumptions about how human beings process language. It also has implications concerning the way in which memory for knowledge about the world is organised and the necessary flexibility that must be associated with such a store if it is to function efficiently. Semantic illusions also provide evidence that much of the cognitive processing that is done every day is subject to trade-offs between speed and accuracy, to allow for the processing of the most information for the least amount of effort. It seems that semantic illusions occur *because* people are used to extracting meaning from language stimuli, even if these are degraded.

Many different aspects of linguistic information affect ease of processing. For example, context, attention, familiarity and semantic/syntactic complexity all play an important part in comprehension and processing and so also in the occurrence of semantic illusions. It is the aim of this thesis to provide a detailed profile of the semantic illusion phenomenon and to explore what semantic illusions can reveal about normal language comprehension.

In this thesis, a number of the issues related to sentence comprehension will be touched upon and the implications of the knowledge gained about semantic illusions will be examined. The first part of the thesis establishes what is known about semantic illusions to date and then explores aspects of semantic illusions as a part of the language comprehension process, including investigations of the surface form of semantic illusion sentences and their relation to other sentences used in day-to-day uses of language. The second part deals with semantic illusions as a kind of problem solving paradigm, describing how the usual strategies of dealing with language information will lead to problems in the case of semantic illusions. As well as this problem solving approach to semantic illusions, the impact of processing limitations of the cognitive system is also examined in the second part of the thesis, relating semantic illusions to working memory research.

In Chapter 2, the literature on semantic illusion research and a few closely related phenomena is reviewed, describing semantic illusion paradigms and methods used, as well as the main findings. Different theories that have been proposed in order to account for semantic illusions are examined, including issues of reader/listener co-operation, sentence focus, semantic and phonological similarity between the target word and its substitution, and the effects of task demands. Theories that potentially account for semantic illusions include the suggestion that there is a failure at the level of encoding, that there is only a partial match between the stimulus sentence and the related memory representation, and that the global goodness-of-fit of the substitution in the context of the sentence is good enough not to trigger more thorough checks of coherence.

The aim of the first four experiments described in this thesis was to make explicit certain assumptions that had been made about semantic illusions, and in this context to replicate some of the findings described in Chapter 2.

Chapter 3 contains a description of an exploratory study that attempted to answer a few questions about semantic illusions that emerged from previous research, but had not been explicitly addressed: do semantic illusions solely occur for substituted proper nouns with a sentence, or can words from other form classes lead to the same effect? Is it possible to prime the substituted word selectively to draw a participant's attention to it and reduce the likelihood of the occurrence of semantic illusions? The effects of selectively priming either the target word or its substitution are examined for target sentences in which words from four different form classes (proper noun, noun, verb and adjective) have been substituted. It is hypothesised that priming the substitution will lead to a greater semantic illusion rate. The second hypothesis is that words from different form classes – which can be shown to have varying processing requirements – will be likely to lead to different semantic illusion rates.

In Chapter 4, Experiment 1 is described. It addresses another previously unanswered question arising from the literature review: do question-answering and statement-verification, the two tasks generally used in semantic illusion research lead to equal semantic illusion rates, or do questions such as that used in the “Moses” example above lead to more semantic illusion responses than corresponding statements-to-be-verified? The same target concept is presented as either a question or as a statement-to-be-verified, to allow a direct comparison between the two semantic illusion tasks to be made. The hypothesis is that question format will lead to more semantic illusion responses because of differences in task demands for questions and statements-to-be-verified.

Chapters 5 and 6 continue with the exploration of the effects of sentence structure upon semantic illusions. Sentence length and the position of the substituted word in the sentence are examined for English sentences in Chapter 5, in an attempt to replicate and bring together findings from a number of previous studies. Reder and Cleeremans (1990) and Reder and Kusbit (1991) demonstrated that the more items relevant to the ‘illusion answer’ that a question contains, the more likely a semantic illusion is to occur. Experiment 2, attempted to find out if this effect also applies to statements-to-be-verified. Apart from one study (van Jaarsveld, Dijkstra and Hermans, 1997), the position of the substituted word in a semantic illusion sentence has not been investigated systematically for semantic illusion sentences, hence Experiment 2 is also concerned with taking another look at the effects of word position on the semantic illusion rate in statements-to-be-verified.

In Chapter 6 the effects of position are investigated for German sentences, as German is a language less dependent on word order for semantic cues than English. Instead, word order in German is often used to indicate focus and stress information. Several of the previous semantic illusion studies have been carried out in languages other than English and were found to lead to similar results as English studies. But to date, semantic illusion research has not been carried out on a language which relies more heavily on inflections than on word order to convey semantic information. Experiment 3 is designed to test the hypothesis that word order will have a greater effect on German semantic illusions, as in German word order conveys different information than it does in English.

Chapter 7 describes how semantic illusion type sentences are perceived by people in comparison with samples from written discourse taken from real-life samples. It was hypothesised that semantic illusions might require a definite linguistic format in order to occur: semantic illusion sentences are generally lengthy and contain enough information to be responded to without needing the semantic contribution of the substituted word. They also must deal with information that is at least in part familiar to the reader/listener. So how do semantic illusion sentences fit in with normal discourse? A series of rating scales, a focus determination task and a categorisation task will be used to access information about how semantic illusion type sentences are perceived in relation to sentences taken from real-life sources of written discourse.

Chapter 8 is the first chapter of the second part of the thesis. The second part aims to examine the question of how semantic illusions are processed. In Chapter 8 various analogies between problem solving research and semantic illusion research are described in an attempt to enhance an understanding of the semantic illusion phenomenon. Partial matching – a mechanism likely to underlie not only semantic illusions and similar phenomena, but much of human information processing in general – is also discussed and related to the schema theory of memory (Bartlett, 1995[1932]) which provides a suggestion as to how knowledge might be stored in long-term memory and how this structure may contribute to semantic illusions. Evidence for the fact that comprehension of linguistic information requires an active input from the reader/listener is also discussed and related to findings about semantic illusions.

In Chapter 9, two experiments are described that attempt to tackle semantic illusions from a different angle to that used in previous studies, trying first to phrase the instructions differently, to affect participants' expectations about the task at hand, and then changing the way in which semantic illusions are presented, to rule out an 'automatic' component to sentence processing. In Experiment 4a, a simple attempt is made to eradicate the associations related to the words 'true' and 'false', which are thought to have strong connotations related to knowledge testing. It is hypothesised that presenting semantic illusion sentences in a slightly different task – asking participants to judge if each statement presented is 'natural' or 'unnatural' – will remove some of the potentially perceived requirement to perform well on the knowledge retrieval component of the task, allowing participants to concentrate more easily on the surface structure of the sentences.

In Experiment 4b, the task was to rebuild so-called 'sentence puzzles'. This manipulation was developed to eradicate the 'automatic' component of sentence processing, which makes it almost impossible for participants not to read a sentence for meaning before looking at its component parts. The device of presenting the sentences as fragments to be assembled should force participants to pay more attention to each individual word and semantic illusion rate might therefore be reduced.

Chapter 10 continues with the attempt to look at semantic illusions from an information processing perspective and examines the effect of the context provided by 'filler' statements upon semantic illusion rate in statement-to-be-verified tasks. Since no processing is ever done in isolation and much of it is not done logically or systematically, the hypothesis is that the nature of the filler statements-to-be-verified (whether they were generally truthful or not), will affect the way in which semantic illusions sentences are perceived and processed. A set of semantic illusion statements was presented in either of two filler statement contexts. The experimental hypothesis is that statements surrounded by many false fillers will be less likely to lead to semantic illusion responses than those surrounded by many true statements.

Chapter 11 relates semantic illusion research to working memory research and explores the involvement of the different components of working memory on semantic illusion processing. It was thought that from this manipulation more information could be gleaned about how semantic illusions are processed. A semantic illusion task was combined with concurrent tasks designed to affect the sub-components of working memory differentially. The concurrent tasks either affected the phonological loop by preventing rehearsal of linguistic inputs, or the central executive by placing a greater overall processing load on working memory. The pattern of interference between the semantic illusion task and the concurrent tasks should help to establish to what extent the processing of such sentences taxes the cognitive system. One possible explanation for semantic illusions is that the sentences from which they arise place heavy processing demands on working memory, so that the failure to detect substitutions is a direct consequence of capacity-maximising strategies, which allow the most efficient processing given a limited capacity to manipulate incoming information. The hypothesis was that semantic illusion rates will increase as concurrent demands on working memory become greater.

In Chapter 12 the conclusions from the thesis are presented and implications of the findings are discussed in relation to sentence processing research, to general theories of cognitive processing, and to applications of such research to the real world.

It is the aim of this thesis to explore the phenomenon of semantic illusions, to tie the findings from this research in with current theories of cognitive processing, and – if possible – to suggest a more complete theory of semantic illusions. It is hoped that the findings from this research will help to turn semantic illusions into a paradigm which might be used for future research into cognitive mechanisms underlying the storage and retrieval of knowledge and to language. Insights into the phenomenon may also suggest why errors occur in real life in comprehension of texts by students and by academic researchers (cf. Vicente and Brewer, 1993).

Chapter 2:

What is known about Semantic Illusions: A Literature Review

Even though it is likely that generations of school children have tricked each other with the Moses and the Ark question, it was not until relatively recently, that semantic illusions were first documented in scientific research. Erickson and Mattson first described semantic illusions in 1981, and it was not until 1987 that the next relevant paper was published (Baker and Wagner, 1987). Since then the question of semantic illusions has continued to intrigue a number of researchers and papers related to semantic illusions have been published from time to time. But the literature is rather fragmented, with groups of researchers preferring to concentrate on certain favourite topics. This chapter contains an overview of the research that has been done on semantic illusions and a few very closely related phenomena. In an attempt to bring together what is known about semantic illusions to date, eight potential explanations for the semantic illusion are presented and relevant findings are described. Existing positions are summarised and evaluated in view of the empirical evidence provided by semantic illusion research, and important key ideas are established.

2.1 The experimental paradigms most frequently used

As mentioned in Chapter 1, there are two basic tasks used in semantic illusion research: a question-answering task and a statement-verification task. In the question task, participants are asked to give short answers to general knowledge questions addressing familiar topics. In each target question an appropriate word is replaced by a semantically related, but distinct word rendering the question technically meaningless (e.g. “How many animals of each kind did Moses take on the Ark?”). Participants are generally instructed to give an answer to each correctly phrased question, and to reply with “wrong” or “can’t say” if they identify the problem with a target question. In the statement – verification task, participants are asked to verify a series of general knowledge statements (e.g. “Moses took two animals of each kind on the Ark”). This experimental set-up will be referred to as *statements-to-be-verified*. Often the semantic illusion task is followed by a knowledge check, so that the experimenter can ascertain that each participant has the necessary knowledge to experience a semantic illusion. This usually requires participants to respond to a series of questions which test the knowledge replaced in the target sentences (e.g. A participant would have to answer a question about who had sailed the Ark).

Semantic illusion researchers have presented their participants with a number of variations on these basic tasks, including spoken and written semantic illusions, as well as semantic illusions in different languages (see Appendix 1 for an overview of task variations used). The target materials are generally compiled on an intuitive basis and have to be adapted to a target audience (Brédart and Modolo, 1988). Many researchers piloted their materials and checked empirically to establish whether they fit the required criteria for the research at hand.

The following sections (2.2 to 2.9) present possible explanations for semantic illusions and evaluate these in the light of semantic illusion research to date.

2.2 Semantic illusions occur because the listener/reader co-operates

The simplest explanation for the occurrence of semantic illusions is that people know what is meant by the question and ignore the substitution, choosing to respond to the task at hand as if there was no problem (cf. Grice, 1975). If this were the case, participants should find it easier (or at least no more difficult) to detect substitutions than to treat the sentence as if it did not contain a substituted term when faced with a substitution detection task. However, empirical data (Reder and Kusbit, 1991; Reder and Cleeremans, 1990) showed that participants found it easier to perform a 'gist' task in which target sentences were to be responded to as though the substitution did not exist, than a 'literal' task in which participants were specifically instructed to look out for substitutions. The 'gist' task was performed both faster and more accurately than the 'literal' task. The findings showed that participants found it much harder to detect substitutions than to ignore them, and people's tendency to fall for semantic illusions cannot be the result of a conscious decision to be 'co-operative'. Instead it is more likely to reflect a default strategy in sentence comprehension (Reder and Cleeremans, 1990).

2.3 Semantic illusions occur because the focus is not on the substitution

Another explanation of semantic illusions could be found in the way that the sentences are phrased. It is possible that the structure of the sentence places little focus on the word that has been substituted, causing the substitution to be overlooked. Erickson and Mattson (1981) suggested that presenting semantic illusion sentences as questions directed participants' attention towards answering the questions and away from the substituted word. They tested this hypothesis by presenting target materials used in a previous experiment as statements-to-be-verified, but a substantial number of semantic illusion responses still occurred for each semantic illusion item used.

2.3.1 Manipulating sentence focus using cleft phrases

Even when a stronger manipulation of the focus of the target sentence is used to direct attention to a specific part of the sentence, semantic illusions still take place. Brédart and Modolo (1988) presented sentences as cleft phrase statements-to-be-verified: some had the substituted name in the initial phrase (e.g. “It was Moses who took two animals of each kind on the Ark”), and others had another item of information in the cleft phrase (e.g. “It was two animals of each kind, that Moses took on the Ark”). The results showed a marked reduction, but not a complete eradication of semantic illusion responses when the wrong name was brought into focus in the cleft phrase. In the other sentence form condition the semantic illusion rate was comparable to that reported by Erickson and Mattson. The data showed that misdirection of focus clearly has an effect on semantic illusion rate, but can only partially explain why semantic illusions occur.

2.3.2 Shifting focus without syntactic change

A serious criticism of the use of a cleft phrase sentence structure to manipulate focus is that it might affect the ease of interpretation of a sentence. Clefting the phrase produces a shift in semantic focus but also requires a change in syntactic structure (Brédart and Docquier, 1989) and the majority of the sentences used in Brédart and Modolo’s study with the incorrect name in focus were cleft-subject sentences (e.g. “It was X, who…”), whereas most of the items in the second condition were cleft-object sentences (e.g. “It was X, that…”). It has been shown that cleft-object sentences are syntactically more complex than cleft-subject sentences, and this difference is likely to affect ease of interpretation (e.g. Waters, Caplan and Hildebrandt, 1987). In order to rule out the possibility that the focalisation effect observed by Brédart and Modolo was caused by a shift in syntactic complexity and not a shift in semantic focus, Brédart and Docquier (1989) repeated the focalisation study with slight modifications: sentences were presented in the same form for both control and ‘focus’ conditions, but in the ‘focus’ condition the substituted item was capitalised and underlined to draw attention to it while in the ‘non-focus’ condition another item of information was underlined and capitalised. It was reasoned that using typographical rather than syntactic manipulation of focus would reveal whether the focalisation effect existed. The results obtained were consistent with those of Brédart and Modolo. When the substituted name was in focus the mean semantic illusion rate was significantly lower than in the non-focus condition. Brédart and Docquier concluded that focus structure has a decisive influence on word meaning analysis, and directly affects the extent to which semantic illusions occur.

2.3.3 Is there a response bias at work?

While the use of paratextual cues such as capitalisation sidesteps the problem of a shift in syntax, it might also bias participants to give an increased number of false detection responses, where a truthful statement is wrongly identified as a semantic illusion. Kamas, Reder and Ayers (1996) attempted to make participants more sensitive to substitutions by influencing their memory structure. Before completing a question-answering task, participants studied statements of the facts to be tested in which either the to-be-substituted word (e.g. "NOAH took two animals of each kind on the Ark"), the answer term (e.g. "Noah took TWO animals of each kind on the Ark") or nothing was presented in capitals. The results indicated improved performance on semantic illusion questions for which the target word had been capitalised during study, but this was offset by an increase in the rate of false detections. Participants did not suddenly improve their ability to detect substitutions, instead they had shifted their response criterion on the basis of the form of the sentence during study.

2.3.4 Detecting false information in logically subordinate sentences

When a sentence consists of a main clause and a subordinate clause, participants are likely to interpret the information in the subordinate clause as presupposed and the information in the main clause as new and in focus. On the basis of normal stress patterns in speech, people also tend to consider information presented earlier in a sentence to be given, and information presented later to be new (Halliday, 1967). Baker and Wagner (1987) demonstrated that logical subordination is a factor in substitution detection by using complex sentences made up of an independent clause and a non-restrictive dependent clause (i.e. a clause that does not identify or limit the meaning of the word it modifies, but rather supplies further details). Sentences with a substitution in the subordinate clause (e.g. "Bloodletting, generally accomplished with the aid of rats, was thought to remove 'poisons' from the blood." The word "rats" has been substituted for "leeches" in this sentence) lead to more detection failures than sentences with the substitution in the main clause (e.g. "Bloodletting, thought to remove 'poisons' from the blood, was generally accomplished with the aid of rats") regardless of the specific context of false information. Baker and Wagner concluded that people are less likely to allocate attention to information when it is conveyed via a linguistic structure which suggests that it is incidental.

In order to control for the possibility of a serial position effect upon error detection causing false propositions at the end of a sentence to be detected with higher probability, Baker and Wagner also compared the detection rate of false information in complex sentences to the detection rate in compound sentences. In compound sentences, the two clauses are joined by the word "and" and neither clause is logically subordinate to the other (e.g. "Bloodletting was thought to remove 'poisons' from the blood and was accomplished with the aid of rats" or "Bloodletting was accomplished with the aid of rats and was thought to remove 'poisons' from the blood."). The detection rate for false information in the subordinate clause in complex sentences was significantly lower than that for either the main clause in the complex or for first or second positions in compound sentences, suggesting that the central/peripheral effect demonstrated for complex sentences is not due to more thorough processing of information at the end of a sentence.

2.4 Semantic illusions occur because the substitution and the target are so similar

It is intuitively obvious that the substituted term in a semantic illusion sentence must be in some way like the target word it replaces or the semantic illusion would not occur. For example, the substitution has to be of the same form class as the target, it has to be able to fill the same role slot, and it has to be similar enough to the replaced word to lead to a ‘feeling of cohesion’. But what makes the terms similar in this way?

2.4.1 Phonological similarity

One possible cause for semantic illusions is that target and substitution sound/look similar enough to be confused. There is little empirical evidence to support this hypothesis, however. Erickson and Mattson (1981) found that phonological similarity alone led to very few semantic illusions. Indeed, Shafto and MacKay (2000) argued that phonological similarity never leads to semantic illusions, but to a different type of illusion termed the ‘Armstrong’ illusion. This refers to detection failures in sentences such as “What was the famous line uttered by Louis Armstrong when he first set foot on the moon?” in which “Louis Armstrong” replaces “Neil Armstrong.” Shafto and MacKay claimed that this is an instance of phonological similarity, since Louis and Neil Armstrong had little in common semantically.

However, while the Armstrong illusion rate is comparable to semantic illusion rate, Shafto and MacKay neglected to take into account the small but likely relevant semantic similarity between the two Armstrongs, such as their shared fame, the fact that they were both American, and the fact that in the context of the first moonlanding

“Armstrong” is a perfect match for the role-slot regardless of the associated first name (cf. Barton and Sanford, 1993; see section 2.9). It appears that phonological similarity can contribute to the causes underlying semantic illusions, but a purely phonological theory cannot account for semantic illusions.

2.4.2 Semantic similarity

A more likely explanation for semantic illusions would be that somehow the target and the substitution mean similar things, by virtue of sharing a number of semantic features or by being otherwise associated. Empirical data shows that semantic similarity clearly affects semantic illusion rates (Erickson and Mattson, 1981; van Oostendorp and de Mul, 1990; Shafto and MacKay, 2000) with more similar substitutions leading to a greater semantic illusion rate. But there appears to be no special type of semantic feature which the substitution must share with the correct name. For example, Moses and Noah are both Old Testament patriarchs; Captain Nemo and Captain Ahab are both fictional sea captains.

2.4.3 The nature of the semantic relationship

Van Oostendorp and de Mul (1990) compared detection rates and response times for substituted names which were highly related to the target name to those for low-related pairings of names. Semantic illusion responses were more frequent in the high-related condition than in the low-related one and participants took longer to respond correctly (detect the substitution) in the high-related condition than in the low-related condition. However, contrary to expectation, the response time data showed that semantic illusion responses were made equally quickly in the high-related condition as in the low-related condition, rather than being made more slowly for low-related pairs. These results imply that in processing semantically low-related sentences, qualitatively different knowledge from that activated in processing high-related sentences is used, since more errors were made in the same length of time for high-related statements than for low-related ones, and it took longer for detections to be made in the high-related condition.

Van Oostendorp and de Mul suggested that readers apply an internal criterion based on perceived semantic cohesion of a representation (how well the ideas within a sentence fit together). Depending upon this monitoring, further processing may, or may not, take place. The cohesion of a representation is defined by the number and strength of relations between the concepts involved. So in the case of a high-related substituted name, the set cohesion criterion is reached sooner or more easily and therefore the activation of information (via inferences) which is crucial for thorough processing might be omitted (den Uyl, 1980; see also section 2.9 on global goodness-of-fit.)

2.4.4 Conceptual relatedness

Readers appear to be able to activate just the required amount of knowledge from memory in order to comprehend a text, possibly by continuously monitoring the text's conceptual cohesion. The initial mental representation is made upon the basis of connections immediately available to working memory rather than the specifics connecting the concepts. The conceptual cohesion then depends upon the relatedness between facts which is determined by shared semantic attributes, and by the strength of the relations in semantic memory (e.g. Anderson, 1983, 1984). Strong conceptual relations can be induced experimentally through the study of arbitrarily connected concepts (as for example in a paired-associate learning task), and lead to inappropriate judgements of sufficient cohesion (e.g. Anderson, 1983). The identification of what each aspect of a text refers to is also an important factor in the establishment of coherence. Once the referents are determined, a mental model can be constructed, but there appears to be a trade-off between conceptual cohesion and underlying details, especially when perceived semantic cohesion is high to begin with.

Thorough processing of a sentence usually takes more than an evaluation of conceptual cohesion – information not directly linked to a specific concept within the semantic network may need to be activated in order to construct a coherent mental representation. Van Oostendorp and Kok (1990) assumed that making such inferences and evaluating their appropriateness would be harder and take more time than the initial conceptual relatedness check. Erickson and Mattson's findings about semantic similarity were seen to support the notion of conceptual cohesion monitoring, as the similarities between names such as Moses and Noah give the impression of sufficient cohesion during initial processing, and thus lead to a failure to spot the substitution in a semantic illusion sentence.

Van Oostendorp and Kok demonstrated that participants made more semantic illusion responses when the names in a target sentence were conceptually highly related, but also when the relations between a substituted name and the sentence context were strengthened as the result of a paired-associate learning task. The results suggest that the relatedness of proper nouns affects the process of knowledge activation in sentence processing: readers are particularly likely to fail to identify referents correctly, when the names in question are conceptually similar or highly related.

2.4.5 Phonological and semantic similarity

Shafto and MacKay (2000) demonstrated that it is possible to get an even stronger version of the Moses illusion (called the mega-Moses illusion), if Moses-factors and Armstrong-factors were combined, i.e. if the target word shared semantic as well as phonological features with the substitution in a semantic illusion sentence. The mega-

Moses illusion is more likely to occur than either the Armstrong illusion or the Moses illusion. For example, in the question "The 1868 impeachment trial involving former vice president Andrew Johnson [*correct*] followed what major American war?" substituting "Samuel Johnson" for "Andrew Johnson" should lead to an Armstrong effect due to shared phonological and lexical-surname similarity. In the same question, the Moses effect should be observed by substituting "Theodore Roosevelt" for "Andrew Johnson", as these names are semantically similar (both became president after the assassination of their respective predecessor), but phonologically dissimilar. But if "Andrew Johnson" was replaced with "Lyndon Johnson" an enhanced illusion effect would be expected, as these names share semantic (again, both became president after the assassination of their respective predecessor) as well as phonological information. (Example taken directly from Shafto and MacKay, 2000).

2.5 Semantic illusions occur because participants prioritise the answering task

Much of the empirical evidence that has been gathered about semantic illusions comes from manipulations of discourse properties like the similarity between target words and correct words, or the focus of the sentence. In all of these investigations an implicit assumption seems to be that the depth of semantic processing is determined only by characteristics of the stimulus materials. But it is highly likely that the depth of semantic processing is also affected by task demands. This in turn suggests that task demands will have an effect upon semantic illusions.

2.5.1 Accuracy vs. speed

Van Jaarsveld, Dijkstra and Hermans (1997) investigated the effect of participants' ability to vary the depth of semantic processing by manipulating specific task demands. In a 'balanced' condition participants were asked to respond to questions as quickly and as accurately as they could, while monitoring for semantic illusions. In the 'accurate' condition, the task was to respond as accurately as possible disregarding time taken to complete the task. It was found that participants are able, to a certain extent, to control their ability to detect semantic illusions. Detection rates were higher when accuracy was stressed in the instructions rather than when both accuracy *and* speed were stressed, but there was a greater false detection rate in the accurate condition than in the balanced one. Presumably, there is more semantic processing in the 'accurate' condition, which results in more extensive checking of a target word's semantic features and of relations between concepts within the sentence (cf. Erickson and Mattson, 1981), but only by tolerating a higher false alarm rate for correct sentences.

Van Jaarsveld, Dijkstra and Hermans' results confirmed indications that the detection of semantic illusions is rather difficult, even when there is little time pressure, or where a warning about the occurrence of semantic illusion sentences is given. People's inability to detect semantic illusions reliably even when explicitly instructed to do so, may be due in part to lapses of attention and falling back to more superficial semantic processing. But it is also likely that participants will have set their own temporal deadline (despite having been told that response times were in no way relevant), which, however lax, is almost certainly shorter than would be required for an exhaustive check of all relations between the constituent concepts in a sentence. As a result some distorted terms will always remain undetected.

2.5.2 Processing tends to be minimal with regard to task demands

From recent research on text processing it appears that the depth of semantic processing is determined at least in part by task demands. Participants process materials to the extent that is needed to perform a given task. Unnecessary elaborative semantic processing is avoided (Foertsch and Gernsbacher, 1994; cf. also McKoon and Ratcliff, 1992). Indeed, participants appear able to adjust the depth of their semantic processing to quite subtle aspects of the task demands (Wilson, Rinck, McNamara, Bower and Morrow, 1993).

In any particular task, semantic processing is likely to continue only until enough information has been gathered to enable a participant to make a response. Response requirements vary for different tasks, with respect to amount or type of information needed, and thus the effects of substitutions may depend upon the task to be carried out. Van Jaarsveld, Dijkstra and Hermans (1997) demonstrated the effect of the minimal semantic processing strategy upon semantic illusion rates by comparing a question-answering task (with no requirement to monitor for substitutions) and a substitution-detection task (with no requirement to answer the question), as the similarity of the substituted word should have different effects for each task, with greater similarity being detrimental in the detection task, but helpful in the question-answering task. This ‘similarity effect’ (first described by van Oostendorp and de Mul, 1990) was found to be significant for both tasks and went in opposite directions as predicted. In the question-answering task, there were longer response times for dissimilar terms than for similar terms, as more similar terms allow for an easier establishment of the gist of the question (e.g. It is easier to respond as if there was no substitution to the question “How many animals of each kind did Moses take on the Ark?” than to the question “How many animals of each kind did Adam take on the Ark?”).

In the detection task, there were longer response times for similar terms than for dissimilar terms, as dissimilar terms disrupted the flow of semantic processing to a greater extent (e.g. It is easier to detect the substitution in the “Adam” question mentioned above, as “Adam” sticks out more). Analyses of the percentages of correct responses for the similarity conditions were carried out separately for each task. For the question-answering task, there was no significant difference in the number of correct responses for similar and dissimilar terms, whereas the effect was highly significant in the detection task, with a much higher rate of correct responses for questions with dissimilar terms than for questions with similar terms. This last finding can be explained in terms of interrupted flow of sentence processing. It also replicated the similarity effect as reported in previous studies (Erickson and Mattson, 1981; van Oostendorp and de Mul, 1990; van Oostendorp and Kok, 1990).

2.5.3 If semantic processing is minimal, the position of the substitution should affect semantic illusion rate

Van Jaarsveld, Dijkstra and Hermans predicted that word order would have an effect upon processing whenever the experimental task could be completed on the basis of processing only part of the sentence. If the distorted word occurred in the part of the sentence that need not be processed to perform the task, detection rates should be lower than if the complete sentence had to be processed. In a statement-to-be-verified task the complete sentence was thought to require processing whereas in a question-answering task only the gist of the question needs to be understood, and the question-answering task could be performed on the basis of part of the sentence only.

2.5.4 The similarity effect and minimal semantic processing

Assuming left to right processing when dealing with written information, Van Jaarsveld, Dijkstra and Hermans reasoned that substitutions at the start of a sentence would have a different effect from that for substitutions at the end of a sentence, depending on task demands. For a statement-verification task, in which the entire sentence needs to be processed, it was thought that a substitution would be equally likely to be detected irrespective of the position it was in and the similarity effect would be equally great in either position. But for a question-answering task, in which processing only part of the sentence might be sufficient to complete the task at hand, the similarity of the substitution should have a different effect depending on the position of the substitution in the sentence. A dissimilar term would, for example, be expected to be more disruptive at the beginning of the sentence, and less so at the end of the sentence. Hence variations in the depth of semantic processing can be assessed by comparing the size of the similarity-effect for substitutions at the start of a sentence with substitutions at the end of the sentence, for a statements-to-be-verified task compared with a question-answering task.

In the verification-task, where the entire sentence needs to be processed, results showed that statements with initial similar terms were responded to more rapidly than statements with initial dissimilar terms, and conversely, statements with final dissimilar terms were responded to more quickly than statements with final similar terms. Contrary to expectation, response times were not significantly faster for dissimilar substitutions than for similar ones. But the number of correct responses covaried with similarity, with more correct detections made for dissimilar terms. Analyses of semantic illusion responses found that these followed the same pattern as correct responses: initial similar terms led to more rapid semantic illusion responses than initial dissimilar terms, whilst final dissimilar terms led to more rapid semantic illusion responses than final similar terms (also, similar terms led to more semantic illusions than dissimilar terms).

In the question-answering task, where only enough of the sentence needs to be processed to understand the gist of the question, answers to items with similar terms at the beginning of the question were given more rapidly than for dissimilar terms, while there was no significant difference in response times for similar or dissimilar terms in questions with the substituted term at the end. Overall, similar terms led to more rapid responses than dissimilar terms, but the position of the substituted term had no significant overall effect upon response times. Fewer correct answers were given to sentences with dissimilar terms than to sentences with similar terms, indicating that there might be a small negative effect of dissimilar terms upon the activation of relevant knowledge for answering the question.

Van Jaarsveld, Dijkstra and Hermans explained the differences in the similarity-effect with respect to the two positions for each task in terms of differences in semantic processing as a result of task requirements. In the verification task, there is no effect of word position, because processing of the entire sentence has to be carried out – thus the two different types of sentences are processed in equal depth. In the question-answering task, semantic processing can be stopped as soon as the gist of the question is understood. Therefore words at the end of a question are processed only superficially or not at all. The absence of a similarity-effect for final substituted terms is in keeping with parallel models of question-answering (Graesser, McMahen and Johnson, 1994; Robertson, Ullman and Mehta, 1992).

2.6 Semantic illusions occur because participants fail to retrieve the correct information

Semantic illusions might be caused by incomplete retrieval of information from memory which means that participants do not have the information necessary to detect discrepancies between the target word and its substitute (Reder and Kusbit, 1991). In order to make sure that people already had the necessary information stored, participants studied (or committed to memory) a series of facts before answering the questions, as exposure to the correct information should eradicate the problem of incompletely retrieved or impoverished knowledge. The sentence studied always contained the correct (unsubstituted) target term. So, for example, a participant might later be asked "Who does Clark Kent become when he changes in a toll booth?" but the studied sentence would be "Clark Kent becomes Superman when he changes in a phone booth". In the question phase, some participants were asked the unmodified correct question, while others were asked the semantic illusion question. Reder and Kusbit looked at performance for both the literal task (in which participants were asked to detect the substitutions) and the gist task (in which participants were asked to ignore the substitutions; see section 2.2).

The basic pattern of the results matched previous findings: the literal task took more time than the gist task, and was less accurate; distorted questions were answered more slowly and less accurately in both tasks, and the effect of substitutions was much more detrimental to accuracy in the literal task than in the gist task. While previously studied items were answered more quickly overall, there was no evidence that priming helped differentially with the detection of substitutions. Thus the improvement in performance caused by the priming task appears to be a knowledge effect, which simultaneously enhances the effects of the other variables which produce the illusion. Even though increased familiarity with the information improved performance with responses being faster and more accurate overall, the pattern of previously established findings concerning semantic illusions was not altered. So imperfect retrieval from memory cannot on its own account for the occurrence of the illusion.

2.7 Semantic illusions occur because participants do not encode the substitution

2.7.1 Reading the sentence aloud

One of the simplest explanations for the semantic illusion is that the substituted name is never encoded, as many semantic illusion victims claim not to have seen the substituted name. Erickson and Mattson (1981) tested this hypothesis by asking participants to read the semantic illusion questions aloud before answering them. This should ensure that the substituted word was encoded at least at a phonemic level. If a failure to encode lay at the root of the illusion, this simple manipulation should eliminate the occurrence of semantic illusions. However, the semantic illusion occurred despite reading aloud, and Erickson and Mattson argued that a failure to encode cannot be the reason for the phenomenon.

2.7.2 A look at target word reading times

Reder and Kusbit (1991) also concerned themselves with the question of encoding. They made the assumption that reading time should vary depending on whether or not a semantic illusion response was made. Participants should take longer to read a sentence when a substitution is detected than when the question is responded to as though it contained no substitution. Contrary to their hypothesis, Reder and Kusbit found that reading appeared to slow down when the target word was substituted and a semantic illusion response occurred, but there was little evidence of slowing-down due to a substitution if the question was answered correctly. There was also no difference between reading time for substituted words when a participant failed to notice the substitution and reading time when the substitution was detected. Reder and Kusbit concluded that an inadequate encoding hypothesis could not be the explanation of the semantic illusion, provided that reading time could be used as a fair indication of the amount of processing or encoding time spent on each word.

2.8 Semantic illusions occur because participants only perform a partial matching check between input and information stored in long term memory

Another possible explanation for the semantic illusion could be that an incomplete memory match is carried out (Reder and Kusbit, 1991). It was suggested that the memory trace – while brought back into working memory in its entirety – is not exhaustively matched to information in the test question before the answer is generated. In most situations a complete match between stored knowledge and a given situation is unlikely to occur, so people are used to tolerating small discrepancies and highly similar but not identical terms are allowed to slip by. Therefore it makes sense for a memory matching process to be carried out at the level of concepts and features rather than words. If there is heavy feature overlap, a mismatch tends to go unnoticed. Such similarity is dependent upon the knowledge state of an individual and also upon their cultural or social context. This explanation of the semantic illusion appears plausible if not quite complete, as in real life situation-matches are not usually exact and most mismatches that occur are inadvertent and unintentional so that, for example, it makes more sense to ignore slight mispronunciations than to check for an error. However, there are always cognitive safeguards for dealing with large discrepancies, such as global goodness-of-fit (see section 2.9).

Reder and Kusbit looked at reading time after priming: participants who had been given a chance to familiarise themselves during the priming phase with the information that was to be questioned, answered semantic illusion questions while the reading time for each word was monitored. Reading times were found to be longer when participants made mistakes than when they answered correctly, except for primed sentences in the literal condition when the question was distorted, where there was a trend for the distorted term to be read faster when a semantic illusion occurred. Priming the materials before the task produced much faster reading overall, as was expected. Reading times for the substitution in the literal task were slower when the question had been primed (as predicted), but priming did not appreciably lower people's tendency to be deceived by semantic illusion sentences. Even though the effect was not statistically significant, Reder and Kusbit drew attention to the fact that more available information (made more available by the method of priming) appeared to draw more attention to the substituted word, so that it took longer to read than when the target was not substituted. Participants read the substitution slightly faster when they noticed it, than when they did not. In the gist task, which did not monitor for semantic illusions, this effect was not seen. In summary: participants in the literal task appeared to be especially thorough in scrutinising experimental questions when they had also been primed.

2.8.1 How partial matching might work

Reder and Cleeremans (1990) proposed a potential model for semantic illusions based on a parallel distributed processing (PDP) framework, because such frameworks are capable of processing incomplete information or information containing small distortions without affecting overall performance. The PDP model consists of a network of three interconnected pools of processing nodes – one represents input information and is connected to a second (hidden) pool. All units in the second pool are connected to the units in the third pool that generates the outputs of the network. The system processes inputs as specific patterns of activation, which spread through the interconnections between the pools of processing units to the output layer. The correct mappings from input layer to output layer are achieved by 'training' the network by repeatedly presenting input/output pairs to be learned. Using a learning algorithm called 'back-propagation' (Rumelhart, Hinton and Williams, 1986), the network can modify the connections between input and output in such a way as to reduce the 'error' between the actual and the target output. Once the error drops below a given threshold, the network will have developed an internal representation in the hidden layer of units, which will allow it to produce the target pattern of activation in the output layer in response to the corresponding input pattern.

Reder and Cleeremans' 'Moses-network' was trained using correct statements like "Noah took two animals of each kind on the Ark". After training, the model can be 'tested' by presenting it with incomplete patterns ('questions') and measuring how well it can complete the patterns ('answer the undistorted questions'). If one now assumes that each bit of input information is represented by a large number of 'microfeatures' that form certain patterns of activation on different subsets of input nodes, a little like semantic features make up certain concepts, then similar concepts (e.g. "Moses" and "Noah") can be represented by overlapping patterns of activation on specific subsets of input units (cf. 'distributed representation', Hinton, McClelland and Rumelhart, 1986). Similar input patterns will lead to similar output patterns, and thus – if the overlap between two concepts is large enough – an incomplete pattern including a substitution will still be able to produce the correct and complete output pattern. Basically, the model responds to the 'question' as if it were not distorted.

The PDP model can be used to account for a number of different findings about the gist task, if one assumes that the 'error' associated with each output is proportional to response time and accuracy. A large error arising from large discrepancies between input and output would lead to a garbled output from the network ("don't know" or wrong responses); a small error would weaken the output, essentially leading to a slower response (as was observed). This can be used to account for the basic finding that participants can answer questions in the gist condition as long as the substitution is similar enough to the target. By the same token, the network model can explain how it is harder to detect substitutions when more related terms are contained in a question: the ratio of mismatched to matched features will be smaller when there are more related terms. Focus effects could be accounted for by assuming that focal items are assigned greater weight, so that distorting this concept would cause a larger degradation of the output than distortions of unfocussed concepts. The difference between expert and novice performances in semantic illusion tasks could be explained by the fact that experts are likely to have richer representations of certain facts than novices, so that the ratio of shared to overall features between two concepts like “Moses” and “Noah” will tend to be smaller for experts than for novices.

However, while the network model can account for the phenomena observed in the gist condition, it is far from obvious how the observations of performance in the literal condition can be modelled. As it stands, the model could monitor for error, but it could not distinguish between “don’t know” and “can’t say” responses. Similarly, quality of retrieval and extent of matching processes would be confounded, unless two different measures of the network’s performance can be found that correspond in behaviour to the finding that strengthening memory traces (i.e. priming certain inputs) affects only the ease of retrieval but not that of the match process. Reder and Cleeremans did not offer a solution to this problem, but concluded that further empirical research would be required before any clearer answers could be given.

2.8.2 Partial matching at the semantic feature level

Kamas, Reder and Ayers (1996) demonstrated that emphasising certain parts of a semantic illusion sentence can cause shifts in response bias (see section 2.3.3). It was concluded that there must be another component to the partial-matching process, at a level lower than the morpheme or word-level – i.e. at a semantic feature level. Kamas, Reder and Ayers reasoned that the substitution would be more readily detected if the semantic features that distinguished the substituted term from the correct term were emphasised by manipulating the salience of particular features of the critical term. Thus, each target question was immediately preceded by a question that either emphasised features shared by the target and substituted terms, or that emphasised features distinguishing the two terms, or that was irrelevant to the target question. The results showed that the substitution was detected more easily when the preceding question drew attention to the features distinguishing target from substitution. When sensitivity and bias measures were calculated, these showed that participants became more sensitive to substitutions when a target question is preceded by a distinguishing question. Bias was not affected by this manipulation, suggesting that participants did not change their response criterion as a result of the manipulation. The results thus indicated that detection rates improve when a preceding question focuses on features that distinguish the substituted term from its original counterpart, and that the emphasis on shared features does not impair detection rates.

2.8.3 Two mechanisms contribute to semantic illusions

Hannon and Daneman (2001) looked at semantic illusions from an individual differences point of view, suggesting that there might be two separate mechanisms which contribute towards the occurrence of semantic illusions. One is related to an individual's ability at accessing and reasoning about knowledge from long term memory. This ability allows participants to detect the substituted target word. The second mechanism is related to an individual's capacity to process and store information in working memory at the same time. This capacity allows participants to avoid being 'led up the garden path' by the context of the sentence containing the substitution.

Hannon and Daneman claim that thinking of the partial matching process as a single mechanism cannot easily be justified, as evidence from research about reading mechanisms suggests that knowledge-based processes and text-based processes are not the same (e.g. Hannon and Daneman, 2001; Kintsch, 1988). There appears to be a distinction between the processes that rely on the incorporation of knowledge about the world with information within a text and the processes that rely on the integration of information that is contained explicitly in a text. The former processes are those needed to detect the difference between the substitution and the information about the correct target that has been retrieved from memory. The latter are those that allow someone to integrate contextual cues and to detect a problem with them. Hannon and Daneman designed an experiment in which they varied the semantic relatedness of target to substitution as well as the number of contextual cues relating to the correct target word. They predicted that people would make most semantic illusion responses when both the semantic relatedness of the substitution to the target was strong and the context was highly related to the target, and that semantic illusions responses would occur least often when the substitution was less related and the context was weak. Hannon and Daneman also predicted that the context effect and the semantic relatedness effect would be additive and not interactive, as two different mechanisms were thought to underlie the respective sources of cognitive error.

Participants' ability to integrate text-based information with information retrieved from long term memory was measured using a knowledge access task that required participants to study three sentence paragraphs that described relations between a number of real and nonsense terms. By using the relations described in the sentences and by integrating them with knowledge about the real world, participants can construct different linear orderings, comparing the real and nonsense terms (e.g. size orderings, weight orderings; see Table 2.1 for an example). After studying the three sentences, participants responded to true/false statements about them. Some of the statements were referred to as 'knowledge access statements', which tested reasoning about prior knowledge.

Table 2.1: An example of materials used in the knowledge access task used to measure participants' ability at integrating text-based information with knowledge stored in long-term memory (adapted from Hannon and Daneman, 2001).

Bird Item			
Study paragraph			
	A MIRT resembles an OSTRICH but is larger and has a longer neck.		
	A COFT resembles a ROBIN but is smaller and has a longer neck.		
	A FILP resembles a COFT but is smaller, has a longer neck, and nests on land.		
Features/relations			
	Size	*mirt > OSTRICH > ROBIN > coft > filp*	
	Neck length	*mirt > OSTRICH > ROBIN; filp > coft > ROBIN*	
	Nests on land	*Filp, OSTRICH*	
	Doesn't nest on land	*coft, ROBIN*	
Test statements			
	Knowledge access		
		A ROBIN lives in CANADA, whereas a PENGUIN typically doesn't.	*(True)*
		A BLUEJAY lives in CANADA, whereas an OSTRICH typically doesn't.	*(True)*
		A PENGUIN lives in CANADA, whereas a ROBIN typically doesn't.	*(False)*
		An OSTRICH lives in CANADA, whereas a BLUEJAY typically doesn't.	*(False)*
	Text memory		
		A MIRT is larger than an OSTRICH.	*(True)*
		An OSTRICH is larger than a MIRT.	*(False)*
	Knowledge integration		
		A MIRT has a longer neck than a ROBIN.	*(True)*
		A ROBIN has a longer neck than a MIRT.	*(False)*

Participants' working memory span was also measured using a version of Daneman and Carpenter's reading span test (1980). Participants read increasingly longer sets of unrelated sentences aloud, made an acceptability judgement about each sentence and then after the end of each set, they were asked to recall the final word of each sentence in the set. The reading span thus determined correlates well with global reading comprehension (e.g. Daneman and Carpenter, 1980; Daneman and Merikle, 1996). This measure was thought to provide an indication of participants' ability to resist being 'led up the garden path' by sentence context.

Hannon and Daneman's results showed that both substitution relatedness and context strength influenced semantic illusion rate, with substitution relatedness having the greater effect upon detection performance. There was no significant interaction between the two effects, and the effects of relatedness and context strength were not correlated, suggesting that the cognitive mechanisms underlying the two effects might be different. The extent to which the theoretically motivated knowledge access and working memory capacity measures contribute to the detection of substitutions was analysed by regression analyses. There was evidence that each measure accounted for independent amounts of the variance. Knowledge access was a better predictor of detection performance when the substituted word was strongly related to the target but in a weak context, and working memory span was the better predictor when the substitution was embedded in a strong context.

Hannon and Daneman further evaluated their theory by using structural equation modelling (using a computer programme called LISREL, Jöreskog and Sörbom, 1993), which tested an independent model in which knowledge access and working memory span made separate, additive contributions to detection performance, and a non-independent model in which knowledge access and working memory span both influence the contributions from the sentence context and from the substitution. The independent model was found to provide the best theoretical explanation for Hannon and Daneman's data, and both the structural equation model and the results of the regression analyses supported the theory that knowledge access is important for the detection of the substitution itself, while working memory capacity is important for the integration of the context surrounding the substitution.

2.9 Semantic illusions occur because the substitution does not interfere with global goodness-of-fit

While it is relatively easy to account for people's ability to ignore small discrepancies, it is harder to explain people's ability to make sense of given inputs, and the fact that more often than not, they easily spot problems with a body of text. In order to comprehend a passage of text, a person must be able to construct a coherent mental representation of what is being expressed in the text containing no logical or semantic contradictions. This is a fairly standard view of text comprehension (e.g. Johnson-Laird, 1983; Garnham, 1985; van Dijk and Kintsch, 1983), based upon a schema-like framework (Bartlett, 1995[1932]; c.f. Schank and Abelson, 1977; see also section 8.4.1) in which certain *role-slots* need to be filled adequately in order for coherence to be established. Knowledge activation seems to be subject to the *global goodness-of-fit* of the information contained in a role-slot in any given context. Global goodness-of-fit refers to how well an item is perceived to fit a given context: any filler for a 'slot' has to meet the criteria for the role, and anomalies are easily discovered if the filler fails to do so. For example, in a sentence such as "Mary ate some rocks for dinner" the anomaly is quickly identified – rocks are not edible. It appears that people are extremely good at detecting such anomalies under most circumstances, and this observation leads to the assumption that fillers are checked thoroughly against role-specifications as a part of normal processing. But the existence of semantic illusions calls this view into question.

2.9.1 Local meaning before global meaning?

A common assumption made in theories of sentence processing is that local meaning is established before global meaning (e.g. Kintsch and van Dijk, 1978), so that first the meaning of component morphemes is combined to find the meaning of phrases, which are in turn put together into the meaning of a sentence. Another frequent assumption is that processing occurs incrementally. Carpenter and Just (1983), for example, suggested that each word is analysed as deeply as possible upon being first fixated by the eye. But while there is some evidence for such word-by-word analysis and while this suggestion is coherent with the ideas of local before global meaning establishment, these ideas are not compatible with the evidence from studying semantic illusions, which are basically sentences in which anomalies are clearly not detected and where item processing is shallow or incomplete.

Barton and Sanford (1993) linked the explanation of semantic illusions to a number of other points that had been raised by various researchers regarding sentence verification. Anderson (1983), for example, found that participants would sometimes classify a sentence like “A cat is a snake” as true in a speeded judgement task when they had previously learned arbitrary propositions such as “The cat attacked the snake”. This appears to suggest that under certain conditions participants do not carry out a full semantic analysis of the statement-to-be-verified, provided that the concepts within the statement are highly related, and instead verification judgements are based upon a priming process. Similarly upon verifying general knowledge statements, participants quite often make a judgement on the basis of semantic overlap, so that a proposition such as “A whale is a fish” might be marked as true (Smith, Shoben and Rips, 1974; Reder, 1982, 1987). A good match on a few basic features appears to supersede any further processing.

To explain the observations reported above, van Oostendorp and den Uyl (1984) suggested that the initial evaluation of a representation depends upon the strength and number of connections in working memory (conceptual coherence) rather than on the specific nature of such connections (semantic coherence). Similarly, Sanford and Garrod (1989) suggested that good global fit in working memory may pre-empt more detailed and time-consuming analysis. These claims also appear to fit with the parallel distributed sentence processing model developed by McClelland, St John, and Taraban (1989), in which comprehension depends upon multiple soft-constraint satisfaction. Contextual constraints imposed by the current representation of a sentence are assumed to determine the extent to which each newly encountered word can affect this current representation. Thus in the case of semantic illusions the constraints associated with the word “Moses” are not strong enough to override the constraints imposed by the context. This model does not rely upon a view in which the meanings of each word have to be combined to get to the meaning of a sentence, but instead the contributions of words to sentence meaning are seen as graded. The model also explicitly breaks the distinction between word meaning and more general aspects of significance – that is, local meaning is not assumed to precede global meaning. Instead global meaning may influence the contribution of a local word meaning to the entire text.

2.9.2 The effect of global goodness-of-fit upon subsequent analysis

Barton and Sanford (1993) examined the claim that global goodness-of-fit influences the extent of subsequent analysis. They explored the type of contextual agreement necessary for participants to detect an anomaly in a sentence. Using a cognitive illusion similar to the semantic illusion based on the premise of a plane crash: "If an air plane crashes, where do you bury the survivors?" Many people failed to notice the anomaly in this proposition (survivors would generally object to being buried!), even when it was presented under straightforward reading conditions rather than as an aspect of testing statements against knowledge stored in memory.

Barton and Sanford tested whether making the information "is alive" more available through the semantics of the target word would improve anomaly detection rate. To this end they compared the detection rate of "survivors" – which implicitly contains the information "is alive" – to the detection rate of various "injured"-terms ("injured", "wounded", "maimed") which they argued do not implicitly contain "is alive". Significantly fewer detections occurred for injured-terms than for "survivors", suggesting that presupposed information enhanced anomaly detection. Barton and Sanford also looked at the effect of adding the qualifier "surviving" to the injured terms upon detection rate: detection levels were brought back up to the level of "survivors" itself.

But even though the term “survivors” has “is alive” as a central part of its meaning, there were still a significant number of detection failures in sentences where “survivors” was used to replace “dead”. Barton and Sanford suggested that a possible explanation for this finding was that the close association between “dead” and “survivors” was sufficient to satisfy the processing needs for the establishment of cohesion and thus for comprehension of the sentences in question.

2.9.3 A partial match is sufficient to satisfy role-slot demands

Based on the partial match theory (Reder and Kusbit, 1991), it seems likely that a good partial match will inhibit further analysis, and a term will be accepted as a role filler. Barton and Sanford used the anomalous noun phrase “surviving dead” to investigate this assumption: “dead” offers a perfect fit to the role-slot in the scenario used, and thus Barton and Sanford assumed that “surviving dead” might produce a much reduced detection rate. The results showed this to be the case – the detection rate went down to 23% compared with the 66% base rate for previous experiments. This extremely low detection rate suggests that the local semantics of the noun phrase are not processed before a more global representation of the text is established (cf. the Armstrong illusion, see section 2.4.1).

2.9.4 Detection rate is affected by scenario-based expectation

The detection rate can also be affected by the expectations of the participants with respect to the likelihood of the victims of an accident being dead. Instead of using an elaborate scenario in this experiment participants were asked a single stand-alone question of the type: “When an aircraft crashes / a bicycle accident occurs, where should the survivors be buried?” The sentences were presented with the verb phrase referring to burying the survivors appearing either in passive (“where should the survivors be buried”) or in active (“where should you bury the survivors”) formats, and with the position of the scenario (bike accident / air crash) being presented either early or late in the question. This was done to allow for a test of any effects of the order in which constraining information was presented (see Table 2.2).

The results showed no reliable difference between the four question types, but the scenario type effect was significant, as the bicycle accident was thought less likely to result in deaths. Detection rate was considerably lower in the air crash scenario than in the bicycle accident scenario. Thus no evidence was found to support a strictly incremental interpretation of sentence comprehension. There was a trend towards higher detection rates with sentences in which the verb phrase was active, but there was no difference between early and late scenarios with respect to detection rate.

Table 2.2: Variations of scenario type and verb voice as used by Barton and Sanford (1993) in experiment 3.

1. early scenario	passive verb	When (an aircraft crashes / a bicycle accident occurs), where should the survivors be buried?
2. early scenario	active verb	When (an aircraft crashes / a bicycle accident occurs), where should you bury the survivors?
3. late scenario	active verb	Where should you bury the survivors, when (an aircraft crashes / a bicycle accident occurs)?
4. late scenario	passive verb	Where should the survivors be buried, when (an aircraft crashes / a bicycle accident occurs)?

Barton and Sanford argued that these results showed a strong global effect of scenario type, which is in keeping with the idea that detection rate is affected by schema or scenario-based expectation. Barton and Sanford also took the fact that the early scenario is no different from the late scenario as evidence that the target word is not fully analysed as it is encountered before processing moves onto subsequent material, and thus later input can influence the impact of any word upon the way in which the sentence is understood.

2.9.5 Adding 'relevant' information affects detection rate by shifting the focus of the sentence

Barton and Sanford tested the idea that if information relevant to the question of place of burial was easily accessed, then a deep analysis of the term "survivors" would not take place. A two-sentence version of the questionnaire format was used. The basic control version of this questionnaire was: "Suppose that there was an air crash with many survivors. Where should they be buried?" The scenario was then elaborated by adding relative clauses, which were either relevant to answering the question (e.g. "… with many survivors who were mostly European"), irrelevant to answering the question (e.g. "…with many survivors, which happened last week"), or focusing on the substituted word (e.g. "… with many survivors who were mostly unhurt"). The results showed that there was a drop in detection rate for question-relevant information clauses. There was no significant difference between basic control and question-irrelevant information clauses. The target item-focusing question led to a 100% detection rate, which was significantly different from the other two groups. The findings supported the relevance hypothesis: if information relevant to answering the burial question is made easily available, this provides a level of coherence satisfactory to the comprehension system, and the critical item "survivors" receives only cursory analysis. Thus detection rates are low. If the information is not as readily available other sources of information for example word meanings will be explored more fully and detection rates are higher.

Barton and Sanford claim that the strength of this substitution detection paradigm lies in the fact that it reveals shallow underlying processing, that ties in with previous claims that exhaustive checks are not carried out in normal comprehension, and that such exhaustive checking of attributes is not even feasible (e.g. Erickson and Mattson, 1981; McClelland, St John and Taraban, 1989). From other research it is also apparent that the failures to notice substitutions are genuine, and not just a case of failure to report.

Barton and Sanford supposed that the comprehension system would accept partial matches and assume coherence as default. Therefore the failure to detect an anomaly can be taken as evidence for relatively shallow semantic processing. From Barton and Sanford's experiments it appears that the establishment of coherence goes no further than a level that is adequate to allow the sentence to be dealt with. Thus processing of terms during reading seems to be rather shallow, and if there is a good semantic match between a role-slot and a role-filler, further analysis does not necessarily take place.

2.10 And now?

In this chapter the research that has been carried out upon semantic illusions to date has been reviewed. It was found that semantic illusions are a very robust phenomenon, and their detection is not under reliable conscious control of the people who experience the semantic illusion. There are a number of ways in which the occurrence of semantic illusions can be affected, such as through varying the point of focus of the sentence by using cleft phrases, by capitalisation of the substituted term, or by providing information that distinguishes between the correct word and the substituted word before the participant makes a response. Asking participants to be particularly careful about spotting substitutions in semantic illusion sentences also affects semantic illusion rate compared with a task that stresses speed and accuracy. But while these manipulations reduce semantic illusion rate, they do not eradicate it completely.

The nature of the semantic relationship between the correct word and the substituted word also affects semantic illusion rate. If the two terms have a lot in common, a semantic illusion is more likely to go unnoticed than if they do not share as many semantic features. But in a task where participants are required to respond to the gist of a sentence including a substitution, a term that is semantically highly similar to its replacement aids performance of the task, while a more dissimilar term disrupts processing.

All the research to date suggests that there is essentially one mechanism which provides a plausible explanation for the occurrence of semantic illusions: when people deal with incoming information they carry out a matching procedure of incoming data against knowledge already stored in memory. Such a matching process is necessarily incomplete, as in real life people do not usually deal with information in exactly the same format time after time. So the very mechanism of partial matching, which allows people to process information quickly and efficiently in their day to day lives, trips them up in the case of semantic illusions. It is not clear exactly how this partial matching process works, but there is evidence to suggest that the sentence context and the substitution in a semantic illusion sentence make separate contributions to the likelihood of semantic illusions occurring. This suggests that partial matching consists of two processes: one which integrates the information contained in the sentence, the other accessing relevant knowledge stored in long-term memory.

There are still a number of aspects of semantic illusions that have not been explored, such as the question of how semantic illusions fit in with normal discourse processing. Another issue is the question of the nature of the semantic illusion task: is it a question to be answered, or a kind of problem to be solved in which a participant's attention is deliberately misdirected? And are semantic illusions in questions theoretically indistinguishable from semantic illusions as part of a statement-to-be-verified? There are some questions about distributional issues: do semantic illusions occur only in relatively long and convoluted sentences? Do they work only for names, or can they also be found for verbs, adjectives and other nouns? Finally there are questions about the cognitive mechanisms involved in semantic illusion responses, such as how working memory restrictions affect semantic illusions.

Some of these issues have been touched upon in the research to date, but it would be useful to try to tie together and tidy up the theoretical understanding of semantic illusions into a more complete and coherent profile. In the next four chapters issues relating to the surface structure and linguistic make-up of semantic illusion sentences are explored in an attempt to answer a few questions that have not yet been raised in semantic illusion research.

Chapter 3:

The pilot study: Form class, priming and semantic illusions

Whilst the various studies of semantic illusions that have been carried out to date range widely over different aspects of the phenomenon, several questions have been glossed over and there are still a few gaps, even in the basic description of the semantic illusion vehicle. In Chapter 3, a pilot study is described that deals with the issues of form class of the target words and selective priming of either the correct word or its substitute. In previous semantic illusion research most target words used were names, but sometimes words of other form classes were used, without this ever being commented upon as a potential source of variability in semantic illusion rate. In this exploratory study an attempt was made to compare the effect of substitutions from different form classes on semantic illusion rate. While the question of priming comes up on occasion in the semantic illusion literature, nobody has tried to prime selectively the substituted word as opposed to the word that would have been correct in the context. This kind of manipulation was also attempted in the study reported below.

3.1 On form class and sentence processing

Almost all research on semantic illusions to date has used names as target words. Reder and Kusbit's (1991) work is an exception, in that they use semantic illusion sentences with dates, nouns, verbs and even phrases replaced, but they did not comment upon this difference between the stimuli in their study and those used in other studies. However, the use of proper nouns as target words might mean that the findings are restricted to a category of stimuli for which the semantics are predominantly extensional (i.e. dependent on their referent; van Oostendorp and de Mul, 1990), and an empirical clarification of whether or not semantic illusion findings do extend across form class boundaries would be of interest.

The distinction between names and words of other form classes is not merely syntactic, but empirical evidence shows that there are processing differences. For example, McWeeny, Young, Hay and Ellis (1987) demonstrated that names were harder to remember than occupations in an experiment, in which participants were shown photographs of middle-aged men's faces, and told the occupation and name for each. Recall performance for occupations was much better than that for names. It also appears that memory for names is stored in a different part of the brain than memory for other meaningful information about individuals (Harris and Kay, 1995). Moreover, there is evidence that cross-modal priming (in which an auditory stimulus prepares for a visual stimulus or vice versa), which cannot be reliably observed for lexical decision tasks, does occur for familiarity decisions primed by a celebrity's name presented first auditorily and later visually (Valentine, Hollis and Moore, 1998).

Processing differences can also be shown to exist between other form classes. For example, people find it easier to remember nouns than verbs when learning word lists (Fillenbaum, 1970; Engelkamp, Zimmer and Mohr, 1990; Langenmayr, 1997). Similarly the effects of omitting words from a sentence are different depending on the form class of the omitted word. The omission of a verb has a greater effect than the omission of a noun on both understanding (Reynolds and Flagg, 1976) and on remembering (Raeburn, 1979).

Since the findings described above suggest that there is a difference in the way in which words from different form classes are processed, it seems likely that the form class of the substituted word in a semantic illusion sentence will have an effect upon semantic illusion rate. As a result of this line of reasoning, the stimuli for the pilot study were constructed with a view to exploring this issue further.

3.2 On selective priming and semantic illusions

Various theories of semantic illusions are based on the assumption that the requisite information to detect the substitution is not at hand. Simply allowing participants to study the relevant facts, or even getting them to commit relevant facts to memory does not increase the semantic illusion detection rate (Reder and Kusbit, 1991). If information relevant to 'answering' the question is added to the semantic illusion sentences, participants are more likely to experience the semantic illusion (Barton and Sanford, 1993). But preceding the question with information that emphasises the features distinguishing between the target and the substitution improves detection performance (Kamas, Reder and Ayers, 1996). The following study examines whether semantically priming the target word before the presentation of the semantic illusion sentence leads to smaller detection rates than semantically priming the substitution – i.e. the word the participants actually see in the target sentence during the test phase.

Priming the correct word rather than its substitute could affect the allocation of mental resources for recognising and dealing with a sentence following the prime. How the allocation of mental resources varies with expectation was demonstrated in an ingenious experiment carried out by Posner and Snyder (1975). Posner and Snyder presented their participants with pairs of letters and asked them to decide whether the letters were identical to or different from each other. Before the presentation of each letter a fixation point was displayed. This fixation point was either neutral (a plus sign "+"), a stimulus that appropriately primed one or both of the upcoming letters (e.g. “A” followed by “AA” or “AB”), or a stimulus that misleadingly primed the upcoming letters by displaying a letter that was unlike both upcoming ones (e.g. “F” followed by “AA” or “AB”). Posner and Snyder used two different presentation distributions. For one (low-validity), only 20% of all cases were primed appropriately. For the other (high-validity), only 20% of all cases were misleadingly primed, while the rest were primed appropriately. In the low-validity condition, response times showed that there was a small benefit of priming and no cost to being misled. In the high-validity condition there was much greater benefit from priming, but also a significant cost in terms of response time to being misled.

If a correct word for a given context is primed, a semantic illusion effect seems more likely, because the items priming the relevant word will also be relevant to the context of the sentence, so that attention will not be drawn to the substituted word in particular. On the other hand, if the substituted word is primed, a semantic illusion response will be less likely – the context of the sentence will not have been primed, whereas the substituted word and its context will have been. Preparing for a specific input usually means that the expected input is much more easily perceived and processed than unexpected inputs (cf. Posner and Snyder, 1975). Thus it seems likely that special attention would be paid to a primed substituted word, as its appearance in the sentence would have been anticipated, whereas the rest of the semantic illusion sentence would be unexpected.

In the light of this reasoning, it was hypothesised that participants would be better at detecting substitutions when the substituted word is primed than in a non-primed condition. By the same token, it was thought that priming the correct word – and with it the rest of the semantic illusion sentence – would lead to a greater semantic illusion rate than in the non-primed condition.

3.3 Method

Participants: 15 participants volunteered to take part in the pilot study. All were either undergraduates or educated to university level.

Materials: Forty sentence frames were constructed for statements-to-be-verified. Each sentence frame could be transformed into any of three versions: ‘correct’ in which the statement was true and no word was replaced; ‘semantic illusion’ in which the target word was replaced by a semantically similar word; or ‘incorrect’ – where the target word was replaced with an obviously incorrect word. Ten of the statements were designed to have proper nouns (i.e. names) as their target words; ten had common nouns, ten had verbs and ten had adjective target words (see Table 3.1 for examples; full set of sentences used in Appendix 2).

There were three priming conditions: ‘not primed’, in which no priming words preceded the presentation of the semantic illusion sentence; ‘target primed’, where a number of words relevant only to the correct word preceded the sentence (for example, “wicked witch, uninvited guest, spinning wheel” would be used to prime “Sleeping Beauty”); or ‘substitution primed’, where the primes specifically referred to the substituted word (for example, “seven dwarfs, glass coffin, poisoned apple” would be used to prime “Snow White” in the example above. Full list of primes used in Appendix 2). Primes were displayed for two seconds before each sentence appeared in the centre of the computer screen. The statements-to-be-verified stayed on the screen until the participant made a response. Sentences and primes were selected on the basis of agreement between two independent judges on the appropriateness of an item to the relevant category.

Table 3.1: Examples of the three sentence variations for each form class.

Form class	Version	Example Sentence
Name	Correct	*After Sleeping Beauty pricked her finger, she slept for 100 years.*
	Semantic illusion	*After Snow White pricked her finger, she slept for 100 years.*
	Incorrect	*After Cilla Black pricked her finger, she slept for 100 years.*
Noun	Correct	*Whales are the largest aquatic mammals on Earth.*
	Semantic illusion	*Elephants are the largest aquatic mammals on Earth.*
	Incorrect	*Camels are the largest aquatic mammals on Earth.*
Verb	Correct	Wearing two pairs of socks can help to avoid blisters.
	Semantic illusion	*Washing two pairs of socks can help to avoid blisters.*
	Incorrect	*Knitting two pairs of socks can help to avoid blisters.*
Adjective	Correct	*Gas central heating helps to keep the house warm.*
	Semantic illusion	*Gas central heating helps to keep the house rosy.*
	Incorrect	*Gas central heating helps to keep the house dark.*

Apparatus: The materials were presented on a PC using a *Superlab 2.0* experimental script to display each sentence in the centre of the computer screen. The computer was set up to record the responses made by each participant consisting of pressing one of three response keys marked ‘true’, ‘false’, and ‘don’t know’ on a *Cedrus-600* 6-button response box. The other three response keys were covered with a cardboard sleeve to prevent any confusion about which keys to use.

Procedure: Participants were asked to read the sentences that appeared in the middle of the screen and to decide as quickly as possible if each sentence was true or false. There were no instructions regarding the primes. Participants were given an opportunity to practice the experimental procedure, before the experiment proper commenced. The experiment was broken down into three subsections of forty trials each. Each statement was shown to each participant three times, once in each of the three versions described above. The idea was that repeating the statements would preclude the need for a separate knowledge check, as participants who answered ‘true’ to both the ‘correct’ and the ‘semantic illusion’ versions of a sentence could be considered to have experienced the semantic illusion. Type of prime was counterbalanced across the statements of each form class, so that each participant saw examples of each combination of form class, sentence type and type of prime. The whole procedure took about half an hour to complete. At the end of the experiment, participants were debriefed and asked a few questions about their subjective experiences.

3.4 Results and discussion

Since the experiments described in the literature can be subdivided into two groups, depending on whether the experimenters report the correct response rate or the semantic illusion rate, both of these are presented here.

3.4.1 Semantic illusion rate

A semantic illusion was considered to have occurred when a participant judged the 'semantic illusion' version of a statement to be true *and* also judged the 'correct' version of the same statement to be true. Two sentences were excluded from the analysis, because they were found not to produce semantic illusion effects. In the sentence "The eruption of a volcano caused Pompeii to be destroyed (*correct*)/evacuated(*SI*)" the substitution of the intended 'semantic illusion' target also leads to a truthful statement, and the sentence was excluded from the analysis. A similar problem arose with the sentence "The phone directory is usually bound in a soft(*correct*)/yellow(*SI*) cover." It was expected that people would make a distinction between the standard telephone directory and the 'Yellow Pages' commercial directory. This was not found to be the case and the sentence was dropped from the analysis.

The semantic illusion rate for this study was found to be very low overall (see Table 3.2). An analysis of variance was carried out to explore the effects of form class and type of priming on semantic illusion rate using participants as subjects. There was no effect of form class (F=0.641; d.f.=3, 168; p=0.590) or of type of prime (F=1.177; d.f.=2, 168; p=0.311). Nor was the interaction between form class and type of prime significant (F=0.778; d.f.=6, 168; p=0.588). It appears that people are equally likely to experience semantic illusions for words from any of the form classes manipulated, and that the priming provides neither a handicap nor a bonus to processing the statements-to-be-verified.

Table 3.2: Mean semantic illusion rate (percent) for each form class and type of priming (standard deviations in parentheses).

	Not primed	*Target primed*	*Substitution primed*	*Total*
Name	*8.9 (15.2)*	*4.4 (11.7)*	*8.9 (15.2)*	*7.4 (14.0)*
Noun	*2.2 (8.6)*	*5.6 (11.6)*	*4.4 (11.7)*	*4.1 (10.6)*
Verb	*1.7 (6.5)*	*12.8 (19.6)*	*8.9 (18.8)*	*7.8 (16.4)*
Adjective	*3.3 (12.9)*	*7.8 (16.5)*	*6.7 (14.8)*	*5.9 (14.6)*
Total	*4.0 (11.4)*	*7.6 (15.2)*	*7.2 (15.1)*	*6.3 (14.0)*

3.4.2 Correct response rate

A second analysis of variance was carried out upon the correct response data. As with the semantic illusion rate there was no effect of either form class (F=1.622; d.f.=3, 168; p=0.186) or priming (F=0.859; d.f.=2, 168; p=0.426), nor was there a significant interaction (F=1.028; d.f.=6, 168; p=0.409). Overall people performed equally well for each condition (see Table 3.3).

Table 3.3: Mean correct response rate (percent) for each form class and type of priming (standard deviations in parentheses).

	Not primed	*Target primed*	*Substitution primed*	*Total*
Name	*77.2 (27.9)*	*81.7 (28.6)*	*78.3 (23.9)*	*79.1 (26.3)*
Noun	*90 (19.5)*	*92.8 (15.7)*	*86.7 (21.1)*	*89.8 (18.6)*
Verb	*95 (10.4)*	*73.3 (31.2)*	*83.3 (21.8)*	*83.9 (24.0)*
Adjective	*87.8 (21.3)*	*82.8 (23.7)*	*82.8 (24.7)*	*84.4 (22.9)*
Total	*87.5 (21.2)*	*82.6 (25.8)*	*82.8 (22.5)*	*84.3 (23.2)*

3.4.3 Presentation order

The order in which the statement-variations had been presented was recorded, so as to examine whether more semantic illusions would occur if the semantic illusion version of a particular sentence frame was seen first. It was found that more than half (51.5 percent) of the semantic illusions occurred on a first presentation of that particular sentence frame, while the rest of the semantic illusion responses were spread evenly across the other two groups (see Table 3.4). This decline in semantic illusion rate together with reports from participants led to the conclusion that for future experiments repetition of individual sentences should be avoided. Participants thought that repeated sentences were often recognised at a glance so that only the target word was given full attention in context. Repeated sentence frames were thus processed more quickly. A number of the participants remarked that the repetitions led to speculation about previous responses to the same sentence frame or to a sense of impatience with the task at hand. Some participants also reported that they found the repetition of sentences confusing or distracting. Others commented that the repeated sentences were easier to process when the correct sentence had been presented first, and that repeated sentences did not have to be read as carefully (or even read at all), as each sentence frame was rapidly recognised and attention was easily turned to the word which differed in the re-presentation (i.e. to the target word). Mean response times became shorter as a function of the number of times a participant had seen each sentence.

Table 3.4: Distribution of semantic illusions occurring for consecutive presentations of statements (percent).

	First	Second	Third
Percentage of SIs	51.5	27.3	21.2

3.4.4 Stimulus design

The variability of semantic illusion rate across sentences in this study was very large, and there was a ceiling effect for a number of the individual sentences which were responded to very accurately. This may in part be due to some supposedly confusable pairs of words not being particularly confusable at all and the reported semantic illusion rates are unlikely to reflect a successful manipulation of people's sensitivity to semantic illusions.

The lack of confusability between target and substitution could be due to the subject matter being too familiar to the participants, making them 'experts' for these particular sentences and therefore more likely to spot any substitutions (cf. Reder and Cleeremans, 1990). Another possibility is that in some cases the semantic illusion version of a sentence was not regarded as unambiguously wrong, once the substitution was pointed out. For a semantic illusion to occur the substitution has to be a similar but distinct term, so that the sentence appears to be coherent and makes global sense, with all role-slots adequately filled. This is easiest to achieve with names, for which the role-slot is very obvious. For example, in the sentence "After Sleeping Beauty/Snow White/Cilla Black pricked her finger, she slept for a hundred years", Snow White and Sleeping Beauty clearly have a lot in common. They are both female fairytale characters associated with princes, magic and fairy godmothers, but they are two distinct characters. Cilla Black, meanwhile, is a TV presenter, who is known for hosting "Blind Date" and not for relationships with fairies and princes. As the results of this study have shown, it is possible to do much the same with words from other form classes, although stimuli have to be chosen very carefully since verbs and adjectives that are perceived as associatively similar are often synonymous (for example, "laugh" and "giggle") or antonymous ("light" or "dark").

Even though there was no significant effect of priming upon semantic illusion rate, many participants commented in a debriefing session after the experiment, that they had found the primes distracting and had not made the connection between the primes and the sentences. Since the participants had not been given any instructions about the primes, it is possible that they found it harder to make sense of them because they did not know what to expect (cf. Bransford and Johnson, 1972; Dooling and Lachman, 1971), and thus received neither benefit nor handicap from the primes.

3.5 Conclusions

Though the semantic illusion rate for this study was extremely low overall, it was found that semantic illusions occur not only for names, but also for words belonging to other form classes, such as verbs, adjectives and nouns. There was no significant difference between the rate at which participants gave semantic illusion responses, and the data from this study did not provide evidence for a form class effect. It was decided that in future research semantic illusion sentences containing non-name target words could be used without affecting the overall results, provided the target and substitution words were sufficiently similar but distinct to meet the criteria required to lead to a semantic illusion.

The priming procedure used in this experiment did not cause a significant priming effect and since priming in various other semantic illusion studies (e.g. Reder and Kusbit, 1991; Kamas, Reder and Ayers. 1996) had not led to a differential effect of priming on semantic illusion rate either, it was decided not to continue with this line of research.

Since the repetition of sentences led to a drop-off in semantic illusion rate after the first presentation of a sentence, future studies should not use the design of having two statements-to-be-verified – one containing a substitution and one not – as a way in which to circumvent the need for a post-experimental knowledge check. A specific knowledge check also makes it possible to identify with greater certainty whether a semantic illusion occurred, or an error due to guesswork.

Chapter 4:

Experiment 1: Questions versus Statements

Just as the question of the form class of the target word had never been systematically examined in previous semantic illusion research, another issue that had not been explained was whether semantic illusions were more likely to occur in response to a target question or a statement-to-be-verified. The first full-scale study set out to deal with this issue and compared the semantic illusion rate for questions to that for statements-to-be-verified with the same semantic content. When Erickson and Mattson (1981) attempted to research whether focus had an effect upon semantic illusions, they turned the semantic illusion questions they had been using into statements-to-be-verified. Though they noted a decrease in semantic illusion rate, semantic illusion responses continued to occur. Erickson and Mattson concluded that focus was not solely responsible for the occurrence of semantic illusions. However, Erickson and Mattson did not comment upon the fact that they had not only changed the format of the target sentences, but they had also changed the presentation of the stimuli from a five-second-per-sentence display on a computer screen to a one-printed-sentence-per-page paper booklet for a self-paced pencil and paper task.

Since Erickson and Mattson's study, questions (e.g. Reder and Kusbit, 1991) and statements-to-be-verified (e.g. Brédart and Docquier, 1989; Brédart and Modolo, 1989; van Oostendorp and de Mul, 1990) have been treated as essentially equivalent to each other in their ability to elicit semantic illusion responses. But it has not been explicitly investigated whether there is a difference in semantic illusion rate between question and statement formats.

When dealing with questions people have to meet slightly different processing demands, than when dealing with statements of similar semantic content. In the processing of questions the same presuppositions about content are made as for declarative statements, but questions then also require further presuppositions to be made before successful processing can take place. For example, *wh*-questions (those starting with *who*, *why*, *what*, *how*, and so on) presuppose the concepts obtained by replacing the *wh*-word by the appropriate ‘existentially quantified variable’ (Levinson, 1983): so *who* would be replaced by *someone*, *how* by *somehow* and so on. For example, the question “How many animals of each kind did Moses take on the Ark?” presupposes that Moses took *a certain number of* animals on the Ark. These ‘placeholding’ presuppositions take up some of the processing resources available to a reader: there is neuropsychological evidence from event-related potential studies and from functional magnetic resonance imaging, that wh-questions cause distinct patterns of activation during processing (Fiebach, Schlesewsky and Friederici, 2001).

People also generally work on the assumption that certain standard conditions are met when they deal with language (e.g. Thomas, 1995). For example, when faced with questions, people tend to assume that Grice’s maxim of quality (1975, 1978) applies: if someone asks a question, they will usually be taken to be asking sincerely and because they are interested in receiving a relevant answer to the question asked. This type of linguistic interaction is not based on logic or semantics but on pragmatics: generally we would attempt to interpret whatever follows a question to be a relevant answer to the question (e.g. Mey, 2001). This pragmatic heuristic is violated in semantic illusion questions, as the perceived task (to answer the question) will not lead to the correct response (to identify the substitution).

With the verification of declarative statements, the task at hand is much more straightforward. While the same presuppositions have to be made regarding the semantic content of the statement as for the corresponding question, there is only ever one other presupposition necessary: the statement at hand is either true, or it is not. Thus participants are required to fulfil the task that they perceive: trying to decide whether the statement at hand is true or false. When young children are asked to complete a verification task and a question-answering task in which the questions correspond to the statements in the verification task, they perform better on verification than on question-answering (Akiyama, Takei and Saito, 1982).

The hypothesis of Experiment 1 was that questions are more likely to cause semantic illusions than statements-to-be-verified due to the difference in task requirements. Statements-to-be-verified require the participant to deal with one straightforward task: to verify the statement at hand. Strictly speaking a statement with one word substituted is no more than an untrue statement. For example, the statement “Moses took two animals of each kind on the Ark” is simply not true: Moses did not sail the Ark in the first place. Provided the participant carries out precisely the task they are instructed to perform, they have a chance of unmasking the semantic illusion. In the case of a question, however, the task requirements are different and slightly more complex. Not only do the participants have to access their knowledge store to retrieve the information relevant to answering the question, but they are also required to scan the question for any internal inconsistencies at the same time in a very unusual task. From schema theory (e.g. Bartlett, 1995[1932]), which suggests that overall meaning is likely to appear more important than details such as the inconsistencies caused by semantic illusion substitutions, and from the results of previous semantic illusion research, which showed that gist is more easily dealt with than the literal words (e.g. Reder and Kusbit, 1991), it seems likely that question answering will take precedence over the monitoring task. As a result, questions were expected to lead to more semantic illusions than statements-to-be-verified.

4.1 Method

Participants: 40 Oxford Brookes undergraduates volunteered to take part in this study when approached at the main entrance area of the university's campus. They were split into two groups of 20: 'questions first' or 'statements first'.

Stimuli: 8 target sentences were constructed from 'general knowledge' facts and piloted by presenting them to 10 volunteers. Each target could be presented as either a question to which a short (one-word) answer could be given (e.g. "What many-coloured garment was Jacob given by his father?") or as a statement for which the participant was required to make a true or false judgement (e.g. "Jacob was given a coat of many colours by his father"). 4 targets were presented to each participant as questions embedded among 16 filler questions, and 4 targets were presented as statements-to-be-verified, also embedded among 16 filler statements (12 true and 4 false). Each target was presented equally often as a question or as a statement-to-be-verified. The order of presentation (questions first or statements first) was counterbalanced across participants. (Examples of questionnaires used are presented in Appendix 3.)

Procedure: Each participant was given a leaflet consisting of 3 sheets of A4 paper, with instructions printed in bold above a grid containing the fillers and target items. For questions the instructions read: 'Questions: Please read each question and write down the answer to it as fast as you can. Some questions don't actually make sense, answer those with “can’t say”.’ The instruction ‘Please do not turn the page until you have completed it’ was printed at the top of each page. For statements the instructions were: ‘Statements: Please read these through as fast as you can and circle for each statement whether you think it is true or false.’ Participants proceeded at their own pace. The final page of the booklet was there to provide a knowledge check: 8 multiple-choice questions, which tested whether participants had the prerequisite knowledge to experience a semantic illusion. (Using the Jacob/Joseph example, a knowledge check question could take the following form: the question “Who was given a coat of many colours by his father?” followed by the response choices “a) Jacob, b) Joseph, c) Benjamin, d) Isaac”.) Instructions for this section were ‘Multiple choice questions: Take as much time over these as you like, and answer them as accurately as possible, please.’

4.2 Results

A response was coded as a semantic illusion response only for statements where the participant judged the statement to be true *and* gave the correct answer in the multiple choice knowledge check. For questions, a semantic illusion was considered to have occurred if the participant responded correctly to the gist of the question and then gave the correct answer in the knowledge check (i.e. if a participant had written down a correct response to a question treating that question as though there was no substituted name at all. Using the Jacob/Joseph example again: a participant had to reply with “Coat” to the question “What many-coloured garment was Jacob given by his father?” and then during the knowledge check they had to choose “Joseph” (option b) in the example above, thus showing that they had the correct knowledge.) A response was considered to have been correct only if the participant responded with “can’t say” for questions and also gave the correct answer for the multiple-choice knowledge check. For statements, participants had to indicate that they thought the statement was false and give the correct response in the knowledge check. All other responses were classed as "don't knows". The number of semantic illusion responses made for each target sentence was recorded for both the question form and the statement form of the sentence as was the number of correct responses and "don't know" responses. Then the semantic illusion rate was calculated for each individual sentence in both its forms. The same procedure was repeated for the correct response rate and for the rate of "don't know" responses.

4.2.1 Semantic illusion rate

A t-test for paired samples was carried out comparing semantic illusion rate for statements-to-be-verified to semantic illusion rate for questions for each sentence. As predicted, the mean semantic illusion rate for questions was significantly higher than that for statement responses ($t = -3.05$; d.f.= 7; $p = 0.019$; see Figure 4.1 and Table 4.1).

Figure 4.1: Mean semantic illusion rate (percent) for question format and statement-to-be-verified format.

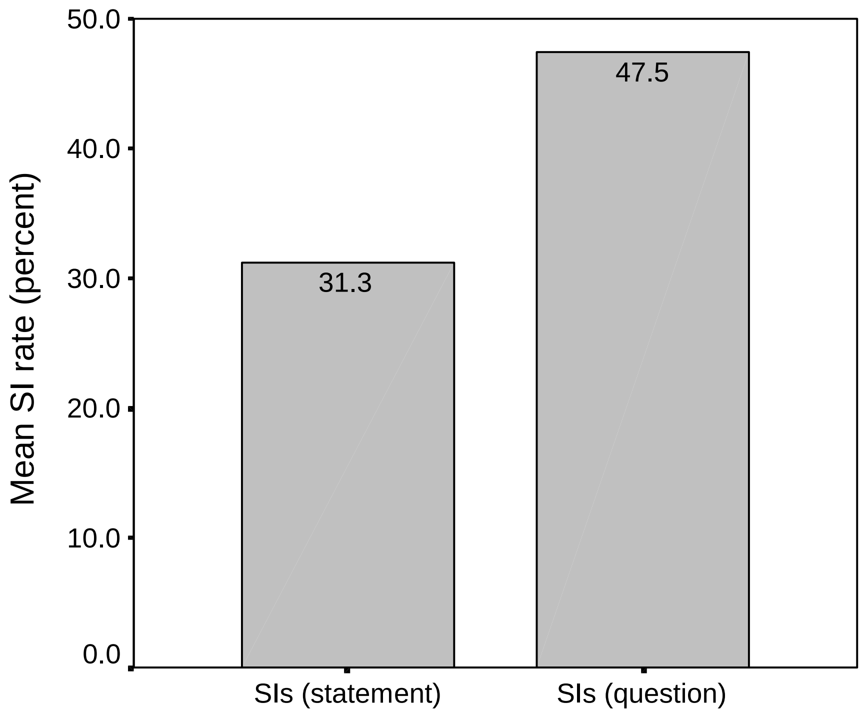


Table 4.1: Mean semantic illusion rate and standard deviations (percent) for questions and statements.

	N	Mean	Std. Dev.
Statements	8	31.3	18.9
Questions	8	47.5	14.4

4.2.2 Correct responses

A paired samples t-test was also carried out upon the rate of correct responses made for each sentence. It was found that the number of correct responses was significantly lower for questions than for statements (t= 4.57; d.f.= 7; p= 0.003; see Figure 4.2 and Table 4.2.).

Figure 4.2: Mean correct response rate for question format and statement-to-be-verified format.

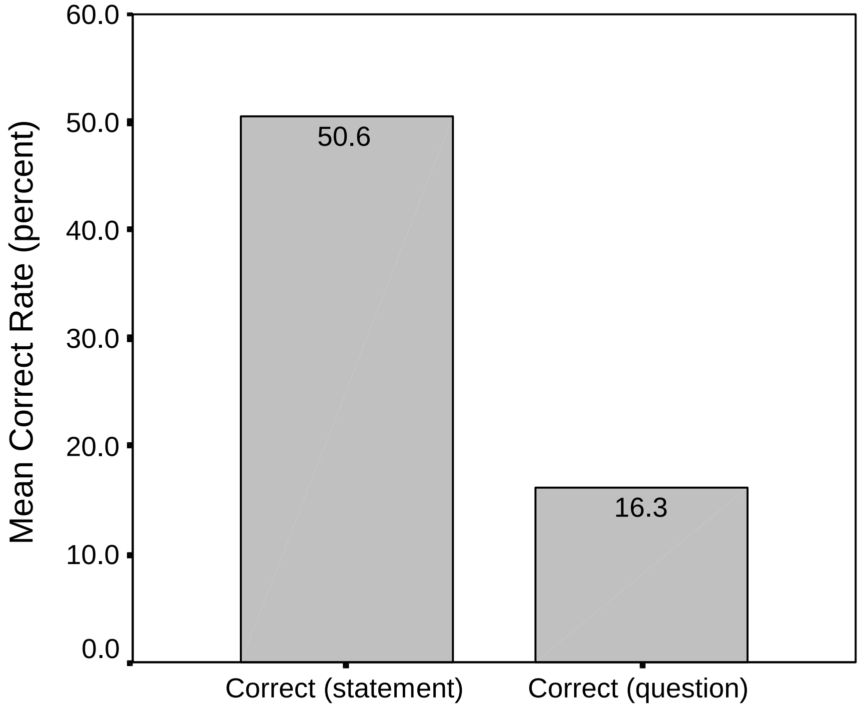


Table 4.2: Mean correct response rate and standard deviations (percent) for questions and statements.

	N	Mean	Std. Dev.
Statements	8	50.6	18.6
Questions	8	16.3	7.9

4.2.3 "Don't know" responses

The "don't know" response rates for questions and statements were also analysed, and a paired samples t-test was carried out for the sentences. As with semantic illusion responses, it was found that "don't know" responses were significantly more likely to occur for questions than for statements (t= -4.529; d.f.= 7; p= 0.003; see Figure 4.3 and Table 4.3.).

Figure 4.3: Mean "don't know" response rate (percent) for question format and statement-to-be-verified format.

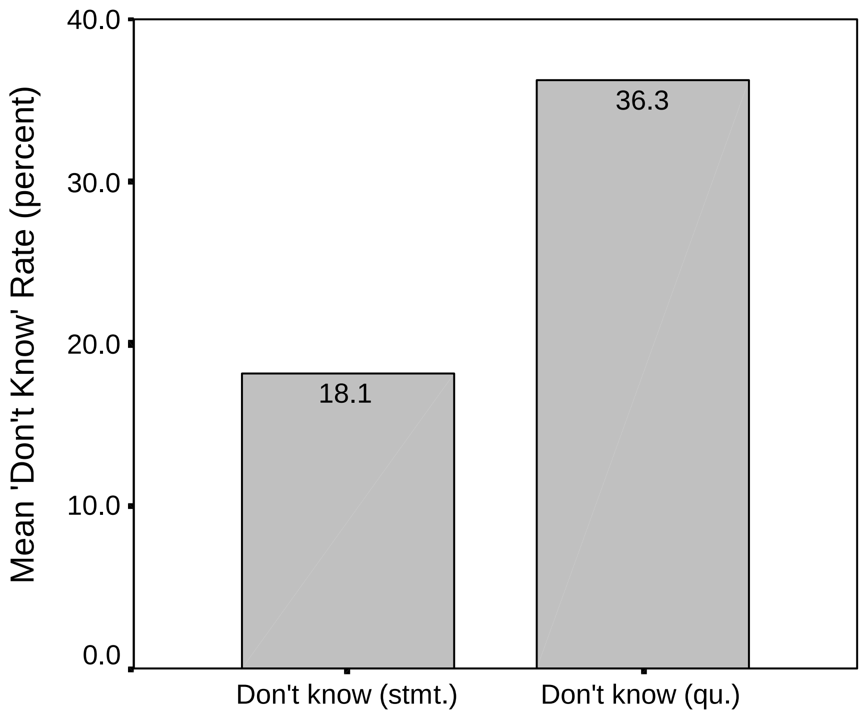


Table 4.3: Mean "don't know" response rate and standard deviations (percent) for questions and statements.

	N of sentences	Mean	Std. Dev.
Statements	8	18.1	13.1
Questions	8	36.3	12.2

4.3 Discussion

The results for the semantic illusion rates for each sentence type were in keeping with the hypothesis, suggesting that semantic illusion rates vary with the tasks participants are required to carry out. In a question-answering task, the emphasis is placed on the completion of a schema-like structure, which each participant is expected to be able to do, given that the stimuli are based on 'general knowledge'. This aspect of the task appears to take precedence over the second task set out in the instructions to monitor each question for internal consistency. In the case of the statement-verification task, the emphasis is placed upon checking the internal consistency of the schema-like structure of each statement. So essentially the difference in performance can be viewed as a question of multiple versus single task performance. This hypothesis is also corroborated by the observations about the other types of response analysed here. In the statement-verification task, the number of correct responses outnumbered both the "don't know" and the semantic illusion responses, with just over fifty percent of all target statements being correct. For the questions, the single largest group of response type is in the semantic illusion category, in fact, almost fifty percent of the responses made to target questions were semantic illusions. The next largest group of responses were the "don't knows" for target questions. This suggests that the task demands for the statement verification task are more straightforward than those for the question-monitoring task. For the question task, participants found it easy enough to respond to the gist of the question, but not to its literal form (cf. Reder and Kusbit, 1991). This may be partially due to the fact that a different type of memory search is required to complete the question answering task compared to the verification task. In the verification task, all the material making up a coherent (except for the substitution) story is present and a recognition check can be carried out. But in the question answering task, a part of the schema has to be recalled from long-term memory. This process is likely to be mediated by a content-addressable memory system, which is not dependent on exact matches for an answer to be retrieved (e.g. McElree, 2000).

4.4 Conclusions

From the results of this experiment it seems quite clear that the outcome of any semantic illusion study will be affected by the form in which the target sentences are presented, and so comparisons between semantic illusion rates in studies that have used statements as presentation form and those in which the target materials have been presented as questions have to be made with care. Since the overall performance is better for target-statements, with a significantly smaller portion of “don’t know” responses, and since participants found it easier to complete the verification task, it was decided to use only the statement verification task in further experiments.

Chapter 5:

Experiment 2: Sentence length and target word position

The fact that questions and statements-to-be-verified are apparently not processed in the same way raises the question of whether variations within each sentence format also have an effect upon semantic illusion rate. When going over the semantic illusion literature – whether the individual studies use questions or statements-to-be-verified – semantic illusion sentences are generally quite long, containing two or more separate propositions which need to be incorporated into the mental representation of the sentence being processed. Thus “Moses” took two animals of each kind, and he took them on the Ark. Or “Captain Nemo” was after a whale, that whale was white and the story of this is told in the novel *Moby Dick*. The various propositions within each sentence – except for the substituted word – form a coherent picture, which should allow the participant to build a mental representation of each sentence, which they can easily respond to. One question that arises from this idea is whether, therefore, the amount of coherent material combined into a sort of schema has an effect on the occurrence of semantic illusions.

5.1 Shorter sentences are easier to interpret

Studies of aphasics’ error rates in interpretation tasks show that the number of propositions in a sentence affects the ease of interpretation, with two propositions being more difficult to interpret than one (Caplan, Baker and Dehaut, 1985). Similarly, Waters Caplan and Hildebrandt (1987, see also section 11.1.2.1) found that interpretation of two proposition sentences was more affected as a result of the performance of a secondary task than interpretation of one proposition sentences. By analogy with these findings, it seems likely that, if the number of propositions pertaining to a given scenario is altered, the semantic illusion rate for such a scenario might change also. Thus a longer sentence might cause the semantic illusion rate to increase, whilst true and false judgements can be made quickly if only one proposition needs to be checked against background knowledge and a replaced name might be relatively easy to spot. In the case of a longer sentence, containing three or four propositions more information contributes to the overall schema, and thus there is more information with which the replaced name needs to be integrated. Semantic processing also needs to proceed in parallel with more complex syntactic processing requirements. It has already been demonstrated that participants find it harder to detect substitutions in questions containing a greater number of concepts associated with the answer (Reder and Cleeremans, 1990; Reder and Kusbit, 1991).

5.2 The substitution is more disruptive in a shorter sentence

Shafto and MacKay (2000) assumed that the target word received semantic-level priming from the sentence context, while the substituted word receives sensory priming only. If this assumption holds true, a target word would receive more priming from a longer sentence which provides more context, thus leading to a greater chance of a semantic illusion occurring. Another way of looking at it is this: there is an increased chance of a semantic illusion, because the information contained in the substituted word, which is required for the comprehension and processing of the semantic illusion sentence, contributes proportionally more to the understanding of a short sentence than to that of a long sentence. Thus processing of a short sentence will be more disrupted by the presence of a substitution. Additionally, sentences with fewer propositions are easier to interpret (Caplan, Baker and Dehaut, 1985; Waters, Caplan, and Hildebrandt, 1987). As a consequence of this line of reasoning, longer sentences are expected to lead to more semantic illusion responses than shorter ones, as the inconsistencies caused by a substitution would be more obvious in the contexts of a shorter sentence.

5.3 The position of the substituted word within the sentence

Another aspect of sentence format that has received little systematic attention is the position that a substituted word occupies with in a sentence. English is not a free-word-order language and many grammatical functions are indicated by word order, including the topic of discourse. It can be demonstrated that English listeners rely more heavily on word order cues than, for example, Italian listeners who rely more heavily on semantic cues (Bates, McNew, MacWhinney, Devescovi and Smith, 1982). In relatively fixed word-order languages like English, word order is sometimes used to denote a change in the topic of a sentence. For example, ‘left-dislocated’ sentences (Ross, 1967) mark the topic by shifting it to the initial position of a sentence (e.g. ‘That blouse, it’s simply stunning.’) Studies of the actual usage of left-dislocation show that there is a correlation between what participants are talking ‘about’ and the words appearing in that position, but the connection is not always straightforward (Geluykens, 1992). Even in longer and more complex sentences, the processing cost of non-canonical word order tends to be greater in English than, for example, in Italian (Bates, Devescovi and D’Amico, 1999).

In keeping with these pragmatic findings, Erickson and Mattson (1981) claimed that locating the semantic illusion term at the start of the sentence would bring it into focus. Since their claim is based on only two example sentences which had been turned from question form to statement-to-be-verified form, confounded with a shift in paradigm from using a time-limited computer presentation to a self-paced pencil-and-paper task, Erickson and Mattson’s results are far from conclusive. Consequently it would be of interest to investigate whether there really is an effect of word position upon semantic illusion rate in statements-to-be-verified. Other research indicates that word position does not have much effect on semantic illusions. Baker and Wagner (1987) found no significant effect of word position on semantic illusion rate in compound sentences; neither did Barton and Sanford (1993): positions of the anomalous word in their research before or after the scenario was presented had no impact on whether or not people spotted the anomaly in a sentence. Reder and Kusbit (1991) carried out a post-hoc analysis on word position and found no position effect, but this could have been due to variations in how 'good' their particular stimulus sentences were.

Van Jaarsveld, Dijkstra and Hermans (1997) set about investigating word position more systematically and found that the similarity-effect differed at different word positions, but only for a question-answering task. In a question, a dissimilar substitution presented before the scenario caused more disruption to the flow of processing than a similar substitution in the same position, or a dissimilar substitution presented after the scenario. Van Jaarsveld, Dijkstra and Hermans explained their finding by suggesting that word position had no effect in a verification task, because the entire sentence would need to be processed. Nevertheless, it was thought worthwhile to test for an effect of word position upon semantic illusion rate in a study specifically designed with this objective in mind, as there often is a pragmatic link between word position and meaning. It was hypothesised that words at the front of a sentence would often be taken to be in focus, and central to the meaning of the sentence. As a consequence it was thought likely that words at the front of a sentence would undergo more careful checks than words in the middle or at the end of a sentence.

5.4 Verification is likely to be easier for true sentences than false ones

A final consideration for Experiment 2 was the question of whether semantic illusion sentences were more difficult to judge correctly than other false statements. According to Chomsky's Standard Theory (1965), the negation of a sentence is achieved by placing a transformational marker in the sentence's deep structure, so that a negative sentence has to be processed similarly to a positive one and then the transformation is processed. Generally, negative sentences take longer to process and are less accurately recalled and processed relative to a given state of affairs than a corresponding positive sentence would be (e.g. Horn, 1989; Glenberg, Robertson, Jansen and Johnson-Glenberg, 1999). By analogy, it was hypothesised that participants would find it easier to verify true statements compared to semantic illusion statements or false statements, with a potential effect upon the rate of correct responses made for each response category.

From research on word recognition there is evidence (e.g. Rubenstein, Garfield and Millikan, 1970; Stanners and Forbach, 1973) that non-words that are not in keeping with the rules of English orthography (e.g. *xgyz*) produce very fast 'no' responses in a lexical decision task, while non-words that could be, but happen not to be, real words such as *nint* take even longer to be rejected than it takes to accept infrequent words. There is also evidence that the greater the similarity between non-words and real words, the harder it becomes to reject the non-word (e.g. Coltheart, Davelaar, Jonasson and Besner, 1977). A similar process might affect the verification of entire sentences. By analogy with the lexical decision paradigm, the false-filler statements in semantic illusion research can be regarded as like obvious non-words. They should be more readily recognisable as false than semantic illusion statements, which as 'almost true' sentences resemble plausible non-words, being closely similar to true statements with much the same content. While for this experiment, response times were not considered, it was thought that the correct response rate for each group of statements might reveal the existence of such a potential ‘sentence-similarity’ effect. It was expected that semantic illusion statements would be associated with the smallest number of correct responses after false-fillers, with true-fillers leading to the greatest correct response rate.

5.5 Method

Participants: 324 participants completed a questionnaire for this study. The age of the participants ranged from 16 to 76 years, with most participants being university students or having received university education.

Stimuli: A general knowledge ‘quiz’ consisting of 45 true/false statements was presented to each participant. Nine of the statements were target statements, which included a substituted name. The other 36 sentences were fillers constructed along similar lines to the target sentences, varying in length and the position of names within the sentence. Ten of these statements were false; the other 26 were true. Each of the target sentences could appear in one of nine different versions: target word at the front, in the middle or at the end of the sentence; and as a short, medium or long sentence. The length of a sentence was measured in propositions, so that for a short sentence the potential illusion name had to be matched only on one fact. For a medium length sentence, two facts had to be matched up; and for a long sentence, three facts had to be checked (see Table 5.1 for an example).

Each sentence was presented equally often for each potential combination of length and position and each participant saw one of each type of combination (i.e. each participant saw one sentence as front short, one as front medium, one as front long, one as middle short, etc.) The presentation order of the target sentences was varied systematically, so that each sentence variation appeared equally often in each of 9 designated target sentence slots within the questionnaire. (An example of the questionnaire used and a full list of target sentences are presented in Appendices 4 and 5.)

Table 5.1: Example of the 9 possible variations in which each semantic illusion sentence could appear.

Length→ Position↓	**SHORT**	**MEDIUM**	**LONG**
FRONT	Wellington's victory is commemorated in Trafalgar Square.	Wellington's victory is commemorated by his column in Trafalgar Square.	Wellington's victory over the Franco-Spanish navy is commemorated by his column in Trafalgar Square.
MIDDLE	In Trafalgar Square, Wellington's victory is commemorated.	In Trafalgar Square, Wellington's victory is commemorated by his column.	In Trafalgar Square, Wellington's victory over the Franco-Spanish navy is commemorated by his column.
END	Trafalgar Square commemorates the victory won by Wellington.	The column in Trafalgar Square commemorates the victory won by Wellington.	The column in Trafalgar Square commemorates the victory over the Franco-Spanish navy won by Wellington.

Procedure: Each participant was given a leaflet of two sheets of A4 paper, with the verbal instruction: "On the first sheet are 45 general knowledge statements. Read each through and respond as quickly as possible by indicating whether you think the statement is true or false by marking the appropriate response on the sheet. The statements are not intended to be difficult, but if you are not entirely sure, feel free to guess. This task is not really about what you know but about how you process sentences. When you are finished with page 1, let me know and I will tell you what to do for page 2." Page 2 was a knowledge check: the 9 targets were presented as multiple choice questions. Each such question was matched in length to the version of the target statement that each participant had seen. Participants were encouraged to respond to the knowledge check questions as accurately as possible and were told that there was no time constraint for this part of the experiment.

5.6 Results

A semantic illusion was considered to have taken place when a participant judged a target statement to be true *and* showed the correct knowledge in their response to the corresponding multiple choice question in the knowledge check. The semantic illusion rate was determined for each target statement variation. The rate of correct responses for each target statement was also recorded. A correct response required both the identification of the relevant target statement as false and a correct response in the knowledge check. All other responses to target questions were classed as don't know responses. Finally the rate of correct responses for true-fillers and false-fillers was determined, to compare the rates of correct responses for all three types of statements-to-be-verified.

5.6.1 Semantic illusion rate

A repeated measures analysis of variance was carried out on the semantic illusion data using sentences as subjects. Mauchly's Test was used to check the sphericity assumption. Mauchly's test of sphericity was found to be significant only for the sentence length effect, and therefore Pillai's Trace was used to determine the significance of this effect. Sentence length was found not to affect the semantic illusion rate (F= 0.703; d.f.= 2, 7; p= 0.527). For the word position effect and for the interaction between length and word position, Mauchly's test of sphericity was insignificant. The interaction between word position and sentence length was not significant (F= 1.413; d.f.= 4, 32; p= 0.252). The word position effect was found to be marginally significant (F= 3.455; d.f.= 2, 16; p= 0.057), and post-hoc t-tests for paired samples were carried out using the Bonferroni correction. This means that an adjusted criterion for significance was set at p= 0.0167 (see Table 5.2 for means).

Table 5.2: Mean semantic illusion rates and standard deviations (percent) for each word position.

Position	N	Mean	Std. Dev.
Front	27	29.1	2.5
Middle	27	32.2	3.0
End	27	24.7	2.7

Only the difference between middle and end positions proved to be significant (t= 2.853; d.f.= 26; p= 0.008; 0.008<0.0167, hence significant by adjusted criterion), with target words at the end of the sentence being less likely to cause a semantic illusion response. There was also a non-significant trend (t= 2.449; d.f.= 26; p= 0.021; 0.021>0.0167, hence not significant by adjusted criterion) suggesting that substitutions at the end of a sentence may also be more easily unmasked than those at the start of the sentence. But there was no significant difference between front and middle positions (t= -1.263; d.f.= 26; p= 0.218). Finally the mean semantic illusion rates were calculated for each length and word position combination (see Figure 5.1 and Table 5.3).

Figure 5.1: Mean semantic illusion rate (percent) at all different combinations of sentence length and word position.

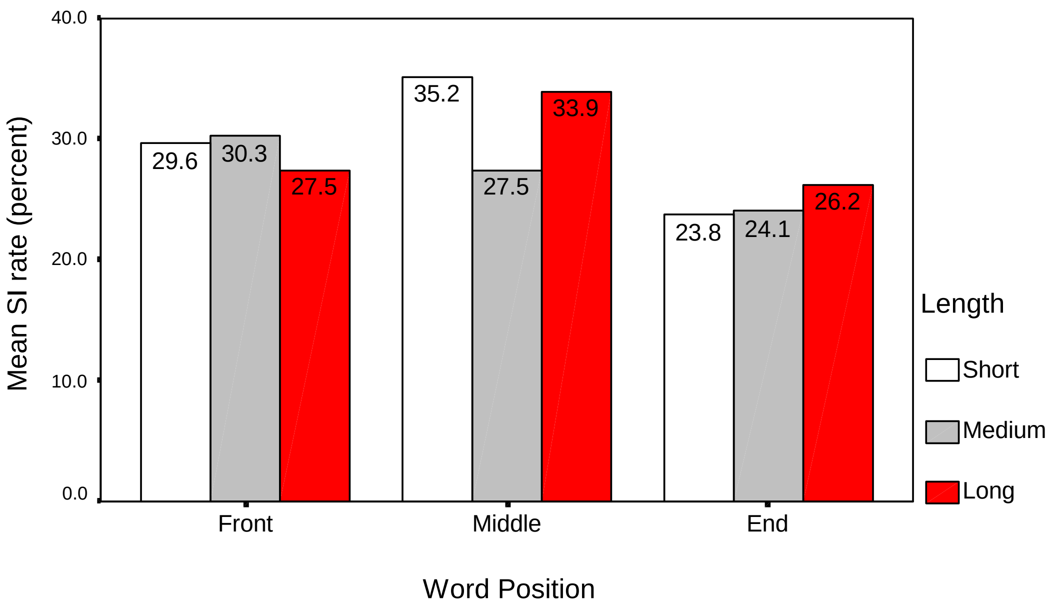


Table 5.3: Mean semantic illusion rates and standard deviation (percent) for each combination of sentence length and target word position.

Position → ↓ Length	**Front**	**Middle**	**End**
Short	29.6 (16.0)	35.2 (20.6)	23.8 (17.9)
Medium	30.3 (14.8)	27.5 (12.8)	24.1 (11.7)
Long	27.5 (9.0)	33.9 (11.6)	26.2 (13.3)

5.6.2 Correct responses to semantic illusion statements

In order to compare the results from this experiment with the results of previous experiments described in the literature, correct responses to semantic illusion questions were also analysed in a repeated measures analysis of variance using sentences as subjects. The sphericity assumption was not violated for any of the effects measured. There was no effect of either sentence length (F= 1.129; d.f.= 2, 16; p= 0.348), or word position (F= 0.505; d.f.= 2, 16; p= 0.613); and there was no significant interaction between sentence length and word position (F= 1.074; d.f.= 4, 32; p= 0.386). The mean correct response rates were calculated for each length and word position combination (see Table 5.4).

Table 5.4: Mean correct response rates and standard deviations (percent) for each combination of sentence length and target word position.

Position → ↓ Length	**Front**	**Middle**	**End**
Short	59.4 (16.6)	53.0 (19.3)	65.8 (17.1)
Medium	60.2 (15.8)	63.0 (11.4)	64.8 (12.5)
Long	60.5 (11.4)	58.7 (12.3)	62.3 (16.0)

5.6.3 Correct responses for fillers and targets

A repeated measures analysis of variance was carried out upon the correct response rate data for correct target responses, correct true-filler responses, and correct false-filler responses using participants as subjects. A true-filler response consisted of a participant correctly judging a true statement to be true and a false-filler response consisted of a participant judging a false statement to be false.

Mauchly's test of sphericity was found to be significant (W= 0.836; d.f.= 2; $p<0.001$) and Pillai's Trace was used to determine the significance of the analysis of variance. The analysis of variance was found to be significant (F= 574.099; d.f.= 2, 322; $p<0.001$), and a series of post-hoc comparisons were carried out, using t-tests for paired samples with a Bonferroni correction. As predicted, the highest rate of correct responses was achieved in the true-fillers category, but false-fillers and targets led to equal correct response rates (t= 0.690; d.f.= 323; p= 0.491). There were significantly more correct responses made for true fillers than for false fillers (t= 31.342; d.f.= 323; p= 0.001), and significantly more correct responses for true fillers than for targets (t= 22.861; d.f.= 323; p= 0.001; see Table 5.5 for means).

Table 5.5: Mean correct response rates and standard deviations (percent) per participant for the three types of statement.

Type of statement	N	Mean	Std. Dev.
True-fillers	324	89.8	9.3
False-fillers	324	59.9	14.9
Targets	324	59.0	24.3

5.7 Discussion

It is surprising that there is no effect of sentence length upon the semantic illusion rate, as it would seem that a longer internally consistent sentence (except for the target word) would match more closely overall to a schema stored in memory because a greater number of propositions related to the schema would match perfectly, leading to greater goodness of global fit (Barton and Sanford, 1993). In a shorter sentence where only one proposition needs to be matched up with the target, the overall goodness of fit would presumably be proportionally smaller. Reder and Kusbit's (1991) partial matching hypothesis of semantic illusions essentially describes such a theory at the word level. According to this theory semantic illusions occur because the partial match process is sensitive to the similarity of all the concept words in the sentence to the representation of the relevant knowledge stored in memory. If the features of the target word and the correct but substituted word overlap, the semantic illusion word tends to go unnoticed. A similar matching process could be envisaged for the entire schema accessed by the statement-to-be-verified in the experimental task: the more propositions in the sentence match the stored schema, the more likely the semantic illusion word is to be overlooked.

The fact that there is no effect of length suggests that once a schema is accessed through a sentence referring to information stored in the schema, the subsequent judgement of truth is relatively automatic and independent of the details of the sentence. This suggestion is in keeping with other observations made about semantic illusion responses: they are made very rapidly and with great confidence (e.g. Erickson and Mattson, 1981); often the participant will realise they have just made a mistake immediately after they have given the semantic illusion response. Automatic actions are generally characterised as stereotyped, rapid, hard to inhibit (e.g. Shiffrin and Schneider, 1977).

It has previously been shown that information from different sentences is rapidly integrated into a schema-like memory structure, to the extent that participants will recognise the gist of the sentences in recognition tasks, but cannot distinguish between sentences they have actually seen and those that are new but in keeping with the semantic content of the sentences seen (Bransford and Franks, 1971). Given this 'abstraction of linguistic material', it is possible to explain the lack of an effect of length upon semantic illusion rate by thinking of the fact that the content of a target sentence has to be familiar to the participant in order for an semantic illusion to occur at all. One can assume that the semantic content of the target sentence has previously been integrated into long-term memory and is there contained within a schema, which can be accessed by shorter sentences as well as by longer ones. Thus the longer sentences might not be very different in terms of processing demands, because the information contained within such sentences has already been integrated at a previous point in time.

Word position is associated with an effect that was not anticipated: target words at the end of a sentence are spotted significantly more accurately than target words at the beginning or embedded in the middle of a semantic illusion sentence. A possible explanation for this effect could be that words at the end of a sentence are subject to a recency effect (e.g. Glanzer and Cunitz, 1966). When a sentence is processed the actual words are rapidly forgotten (e.g. Sachs, 1967), but the semantic content is retained to a greater degree. This appears to be associated with the shift from building one sub-structure to starting with the next (Gernsbacher, 1985). In the case of a word at the end of a sentence it might be the case that the word has not completely vanished from working memory by the time a true/false judgement is made. Thus this particular word is taken at face value as opposed to having already been processed for meaning by the time the judgement is made, and hence the 'intruder' stands a better chance of being discovered. It could also be that the substituted word at the end of a sentence stands out, because the placing of a name at the end of a sentence usually requires the use of slightly unusual and more complex syntactic structures, such as phrasing the sentence in the passive voice (Chomsky, 1965), which might draw attention to the name in question.

The comparison of correct response rates showed that participants found it generally easier to verify true statements. This is in keeping with findings showing that processing negatives is more difficult than processing positive statements. The proposed analogy between target statements and word-like non-words did not hold and participants performed equally well on false-fillers verification as they did on target verification. On the whole, however, both types of false statements led to a much smaller correct response rate than the true statements, suggesting that people have a fundamental problem at dealing with false information.

5.8 Conclusions

Even though there were some (marginally significant) variations in semantic illusion rate observed in this study, the semantic illusion rates fall within the usual range for semantic illusion responses reported in previous studies. It appears that Brédart and Modolo (1989) were justified in their criticism of Erickson and Mattson's claim that putting the substituted name at the front of the sentence would bring it into focus: target words at the front of a semantic illusion statement led to no fewer semantic illusion responses than target words embedded in the middle of a statement. The length of a semantic illusion statement was also shown not to affect the semantic illusion rate, suggesting that not much information is required for a schema to be successfully activated in long-term memory. This is in keeping with Reder and Kusbit's partial match hypothesis (see section 2.8.). Another point arising from this experiment comes from the comparison of the correct response rate for false filler statements compared to targets. It would appear that part of what affects the processing of semantic illusion statements is to do simply with the fact that these statements are false, and as such harder to verify, as a falsehood can only ever be established on the basis of absent confirmatory evidence. This finding raises the question of how semantic illusion sentences might differ from other forms of discourse and how they fit into ‘normal’ language usage.

Chapter 6:

Experiment 3: Semantic illusions in German

Many previous semantic illusion studies have been carried out in different languages, including English (Erickson and Mattson, 1981; Reder and Kusbit, 1991; Reder, Kamas and Ayers, 1996; Shafto and MacKay, 2000; Hannon and Daneman, 2001), French (Brédart and Modolo, 1989; Brédart and Docquier, 1990), and Dutch (van Oostendorp and de Mul, 1989; van Oostendorp and Kok, 1990; van Jaarsveld, Dijkstra and Hermans, 1997). This observation led to further considerations of the potential effect of word order upon semantic illusion rate. Even though little effect of word-order *per se* has been found, there is evidence that at least in Dutch, the similarity effect (see section 2.5.2) changes in magnitude for semantic illusion questions depending on whether the substitution is more or less similar to the target it replaces (van Jaarsveld, Dijkstra and Hermans, 1997). Furthermore, it was found in Experiment 2 (see chapter 5), that substitutions at the end of a statement are marginally easier to detect.

However, the three languages used in semantic illusion research to date are all more 'configurational' languages that make much use of word order as a syntactic indicator. In less 'configurational' languages, such as German, word order is of smaller importance, and inflections are used to convey syntactic information. For example, subject and object in German are determined by the case endings attached to the relevant words in combination with specific matching articles, and not only by the order in which they appear in a sentence. For example, MacWhinney, Bates and Kliegl (1984) showed that English speakers relied heavily on word order in a task in which they were required to determine the actor in simple transitive sentences (e.g. “The dog grabs the pencil.”) German speakers, on the other hand, relied more on agreement between the verb and the subject, which is usually indicated by a system of case markings (e.g. “Der Hund greift den Stift.” [transliterated: “The {masculine singular nominative article} dog grabs the {masculine singular accusative article} pencil.”]) As a result of this, it was thought that manipulations of word order in German sentences would be likely to be interpreted as a way of signalling information about the focus of the sentence, with words near the front being perceived to be in focus. Indeed, it can be shown that word order affects processing of sentences starting with transitive verbs in a condition where semantically related distractors are presented, while the same interference is not observed when the transitive verb is in a non-initial position (Schriefers, Teruel and Meinhausen, 1998). There is also neuropsychological evidence from the study of event-related brain potentials (ERPs) recorded while native speakers of German read sentences in which the word order had been systematically varied. The pattern of ERPs shows that sentences deviating from the usual word order of subject, indirect object, direct object were more difficult to process and led to different patterns of brain activity (Roesler, Pechmann, Streb, Roeder and Hennighausen, 1998).

As a result of these findings, a study was designed to investigate the effect of substitution word positioning upon semantic illusion rate in a German statement verification task. It was hypothesised that statements with a substitution near the beginning of a sentence would lead to fewer semantic illusions for the following two reasons: 1) names nearer the beginning of a sentence would tend to be seen to be in focus, and 2) smooth flow of processing of German sentences was thought to be more likely to be disrupted, when difficulties (inconsistencies) were encountered nearer the start of a sentence.

6.1 Method

Participants: 66 participants educated at least to the German equivalent of 'A'-level participated in this study on a volunteer basis.

Stimuli: The nine statements-to-be-verified used in Experiment 2 were translated into German. Three sentences were excluded as they were found to use knowledge that was too specifically 'British' when piloted. The remaining six sentences were manipulated so that they could appear in either of two forms: 'first word substituted' (e.g. "Ali Baba befreite im Märchen in *1001 Nacht* den Geist aus der Wunderlampe." [transliterated: "Ali Baba freed in the tale in *1001 Nights* the genie from the magic lamp."]) or 'substituted word nearer to end' (e.g. "Der Geist aus der Wunderlampe im Märchen in *1001 Nacht* wurde von Ali Baba befreit." [transliterated: "The ghost out of the magic lamp in the tale in *1001 Nights* was by Ali Baba freed."]) The word positions could not be matched exactly with those used in the English language experiment, as it is very difficult to create an idiomatic German sentence of the kind used as semantic illusion sentences in experiment 2 that ends in a name, without putting strong emphasis on that name. Once the targets were prepared they were integrated with 24 filler statements-to-be-verified constructed along the same lines as the targets (17 fillers were true statements and 7 were false) to form a general knowledge 'quiz' consisting of a total of 30 true/false statements to be presented to each participant. (Examples of the questionnaires used are presented in Appendix 6.)

Procedure: Each participant was handed an A4 sheet of paper with 30 statements-to-be-verified printed on it. Instructions were printed at the top of the sheet, asking participants to read each sentence as quickly and carefully as they could and then to make a mark in the allocated column on the paper to indicate whether they thought the sentence was true or false. When participants had completed the 30 judgements in their own time, they were handed a separate sheet on which 6 knowledge check questions were printed with multiple-choice answers to be indicated by circling or otherwise marking. Instructions for the knowledge check were printed at the top of the page. Participants were encouraged to complete this part of the questionnaire as accurately as possible and told that there was no time limit for the completion of this task.

6.2 Results

As in the previous studies, a response was coded as a semantic illusion only if the participant had judged the given statement to be true *and* given the correct answer in the multiple choice knowledge check. The number of semantic illusion responses for substituted words in each of the two positions was recorded for each target statement. In order to compare the results of this experiment to the results of previous studies described in the literature, the rate of correct responses to target questions was also recorded and analysed. A response was considered to be true if the participant judged the statement as false *and* gave the correct answer in the knowledge check. The rate of correct responses for each participant was also recorded for targets and both types of filler statement (true fillers and false fillers).

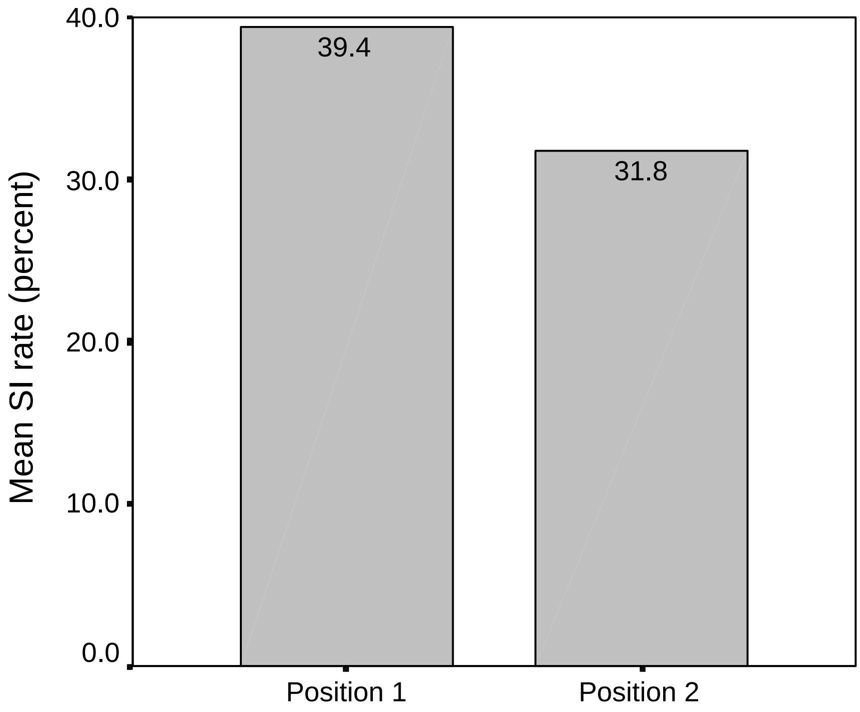
6.2.1 Semantic illusion rate

A t-test for paired samples was carried out to compare the semantic illusion rate for statements with the substituted word at the front of the sentence (position 1) with that for statements with the substitution nearer the end of the sentence (position 2), using the statements as subjects. There was no significant difference between the mean semantic illusion rate in positions 1 and 2 ($t = 1.480$; d.f.= 5; $p = 0.0995$, one-tailed; see Table 6.1 and Figure 6.1.)

Table 6.1: Mean semantic illusion rate and standard deviations (percent) for substitutions in position 1 and position 2.

Position of substitution	N	Mean	Std. Dev.
Position 1	6	39.4	11.3
Position 2	6	31.8	13.4

Figure 6.1: Mean semantic illusion rate (percent) for position 1 (at the start of the statement) and position 2 (nearer the end of the statement).



6.2.2 Correct responses to semantic illusion statements

A t-test for paired samples was carried out to compare the correct response rate for statements with the substitution at the front of the sentence (position 1) to the correct response rate for statements with substitutions nearer the end of the sentence (position 2). Again, statements were used as subjects. There was a non-significant trend in the opposite direction to that predicted, with more correct answers given to semantic illusion statements with the substitution in position 2 (t= -1.909; d.f.= 5; p= 0.0575, one-tailed; see Table 6.2. for means).

Table 6.2: Mean correct response rate and standard deviations (percent) for substitutions in position 1 and position 2.

Position of substitution	N	Mean	Std. Dev.
Position 1	6	48.5	7.7
Position 2	6	57.7	14.6

6.2.3 Correct responses for fillers and targets

The rate of correct responses in each statement category was determined and compared, and a repeated measures analysis of variance was carried out upon the correct response rate data for correct target responses, correct true-filler responses, and correct false-filler responses, using participants as subjects. A true-filler response consisted of a participant correctly judging a true statement to be true and a false-filler response consisted of a participant judging a false statement to be false, as in Experiment 2.

Mauchly's test of sphericity was significant (W= 0.805; d.f.= 2; p= 0.001) and so Pillai's Trace was used to determine the significance of the analysis of variance. The analysis of variance was significant (F= 92.253; d.f.= 2, 64; p= 0.001), and a series of post-hoc comparisons were carried out, using t-tests for paired samples with a Bonferroni correction (adjusted criterion for significance p= 0.0167). The highest rate of correct responses was observed for true fillers, but false fillers and targets led to equal correct response rates (t= 2.239; d.f.= 65; p= 0.029). There were significantly more correct responses made for true fillers than false fillers (t= 11.800; d.f.= 65; p= 0.001), and significantly more correct responses for true fillers than for targets (t= 10.327; d.f.= 65; p= 0.001; see Table 6.3 for means).The pattern of results followed that of the results of the English language study carried out in Experiment 2 (see Chapter 5, and see Table 6.4 for means).

Table 6.3: Mean correct response rates and standard deviations (percent) per participant for the three types of statement.

Type of statement	N	Mean	Std. Dev.
True-fillers	66	84.9	9.9
False-fillers	66	60.2	14.5
Targets	66	53.2	24.6

Table 6.4: Mean correct response rates (percent) for the three types of statement for Experiment 2 (English statements) and for Experiment 3 (German statements).

Type of statement	Mean (English)	Mean (German)
True-fillers	89.8	84.9
False-fillers	59.9	60.2
Targets	59.0	53.2

6.3 Discussion

The results were not in keeping with the hypothesis that a substituted word at the beginning of a semantic illusion sentence would be more easily detected in a German sentence than in an English one, as in German word order was thought to be an indicator of sentence focus, but not of the syntactic function of the word within the sentence. Nor did a substitution near the beginning of a statement disrupt the flow of processing in such a way that more semantic illusions were detected. Instead, the direction of the difference in semantic illusion rate goes in the same direction as the findings for the English sentences, with slightly more semantic illusions occurring for sentences which start with a substitution. It is possible that the way in which sentences were manipulated in order to place the target word in either position 1 or position 2 had an effect on semantic illusion rate. In order to place the target word in the desired position, the usual word order (in German) of subject, indirect object and direct object has to be varied. It has been shown that sentences deviating from this canonical word order are harder to process (Roesler, Pechmann, Streb, Roeder and Hennighausen, 1998). Hence it is possible that statements with the substitution (which in this study was always a name) in position 1, where it occupied the role of the subject of the sentence, were easier to process overall, and thus less disruption was caused. Conversely, statements with the substitution in a non-initial position, might generally have been harder to process overall, allowing the substitution to cause greater disruption of the flow of processing, and consequently the substitutions were detected more easily.

When the overall rate for correct responses for each type of statement used in this experiment was examined, the same pattern of results emerged as in Experiment 2: participants found it generally easier to verify true statements. In fact the actual correct response rates are very close to those in the English language study. This suggests that the extent to which these languages are 'configurational' does not affect semantic illusions.

6.4 Conclusions

As in van Jaarsveld, Dijkstra and Hermans' (1997) and Barton and Sanford's (1993) findings, word position was not found to have a significant effect upon the semantic illusion rate for semantic illusions in German. It therefore seems that even in a language that is more highly inflected than English, the semantic illusion phenomenon follows much the same pattern that has been reported in previous studies, and it appears that the semantic illusion phenomenon is – at least partially – language independent.

From the results of Experiments 2 and 3 it appears that the actual order or number of words used in a semantic illusion sentence has little to do with the likelihood of a substituted target word being detected. So what is it about semantic illusion type sentences that makes them what they are? Are they in some way different from other sentences that are encountered in day-to-day processing? The next chapter is concerned with the question of how semantic illusion type sentences fit into ordinary discourse.

Chapter 7:

Semantic illusions as part of normal discourse

When looking at semantic illusion research, it quickly becomes apparent that the sentences used to elicit semantic illusions are restricted to a definite format. They have to contain information already known to the reader in order for the illusion to occur. Semantic illusion sentences also tend to be fairly long and to contain redundant information, presumably so that a coherent mental representation can be established, in which the substituted word fills a specific role-slot sufficiently well not to be noticed. The question that arises out of such observations is how sentences which are constructed identically to semantic illusion sentences – but lack the substituted target word – compare to other sentences used in the kind of written discourse that is seen on a day-to-day basis. Even though written discourse varies widely, so that it is almost impossible to establish whether in some absolute sense semantic illusion type sentences (referred to in this chapter as SITS) are different from 'other' sentences, it would be of considerable interest if SITS can be shown to resemble each other more than they resemble other types of sentences.

In order to compare SITS with other sentences, participants were asked to rate/classify each of a set of fifteen sentences on a series of semantic description tasks (see Table 7.1 for an overview). Eight of the sentences used were SITS based upon examples of semantic illusions used in the literature but presented in their 'correct' version. For example, the statement "Noah took two animals of each type on the Ark" was used as

Erickson and Mattson used a very similar sentence in their original experiment. The control sentences were chosen at random from a number of different sources, such as novels, textbooks, academic journal articles and newspapers. Each task was designed to be relatively intuitive and to elicit participants' subjective evaluation. Instructions were supplied for each task, and for the rating scales examples of the extremes of the scales were provided.

Table 7.1: Overview of the sentence description tasks to be carried out by the participants.

1.	*What kind of text is it?* ***(categorisation task)***
2.	*How great is the quantity of background knowledge required to understand the sentence?* ***(rating scale)***
3.	*How important is it to know the context of the sentence in order to understand it?* ***(rating scale)***
4.	*How specific are the meanings of the individual words used in the sentence?* ***(rating scale)***
5.	*How familiar is the proposition stated in the sentence?* ***(rating scale)***
6.	*How grammatical is the sentence?* ***(rating scale)***
7.	*How complex is the sentence?* ***(rating scale)***
8.	*How formal is the style of the sentence?* ***(rating scale)***
9.	*How natural* ***(i.e. likely to be encountered in normal discourse)*** *is the sentence?* ***(rating scale)***
10.	*How literally is the sentence meant?* ***(rating scale)***
11.	*Underline the part of the sentence that the main focus is placed upon.*

7.1 The categorisation task

Participants were first asked to classify each sentence by deciding which semantic/ ontological category of text it might belong to. People approach written material with certain genre-specific expectations and processing strategies acquired during past encounters with other examples of text from the same category. It can be shown that people remember more surface information about the same text if the text's ostensible genre is fiction, than they do if the text is presented as news (Zwaan, 1994). Similarly a different perspective can easily affect the way in which a participant interprets and remembers a 'story'. If, for example, a story about a house were to be read with a view to breaking in as opposed to buying, it can be shown that participants are likely to recall different aspects of the same text according to the context they were initially provided with (Anderson and Pichert, 1978). Hence the semantic/ontological category to which a participant assigns a sentence is likely to play an important role in the interpretation of the sentence. This task was included to see how participants would class the SITS: would they all be seen as belonging to a certain category, and if so, which one? From the observations made in semantic illusion research one could assume that SITS would generally be seen as 'factual', snippets of knowledge usually known to each participant, and thus might be processed in a similar way as news items, with little attention to surface details (cf. Zwaan, 1994).

7.2 Background and context

The first two rating scales were concerned with the question of background and context. 'Background' was defined as the amount of extraneous knowledge required to comprehend the ideas within a sentence. In this chapter the term 'background' is used to refer to 'that which is assumed or taken for granted'. This idea is often referred to as 'pragmatic presupposition' (e.g. Stalnaker, 1977 [1974]). Specifically, 'background' here denotes information that is assumed by the speaker/writer to be shared by them and the hearer/listener (e.g. Jackendoff, 1972).

'Context', by contrast, is used to refer to the specific context in which a sentence might be encountered and its potential effect upon the interpretation of that sentence. Thus context is defined very narrowly in this chapter, referring only to the immediate situation in which each sentence might be encountered. Specifically, the notion of context here is defined negatively in terms of how 'understandable' a sentence would be if it appeared on an anonymous postcard (cf. 'null-context'; Katz, 1977).

Another way of describing the difference between background and context in this chapter would be to think of background as analogous to ‘global’ or ‘discourse’ setting, while the word ‘context’, is analogous to ‘local’ or ‘sentence’ setting. Both background and context were considered to be of relevance because there is a lot of variation in how much background knowledge different types of written discourse presuppose of their readers. There is also evidence that global context has more effect on naming times in lexical processing than local context (Hess, Foss and Carroll, 1995). The distinction between background knowledge required and the effect of context is also of interest, as semantic illusion sentences appear to require a certain level of background knowledge, but are relatively self-contained apart from that, and thus have meaning independently of the specific context in which they may be found. It was assumed that SITS would therefore score relatively low ratings on both these scales.

7.3 Specificity of individual words

The third scale was concerned with how participants perceived the specificity of individual words within each sentence. This referred to the extent to which a given word was used in a restricted, context-specific and unambiguous way. For example, depending upon context the meaning of the word “paper” can refer to anything made of that material, but also to a piece of academic writing. This is of relevance to processing of semantic illusion sentences, as it can be shown that the central meanings of words (e.g. “firm” as in “hard” as opposed to “firm” as in “strict”) are easier to process (Williams, 1992). The notion of the specificity of individual words is also closely related to the notion of context as defined above and it requires participants to look individually at each word within a sentence to see how context-free or context-dependent each word is. Participants were asked to rate each sentence according to how specifically they felt the words were being used. The assumption was that ambiguously used words would require more processing in order to extract meaning from the sentence (cf. ‘garden path’ sentences: sentences such as *The horse raced past the barn fell* that leads a reader down a ‘garden path’ to a momentarily wrong conclusion (e.g. Crain and Steedman, 1985; Milne, 1982). It was hypothesised that SITS would be considered to be particularly specifically phrased compared to other sentences, using central meanings of the words involved, therefore making SITS appear easier to comprehend, as the meaning of the individual words used would be easily accessible.

7.4 Ease of processing

Semantic illusions only occur under circumstances where participants already know the information contained within the sentence being processed. Therefore a fourth scale was constructed asking participants to rate the ‘familiarity’ of the ideas expressed in each sentence. ‘Familiarity’ here refers to the participant’s specific knowledge of the ideas expressed in each sentence rated and is as such related to the background required to understand the sentence. A familiar sentence was thought to contain a number of pragmatic presuppositions that are shared by the participants, and that the participants are likely to believe are shared by other people. This is a similar view to that espoused in Koriat’s (1993) model of the feeling of knowing. The model predicts that feeling of knowing will be higher for questions that many people know the answer to, as more partial information would usually be available for these questions. It was thought that most forms of written discourse would be used to convey novel information, while familiar information would generally be assumed rather than spelled out as in SITS, and thus SITS would be rated as more familiar than control sentences.

In order to establish whether there was something about the way in which SITS are constructed that participants found inherently odd or even just noteworthy compared to how control sentences were constructed, a fifth rating scale was concerned with the perceived ‘grammaticality’ of each sentence. Native speakers of a language are able to make judgements about the ‘well-formedness’ of sentences intuitively in so-called ‘grammaticality judgements’, and neuropsychological evidence shows that people are very sensitive to grammatical violations (e.g. Meyer, Friederici and von Cramon, 2000). Grammaticality judgements reflect the linguistic knowledge of an individual and are part of that person’s linguistic competence (cf. Chomsky, 1965). There is empirical evidence that ungrammatical sentences lead to longer response times compared to grammatical sentences (Vos, Gunter, Kolk and Mulder, 2001), suggesting that they are more difficult to process. It was hypothesised that SITS would be regarded as inherently odd in their grammatical construction as a result of the built-in redundancy of information, and thus it was thought that they would be rated as less grammatical and therefore harder to process than control sentences.

Another point of interest is the perceived complexity of SITS. Complexity in sentence processing is made up of syntactic complexity (to do with the way in which a sentence is structured, e.g. Chomsky, 1957) and of semantic complexity (to do with the number of inferences required to comprehend the sentence). SITS are usually quite lengthy, often consisting of a number of different but highly familiar sub-clauses. It was thought that perceived complexity – as rated by the participant – would be highly likely to be a combination of syntactic and semantic components, since semantic cues play an important role in comprehending syntactically complex sentences. For example, Schlesinger (1968) demonstrated that nested (repeatedly embedded) sentences are much easier to process when semantic cues are available. For example, participants found it much easier to grasp the content of a sentence like “This is the hole, that the rat, which our cat, whom the dog bit, made, caught” as compared with a sentence with minimal cues about which noun goes with which verb, such as “This is the boy, that the man, whom the lady, whom our friend saw, knows, hit”. This is true, even though the cues in the rat-cat-dog-example are incongruous, and participants nearly always give an interpretation in line with semantic expectations and not with correct syntactic analysis, much as seems to happen in semantic illusion sentences. Semantic illusion sentences were therefore expected to be perceived as simple rather than complex, as they contain much highly familiar material. Thus the interpretation of the sentence would be in line with semantic/referential expectations – a strategy that participants can be shown to employ when resolving structural ambiguities in garden path sentences (Ni, Crain and Shankweiler, 1996).

7.5 Style of the sentence

The last three rating scales were devised to determine how the wording of SITS was generally perceived in comparison with that of other sentences found in written discourse. Participants were asked to rate each sentence for how formally it was phrased in terms of the vocabulary and syntactic structures used ('formality'). The term 'formality' was used to refer to the linguistic resources that speakers have at their disposal to mark their attitude towards the people they communicate with. This is also sometimes called the 'register' of an utterance (e.g. Brown and Gilman, 1961; Lyons, 1977) and it refers to the way in which 'the language we speak or write varies according to the type of situation' (Halliday, 1978). In certain situations, such as formal meetings, or for certain types of language use (e.g. report-writing versus writing a note to a close friend) a more formal register is needed and expected (e.g. Thomas, 1995). It can be shown empirically that some sentence structures are seen to be more formal than others and will be attributed to formal rather than informal listeners (Levin and Garrett, 1990). It was thought that more formal sentences would be seen to convey information with greater authority and would thus be less likely to be questioned. Following this line of reasoning, it was hypothesised that SITS would be regarded as more formal than other sentences and would generally be taken as read.

Participants were also asked to rate how likely they thought they might be to come across each sentence in normal discourse (‘naturalness’). People are very good at judging the naturalness of conversational structure, choosing original naturally occurring conversations over reconstructions of the same conversation as most natural in a naturalness judgement task (Goldthwaite, 1997). This ability is likely to be due to people’s implicit knowledge about conversational pragmatic structure. The naturalness of a sentence is connected with its ease of use – more natural sentences being easier to process (e.g. Liberman, 1995). Hence it was hoped that this rating scale would provide a measure of the ease with which participants could imagine using each of the sentences. It was hypothesised that SITS would seem rather artificial, due to the degree of informational redundancy in each sentence, and thus be rated to be less natural.

The last rating scale was concerned with how literally each sentence was perceived to be meant (‘literalness’). The basic assumption about a sentence is that it is meant literally, but in language use, many utterances perform an action over and above the meaning of the words used. For example, in the exchange “Let’s go to the movies tonight” – “I have to study for an exam” the second sentence primarily performs the function of rejecting the proposal of the first sentence (e.g. Searle, 1975). Literalness is closely related to the familiarity scale, as non-literal uses of language – in which the words used have taken on a meaning over and above their lexical definitions (e.g. legal, military or religious language) – implies that little or no attention is paid to individual words in such an utterance. Literally meant sentences would be expected to take longer to read (cf. Cronk and Schweigert, 1992; Cronk, Lima, and Schweigert, 1993). It was thought that SITS would generally not be seen to be performing another function over and above conveying the meaning of the words within the sentence, and thus they were expected to be rated as more literal than sentences from other sources.

7.6 The focus determination task

The final task that participants were asked to perform was to underline what they perceived to be the focal point of each of the fifteen sentences. This task was included to build on the work by Erickson and Mattson (1981) and Brédart and Modolo (1988). It was assumed that this task would help to establish if substituted words in SITS were generally thought to be in focus. The hypothesis was that SITS would have more clearly defined focus points than most control sentences, and that these focus points would generally not coincide with the position of the target word.

7.7 Method

Participants: 68 volunteer participants with at least undergraduate level education completed self-timed questionnaires consisting of the 11 tasks described above to be carried out upon 15 sentences. The participants were aged between 18 and 45 years with a mean age of 22 years.

Materials: Fifteen sentences were rated: 8 SITS and 7 control sentences. The 8 sentences that were used as SITS were based upon examples of SITS used in the literature. One sentence was also based upon the style of questions of the experiment described in Chapter 4 (see Table 7.2).

Table 7.2: SITS used and the experiments in which similar SITS were used for semantic illusion research (the potential target word is underlined in each sentence).

Sentence:	Based upon:
1. Noah took two animals of each type on the Ark.	Erickson and Mattson, 1981; experiment 2; substitution: Moses. (N.B. the original sentence reads “two animals of each *kind*” – the replacement of “type” for “kind” is accidental.)
2. Bloodletting, generally accomplished with the aid of leeches, was thought to remove “poisons” from the blood.	Baker and Wagner, 1987; experiment 1, subordinate clause condition; substitution: rats
3. It was President Kennedy, who was killed in Dallas in 1963.	Brédart and Modolo, 1988; Condition 1; substitution: Luther King
4. Snow White was sheltered by seven dwarfs before marrying her prince.	Brédart and Docquier, 1989; substitution: Cinderella
6. In the October revolution of 1917 the Bolsheviks took under the leadership of Lenin the power in Russia.	van Oostendorp and de Mul, 1990; substitution: Stalin (high-related) or Marx (low-related)
7. In what mythology was Venus known as the Goddess of Love?	Reder and Kusbit, 1991; substitution: War
9. Inspector Morse who does his policing in Oxford was invented by Colin Dexter.	Sentence length and word position study (see Chapter 4); target at start; substitution: Inspector Clouseau (N.B. this example was not actually used)
13. The archaeologist Schliemann discovered the ruins of the ancient city of Troy.	Sentence length and word position study (see Chapter 4); target at end; substitution: Pompeii (N.B. this example was not actually used)

The 7 control sentences were chosen at random from novels, newspapers, academic journal articles, and textbooks (see Table7.3).

Table 7.3: Control sentences used and the type of printed material they were taken from.

Sentence:	Taken from:
5. Calcium phosphate is a mineral salt and is the principal mineral constituent of bones and teeth.	Textbook
8. The future of Barclays, one of Britain’s biggest banking groups, was thrown into doubt yesterday with the shock resignation of its chief executive, Martin Taylor.	Newspaper
10. This particular pool of light moving in a mesmeric manner backwards and forwards picked out from time to time a long red island of spilt wine.	Novel
11. Subjects averaged 47 seconds longer to name ink colors of incongruent words than solid-color squares.	Academic journal article
12. Of course a certain number of scientists have to go mad, just to keep the tradition alive.	Modern novel
14. These results are consistent with the current literature on focalization and question the Erickson and Mattson (1981) claim that the Moses illusion is not dependent on a misdirection of focus.	Academic journal article
15. Away ran the girls, too eager to get in to have time for speech.	Novel

Procedure: Participants were given questionnaire booklets consisting of 7 sheets of A4 paper printed on both sides. Apart from the first task (which spread out over two facing pages), each task was laid out in such a way that all titles, instructions, examples and response ‘tick-boxes’ were printed on a single page. On the front-page of the questionnaire booklet the following general instructions were printed:

This programme of sentence property rating scales has been designed as part of a research study about sentence processing. Each different scale attempts to tap into an aspect of what makes English sentences comprehensible, readable or even what makes them “comfortable” to process.

Please read the instructions for each of the following scales carefully and rate each sentence according to these instructions on the scale provided.

Please make a response for every sentence on each of the eleven rating scales – it should take about 25 - 30 minutes to complete the programme – and return the booklet using the envelope provided.

The first task (the categorisation task) was laid out on a double spread of pages. At the top of the left-hand page task-title and instructions were printed:

1.) What kind of text is it?

Please classify each sentence by indicating the categories you think it fits. You may choose more than one category if you feel that this is appropriate. Also feel free to add other categories if needed: the categories presented are not exhaustive.

Each sentence to be categorised was printed in full above a tick-box table with eleven categorisation options plus a space for ‘other’ possible categories, for example:

1. Noah took two animals of each type on the Ark.

News	Opinion	Fact	Fiction	Description	Speech
Narrative	Textbook	Academic	Response to a question	Magazine	Other: ______________

The rating scales (tasks 2 to 10) were all designed to have the same layout: the scale-title was printed at the top of the page followed by a paragraph of instructions including example sentences with appropriate ratings. (See Appendix 7 for the instructions and examples for each individual scale.) Underneath this each sentence was printed directly above five possible rating tick-boxes (ranging from '0: not at all' to '4: extremely'). The instructions for each of the nine scales had been phrased in such a way that each of the scale points could always be referred to by the same label, for example:

1. Noah took two animals of each type on the Ark.
not at all ☐ 0 **slightly** ☐ 1 **moderately** ☐ 2 **very** ☐ 3 **extremely** ☐ 4

The focus-determination task was printed on the last page of the questionnaire booklet. Again instructions and examples were printed at the top of the page:

> **11.) Underline the part of each sentence that the main focus is placed upon.**
> For example, "Newton discovered gravity by having an apple fall on his head" or "The mist has dispersed a bit, but it is still a very gloomy looking swamp".

Then each sentence was printed with generous spacing between them. A thank-you note to the participants was printed at the bottom of the page.

7.8 Results and discussion

7.8.1 The categorisation task

The number of times that each sentence was classified as a particular type of text was recorded. Since participants were allowed to classify each sentence as more than one specific kind of text, as they saw appropriate, the number of total categorisation responses varied for the different sentences. In order to be able to compare which categories were chosen most frequently for each individual sentence, percentages were determined for the sentences. Most sentences tended to be classed fairly consistently by all participants as either one or two particular types of text see Table 7.4). The percentages of classification responses for each category were also determined for the two sentence groups (SITS and controls) (see Table 7.5).

Table 7.4: Preferred description for each sentence (percentage of count > 20)

Sentence:	Modal categorisation:
1. Noah took two animals of each type on the Ark.	Narrative (34%)
2. Bloodletting, generally accomplished with the aid of leeches, was thought to remove “poisons” from the blood.	Textbook (33%); Fact (25%)
3. It was President Kennedy, who was killed in Dallas in 1963.	Fact (39%); Response (35%)
4. Snow White was sheltered by seven dwarfs before marrying her prince.	Fiction (50%); Narrative (34%)
5. Calcium phosphate is a mineral salt and is the principal mineral constituent of bones and teeth.	Fact (33%); Textbook (33%); Academic (21%)
6. In the October revolution of 1917 the Bolsheviks took under the leadership of Lenin the power in Russia.	Fact (40%); Textbook (27%)
7. In what mythology was Venus known as the Goddess of Love?	Other: Question (43%); Speech (22%)
8. The future of Barclays, one of Britain’s biggest banking groups, was thrown into doubt yesterday with the shock resignation of its chief executive, Martin Taylor.	News (58%)
9. Inspector Morse who does his policing in Oxford was invented by Colin Dexter.	Fact (44%)
10. This particular pool of light moving in a mesmeric manner backwards and forwards picked out from time to time a long red island of spilt wine.	Description (38%); Narrative (29%); Fiction (24%)
11. Subjects averaged 47 seconds longer to name ink colors of incongruent words than solid-color squares.	Academic (43%); Fact (27%)
12. Of course a certain number of scientists have to go mad, just to keep the tradition alive.	Opinion (48%); Speech (23%)
13. The archaeologist Schliemann discovered the ruins of the ancient city of Troy.	Fact (45%); Textbook (25%)
14. These results are consistent with the current literature on focalization and question the Erickson and Mattson (1981) claim that the Moses illusion is not dependent on a misdirection of focus.	Academic (49%); Opinion (20%)
15. Away ran the girls, too eager to get in to have time for speech.	Narrative (44%); Fiction (34%)

Apart from sentences 1 (Noah), 4 (Snow White) and 7 (Venus) the SITS were generally classed as ‘fact’. Sentences 2 (Bloodletting), 6 (Bolsheviks) and 13 (Schliemann) were also categorised as ‘textbook’ material. Sentence 1 (Noah) appeared to cause some indecision amongst participants: while 33.9% agreed that the sentence could be classed as a ‘narrative’, there was no other single category in which the classification count exceeded the 20% cut-off mark. Nearest came the classification as ‘fiction’ at 18.8%, followed by ‘fact’ at 13.4% and ‘description’ at 12.5%. It seems likely that the division of classifications of sentence 1 partially reflects the fact that this sentence is derived from a religious source: it seems plausible that Christian participants would have chosen to label this sentence as ‘fact’, while non-Christian participants would be more likely to choose ‘fiction’. Sentence 4 (Snow White) was categorised consistently as ‘fiction’ (50.0%) and also as a ‘narrative’. For sentence 7 (Venus) participants consistently thought that an extra category should be applied to class the sentence: 42.9% ticked ‘other’ and added ‘question’ as a kind-of-text category. The next most frequently ticked box for sentence 7 was that for ‘speech’ (22.4%). What is interesting about the choice of the new category ‘question’ for sentence 7 is that it implies that participants believe that questions differ from other types of text. They appear to regard the question as belonging to its own category rather than classifying it as part of a specific genre of text based upon its contents.

Table 7.5: Percentage of individual categorisations for each sentence type.

	SITS	Control
News	2.7	8.6
Opinion	1.3	12.4
Fact	27.7	12.3
Fiction	9.8	8.2
Description	4.8	9.6
Speech	4.3	4.6
Narrative	11.5	11.2
Textbook	13.3	10.8
Academic	6.9	17.3
Response	7.9	2.5
Magazine	2.8	2.4
Other	7.0	2.4

The control sentences were classified in good agreement with their actual sources (see Table 7.6), suggesting that participants generally have a good feel for the kind of text that they are required to process, and thus it is possible that the attributed source of a sentence is taken into account when a sentence is processed. In the case of SITS the single most frequently made classification is 'fact' (27.7). The high proportion of categorisations of SITS as 'fact' is also in keeping with the idea that participants are likely to feel that they know the material expressed in SITS quite well.

Table 7.6: Sources of control sentences and preferred categorisations given (correct source attribution in **bold).**

Sentence:	**Source**	**Preferred categorisation:**
5. Calcium phosphate is a mineral salt and is the principal mineral constituent of bones and teeth.	Textbook	Fact (33%); **Textbook (33%)**; Academic (21%)
8. The future of Barclays, one of Britain's biggest banking groups, was thrown into doubt yesterday with the shock resignation of its chief executive, Martin Taylor.	Newspaper	**News (58%)**
10. This particular pool of light moving in a mesmeric manner backwards and forwards picked out from time to time a long red island of spilt wine.	Novel	Description (38%); **Narrative (29%); Fiction (24%)**
11. Subjects averaged 47 seconds longer to name ink colors of incongruent words than solid-color squares.	Academic journal article	**Academic (43%)**; Fact (27%)
12. Of course a certain number of scientists have to go mad, just to keep the tradition alive.	Modern novel	Opinion (48%); Speech (23%)
14. These results are consistent with the current literature on focalization and question the Erickson and Mattson (1981) claim that the Moses illusion is not dependent on a misdirection of focus.	Academic journal article	**Academic (49%)**; Opinion (20%)
15. Away ran the girls, too eager to get in to have time for speech.	Novel	**Narrative (44%); Fiction (34%)**

7.8.2 The rating scales (tasks 2 to 10)

For each rating scale, a t-test for paired samples was carried out comparing the two sentence types using participants as subjects. Mean rating scores for SITS and for control sentences were calculated for each participant to form the matched pairs, and overall mean ratings and standard deviations were recorded for sentence types for each scale. Additionally histograms were plotted of the percentage of ratings falling into each response category for each sentence type to gain insight into how the rating distributions varied between sentence types. The mean rating for each individual sentence was also recorded for each scale. Results and discussion are presented for each scale in turn.

7.8.2.1 Scale 2: Amount of background knowledge required

SITS were perceived to require significantly less background knowledge than control sentences in order to be comprehended (t= -6.77; d.f.= 67; p<0.001). Mean ratings and standard deviations for sentence type are presented in Table 7.7.

Table 7.7: Mean ratings and standard deviations for Scale 2: Background

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	1.44	0.56
Control sentences	68	1.94	0.61

From the histogram (Figure 7.1) it can be seen that SITS have a distinct modal rating of 1 (slightly), while the control sentences' ratings are evenly distributed across all five rating categories. This observation suggests that SITS form a more homogeneous group than the control sentences where the requirement for background knowledge is concerned. Participants rated most SITS as requiring a low to medium amount of background knowledge, whereas control sentences varied widely on this scale and on average required more background information to be understood.

Figure 7.1: Categorised histogram for SITS and control sentences for Scale 2: Background

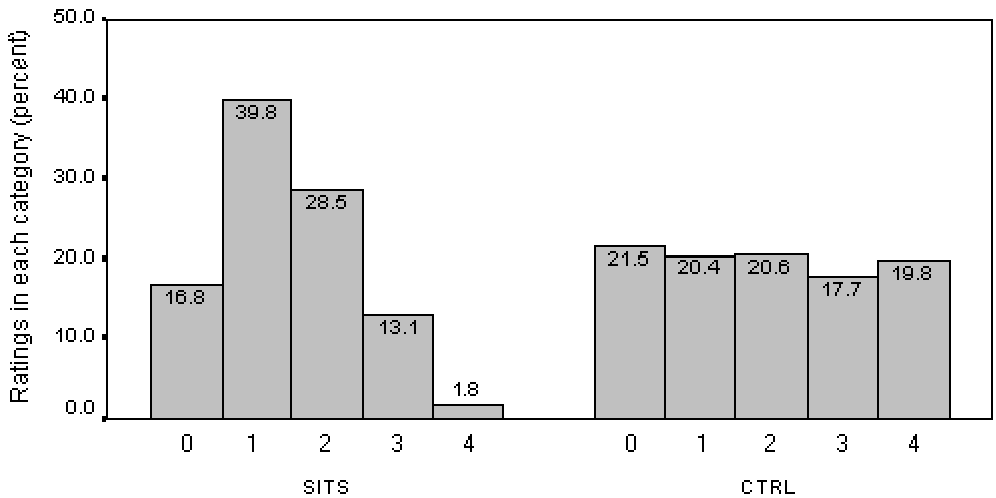


Table 7.8: Means for individual sentences (SITS in **bold) Scale 2: Background**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**1.2**	**1.6**	**1.3**	**1.0**	2.3	**2.3**	**1.6**	1.5	**1.3**	1.7	2.7	0.8	**1.4**	3.7	0.9

These findings were in keeping with the hypothesis that SITS are seen as more self-contained in comparison with most other instances of written language. Participants appear to view SITS to be comprehensible almost in isolation from other background, as they deal with facts that are 'known' (cf. 'remember' vs. 'know' judgements; e.g. Postma, 1999).

7.8.2.2 Scale 3: Importance of the context in which the sentence is presented

A t-test showed that SITS were seen to be much less dependent upon the specific context in which they are encountered than control sentences, and can be understood on their own terms ($t = -16.25$; d.f.= 67; $p < 0.001$). Mean ratings and standard deviations for sentence type are presented in Table 7.9.

Table 7.9: Mean ratings and standard deviations for Scale 3: Context

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	0.79	0.54
Control sentences	68	2.03	0.65

The histogram (Figure 7.2) shows a clear mode of 0 (not at all) for the SITS, while the control sentence ratings are distributed evenly across the scale interval. Again SITS appear to be similar to each other, while the control sentences are seen as coming from a variety of sources.

Figure 7.2: Categorised histogram for SITS and control sentences for Scale 3: Context

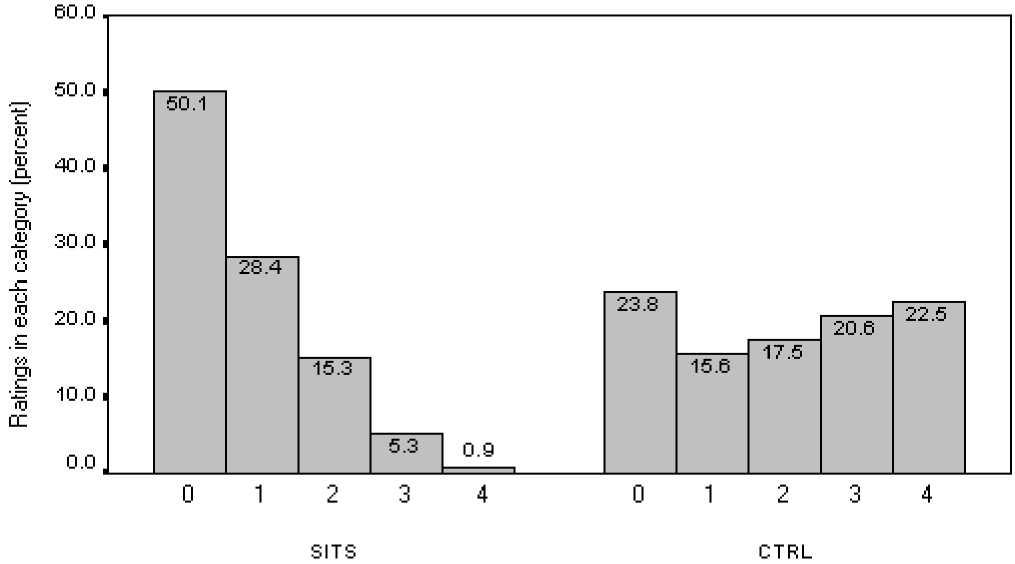


Table 7.10: Means for individual sentences (SITS in **bold) Scale 3: Context**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**0.8**	**0.7**	**0.9**	**0.8**	0.5	**0.9**	**0.9**	1.0	**0.8**	2.8	2.6	1.5	**0.5**	3.1	2.7

As predicted, participants considered the SITS to be more independent of the context in which they were presented than control sentences. Whereas most control sentences were thought to be affected by context, SITS were seen to be comprehensible in isolation. Since SITS seem to deal only with apparently self-contained facts, they are virtually independent of the context in which they are presented.

7.8.2.3 Scale 4: How specific are the meanings of the individual words?

SITS were seen to contain words which appear to be used with more specific meanings than words used in control sentences (t= 3.99; d.f.= 65; p<0.001). Mean ratings and standard deviations for sentence type are presented in Table 7.11.

Table 7.11: Mean ratings and standard deviations for ratings on Scale 4: Words

Sentence Type	**N**	**Mean**	**Std. Dev.**
SI sentences	66	2.37	0.79
Control sentences	66	1.97	0.53

As for scales 2 and 3, the histogram (Figure 7.3) shows a distribution with a peak at 2 (moderately) for the SITS, although it is much less sharp than the respective peaks for the previous two scales. Again the control sentence ratings are evenly distributed.

Figure 7.3: Categorised histogram for SITS and control sentences for Scale 4: Words

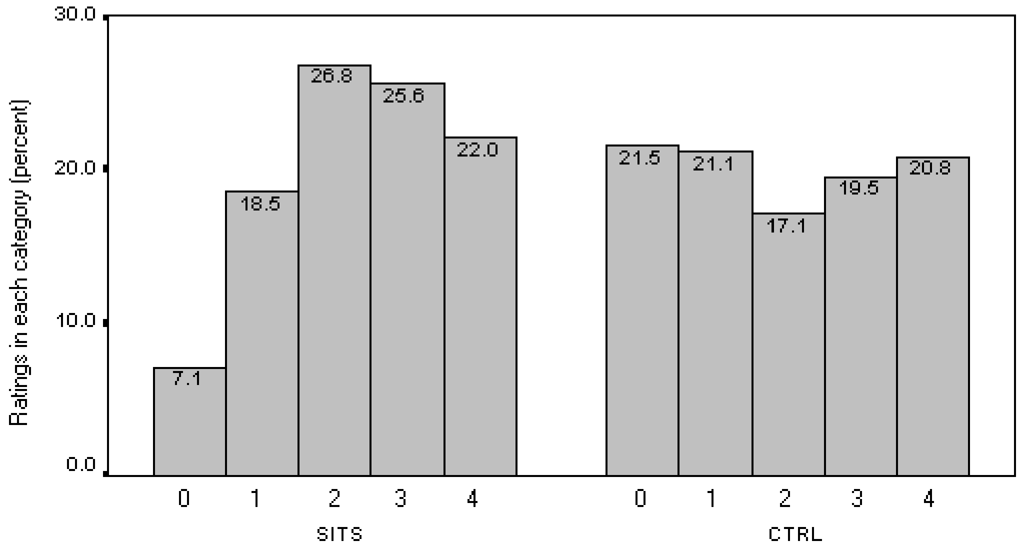


Table 7.12: Means for individual sentences: (SITS in **bold) Scale 4: Words**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**2.3**	**2.1**	**2.7**	**2.0**	2.7	**2.8**	**2.3**	2.6	**2.3**	1.8	2.0	0.9	**2.5**	3.0	0.7

The results of Scale 4 weakly support the hypothesis that SITS use words in a specific and unambiguous way, presumably utilising more central meanings of the words involved. Since central meanings are easier to process, the fact that SITS seem to contain very specifically used words is likely to contribute towards the overall apparent ease of processing of these sentences, which in a semantic illusion situation might lead to a substitution detection failure.

7.8.2.4 Scale 5: How familiar is the content of the sentence?

The t-test results show that SITS are seen as having far more familiar contents than the control sentences (t= 19.37; d.f.= 67; p<0.001). Mean ratings and standard deviations are presented in Table 7.13.

Table 7.13: Mean ratings and standard deviations for Scale 5: Familiarity

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	2.64	0.60
Control sentences	68	0.94	0.49

The histogram (Figure 7.4) demonstrates this dramatically, with the modes of each of the two distributions clearly identifiable at opposite ends of the scale: for SITS the modal rating is 4 (extremely), for the control sentences it is 0 (not at all).

Figure 7.4: Categorised histogram for SITS and control sentences for Scale 5: Familiarity

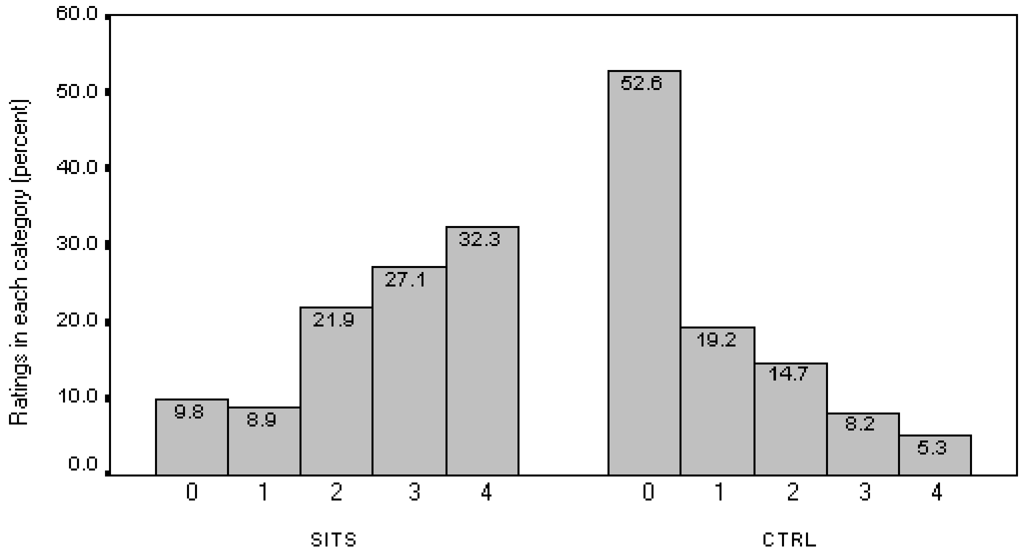


Table 7.14: Means for individual sentences: (SITS in **bold) Scale 5: Familiarity**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**3.8**	**2.5**	**3.2**	**3.3**	1.9	**2.1**	**2.4**	1.3	**2.6**	0.2	0.6	1.9	**1.1**	0.2	0.5

These results support the hypothesis that SITS deal with very familiar ideas, whereas most other uses of written language are imparting new information. This suggests that SITS are not using the same conventions as most other sources of written material. Maybe the fact that material presented in SITS is very familiar to participants causes the sentences to be less carefully read than if there was much new material contained within the sentence. It seems plausible that participants might analyse SITS just sufficiently to allow retrieval of a previously stored version of the relevant facts from memory, judging the truth of the remembered rather than the perceived version.

7.8.2.5 Scale 6: How grammatical is the sentence?

The t-test results showed that SITS are perceived to be more grammatical than control sentences (t= 2.06; d.f.= 67; p= 0.043). Mean ratings and standard deviations are presented in Table 7.15.

Table 7.15: Mean ratings and standard deviations for Scale 6: Grammaticality

Sentence Type	**N**	**Mean**	**Std. Dev.**
SI sentences	68	2.74	0.53
Control sentences	68	2.61	0.66

However, the histogram (Figure 7.5) shows that the distributions of responses are similar for each sentence type, with a modal rating of 4 (extremely) for both, suggesting that the difference between SITS and control sentences is not fundamental as for the other scales.

Figure 7.5: Categorised histogram for SITS and control sentences for Scale 6: Grammaticality

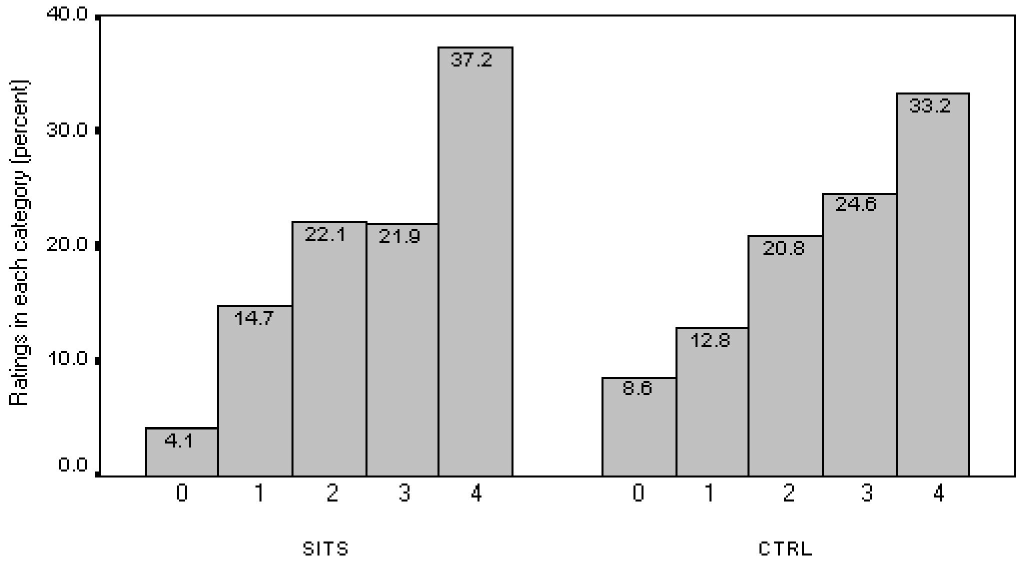


Table 8.20: Means for individual sentences: (SITS in **bold) Scale 6: Grammaticality**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**3.0**	**3.7**	**2.4**	**3.0**	3.4	**1.3**	**3.2**	3.5	**1.9**	1.8	2.3	2.8	**3.3**	2.7	1.8

The findings do not support the hypothesis that SITS might be regarded as ungrammatical because of their redundant phrasing. Instead it was found that SITS are not markedly different from the control sentences, and, if anything, are viewed as slightly more grammatical. It may be that perceived grammaticality facilitates processing and aids a sense of coherence. Maybe this contributes to the feeling of SITS being perfectly sensible sentences, which do not immediately strike a participant as odd.

7.8.2.6 Scale 7: How complex is the sentence?

SITS are seen as significantly less complex than the control sentences ($t= -17.39$; d.f.= 67; $p<0.001$). Mean ratings and standard deviation are presented in Table 7.17.

Table 7.17: Mean ratings and standard deviations for Scale 7: Complexity

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	0.93	0.49
Control sentences	68	2.06	0.51

The histogram (Figure 7.6) shows that the modal rating for SITS at 0 (not at all) and for control sentences at 1 (slightly) are quite close, but the SITS distribution is much less divergent than the control sentence distribution.

Figure 7.6: Categorised histogram for SITS and control sentences for Scale 7: Complexity

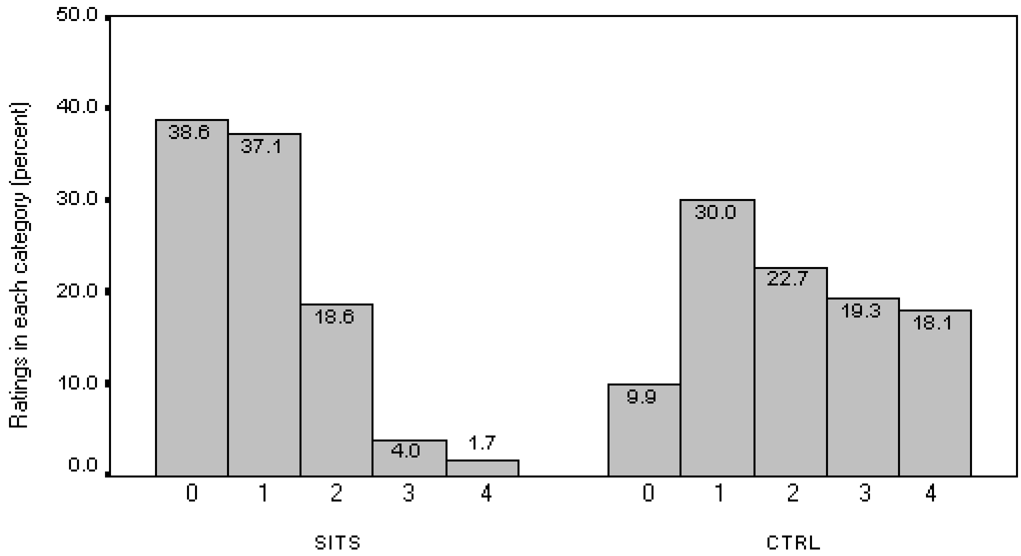


Table 7.18: Means for individual sentences: (SITS in **bold) Scale 7: Complexity**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**0.3**	**1.5**	**0.6**	**0.7**	1.4	**2.2**	**0.6**	2.1	**1.0**	3.2	2.1	0.9	**0.5**	3.4	1.5

As predicted in the hypothesis, SITS were regarded as less complex than controls. This is likely to be due to the fact that SITS deal with highly familiar ideas (see also Scale 5), leading to strong semantic expectations, facilitating processing. If a sentence deals with new information like the control sentences, it is likely to be perceived as more complex because of the requirement to integrate the new information.

7.8.2.7 Scale 8: How formal is the style of the sentence?

SITS are seen as significantly less formally phrased than other sentences (t= -9.93; d.f.= 67; p<0.001) Mean ratings and standard deviations are presented in Table 7.19.

Table 7.19: Mean ratings and standard deviations for Scale 8: Formality

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	1.72	0.54
Control sentences	68	2.27	0.43

The histogram (Figure 7.7) shows a modal rating of 2 (moderately) for SITS, and a modal rating of 3 (very) for control sentences. The control sentence ratings are distributed more evenly across all ratings compared to the SITS ratings which peak sharply.

Figure 7.7: Categorised histogram for SITS and control sentences for Scale 8: Formality

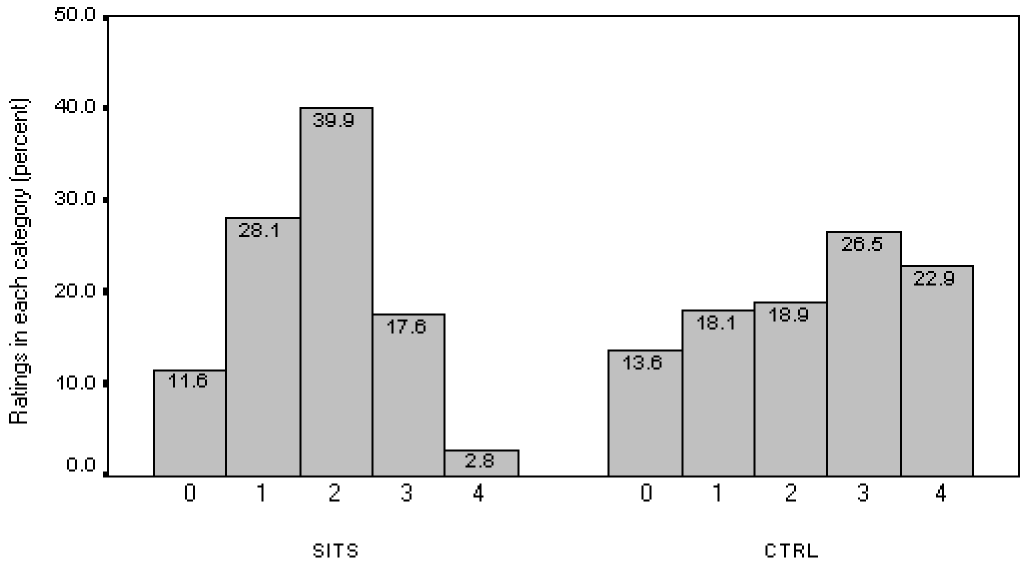


Table 7.20: Means for individual sentences: (SITS in **bold) Scale 8: Formality**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**1.5**	**2.7**	**1.2**	**1.4**	2.8	**2.3**	**1.5**	2.5	**1.2**	2.4	2.9	0.6	**2.0**	3.7	0.9

The findings are in opposition to the predictions made about the perceived formality of SITS. It was thought that SITS would be seen as very formal as in very formal uses of language little or no attention is paid to the actual words used (cf. speech acts, e.g. Mey, 2001). However, since semantic illusion research derives originally from a children's game, the language used in SITS may have an overall tone of informality, leading to a more casual inspection of the individual words.

7.8.2.8 Scale 9: How natural is the sentence?

SITS seem to strike people as significantly more likely to come up in normal discourse than any of the control sentences (t= 6.96; d.f.= 67; p<0.001). Mean ratings and standard deviations are presented in Table 7.21.

Table 7.21: Mean ratings and standard deviations for Scale 9: Naturalness

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	1.88	0.71
Control sentences	68	1.30	0.57

The histogram (Figure 7.8) shows a dramatic difference in the distribution of the ratings for SITS (mode = 2, moderately) and control sentences (mode = 0, not at all).

Figure 7.8: Categorised histogram for SITS and control sentences for Scale 9: Naturalness

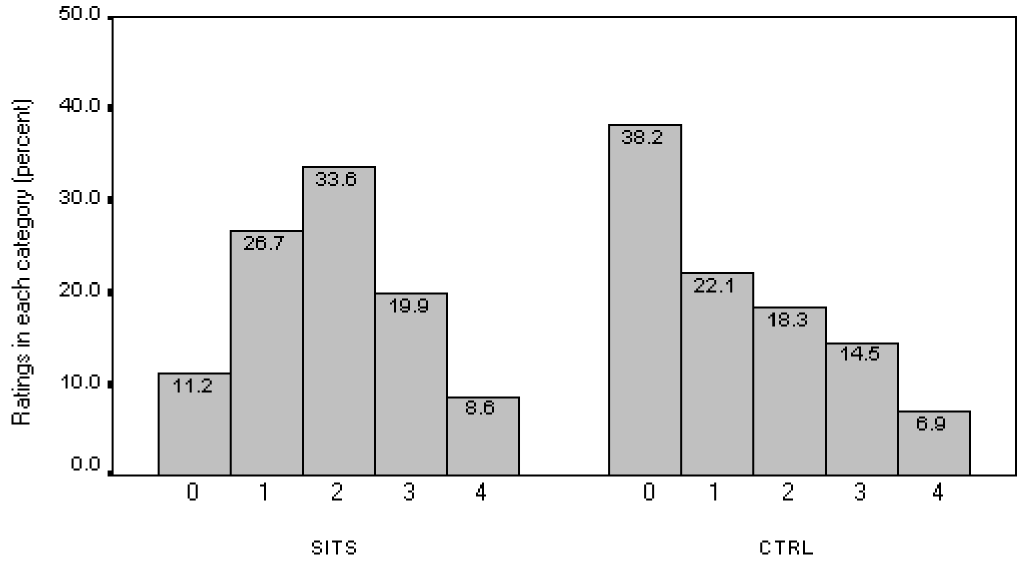


Table 7.22: Means for individual sentences: (SITS in **bold) Scale 9: Naturalness**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**2.1**	**1.6**	**2.4**	**1.8**	1.8	**1.3**	**2.0**	1.9	**2.0**	0.3	1.1	2.5	**2.0**	0.6	1.0

These results were in opposition to the hypothesis that SITS would be seen as unnatural, but in keeping with other findings which indicate that SITS are generally viewed as easy to interpret, such as the high degree of familiarity (Scale 5) and the low level of formality (Scale 8). It seems that part of the causes underlying semantic illusions can be explained by the way in which SITS appear easy to process.

7.8.2.9 Scale 10: How literally is the sentence meant?

SITS are perceived to be meant significantly more literally than other sentences ($t = 9.51$; $d.f. = 67$; $p<0.001$). Mean ratings and standard deviations are presented in Table 7.23.

Table 7.23: Mean ratings and standard deviations for Scale 10: Literalness

Sentence Type	N	Mean	Std. Dev.
SI sentences	68	2.90	0.58
Control sentences	68	2.25	0.44

While the modes for both SITS and control sentences are the same at 4 (extremely), the histogram (Figure 7.9) still shows clearly that SITS are considered to be more literal. The distribution of ratings for the control sentences is more even: control sentences are seen to vary more widely than SITS in terms of literalness.

Figure 7.9: Categorised histogram for SITS and control sentences for Scale 10: Literalness

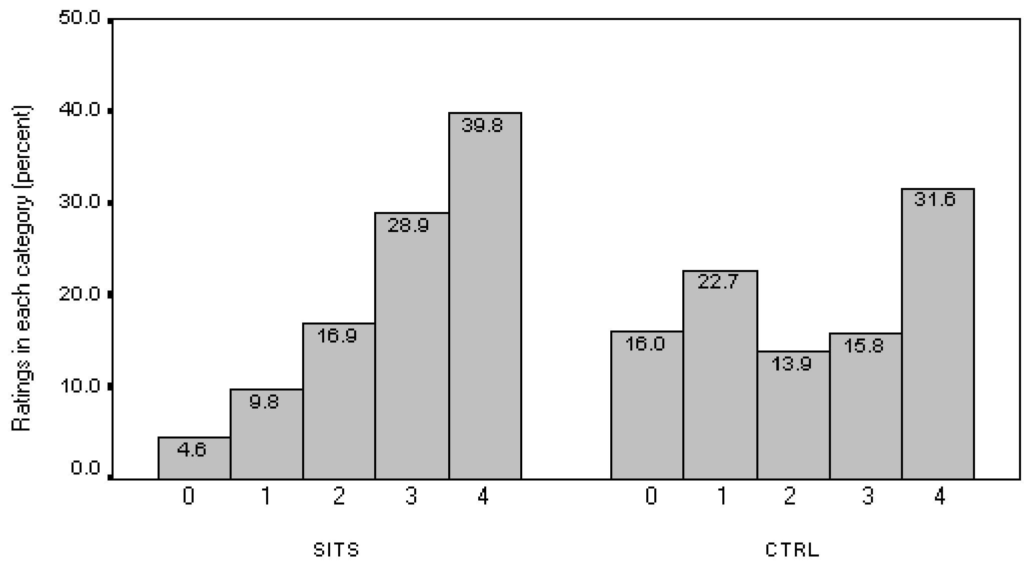


Table 7.24: Means for individual sentences: (SITS in **bold) Scale 10: Literalness**

Sentence:	**1**	**2**	**3**	**4**	5	**6**	**7**	8	**9**	10	11	12	**13**	14	15
Mean rating:	**2.7**	**3.1**	**3.7**	**2.0**	3.8	**3.0**	**2.8**	2.2	**2.5**	0.7	3.3	0.9	**3.5**	3.0	1.7

The results are in keeping with the predictions made about how participants would rate the literalness of SITS compared to control sentences. Because SITS are seen to be quite straightforwardly phrased and as presenting familiar ideas, people think of them as meant literally and not as performing another function.

7.8.3 The focus determination task

A record was made of which part of each sentence was underlined as focus point by every participant. When all these data points were compiled, a number of potential focus points was chosen for each sentence. For example, sentence 1 (Noah) was divided as follows:

> $[\text{Noah}]_1$ took $[\text{two animals}]_2$ of each type on the $[\text{Ark}]_3$.

Focus point 1 was said to have been chosen whenever a participant underlined the name “Noah”, or when they underlined “Noah took” (“took” on its own was never underlined as sentence focus). Focus point 2 was said to have been chosen whenever “two”, “animals”, “two animals” or even “two animals of each type” had been underlined. Focus point 3 covered “Ark”, “the Ark” and “on the Ark”. The number of responses was recorded for each identified focus point. For most sentences one major focus point was found to emerge with half or more of the 68 participants choosing that particular focus. Only sentence 11 (Stroop) did not have a clear focus. Nevertheless the modal choice of focus for this sentence came to a count of 32 out of 68 (see Table 7.25).

The SITS were then looked at separately from the control sentences, to determine if the hypothesis held that the point of focus would generally be perceived to fall on a word other than the target word. It was found that the point of focus chosen by participants coincided with the word that would be the target word when the relevant sentences had been used as semantic illusion sentences for 3 SITS (sentences 1, 3 and 4). For the other 5 SITS, the focus and the target word were at two different positions in the sentence (see Table 7.26).

Table 7.25: Focus points for each sentence (including number of times each focus point was chosen for each sentence out of 68).

Sentence	*Focus point 1*	*Focus point 2*	*Focus point 3*	*Focus point 4*	*Focus point 5*	*Focus point 6*
1	**Noah (45)**	two animals (17)	Ark (6)			
2	**Bloodletting (60)**	leeches (2)	"poisons" (6)			
3	**President Kennedy (64)**	was killed (4)				
4	**Snow White (54)**	sheltered (4)	seven dwarfs (9)	*(missing data: 1)*		
5	**Calcium phosphate (61)**	mineral salt (4)	principal mineral constituent (3)			
6	October revolution (11)	**the Bolsheviks (42)**	Lenin (2)	power in Russia (11)		
7	what mythology (12)	**Venus (45)**	Goddess of Love (11)			
8	**future of Barclays (56)**	was thrown into doubt (2)	shock resignation (6)	chief executive, Martin Taylor (4)		
9	**Inspector Morse (50)**	Oxford (2)	was invented (4)	Colin Dexter (12)		
10	**This particular pool of light (44)**	mesmeric manner (2)	long red island of spilt wine (20)			
11	Subjects (21)	**47 seconds longer (32)**	ink colors of incongruent words (15)			
12	Of course (6)	**certain number of scientists (37)**	have to go mad (21)	the tradition (4)		
13	**archaeologist Schliemann (58)**	discovered (2)	ruins (1)	ancient city of Troy (7)		
14	**These results (36)**	consistent (4)	current literature on focalization (11)	Erickson and Mattson (1981) (4)	Moses illusion (9)	misdirection of focus (3)
15	Away ran (19)	**the girls (34)**	too eager to get in (14)	speech (1)		

Table 7.26: Focus points and target words for SITS.

Sentence	*Focus point*	*Target word*
1. Noah took two animals of each type on the Ark.	**Noah**	
2. Bloodletting, generally accomplished with the aid of leeches, was thought to remove "poisons" from the blood.	**Bloodletting**	**leeches**
3. It was President Kennedy, who was killed in Dallas in 1963.	**President Kennedy**	
4. Snow White was sheltered by seven dwarfs before marrying her prince.	**Snow White**	
6. In the October revolution of 1917 the Bolsheviks took under the leadership of Lenin the power in Russia.	**the Bolsheviks**	**Lenin**
7. In what mythology was Venus known as the Goddess of Love?	**Venus**	**Love**
9. Inspector Morse who does his policing in Oxford was invented by Colin Dexter.	**Inspector Morse**	**Colin Dexter**
13. The archaeologist Schliemann discovered the ruins of the ancient city of Troy.	**archaeologist Schliemann**	**Troy**

The results of the focus determination task showed that there is good overall agreement about which part of a sentence is in focus. There appears to be little difference between control sentences and SITS – participants do not appear to be more consistent in determining the focus of SITS compared to control sentences. It was also found that the hypothesis that SITS would usually have a different focus point to the target word did not hold for three of the SITS used in this study. This is in keeping with previous findings (e.g. Brédart and Modolo, 1988; Brédart and Docquier, 1989) that focus alone is not responsible for the occurrence of semantic illusions.

7.8.4 Factor Analysis

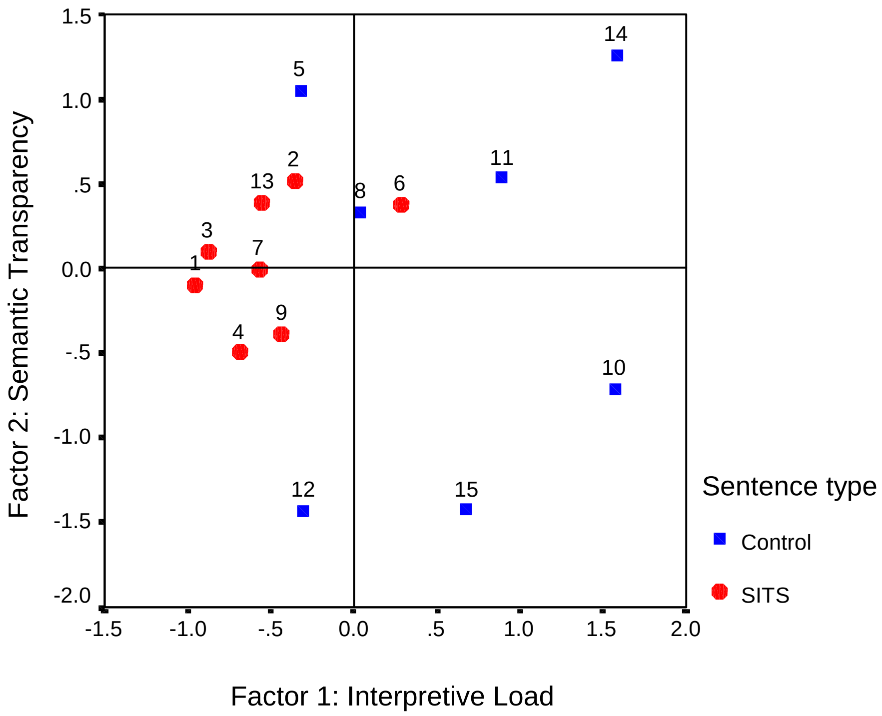
In an attempt to see if the rating scales (Tasks 2 to 10) as a group could reveal anything about SITS as opposed to control sentences, a factor analysis was carried out using principal component extraction. After the factors had been rotated using varimax rotation with Kaiser normalisation, two factors emerged which explained 51 percent of the variance observed. The first factor, which accounts for 31 percent of the variance, shows a strong positive association with complexity and a strong negative association with familiarity. There is also a positive association with the need for context in order to understand each sentence, and there is a negative association with naturalness, suggesting that participants do not expect usually to find such sentences in normal discourse. This factor will be referred to by the label 'Factor 1: Interpretive Load' (Factor 1: IL). The second factor extracted accounted for a further 20 percent of the variance. This factor has strong positive associations with literalness, with formality and with the specificity of individual words used. This factor has been labelled 'Factor 2: Semantic Transparency' (Factor 2: ST). There is a positive association with background for both factors: people appear to require a certain amount of background knowledge whether they are dealing with the information content of a sentence or with the way in which it is phrased. Grammaticality does not load onto either factor (see Table 7.27 for factor loadings)

Table 7.27: Factor loadings for each scale (loadings < 0.5 only)

	Factor 1: IL	Factor 2: ST
complexity	.785	
familiarity	-.739	
context	.720	
naturalness	-.667	
background	.525	.533
literalness		.663
formality		.690
individual words		.665
grammaticality		

The mean loadings for each of the individual sentences were then plotted against the two major factors. Most of the semantic illusion sentences fell into roughly the same area of the plot, whereas the control sentences formed a much more widely spread group (see Figure 7.10).

Figure 7.10: Distribution of sentences depending upon individual mean loadings upon factors 1 and 2. (Labels correspond to sentence numbers).

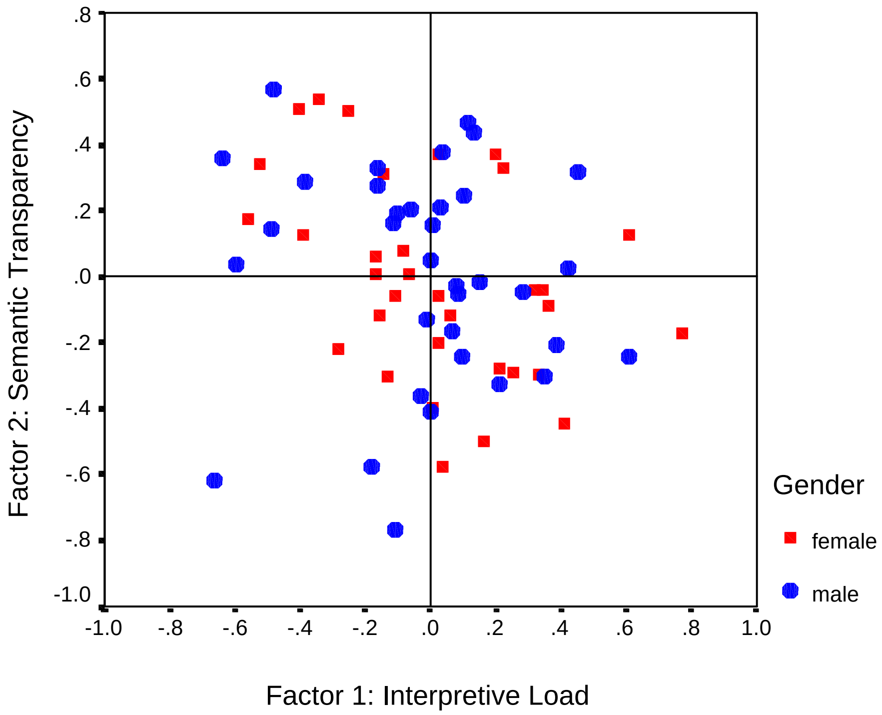


SITS appear in general to convey relatively little new information (Factor 1 loadings are low), and they appear to be neutral with regards to the transparency of their phrasing (Factor 2 loadings are distributed around zero.) The control sentences used in this study, on the other hand, varied widely across both dimensions, but provide much more novel information to the reader on average. It is worth noting, that participants did not feel that there was anything particularly odd about SITS compared with sentences taken at random from actual printed sources – if anything participants felt that some of the control sentences were more unnatural than SITS, which were each constructed specifically for the purposes of semantic illusion research.

The scatter plot (Figure 7.10) shows that SITS and control sentences appear to come from two different populations of sentences. SITS are part of a distribution with a much lower standard deviation for both Factor 1 and Factor 2 loadings than the control sentences, and SITS clearly draw on familiar information rather than being used to convey novel material: SITS tend to score much lower on Factor 1: IL. On Factor 2: ST the SITS are much closer to zero than the control sentences. Overall, these findings appear to be in keeping with the hypothesis that semantic illusions can only occur where sentences are used which draw upon already established knowledge – indeed, they must do this if the type of confusion at the root of a semantic illusion is to occur. However, the results of this factor analysis are indicative only as the number of sentences used was very small.

In order to ascertain that the clustering of the two sentence types is not an artefact of the way in which the factor analysis was carried out, mean factor loadings were also plotted for each participant and coded by gender: there is no evidence of clustering due to gender (see Figure 7.11), suggesting that SITS and control sentences load differently on Factors 1 and 2 due to different processing requirements and not due to chance.

Figure 7.11: Distribution of participants depending upon individual mean loadings upon factors 1 and 2.



7.9 General discussion

The study described in this chapter was designed to help identify how people perceive SITS when they encounter them and how such a perception might vary from the expectations that people have about how language is generally used. It is quite clear from the results of the categorisation task (Task 1) that participants have no difficulty in correctly assigning a kind of text to a likely source. For all the control sentences used in this study, the task was completed with remarkable accuracy. SITS appeared to be equally confidently pigeonholed. Most often they were seen as mere statements of fact. The fact that participants completed the classification task so competently – and especially the fact that an extra category was consistently added when the categories presented as options were not perceived to be adequate descriptors for a particular sample of text (for sentence 7 (Venus) 42 out of 68 participants added the category 'question') – suggests that participants are likely to come to each sample of text they are required to deal with expectations specific to how the sample of text is perceived (cf. Zwaan, 1994). Part of the reason why semantic illusions take place might therefore be due to the fact that most SITS are thought of as factual statements, and thus are treated as such.

The results of the rating scales (Tasks 2 to 10) show that SITS are not just standard sentences but are highly atypical. SITS were judged to be significantly different from control sentences on all nine rating scales. For six of these scales the differences found between SITS and control sentences were in keeping with the predictions made. SITS were found to be less dependent upon background knowledge and specific context. They were seen to contain words with more specific (central) meanings and to contain material more familiar to the participants than control sentences. SITS were also seen to be less complex than control sentences and were thought to be meant more literally. But on the other three rating scales SITS and controls differed in the opposite direction to that predicted. Participants perceived SITS to be more grammatical than control sentences (although the difference on this scale is small). This may be indicative of the fact that participants fail to distinguish between syntactic and semantic cues when rating sentences on the grammaticality scale, because SITS contain highly familiar information (cf. Schlesinger, 1968). The phrasing of SITS was seen to be less formal rather than more formal as predicted. The observed pattern of results on the formality scale might be explained by the origin of SITS in a children's game. Finally, SITS were also perceived as more natural (and likely to occur in normal discourse) than control sentences. Once again the high level of familiarity of SITS content may help explain these results: if information in a sentence feels very familiar it seems likely that it has been heard or read before, suggesting that it has already occurred in natural discourse at an earlier point in time.

The overall pattern of results from the rating scales suggests that participants appear to feel much more confident and comfortable in dealing with SITS than they do with the control sentences. This is in evidence both from the ratings given to these sentences on the rating scales in this study and from the fact that semantic illusions occur. The occurrence of semantic illusions may be in part explained by the fact that SITS feel straightforward to deal with, through phrasing, familiarity of context, choice of words and so on.

When the data from the rating scales was examined by means of a factor analysis, the main conclusion drawn from the ratings scales that SITS were very atypical sentences was confirmed. It became clear that the main difference between SITS and control sentences was in the amount of new information that the different kind of sentences conveyed to a reader. SITS score lower on the major factor that emerges (Factor 1 : Interpretive Load) than control sentences and SITS score closer to zero for the second factor (Factor 2: Semantic Transparency). A plot of the factor scores showed SITS and control sentences to be parts of two different distributions.

The focus determination task showed that generally there is good agreement about which part of a sentence is in focus. It was found that only five of the eight SITS investigated in this study had a definite focus on a word other than the target word. It was expected that generally the target word in a SITS would not be in focus (except in the case of sentence 3 (President Kennedy) which had been designed to have the target in focus; Brédart and Modolo, 1989). This lends further support to the claim that focus alone cannot account for the occurrence of semantic illusions.

7.10 Conclusions

Overall the results of this study showed that SITS are highly atypical sentences that are nevertheless perceived to be part of normal discourse – and not just that – they are perceived as more ‘normal’ than many examples of written discourse taken from sources such as books, journals and newspapers. SITS are seen to be comfortable to process and participants do not appear consciously to notice anything unusual about them. It seems that asking participants’ opinions of SITS does not help in trying to find an explanation for semantic illusions, but it does provide evidence for the claim that the same processes used in ordinary processing lead directly to semantic illusion responses (see Chapter 1).

In the second part of this thesis, semantic illusions are approached via the possible mechanisms which may be involved in processing them. The next chapter provides an overview of findings related to the way people think, store knowledge and process linguistic inputs, and suggestions are made about how these processes could link in with semantic illusion research.

Part Two:

The problem with semantic illusions

Chapter 8:

A problem solving approach to semantic illusions

The experiments described in chapters 3 to 6 were all concerned with semantic illusions as a part of sentence processing, and with how the structural make up of the sentences might contribute to the occurrence of semantic illusions. It was found that neither priming, form class, sentence length, target word position nor language had much of an effect upon semantic illusion rate. Indeed only the syntactic form of the sentences either as a question or a statement (Chapter 4) had any appreciable effect at all, with question format leading to more semantic illusion responses than statements. Furthermore, the study presented in Chapter 7 which compared semantic illusion type sentences to sentences taken from sources of written discourse showed that SITS are not typical of other sentences, even if they are seen as easy and straightforward to process. But *per se*, none of these findings come much closer to explaining semantic illusions, and so far partial matching (see section 2.8) is still the best candidate for the mechanism underlying the semantic illusion phenomenon.

But maybe a clue about semantic illusions can be gleaned from the facts that a) questions lead to more semantic illusions that statements (Chapter 4) and b) participants will consistently classify a sentence in question format not by its genre, but as a question (see section 7.8.1). Evidently people perceive questions differently from statements. This difference in perception may extend to processing, as questions appear to fulfil a different function to statements. Perhaps they are viewed as a kind of problem to be solved according to a given set of rules (c.f. Grice, 1975). In order to explore this possibility, the second part of this thesis attempts to view semantic illusions from a problem solving perspective, including a consideration of the implications of partial matching. A number of observations about cognitive performance in syllogistic reasoning and other kinds of problem solving are described and related to semantic illusions to show that semantic illusions are not simply an isolated phenomenon arising from the mechanisms of language processing, but need to be viewed as similar to a wide range of cognitive tasks requiring attention, knowledge and insight.

Under most circumstances partial matching is a mechanism which allows people to process incoming real world information quickly and effectively. It is a mechanism, which is by no means restricted to explaining semantic illusions as a cognitive phenomenon. In this chapter the semantic illusion phenomenon is related to a number of non-language specific cognitive mechanisms, in which partial matching strategies are also employed.

8.1 Why does partial matching occur

At the most basic level, the requirements that the environment makes upon the cognitive system are never constant from one instance to the next. Even situations that people class as deriving from the same category vary considerably, quite often in a number of different respects. One important ability that humans have is that they are able to ignore irrelevant distinctions and make approximations, which allow them to deal rapidly and efficiently with incoming data. If exact matches were needed in order to recognise or deal with any situation, real-time processing would be impossible. For example, an individual tree on different occasions is never exactly the same – the wind might be blowing the leaves differently, a new leaf may have grown, or it might be bare in winter. Yet a human observer still concludes that it is the same tree, summer or winter, wind or rain. Similarly people can easily identify different chairs as being just that, never mind whether they have three legs or four, are covered in upholstery, made from moulded plastic or wood. From common sense and experience, people know that they cannot take in everything that goes on around them, and they also know that in order to deal with a specific input they need to 'pay' attention to it. In day-to-day usage of language expressions regarding attention, there is an implication that there is a cost to allocating attention to something (one 'pays' attention to something) – there is a distinct effort involved, and attention is a scarce resource to be deployed, and if one thing is attended, another must go unattended.

The observation that the environment does not usually confront people with identical processing demands, leads to questions about the way in which pre-existing knowledge is stored in memory. People can usually generalise from one situation to the next without apparent effort, suggesting that human memory is not made up of static images, but that it possesses adaptive capability. One theory describing such active memory constructs is Bartlett's 'schema' theory (1995[1932]), which essentially implies that all that can ever be hoped for is a partial match between a new situation and the representation of similar situations experienced in the past, which are stored in memory.

8.2 Schema theory

One way in which previous knowledge is stored and how it affects people's responses and recall of certain situations is described in Bartlett's book *Remembering* (Bartlett, 1995[1932]), in which he presents a vast number of experiments about perception, images, recognition and recall. The experiments were designed to be as naturalistic in content as possible, while still providing a certain level of laboratory control, because Bartlett was interested in how human memory worked in the real world outside the laboratory. Most famously, Bartlett presented participants with a story called 'The War of the Ghosts' (see Appendix 8) which was designed to be outside the usual frame of experience of the literature that would have been predominant at the time of Bartlett's research. It was also written in a culturally unfamiliar style (since the story was based on a North American folk tale), which meant that participants generally had some difficulty in finding the relevant connections between different parts of the story.

The experiment was very simple: participants were asked to read the story and then to reproduce it repeatedly at different time intervals, ranging from 15 minutes to 10 years after first reading the story as opportunity arose. Even though Bartlett made no effort to control the number or time intervals between recalls of the story (he felt that people were so different from each other anyway that there would be little benefit from imposing such controls), he nevertheless observed that certain predictable errors crept into the reproductions of the story. Overall he found that recall tended to be very inaccurate. Details would be omitted, in particular if the details in question were at odds with a participant's understanding of the story. Any perceived inconsistencies in the story tended to be rationalised, in that participants would invent possible ways in which the details of the story could fit together to make the story seem more sensible and coherent. Sometimes transformations of the order of events in the story occurred – if, for example, something was of particular interest to a participant it might be brought forward in time in the retelling of the story – but generally the order of events was not much affected in participants' retellings of the story. Participants' attitudes and emotional reactions to the story also led to some distortions. Finally, the perceived importance of different events in the story sometimes led to changes of importance in the retelling, so that, for example, the ghosts in 'The War of the Ghosts' might be made more central to the plot upon recall than they were in the original. In short, participants appeared to reconstruct rather than remember the material they had been presented with, and they did this in such a way that the material fitted meaningfully into their experiences of the world. They were making what Bartlett referred to as an ‘effort after meaning’.

In order to explain his findings, Bartlett suggested that our memory did not consist of static images that could be reproduced via the re-excitation of a memory trace, but instead that it was likely to be made up of active knowledge structures called *schemata*. A *schema* was defined as an ‘active organisation of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response’. Bartlett proposed that if responses appear to be in any way organised or regular, that is because each of these responses is related to other similar responses, which are organised serially and could be drawn from, when a new similar response is called for. Thus every new instance of a similar event to that stored in the schema would contribute towards the schema, changing it subtly.

In terms of remembering, which involves determination by the past, the influence of such schemata could be viewed as one mechanism by which past experiences influence present processing. In the schematic form, past experiences operate together as a whole: all the instances of input relevant to a situation contribute to the schema relating to the knowledge of that situation. The latest inputs into each schema have a predominant influence upon the next response. Schemata are built more or less chronologically, but based upon the observations Bartlett made about recall, it seems that schemata can somehow be used ‘backwards’ to construct a probable past sequence of events from a present situation. In Bartlett’s own words: ‘It would then be the case that the organism would say, if it were able to express itself: ‘This and this and this must have occurred, in order that my present state should be what it is’.’

In other words, schemata are interactive subconscious mental structures that are composed of old knowledge and added to and modified by newly acquired knowledge. A schema contains the information pertaining to certain situations, which are familiar to us, and thus they can be used to draw relevant conclusions about how these kinds of situations generally unfold. A similar idea has been used by Schank and Abelson (1977) who referred to the memory structures in question as 'scripts'. For example, a restaurant script would include all the aspects that in our experience belong to a restaurant setting: waiters, food, chairs, tables, dishes, ordering a meal, paying the bill, the order in which events usually unfold… People possess schemata or scripts about many aspects of life – they help them deal with situations quickly and efficiently, because the regularities of the world and routine interactions with them are internally represented, ready to be called upon.

But there is a price to pay for this largely automatic processing of information – the various cognitive processes always tend to drift in the direction of the familiar and the expected, because a schema only contains evidence of how a particular situation or input should appear. But since there is no representation within a schema of what it should *not* be like, errors such as semantic illusions can occur quite easily. Generally speaking many attempts at explaining human error incorporate the concept of schemata, by making a distinction between conscious, controlled processing and automatic, unconscious processing (e.g. Reason, 1990). The first could be referred to as processing under the attentional mode of control, which is limited, time-consuming and effortful. The second type of processing could be said to take place under a schematic mode of control, which deals very rapidly with familiar information, without any known limits and without conscious effort. But this mode of control cannot deal with unexpected differences between a situation at hand and the way the situation ought to be. This could explain why an 'almost-right' sentence such as a semantic illusion sentence passes as perfectly unobjectionable. The sentence fits the relevant schema, the substitution fits the relevant role-slot well enough, and in an effort after meaning, the substitution goes unnoticed.

Schema theory provides a plausible suggestion about how people store knowledge in long-term memory, but it isn't explicit about how people think and deal with incoming data. How people think – or rather how they don't think – is addressed in the next section.

8.3 How people think: Syllogistic reasoning and semantic illusions

The question of how the rules of thought can be described has concerned philosophers for a very long time, one of the most enduring suggestions being that logic is a condensed version of what goes on in people's brains (e.g. Boole, 1847). But it seems quite clear from studying human cognition empirically that people have tremendous difficulty in thinking or reasoning logically. A good demonstration of this inability to cope with formal logic comes from studies of people's performance when solving categorical syllogisms.

Syllogisms are based on Aristotelian logic and involve reasoning about the relation between categories. A syllogism starts with two statements called the 'premises' that are to be treated as true and which are to be combined to draw a 'conclusion'. For example,

> *All psychologists are wine drinkers.*
>
> *All Italians are psychologists.*

One premise relates the subject of the conclusion (*Italians*) to a middle term (*psychologists*). The other premise relates the middle term to a predicate (*wine drinkers*). The task is to draw a conclusion about the relationship between subject and predicate, if this is possible. In the example a valid conclusion would be *All Italians are wine drinkers*.

The subject (S), predicate (P) and middle term (M) can be combined in a variety of ways using relations defined by the terms *all*, *some*, *no* and *some…not*. Each of the two premises of a syllogism can take the form of each of these four logical relations, so that the number of possible arguments is large. The four basic ways in which S, M and P can appear are known as 'figures':

Fig 1.	Fig 2.	Fig 3.	Fig 4.
M – P	P – M	M – P	P – M
S – M	S – M	M – S	M – S

Combining the 16 logical arguments with the 4 figures means that there are 64 logically distinct variants in which a syllogism can appear. Some syllogisms – such as the example above – allow people readily to draw valid conclusions, but others such as: *Some Italians are not wine drinkers; All Italians are psychologists; What follows if 'psychologists' is the subject of the conclusion?* hardly ever allow the correct conclusion to be drawn: *Some psychologists are not wine drinkers* (e.g. Johnson-Laird and Bara, 1984). Not all syllogisms have valid conclusions, for example, *All A are B; Some B are C* does not have a valid conclusion, if *C* is the subject of the conclusion.

8.3.1 The atmosphere effect and semantic illusions

When people are given syllogisms to solve, they usually display a high error rate, and generally find it very hard to deal with this form of reasoning exercise. Woodworth and Sells (1935) proposed an explanation for people's problems in dealing with syllogistic reasoning: people are likely to draw conclusions about a syllogism on the basis of the ‘atmosphere’ of the premises. A ‘negative’ premise (*no/some…not*) is said to create a negative atmosphere, even when the other premise is ‘affirmative’ (*all/some*), and a negative conclusion is likely to be drawn. Similarly a ‘particular’ premise (*some/some…not*) is said to create a particular atmosphere, even when the other premise is ‘universal’ (*all/no*), resulting in a particular conclusion. In addition to the atmosphere effect, Woodworth and Sells also used a ‘principle of caution’ to predict the way that conclusions are drawn: people are more likely to accept weak and guarded conclusions rather than strong ones (i.e. *some/some…not*, rather than *all/no*). The atmosphere effect accounts well for people’s responses in a number of studies (e.g. Begg and Denny, 1969; Dickstein, 1978; Revlin; 1975; Sells, 1936).

While the atmosphere effect implies that participants combine information from the two premises in a syllogism in order to draw their conclusion, there is an even simpler strategy for dealing with syllogisms, that also accounts well for the results observed: participants ‘match’ the conclusion to the more conservative of the premises (Wetherick and Gilhooly, 1990).

Both matching and atmosphere strategies for syllogistic reasoning provide evidence for the fact that participants approach these kinds of problems in a superficial manner and with little understanding of the underlying relations (e.g. Gilhooly, Logie, Wetherick and Wynn, 1993). In a similar way participants seem unable to pay attention to local details in semantic illusion sentences. Semantic illusion sentences generate an impression of coherence, and instead of processing the actual content of a sentence in depth, responses are based on the global perception of the sentence.

8.3.2 The belief bias effect and semantic illusions

Another problem participants have when dealing with syllogisms is that they are generally unable to reason locally in the way such logic problems require. Instead, participants are influenced by the semantic content of a problem. This is called a 'belief bias' (Revlin, Leirer, Yopp and Yopp, 1980): conclusions that happen to coincide with participants' real-life beliefs are more likely to be judged true, and conclusions that happen to coincide with something that participants do not believe anyway are more likely to be rejected. Instead of processing the local specifics of such a problem, people reason based on the totality of their knowledge, which is not what syllogistic reasoning requires. A similar processing strategy seems to underlie semantic illusions and similar phenomena – participants respond to the perceived overall meaning of a semantic illusion sentence and find it very hard to concentrate on local semantics. For example, "surviving dead" – a locally anomalous phrase – escapes notice remarkably often (Barton and Sanford, 1993; see section 2.9).

8.3.3 Conversion errors and semantic illusions

A further way in which syllogistic reasoning and semantic illusions are similar is that both are instances of processing in which a usual processing strategy from ordinary conversation is employed inappropriately. Many errors in syllogistic reasoning come from the fact that participants treat premises such as *All A are B* and *All B are A* as though they were identical. This is, however, not warranted: for example, *All apples are fruit* is not equivalent to *All fruit are apples*. Chapman and Chapman (1959) suggest that these 'conversion errors' are a result of the inappropriate use of a strategy that works quite often in day-to-day conversation – participants often hear and use statements like *All three-sided figures are triangles* which are reversible, and thus being able to convert statements from one form into another is a useful and sensible habit in ordinary conversation. In processing semantic illusion sentences a similarly useful and well-practised habit – the rapid extraction of semantic content from the surface form of a sentence – also leads to an inappropriate response and a semantic illusion occurs.

8.3.4 Expectations in solving syllogisms and responding to semantic illusions

Finally it is also possible that people's expectations about the form of a task affect performance. Perhaps people perform very poorly at many syllogistic reasoning tests, because these often contain many problems for which no valid conclusion can be drawn, when from their normal life experience people would expect most if not all problems in a test to have solutions (Chapman and Chapman, 1959). A similar sense of expectation may play a part in people's processing of semantic illusion sentences.

8.4 How semantic illusions are like problems: expectations and processing

Expectations in general have a demonstrable effect upon people's ability to process information and solve problems. These expectations arise out of people's knowledge of the world and from their specific experiences in similar situations. In perception tasks, expectations allow participants to recognise degraded stimuli (e.g. Biederman, 1987). However, in problem solving, expectations can get in the way of a successful solution. People often adopt a certain attitude towards the task situation – often referred to as *Einstellung* (German for 'attitude') – which is influenced by their previous experience either of solving similar problems, or of the world in general. Such an Einstellung tends to be fairly rigid and can make it very hard for people to deal with the real characteristics of the problem.

8.4.1 Prior experience with day-to-day objects and problem solving

One manifestation of Einstellung can be observed when a person finds it difficult to dissociate an object from its conventional function in the environment. This is sometimes called *functional fixedness* (Duncker, 1945). For example, when presented with a box of tacks, a book of matches and a candle, and given the task: mount the candle on the wall and light it, participants find it very difficult to dissociate the tack box from its containing function in order to use it as a part of the solution of the problem. (The model solution requires the participant to mount the box on the wall using the tacks and to place the candles on the box.) If presented with the same items, but with the tacks placed in a pile and the box empty, participants are less likely to think of the box as exclusively being a container, and are consequently more likely to solve the problem (Duncker, 1945; Adamson, 1952).

Another example of functional fixedness can arise when participants are required to change their perception of what an object 'can do'. For example, Maier (1931) presented participants with the following problem: in a large room, which contained many objects such as poles, ringstands, clamps, pliers, extension cords, tables and chairs, two strings were hung from the ceiling. The strings were long enough to reach the floor, but not long enough for a participant simply to hold onto the end of one string and walk over to the other string with it. The task was to tie the ends of the strings together.

There were a number of ways in which this problem could be solved, but one solution was more insightful than others and not easy to come up with: to use a weight to set one cord swinging like a pendulum, while the other cord could be held and brought close to the swinging string, thus making one string 'do' part of the work. Maier suggested that the reason for the difficulty that participants experienced with this pendulum solution was partially due to the ‘direction’ from which the problem was approached. In this solution, the second rope is required to move towards the other rope, which is being held by the participant, and – as opposed to direct action on part of the participant as in various other solutions – a way must be found in which the rope ‘performs an action’. Another difficulty was that the participant must find a way in which to see the rope as a pendulum, that is, to change their mental representation of the essentially static string into a mobile object.

Maier’s experiment also showed other evidence of Einstellung – participants displayed a strong tendency to attempt variations on previously successful solutions. Maier stated that under the circumstances where an old solution to a similar but new problem does not work, the memory of the previous solution turns into an obstruction. He concluded from his studies that neither trial and error nor solving the problem by similarity to another one could adequately explain the sudden organisational shift in participant’s approach to the strings problem. ‘Direction’ appears to be a crucial factor in finding a solution.

8.4.2 Specific prior experience and problem solving

Functional fixedness does not have to be the result of long years of experience: even a single experience with an object can blind people to the object’s other properties. Birch and Rabinowitz (1951) carried out an experiment which was designed to determine what effect specific prior experience with objects had upon their usefulness as problem-solving tools. Using the two-strings-problem described above, three groups of experimental participants were required to perform a pre-test task: one group was required to complete an electrical circuit on a board using a switch, another group had to complete a similar task – they were required to install a relay. The control group was given no pre-test task.

In the problem solving test only two objects could be used as weights to turn one of the strings into a pendulum: the switch or the relay used in pre-training. The results showed that for individuals who had previously used one of the objects on the pre-task, it was likely that they would choose the object not previously encountered as the pendulum weight. (For the control group, 50% used one object; the other 50% used the other item). When participants were asked why they had chosen either object as the pendulum weight, their replies further indicated that specific prior experience with an object had a noticeable effect upon their problem solving strategy. Their responses were identical – ‘easier to attach’, ‘more compact’ – for both objects. There were no objective reasons for the participants’ preferences for either object, so it would appear that their choices were influenced by the effects that their previous experience with either object had upon their perceptions.

Birch and Rabinowitz suggested that the type of prior experience with an object appears to limit the properties of that object that could be perceived by a person. Participants in each group appeared to see the pre-test object as an electrical component, whereas the ‘new’ object could be seen also in terms of its mass, shape, attachability. It would appear, therefore, that there are two types of learning going on: a person acquires certain broad, non-specific, general notions about an object. This type of experience seems to provide what it takes to think productively – in a way that allows problems to be solved. The second type of learning involves much more specific limited perceptions about the functions of an object. It is this type of learning which leads to functional fixedness.

8.4.3 Prior experience and semantic illusions

While the analogy between sentences and objects with specific functions may not be immediately obvious, it seems safe to say that people are certain to make assumptions about the format of the sentences they process. Generally speaking, uttered sentences would be expected to make sense, be truthful according to certain standards, and convey some new information or refer to already known information, to be relevant to the topic at hand and to be kept as short as possible (cf. Grice's conversational maxims, 1975). Such a description covers most kinds of written material people are usually faced with in everyday life. Specific types of texts, such as tests and quizzes, conform to certain conventions and fulfil certain functions.

According to such conventions, questions generally require an answer to what they are asking about, which can be retrieved from memory. Statements-to-be-verified generally require a check of truthfulness on a coherently expressed idea and would be perceived as a check of an individual's state of knowledge. Truthfulness could be tested by forming a mental representation of such an idea and checking it against stored general knowledge. Semantic illusions fall outside these conventions and therefore they can be seen to be affected by a kind of functional fixedness. People perceive semantic illusion sentences as simply being questions or facts, whose function it is to tap into their general knowledge stores. They do not strike people as possibly having another 'use'.

This suggests an analogy with the effect of prior experience upon problem solving: Participants' prior experience with sentences affects the way in which the sentences are approached. But in the case of semantic illusions the usual direction of approach does not lead to the desired response. In the case of questions, the usual task of answering the question will lead to an error. In the case of statements, the mental representation checked against knowledge is likely to withstand a quick check, as the error with such a statement is designed to 'fit' the context. As with the two strings problem it is not easy for participants to change their direction of approach. Simply telling participants about the presence of strangely phrased questions does not improve their performance at detecting the semantic illusion questions (e.g. Erickson and Mattson, 1981; Reder and

Kusbit, 1991). So maybe viewing semantic illusions as a kind of insight problem, rather than a processing error due to lack of attention and partial matching would lead us to a better understanding of the phenomenon? If participants were instructed to ‘solve’ each semantic illusion sentence as a problem, they might find it easier to detect the substitutions.

8.4.4 Tried and tested methods: Luchins’ waterjars and semantic illusions

An Einstellung can also be observed in people’s preference to solve problems using a tried and tested method, rather than continually searching for new solutions. Luchins and Luchins (1950) carried out a number of investigations on this form of Einstellung: they looked at the tendency towards developing a mechanised response in solving certain types of problems and how such a habitual response can stop people from seeing more direct solutions.

The Luchins’ basic task (Luchins, 1942; Luchins and Luchins, 1950, 1959) consisted of volume-measuring problems. Participants were required to use a number of containers of different sizes (*a*, *b*, *c*) to figure out how to obtain a specific volume of fluid. Eleven problems were designed in such a fashion that problems 2 to 6 could all be solved by adhering to the formula *b - a - 2c = [required volume]*, called the ‘Einstellung method’. Then there were two problems (7 and 8) which could be solved both by the established formula, but also by simpler (‘direct’) methods: *a - c* and *a + c*. The ninth problem could not be solved by the Einstellung method, but only by *a - c*. And the final two problems (10 and 11) were like 7 and 8 above. Problems 7 to 11 were called the ‘test problems’ or ‘criticals’, and Einstellung was measured by the percentage of participants who solved problems 7 and 8 by the Einstellung method, by the percentage of those who failed to solve problem 9 because they attempted to use the Einstellung method, and by the percentage of participants who solved problems 10 and 11 by that method, rather than the direct method. Recovery from the Einstellung was measured by comparing the direct solutions achieved for problems 10 and 11 with the number of direct solutions for 7 and 8. For the more than 9000 participants who completed the basic task, the Luchins found that most showed Einstellung, and that recovery from mechanisation was not large. (See Table 8.1 for examples of the problems used.)

Table 8.1: Luchins' waterjars: examples of problems and solutions.

Problem number	Containers given (capacity in quarts)			**To get**	**Problem type**	**Solution**
	a	**b**	**c**			
2	21	127	3	100 quarts	Ein-stellung problems	b-a-2c
3	14	163	25	99 quarts		b-a-2c
4	18	43	10	5 quarts		b-a-2c
5	9	42	6	21 quarts		b-a-2c
6	20	59	4	31 quarts		b-a-2c
7	23	49	3	20 quarts	test problems (criticals)	a-c (b-a-2c)
8	15	39	3	18 quarts		a+c (b-a-2c)
9	28	76	3	25 quarts		a-c
10	18	48	4	22 quarts		a+c (b-a-2c)
11	14	36	8	6 quarts		a-c (b-a-2c)

(Table adapted from Luchins and Luchins, 1950, as cited in Wason and Johnson-Laird, 1968)

8.4.5 Relating the waterjars to semantic illusions

When an individual is faced with the usual kind of semantic illusion task, they are dealing with a situation closely analogous to a Luchins' waterjars task: they are asked a series questions (or required to verify statements) of a particular format. But amongst the fillers are semantic illusion sentences, which, like the critical problems in the Luchins' experiments, are superficially identical to all the other instances. However, there is now a different kind of problem to be solved, and the tried and tested method does not lead to the correct solution. In the case of semantic illusion questions, the participant is required to complete a totally different task – detect a word substitution, when on all other questions the task was to supply an answer.

For statements-to-be-verified the situation is similar, but less obvious. In the case of a verification task, the participant strictly speaking does not suddenly have to carry out a totally different task overall, as a semantic illusion statement is – on one level – no more than a false statement. But there are a number of ways in which a statement can be made false. Usually most people would expect a wrong basic fact to be detectable by virtue of having the wrong semantic associations (e.g. “The Eiffel Tower stands in Berlin.”). In semantic illusion sentences, however, the substitution is designed to fit adequately into the role-slot it occupies in a mental representation of the ideas contained within the statement. So again, it could be argued that there is a subtle change in task. Participants now do not need to find a wrong-seeming fact, but they need to find a substitution in a right-seeming statement. Similar to the Luchins’ critical problems, one solution seems to fit, when another is required.

8.4.6 Trying to prevent Einstellung

Luchins and Luchins carried out a number of variations upon the basic task of the waterjars task in an attempt to get participants to use the direct methods of solution as opposed to the Einstellung method. However, the results showed that it is virtually impossible to prevent Einstellung.

8.4.6.1 Speed of responses

For example, participants performed worse (more Einstellung solutions) when asked to solve the problems under speeded conditions, oftentimes even when they were instructed to solve the problems a second time. Luchins and Luchins suggested that this may in part be cause by ‘ego-motivated forces and social pressures’ experienced by the participants.

In van Jaarsveld, Dijkstra and Hermans' (1997) study, a comparison between substitution detection performance under instructions stressing accuracy and performance under instructions stressing speed and accuracy equally was made. As with the waterjars task, the speeded task was performed less successfully, suggesting that the reading and processing of the sentences in question was carried out mechanically and fairly superficially, superseding any checking of details. If a question appears to be true at first glance, the participant is less likely to detect a substitution under speeded conditions than under conditions where accuracy is stressed.

'Ego-motivated forces and social pressures' may also play a part in processing semantic illusions. The semantic illusion tasks are quite often presented in the form of a general knowledge quiz and consequently participants are likely to want to demonstrate their knowledge as accurately as possible. Since semantic illusion sentences, by necessity, tend to refer to something well-known, which participants are likely to feel they ought to know, it is possible that the same 'ego-motivated forces and social pressures' that the Luchins refer to are at work in participants completion of the semantic illusion tasks.

8.4.6.2 Redundancy in the problem context

Adding a fourth redundant jar to the problem caused a lot of confusion and poorer performance at solving the initial scene-setting problems. Luchins and Luchins comment that the fourth jar made the problem look more complicated, but also made for an unusual type of task, given the nature of school teaching: generally the solution of a problem in a school setting requires all present elements and does not tend to include redundant elements.

While there is no direct parallel between the Luchins' fourth jar variation and semantic illusions, it is nevertheless fair to say that semantic illusions are also an unusual type of problem in a context of familiar problems. There is something misleading in the presentation of a problem which does not require all aspects of it to be involved in meeting the precise demands of the task (the detection of the substitution). So in the jars situation having a fourth unnecessary jar is confusing and distracting, just as the 'answer' to a semantic illusion question has nothing to do with the actual task as it is perceived. A closer analogy between the fourth jar and its effect upon Einstellung and semantic illusions can be observed in the Brédart and Docquier (1989) study, where underlining the substituted word created a slightly unusual problem (sentences in normal discourse do not tend to include capitalised and underlined words to which particular attention needs to be paid). The manipulation also led both to an increase in detections and to an increase in false detections, just as the fourth jar decreased Einstellung responses for both the critical problems and decreased the number of correct solutions for the training problems (problems 2-6).

8.5 How people actually process: Memory for gist

Even if semantic illusions are regarded as a kind of problem to be solved, they nevertheless are also sentences to be processed. Sentence processing is something that people are incredibly good at – they do it quickly and with little apparent effort. People can easily make sense of new sentences that they have never heard or read before and make mental representations of the content of such sentences. But how do people deal with all the linguistic inputs that they are faced with in everyday processing, given that there is potential for infinite variety? Earlier in the chapter, it was stated that there must be a kind of partial matching process in our day to day processing, which leads to errors in the case of semantic illusions. The following section is concerned with a number of experiments which look at how people integrate and process incoming linguistic information in order to make sense of it. Meaning, it would appear, is what is remembered and not surface detail.

8.5.1 Meaning vs. form

Given the idea that our world knowledge is stored in schemata in long-term memory, it seems only natural that only the meaningful content of verbal interactions is retained. When language is comprehended, it seems that the meaning of what is heard or read is remembered, but unless special attention is paid to the specific characteristics of the words, surface details are rapidly forgotten. Sentences appear to be encoded only with respect to their meaning, and the precise grammatical form – and even to a certain extent the precise wording – are important only to support comprehension (c.f. Sachs, 1967).

These findings can essentially be seen as a way of optimising our performance in language processing. They provide a possible explanation for why anything other than a partial matching process would be impossible when dealing with general knowledge questions such as those carried out in semantic illusion type tasks. It is possible that by the time a participant responds to a semantic illusion question, the substituted word will have been instrumental in creating a coherent mental representation which would have been used to reply to the question, while the actual word used will have been forgotten. This assumption is corroborated by the fact that often the question is recalled with the correct name in it, when participant are asked to recall it after the semantic illusion task has been carried out (Erickson and Mattson, 1981).

Similarly, Bransford and Franks (1971) demonstrated that people spontaneously integrated information expressed by a number of semantically related but not consecutively placed sentences into holistic semantic ideas. The results of a series of experiments designed to test the 'abstraction of linguistic ideas' showed clear evidence of what Bransford and Franks referred to as 'productivity' – participants integrated the ideas presented and 'recognised' new sentences made up of a novel combination of propositions that fitted the relations learnt during acquisition, as well as recognising old sentences, but sentences that combined previously unassociated ideas were confidently rejected as not having been seen before.

It is possible that semantic illusions are the result of a similar 'productivity' or integration process: because of the semantic associations between the substitution and the correct word (which are analogous to the learnt relations in Bransford and Franks' experiment), a participant reading a semantic illusion sentence might incorrectly recognise the idea expressed in the sentence and thus overlook the substitution.

8.5.2 Comprehension requires an active contribution

Human beings do a lot of language processing with little conscious control, but in a very efficient way, which allows them to integrate and deal rapidly with new input. This raises the question of how language and general knowledge fit together. In a series of studies Bransford and Johnson and various others investigated the relationship between language comprehension and extra-linguistic knowledge, such as general knowledge of the world and context. Using inference and recall as measures of comprehension, they used a basic strategy of asking participants if they (falsely) recognised information that they had not been presented with during an acquisition phase, and which therefore could only have been inferred.

It can be shown, for example, that participants will 'recognise' different sentences describing the same spatial relations as a previously encountered but differently phrased sentence. Thus participants 'recognise' "Three turtles rested on a floating log and a fish swam beneath *it*” as the sentence they had been presented with previously: “Three turtles rested on a floating log and a fish swam beneath *them*” (Bransford, Barclay and Franks, 1972).

Similarly, participant make inferences about the consequences of input events which are not mentioned in the information studied For example, “The spy threw the secret document into the fireplace just in time since 30 seconds longer would have been too late” would generally lead to the inference that the spy burned the secret document (Johnson, Bransford and Solomon, 1973).

Most people have little trouble when required to process sentences in which two proposition are linked in a way that requires some form of justification (e.g. The floor was dirty *because* Sally used the mop), but they would usually assume additional information (i.e. that the mop was dirty). This suggests that participants’ understanding depends not only upon what is heard/read but also upon the implications of such information in conjunction with their prior knowledge. Such processes of inference and creating justifications are a part of the normal course of comprehension – a listener is essentially required to solve the problem of creating a situation in which, for example, a *because*-structure makes sense. (Bransford and Johnson, 1972; Bransford, McCarrell and Nitsch, 1976).

By analogy with these facts about language comprehension semantic illusions can be regarded as an instance where a participant has understood the context of a given sentence and has identified the relevant referents from their knowledge, while failing to notice that the substituted word does not strictly speaking match the idea that it is seen to refer to. The participant has essentially made the kind of active contribution towards understanding that is often necessary under normal processing circumstances. In a semantic illusion sentence this mechanism leads to an error, when usually it would lead to comprehension.

8.6 Onwards

Chapter 8 was concerned with finding another approach to semantic illusions, treating them as in instance in which usually efficient processing strategies can be demonstrated to lead to errors similar to those encountered in the study of problem solving. An attempt was made to relate these observations to semantic illusions and to partial matching as a theoretical explanation of semantic illusions.

It appears that people's 'rough and ready' partial matching approach may be related to the way in which information is stored in long term memory: as dynamic schemata. In an 'effort after meaning' all incoming information is quickly integrated and much detail about surface-structure is forgotten.

Findings from the study of syllogistic reasoning provided an example of how an inability to focus upon certain specific aspects of incoming information can lead to a faulty approach to a problem. People have a tendency to go with their impression of a situation and to relate it to the totality of their knowledge, finding it hard to step back and analyse things logically and in depth.

The attitude or Einstellung a person has towards a task or problem has a profound effect upon performance. People often have great difficulty in shifting their perception of a given situation away from previous experiences of a similar nature, so that they cannot easily tackle the problem that a semantic illusion task presents: to carefully monitor the details contained within a sentence rather than making sense of its broad content.

Meaning plays a very important part in the human ability to deal with linguistic inputs and it affects how they later remember such inputs. Various studies on language comprehension provide evidence of the fact that a hearer must take an active part in processing language, in order to be able to make sense of language as a symbol system or indeed as a kind of machine which generates understanding by operating (Wittgenstein, 1958). This active contribution is what can lead to failures in detecting substitutions in semantic illusion tasks.

In the next three chapters further experiments on semantic illusions are described, which are concerned not so much with the nature of semantic illusion sentences themselves, but with the way in which they are approached and processed. Chapter 9 describes the effects of instructions and presentation format of semantic illusion sentences on semantic illusion rate; Chapter 10 is concerned with how the context provided by the filler items affects semantic illusions; and Chapter 11 describes two experiments investigating working memory and semantic illusions to explore processing demands and likely depth of processing in dealing with semantic illusion sentences.

Chapter 9:

Experiments 4a and 4b: Changing the point of view

Since neither the overall length nor the position of the target word within the sentence have great effects upon the rate at which semantic illusions occur, but the syntactic form of the sentence (questions or statements) does, it seems possible that the manner in which semantic illusions and semantic illusion tasks are presented is the major determinant of the likelihood of semantic illusions occurring. In Chapter 4 it was shown that questions lead to more semantic illusion responses than statements. It was argued that this was caused by different processing requirements for questions which – by convention – need an answer. This can be seen, for example, by people's consistent classification of the question used in the categorisation task described in Chapter 7. This chapter is concerned with an attempt to prevent the kind of superficial processing that leads to semantic illusions, first by varying the instructions given and then by manipulating the way in which the sentences in a semantic illusion task are presented.

As suggested in Chapter 8, sentences create a specific set of expectations (Grice, 1975) analogous to an Einstellung which stops people from seeing an item as anything other than its usual function (e.g. Duncker, 1945). People appear to find it hard to focus on specific words and are obviously affected by the global meaning of a sentence, and semantic illusion sentences are often responded to depending on how 'right' they feel, a little like the belief bias effect in syllogistic reasoning (Revlin, Leirer, Yopp and Yopp, 1980). The way that semantic illusion sentences are presented encourages this tendency.

A superficially familiar idea is presented in a familiar ‘quiz’ format (true/false choice), so that it appears participants are encouraged to misinterpret the task at hand. Thus a statement-to-be-verified might be seen to test knowledge about a specific fact that is expressed in a coherent sentence. A participant might expect a false statement to be more obviously false than semantic illusion statements usually are. Following this line of argument, it seems likely that participants find it difficult to perceive the semantic illusion statement as having another ‘use’ than simply testing a participant’s knowledge and thus that it is an insight problem of a kind. This suggestion is corroborated by the observation that once attention is directed to the substituted word after the fact, most participants have a type of *Aha!*-experience (cf. Wallas’ (1926) stage of illumination), in which they are suddenly able to see the substitution clearly.

A second analogy would be to compare semantic illusion sentences to the ‘critical’ problems in Luchins and Luchins’s waterjar experiments (Luchins and Luchins, 1950): semantic illusion sentences are superficially identical to the true filler sentences used in semantic illusion experiments, but a different kind of problem has to be solved. For a true statement to be verified, a set of facts has to be recognised and checked against a ‘forgiving’ memory representation, such as a schema (Bartlett, 1995[1932]), which can accommodate a certain degree of discrepancy between a stimulus and the stored information in the process of verification. For a semantic illusion statement, this strategy leads to the wrong response. Here the best strategy to produce a correct response would be to check every word against the schema to identify the extraneous element within the semantic illusion sentence.

Two experiments were designed based on these analogies, to see if participants would be more successful at detecting substituted words in semantic illusion statements, if the presentation of the task was altered. In Experiment 4a, the instructions were presented in a non-standard way to attempt to reassign the task focus and to encourage participants to think differently about the actual phrasing of the statements at hand. Experiment 4b took this idea a step further, by changing the task demands more drastically at the same time as eliminating any possibility of functional fixedness effects created by complete sentences: in this experiment participants were presented with a series of sentence puzzles consisting of the component parts of sentences that they were required to reconstruct before answering the question of whether or not each sentence was truthful. It was thought that this manipulation would force participants to pay more than superficial attention to each component of a sentence.

9.1 Experiment 4a: Varying the instructions

In Experiment 4a, the basic task was essentially the same as that used in many previous semantic illusion studies. Participants were required to read a series of statements and make a decision about whether or not they contained truthful information. However, the instructions were changed in order to avoid the words ‘true’ and ‘false’ which were thought to be ‘philosophically charged’ as a consequence of which participants might start speculating about the ‘Truth’ behind some of the claims they were asked to verify against their background knowledge. In fact, in Experiments 2 and 3 individual participants made comments to that effect about some of the statements-to-be-verified which related to well-known fictional characters. Instead participants were asked to decide if each statement was ‘natural’ or ‘unnatural’ according to the guidelines given in the instructions. It was thought that participants might find it easier to spot more substitutions with these instructions, as they might make the task seem as if it was a more subjective choice than a true or false decision to test participants’ knowledge state.

9.1.1 Method

Participants: 18 adult volunteer participants with at least University level education took part in this study.

Stimuli: A set of 52 true/false statements was presented to each participant. Twelve of the statements were target statements, which included a substituted name. The other 40 sentences were fillers constructed along similar lines as the targets. Nine of these statements were false, the other 31 were true.

Procedure: Each participant was given a leaflet of six sheets of A4 paper, with the following instructions printed on the first page:

> For Section 1.
>
> Below you are going to see a series of 52 statements. Some of these statements are correct and contain everyday trivia or facts. Such sentences are considered to be “NATURAL”. Some of the statements contain elements that are incorrect thus making the sentence “UNNATURAL”. (Please take myths and legends on their own terms: if that’s how the story goes, it’s “natural” for the purposes of this quiz.)
>
> Your task is to read each statement and to judge which category each sentence belongs to and then to indicate (by placing an “X” in the brackets [ ] next to your choice of answer) EITHER “NATURAL” OR “UNNATURAL”.
>
> Please note that this research is as much concerned with your immediate reaction as with your actual answer, so please response promptly. If you should wish to correct an answer you have made, please indicate this by placing a “1” in the bracket by your immediate response, and a “2” in the bracket by you final response.
>
> For Section 2.
>
> Please do not look at part 2 until you have completed section 1.
>
> Section 2. consists of 12 multiple-choice questions. Please enter the letter corresponding to your choice of answer in the space provided.
>
> Thank you very much for taking part!

The final page of the booklet was there to provide a knowledge check: twelve multiple-choice questions, which tested whether participants had the prerequisite knowledge to experience a semantic illusion. Instructions for section 2. were printed at the top of the last page:

Section 2. Knowledge check.
Please indicate for each of the following statements which option (a), (b), (c) or (d) best fits in the gap in order to make the statement "natural". If you do not know the answer at all, please indicate this.

(An example of the questionnaire is presented in Appendix 9).

9.1.2 Results

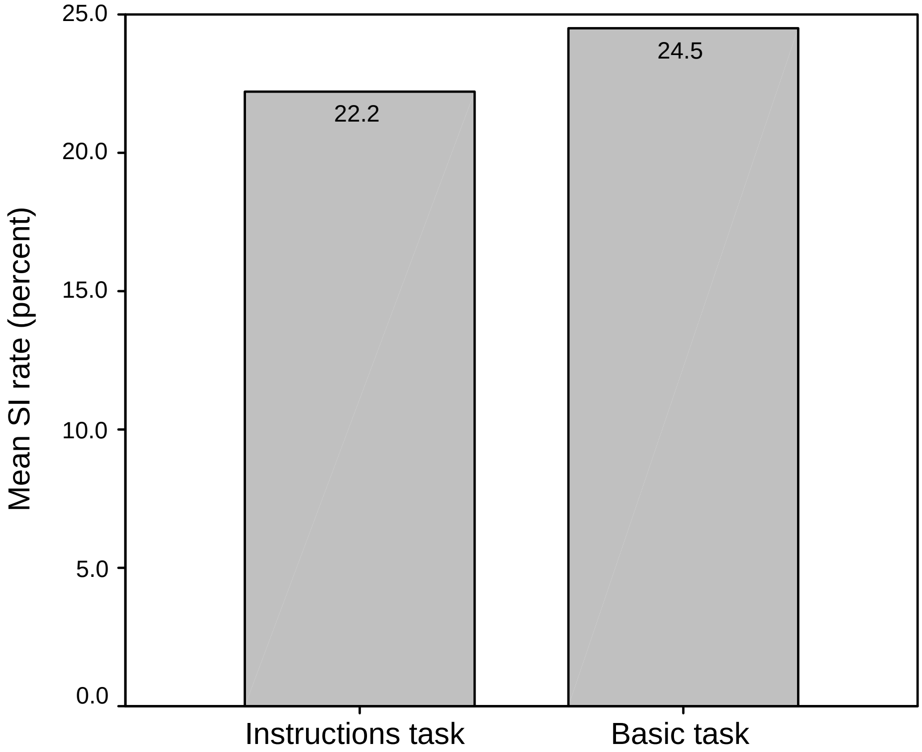
A response was coded as a semantic illusion response only for statements where the participant judged the statement to be natural *and* gave the correct answer in the multiple choice knowledge check. If a participant responded correctly to the statement *and* the corresponding knowledge check, the response was coded as correct. The number of semantic illusion responses made for each target sentence was recorded, as was the number of correct responses. Then the semantic illusion rate and the correct response rate were calculated for each individual sentence.

A t-test for paired samples was carried out using sentences as subjects, to compare the mean semantic illusion rate in this study to the mean semantic illusion rate for the same sentences presented as straightforward statements-to-be-verified (asking participants to make a true/false judgement about the statement). Contrary to the hypothesis, the mean semantic illusion rate for the 'instructions' task was not found to differ significantly from the basic statements-to-be-verified task ($t= -0.410$; $d.f.= 11$; $p= 0.345$, one-tailed. Means and standard deviations are presented in Table 9.1; see also Figure 9.1).

Table 9.1: Mean semantic illusion rate (percent) and standard deviations for statements-to-be-verified in the 'Instructions' task and for the basic statements-to-be-verified task.

	N	Mean SI rate	Std. Dev.
Instructions	12	22.2	14.6
Basic Statements	12	24.5	14.7

Figure 9.1: Mean semantic illusion rate for statements-to-be-verified for the 'Instructions' task compared with mean semantic illusion rate for basic statements-to-be-verified task.

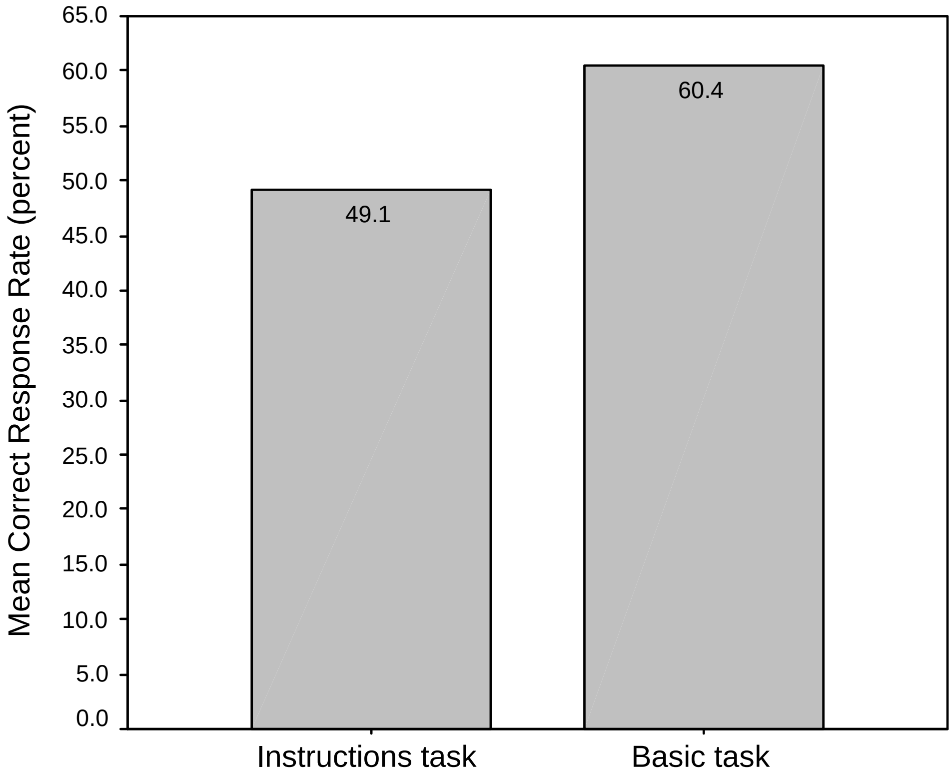


Correct response rates were also compared for the instructions task and the basic task using a t-test for paired samples with sentences as subjects. It was found that significantly fewer correct responses were made in response to the instructions task than for the basic task ($t= -1.937$; d.f.= 11; $p= 0.04$, one-tailed; see Table 9.2 and Figure 9.2).

Table 9.2: Mean correct response rate (percent) and standard deviations for statements-to-be-verified in the 'Instructions' task and for the basic statements-to-be-verified task.

	N	Mean correct rate	Std. Dev.
Instructions	12	49.1	17.7
Basic Statements	12	60.4	16.1

Figure 9.2: Mean correct response rate for statements-to-be-verified for the 'Instructions' task compared with mean semantic illusion rate for basic statements-to-be-verified task.



9.1.3 Discussion

Even though the semantic illusion rate with changed instructions was not significantly different to the semantic illusion rate in the basic statements-to-be-verified task, the correct response rate for the instructions task was significantly lower, suggesting that there is an effect of type of instruction upon processing, even if it was not the anticipated effect. It had been hypothesised that instructions for a slightly less familiar task (deciding the 'naturalness' of sentences according to the provided guidelines) would carry with them fewer task-dependent expectations, from previous experiences with the similar tasks. The results of the experiment confirmed that this was the case, but the effect of the change in instructions was to reduce overall accuracy and not to decrease semantic illusion rates as predicted. It seems likely that participants found the unfamiliar task of judging the naturalness of a statement to be somewhat confusing, leading to fewer correct judgements. The fact that the semantic illusion rate did not differ could be explained by the idea that participants may have mentally converted the task requirements from the instructions to the more familiar and well-practised true/false paradigm, thus eradicating the potential effect of the instructions upon semantic illusion rate. Either way, the change from the true/false task to the natural/unnatural task did not affect the semantic illusion rate or increase the number of correct responses, suggesting that this manipulation did not help participants in the detection of substitutions.

9.2 Experiment 4b: Varying the presentation format

Since simply varying the instructions did not decrease the semantic illusion rate, a stronger manipulation with which to change the perception of the task at hand was used. Turning the sentences into puzzles – by decomposing the sentences into their component parts and asking participants to decide whether or not each sentence puzzle could be turned into a truthful statement – was chosen to provide a sufficiently different task. It was hypothesised that participants would be compelled to pay attention to each individual word within a sentence puzzle, as they had to decide how it fitted into the overall statement. This task was also thought to remove the risk of the statement-to-be-verified structure leading participants into a functionally fixed approach to the verbal materials presented, in which they would respond primarily to how 'right' the individual statements feel. The hypothesis was that in the sentence puzzle paradigm, participants would be more likely to correctly identify substituted words, because the sentence flow was not given from the onset but had to be constructed from the individual elements of each sentence, so that each sentence component had to be considered individually.

9.2.1 Method

Participants: 16 adult volunteer participants with at least University level education participated in this study.

Stimuli: 26 sentence puzzles were presented to each participant. (For examples of sentence puzzles, see description below in the *Procedure* section and Appendix 10.) Six of the sentence puzzles were targets, which included a substituted name. The other 20 sentence puzzles were fillers constructed along similar lines as the target. Four of these fillers could not be made into a complete and truthful sentence, the other 16 could.

Procedure: Each participant was given a leaflet of four sheets of A4 paper, with the following instructions printed on the first page:

Sentence puzzle quiz

For Section 1.

Below you will see a series of 26 sentence puzzles. Each of these puzzles consists of a number of meaningful components represented by one or more words in square brackets (e.g. **[Last year's], [took place], [in Sydney], [Olympic Games]**). Please look carefully at the sentence components. **Without** changing any of these components, is it possible to combine all of the components into a truthful statement? Please circle the appropriate answer. If you answer YES, please write down the sentence, if you answer NO, please note down why it does not work. So, for the example above, the correct answer would be "YES" and the truthful statement would be "Last year's Olympic Games took place in Sydney." (N.B. the first element of each sentence starts with a capital letter.)

And here is another example:
[released], [Ali Baba], [the lamp], [by rubbing], [the genie]
The correct answer in this case would be NO. While the components could be made into a coherent statement ("Ali Baba released the genie by rubbing the lamp."), the statement itself is not true: Aladdin rubbed the lamp, not Ali Baba.

Alternatively the answer could be wrong for another reason:
e.g. **[the Prime Minister], [Tony Blair], [of the United Kingdom]**
The correct answer in this case is NO because there is no verb, and therefore the components cannot be turned into a truthful complete sentence.

For Section 2.

Please do not look at section 2 until you have completed section 1.

Section 2. consists of 6 multiple-choice questions. Please answer these as accurately as possible.

Thank you very much for taking part in this study.

The final page of the booklet was there to provide a knowledge check: twelve multiple-choice questions, which tested whether participants had the prerequisite knowledge to experience a semantic illusion. Instructions for section 2. were printed at the top of the last page:

> **Section 2. Knowledge check.**
> Please indicate for each of the following statements which option (a), (b), (c) or (d) best fits the gap in order to make the statement truthful. If you do not know the answer at all, please indicate this.

9.2.2 Results

Unfortunately not all participants in this study complied with the instruction to make a note of the reason why they classified sentence puzzles as being solvable or not, nor did all participants write down the correct solution as they saw it for each puzzle. Hence a response was coded as a semantic illusion response for sentence puzzles where the participant responded "Yes" to the puzzle *and* gave the correct answer in the multiple choice knowledge check. If a participant responded correctly to the statement *and* the corresponding knowledge check, the response was coded as correct. The number of semantic illusion responses made for each target puzzle was recorded as was the number of correct responses. Then the semantic illusion rate and the correct response rate were calculated for each individual puzzle.

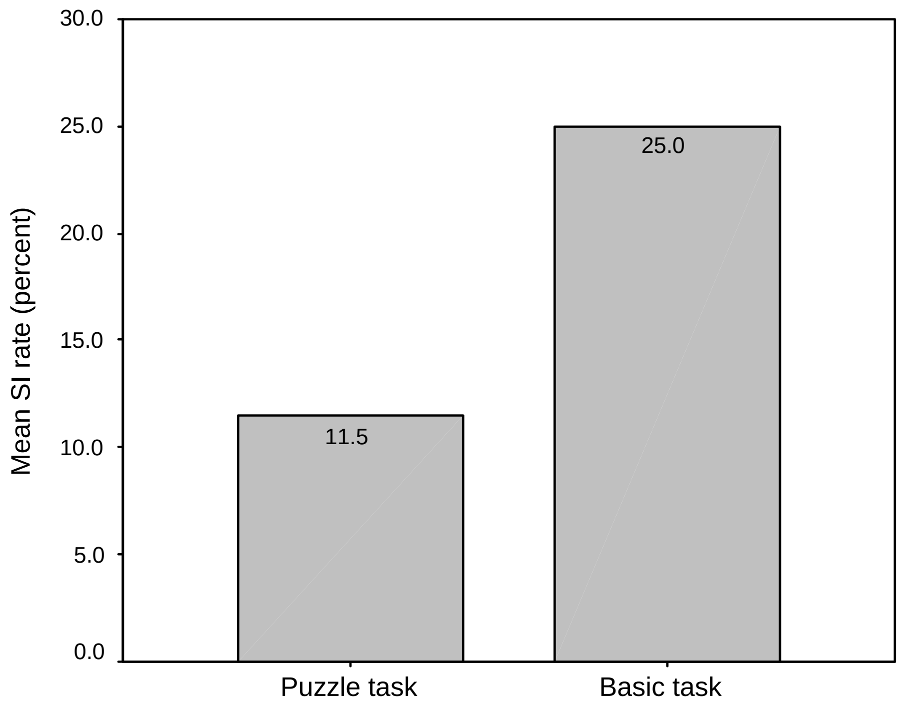
A t-test for paired samples was carried out using sentences as subjects, to compare the mean semantic illusion rate for the puzzle task to the mean semantic illusion rate for the same sentences presented as straightforward statements-to-be-verified (asking participants to make a true/false judgement about the statement). As predicted, the mean semantic illusion rate for this study was found to be significantly smaller than the semantic illusion rate for the basic statements-to-be-verified task (t= -5.398; d.f.= 5; p=

0.002, one-tailed. Mean semantic illusion rates and standard deviation are presented in Table 9.3; see also Figure 9.3).

Table 9.3: Mean semantic illusion rate (percent) and standard deviations for statements-to-be-verified in the 'Puzzle' task and for the basic statements-to-be-verified task.

	N	Mean	Std. Dev.
Puzzles	6	11.5	7.3
Basic Statements	6	25.0	10.5

Figure 9.3: Mean semantic illusion rate for statements-to-be-verified for the 'Puzzle' task compared with mean semantic illusion rate for basic statements-to-be-verified task.

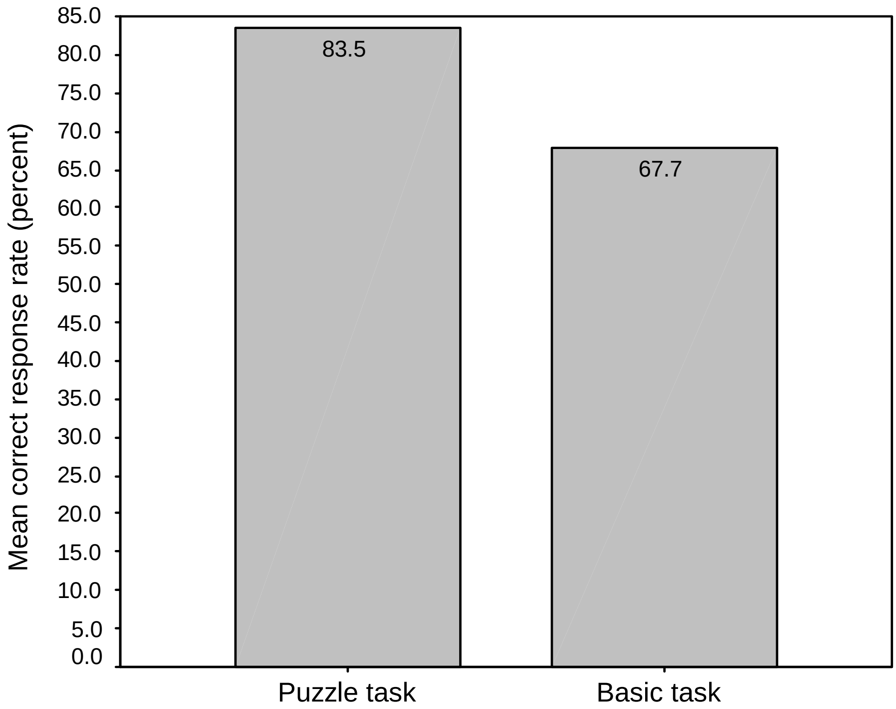


Correct response rates were also compared for the 'puzzle' task and the basic task using a t-test for paired samples with sentences as subjects. In keeping with the findings from the semantic illusion rate analysis, it was found that significantly more correct responses were made in response to the 'puzzle' task than for the basic task (t= 3.532; d.f.= 5; p= 0.09, one-tailed; see Table 9.4 and Figure 9.4).

Table 9.4: Mean correct response rate (percent) and standard deviations for statements-to-be-verified in the 'Puzzle' task and for the basic statements-to-be-verified task.

	N	Mean correct rate	Std. Dev.
Puzzles	6	83.5	6.4
Basic Statements	6	67.7	11.5

Figure 9.4: Mean correct response rate for statements-to-be-verified for the 'Puzzle' task compared with mean semantic illusion rate for basic statements-to-be-verified task.



9.2.3 Discussion

As hypothesised, participants were much more likely to detect substitutions in this task than in the straightforward statement-to-be-verified paradigm. It seems that presenting semantic illusions as problems to be solved rather than as facts to be verified does make them easier to detect, as is also suggested by the increase in correct response rate in the 'puzzle' task. However, a number of semantic illusion responses still occurred even with this task. The semantic illusion rate for the sentence puzzles at 11.5 percent is comparable to that in focus-shifting semantic illusion paradigms for the condition in which the substituted word is placed in focus. For example, Brédart and Modolo's

(1988) focus-condition showed a semantic illusion rate of 9 percent. This suggests that the importance of meaning extraction is greater than the ability to pay attention to specific details, even in situations where that meaning has to be constructed by rearranging different elements of a sentence. It is possible that the semantic illusion rate in this task reflects the strategy used by some of the participants of first constructing a sentence out of the available materials and then deciding if the sentence was true or false. There is some evidence for this in that nine out of the sixteen participants in this experiment did not make any semantic illusion responses and a further four only made one semantic illusion response each.

9.4 Conclusions

The two experiments described in this chapter were concerned with changing the way in which participants approached the statements in a statement-to-be-verified task, as it was assumed that semantic illusions ensued from a type of functional fixedness in which the expectations made about the task and sentences at hand prevented participants from spotting the substitutions made in semantic illusion statements. An attempt to increase participants' sensitivity to substitutions by changing the basic task ('true/false') to a slightly 'unusual' task ('natural/unnatural') had no significant effect upon the semantic illusion rate, possibly because participants reinterpreted the 'unusual' task to be equivalent to the basic task. However, the unusual task lead to fewer correct responses, suggesting that changing the instructions affects overall performance, but not in the way predicted. When the nature of the task was shifted by a stronger manipulation of requiring participants to decide if a number of sentence puzzles could be assembled to form truthful statements, the semantic illusion rate dropped significantly, suggesting that this task forced participants to pay more attention to each component of the sentence. This finding was also supported by the fact that the correct response rate in the ‘puzzle’ task was significantly higher than that in the basic task. But semantic illusions were still not eradicated.

Chapter 10:

Experiment 5: Expectation from context

Section 8.3 described how people's thinking is far from logical, and how it is subject to a variety of biases related to the way in which information is presented (e.g. Woodworth and Sells, 1935; Wetherick and Gilhooly, 1990), and to what is believed about the world (Revlin, Leirer, Yopp and Yopp, 1980). Since it is unlikely that such biases are confined to syllogistic reasoning, this chapter is concerned with the question of how the context provided within each semantic illusion task affects the rate at which semantic illusions occur.

In experiment 2 (see Chapter 5) and experiment 3 (see Chapter 6) the correct response rates for true and false filler statements were compared to examine the assumption that false statements were generally harder to identify than true statements. The results showed that participants gave significantly more correct responses to true fillers than to false fillers. It was also found that the correct response rate for false filler statements was statistically indistinguishable from the correct response rate for target statements (i.e. semantic illusion detections). This could be due to the fact that semantic illusion statements are technically just false statements, but the semantic illusion research described so far provides no empirical evidence to support this assumption. The explanation that was suggested in Chapter 5 to explain participants' performance was that in general positive statements are easier to process than corresponding negative ones, as the negative statements require extra steps in processing (see section 5.4). Similarly it can also be shown that participants take longer to verify an incorrect statement when comparing sentences against pictures (Clark and Chase, 1972). By analogy it seems likely that it is easier to verify a truthful statement against stored knowledge than it would be to disconfirm a false statement, as was found in experiment 2 (see section 5.6.3).

Given the nature of statement-verification tasks used in semantic illusion research to date, there may however be another cause contributing to the greater rate of correct responses to true filler statements than to false fillers: a form of Einstellung (see Chapters 8 and 9). In the case of a statement-to-be-verified task, the target sentences are usually surrounded by many true fillers and only few false statements apart from the target sentences. This set-up might in itself lead to a Luchins and Luchins (1950) –type Einstellung, that is based on the use of a previously successful response strategy which does not work for certain superficially identical problems.

In Chapter 9 (Experiment 4b) it was demonstrated that semantic illusion sentences as *sentences* are subject to a type of functional fixedness (cf. Duncker, 1945; Birch and Rabinowitz, 1951; see Chapter 8). Participants find it virtually impossible to break the habit of processing sentences as holistic units and for meaning, even if they are asked to pay attention to details (cf. ‘literal’ tasks; Reder and Cleeremans, 1990; Reder and Kusbit, 1991). This difficulty in analysing the component parts of a sentence is a little like a global version of the Stroop effect (1935): just as a printed colour word interferes with the task of naming the colour of the ink in which the word is printed, so people cannot help processing the semantic content of a sentence. The ink colour of the word as a physical characteristic of the stimulus in the Stroop effect is analogous to the individual words connected in a sensible fashion in the sentence in a semantic illusion task: attention can be paid to each word individually, but it is much easier to process the gist of the sentence.

Since it can be shown empirically that Einstellung affects the processing of semantic illusion sentences, and since many statement-to-be-verified tasks in semantic illusion research to date have used a majority of fillers that were true, it is possible that the difficulty of processing negative as opposed to positive statements is not the only reason for participants' poorer performance in verifying false and semantic illusion statements. The correct response rates for semantic illusion statements (all plausible and dealing with reasonably familiar topics) in experiments 2 and 3 were not different from the response rates for other false statements (often more obviously incorrect than semantic illusion sentences). This suggests that maybe the overall context of the task, in which the participants were required to make a large number of "true" responses to the true filler statements, had a direct effect upon semantic illusion rate. Perhaps in the context of many true statements the semantic illusion sentences appear more likely to be true at first glance, so that – if there is any doubt – a participant might be inclined to accept the statement as true based on a combination of the feeling of coherence associated with the sentence and the overall context of the task involving many true responses. Conversely, a target sentence surrounded be many false fillers might under similar circumstances by more likely to be rejected.

To test this assumption a questionnaire was designed and distributed via the internet. One version of the questionnaire was designed to consist of primarily false statements, while the other version was similar in design to the format used in earlier statement-to-be-verified studies, with most fillers being true. It was hypothesised that the semantic illusion rate in a questionnaire consisting of a majority of false fillers would be lower than the semantic illusion rate in a questionnaire with a majority of true fillers. Furthermore it was thought that comparing the correct response rates for true fillers and for false fillers for the two versions of the questionnaire would shed light on whether or not the better performance on correct fillers in experiments 2 and 3 was due to an 'expectation' effect, or whether false statements are actually harder to verify than true statements. It was hypothesised that the correct response rate would be greater for the true version of each filler statement than for the false version, if negative statements were harder to analyse than positive ones.

Finally, a context effect could be tested for by comparing the correct response rates for true and false fillers and for targets in each group. If there is a context effect, one would expect to see a different pattern of response rates for each group: a questionnaire consisting of a majority of false fillers should lead to the greatest correct rate for responses to false statements, while for a questionnaire consisting primarily of true statements, the greatest correct rate should be for responses to true statements, regardless of whether or not false statements are harder to verify than true ones.

10.1 Method

Participants: 138 volunteers participated in this study. Participants were recruited by email informing them of the web-address of the questionnaire used in the study. They were invited to forward the address to as many people as possible to find further participants. The website was left active for a week to gather data, and then taken off-line. Participants were assigned randomly to one of the two versions of the questionnaire. 65 participants saw version one ('Mostly false') and 73 participants saw version two ('Mostly true').

Stimuli: 50 true/false statements were presented to each participant. Twelve of the statements were target statements, which included a substituted word. The other 38 sentences were fillers constructed along similar lines as the target. The two versions of the questionnaire were: 1 (‘Mostly false’) and 2 (‘Mostly true’). For the ‘Mostly false’ version 27 filler statements were false and eleven were true; for the ‘Mostly true’ version the same statements used in version 1 were manipulated so that the 27 previously false statements were now presented as 27 true statements and the eleven previously true statements were presented as false statements.

The questionnaire was presented via the internet as an interactive online quiz written in PHP3 script (see Appendix 11 for screen captures of the website used). The data was recorded and sent directly to a secure databank from which it could downloaded for analysis.

Procedure: When a participant opened the questionnaire website, the first screen showed instructions to read each statement as quickly as possible and to respond by using the mouse to click either a true or a false button displayed below the statement. The statements-to-be-verified were displayed one at a time and disappeared only when the participant made a response. The next statement was then displayed. Once all 50 statements had been answered, a knowledge check was displayed. This consisted of a ‘fill-the-blank’ version of each statement used with 4 multiple choice options to complete it. The 50 knowledge check statements were all displayed on the same screen and participants were instructed to respond to these as accurately as possible, by ‘ticking’ one of the response options for each with the mouse. Once a participant had ‘ticked’ one of the multiple choice options for each statement, a button at the bottom of the screen could be clicked for feedback – this displayed the knowledge check-list with correct answers highlighted. Wrong answers that had been given were also highlighted in a different colour. Participants were also presented with a score out of fifty for the first half of the experiment.

10.2 Results

10.2.1 Semantic illusion rate and correct response rate

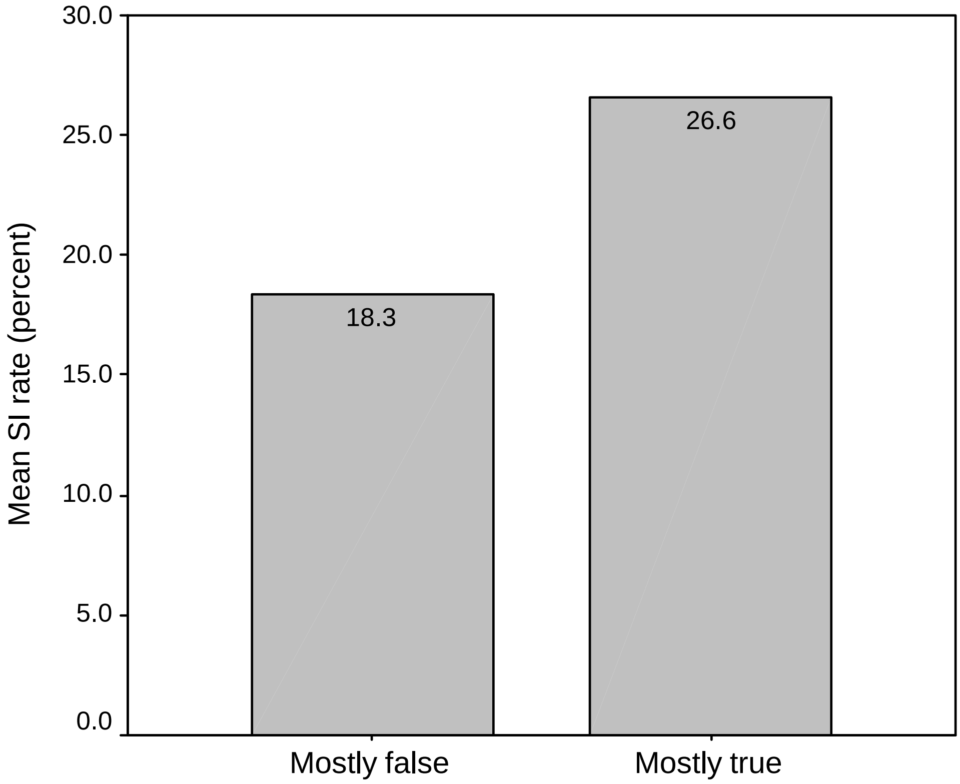
A response was coded as a semantic illusion response for statements where the participant judged the statement to be true *and* gave the correct answer in the multiple choice knowledge check. When a participant gave the correct response to both a statement *and* the corresponding knowledge check, the responses was coded as correct. The semantic illusion rate and the correct response rate were recorded for each target statement for each group.

A t-test for paired samples was carried out comparing the semantic illusion rate for each target sentence in the 'Mostly false' version to that for the corresponding statement in the 'Mostly true' version (i.e. sentences were used as subjects). In keeping with the hypothesis, the mean semantic illusion rate in the 'Mostly false' version was found to be significantly lower than that in the 'Mostly true' version of the questionnaire ($t= -4.296$; $d.f.= 11$; $p= 0.001$, two-tailed). The means and standard deviations are shown in Table 10.1 (see also Figure 10.1).

Table 10.1: Mean semantic illusion rates (percent) and standard deviations for 'Mostly false' and 'Mostly true' versions.

Version	N	Mean SI rate	Std. Dev.
'Mostly false'	12	18.3	9.6
'Mostly true'	12	26.6	13.5

Figure 10.1: Mean semantic illusion rates for 'Mostly false' and 'Mostly true' versions.

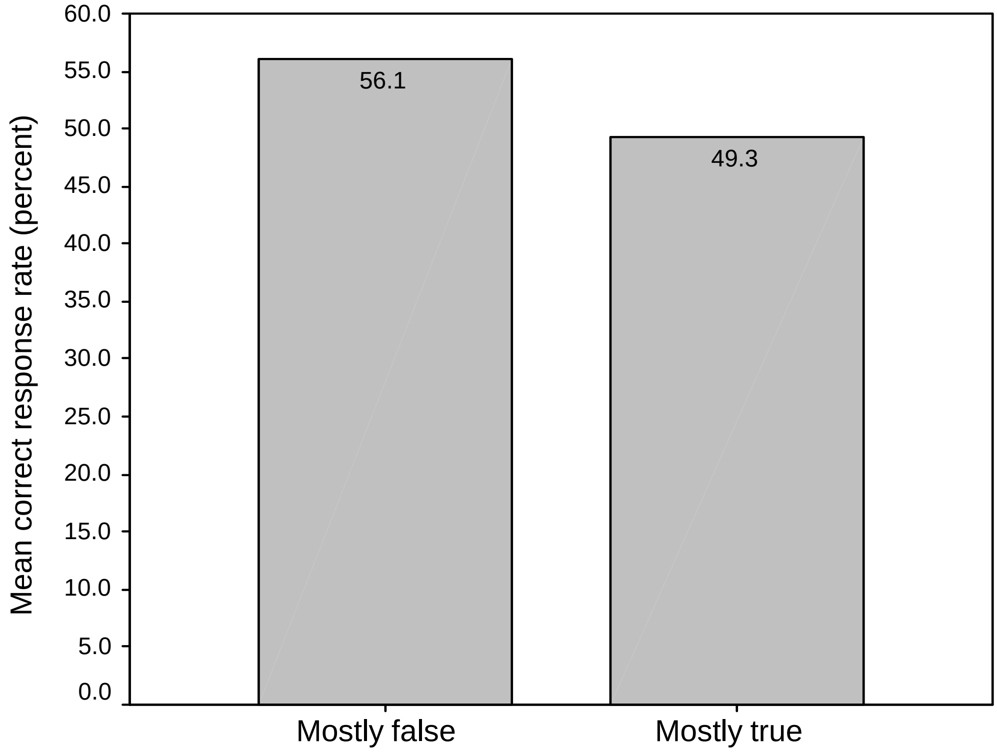


A t-test for paired samples was also carried out comparing the correct response rate in the 'Mostly false' version to that in the 'Mostly true' version, using sentences as subjects. In keeping with the finding that fewer semantic illusion responses were made in the 'Mostly false' group, it was also found that the number of correct responses to target statements was significantly greater in the 'Mostly false' version than in the 'Mostly true' version of the questionnaire (t= 2.524; d.f.= 11; p= 0.028, two-tailed). The means and standard deviations are shown in Table 10.2 (see also Figure 10.2).

Table 10.2: Mean correct response rates (percent) and standard deviations for 'Mostly false' and 'Mostly true' versions.

Version	**N**	**Mean correct rate**	**Std. Dev.**
'Mostly false'	12	56.1	22.8
'Mostly true'	12	49.3	21.0

Figure 10.2: Mean correct response rates to target statements for 'Mostly false' and 'Mostly true' versions.



10.2.2 Correct response rates for true fillers, false fillers and targets

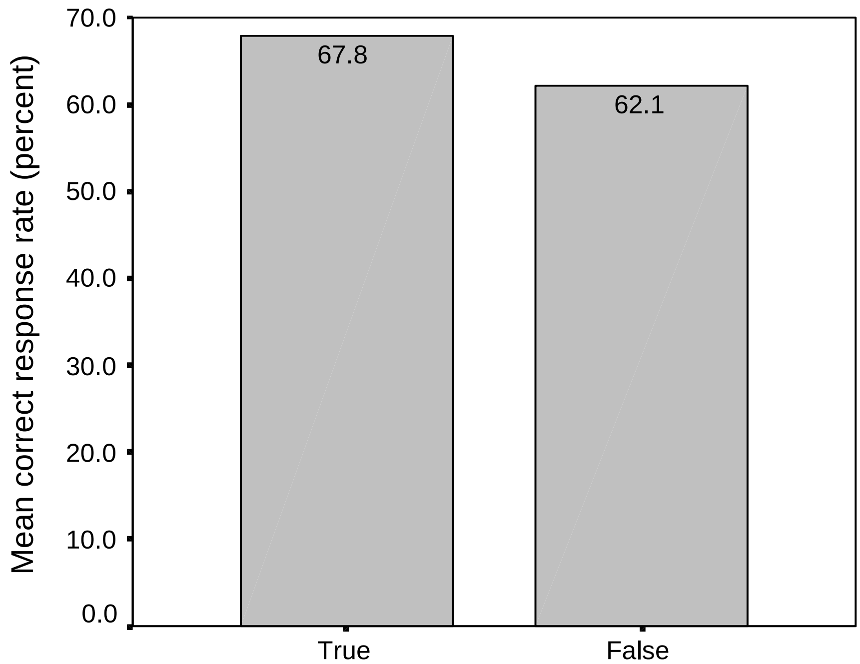
To test the hypothesis that a positive statement is easier to process than a negative statement, correct response rates for each filler statement were also recorded. A filler was coded to have been answered correctly when the correct response had been given during the verification task (i.e. 'true' for true fillers and 'false' for false fillers) *and* the knowledge check had been answered correctly.

A t-test for paired samples was carried out comparing the correct response rate for each filler in its true form to the correct response rate for its false form. As predicted, the truthful version of a statement was found to be significantly easier to judge correctly than the false version of the same statement (t= 1.785; d.f.= 37; p= 0.042). But while significant, the difference in correct response rate for true and false statements is not very large (see Table 10.3 and Figure 10.3).

Table 10.3: Mean correct response rate (percent) and standard deviations for true and false statements.

	N	Mean correct rate	Std. Dev.
True	38	67.8	21.1
False	38	62.1	20.5

Figure 10.3: Mean correct response rates for true and false statements.



A repeated measures analysis of variance was carried out on the correct response rates for correct target responses, correct false fillers and correct true fillers, for both groups ('Mostly false' and 'Mostly true') combined, to determine if there was an overall effect of statement type on correct response rate. (Participants were used as subjects). Mauchly's test of sphericity was found to be significant (W= 0.931; d.f.= 2; p= 0.008) and Pillai's Trace was used to determine the significance of the analysis of variance. The effect of type was found to be significant (F= 18.108; d.f.= 2, 135; p<0.001) and there was a significant interaction between type and group (F= 57.505; d.f.= 2, 135; p<0.001). A series of post-hoc comparisons were carried out, using t-tests for paired samples with a Bonferroni correction (p<0.0167 for 0.05 level of significance). It was found that false fillers led to fewer correct responses than true fillers (t= -2.648; d.f.= 137; p= 0.009, two-tailed). True fillers led to more correct responses than targets (t=5.533; d.f.=137; p<0.001, two-tailed). False fillers also led to more correct responses than targets (t= 3.572; d.f.=137; p<0.001, two-tailed). These results were consistent with the hypothesis that true statements are easier to verify than false statements (see Table 10.4 for means and standard deviations).

Table 10.4: Mean correct response rate (percent) and standard deviations for each type of statement.

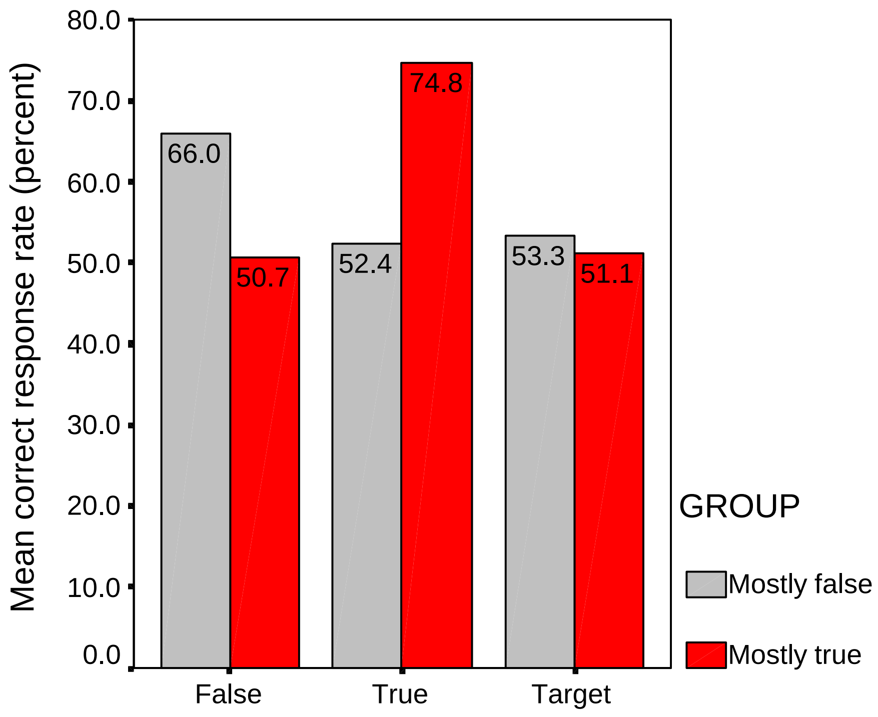
Type of statement	N	Mean correct rate	Std. Dev.
True filler	138	64.2	17.4
False filler	138	57.9	17.7
Target	138	52.2	20.6

To explore the interaction, the mean correct rates were calculated for each combination of group and type of statement (see Table 10.5 and Figure 10.4).

Table 10.5: Mean correct response rates (percent) and standard deviations (in parentheses) for each combination of group and statement type.

	Statement type		
Group	**False**	**True**	**Target**
'Mostly false'	66.0 (16.5)	52.4 (15.1)	53.3 (20.1)
'Mostly true'	50.7 (15.7)	74.8 (11.5)	51.1 (21.1)

Figure 10.4: Mean correct response rates for each combination of group and statement type.



A separate repeated measures analysis of variance was carried out for each group to examine the effect of statement type. For the ‘Mostly false’ group, Mauchly’s test of sphericity was significant (W= 0.852; d.f.= 2; p= 0.006) and Pillai’s Trace was used. There was a significant effect of type of statement on correct response rate (F= 23.489; d.f.= 2, 63; p<0.001) and a series of post-hoc t-tests was carried out using a Bonferroni correction. It was found that in the ‘Mostly false’ group, participants made significantly more correct responses to false statements than to either targets (t= 5.987; d.f.= 64; p<0.001, two-tailed) or true statements (t= 4.904; d.f.= 64; p<0.001, two-tailed). There was no statistical difference between the correct response rate for true statements and targets (t= -0.313; d.f.= 64; p= 0.755, two-tailed).

For the ‘Mostly true’ group, the sphericity assumption was not violated. There was a significant effect of type of statement on correct response rate (F= 73.699; d.f.= 2, 144; p<0.001) and a series of post-hoc t-tests was carried out using a Bonferroni correction. It was found that in the ‘Mostly true’ group, participants made significantly more correct responses to true statements than to either targets (t= 9.738; d.f.= 72; p<0.001, two-tailed) or false statements (t= -10.909; d.f.= 72; p<0.001, two-tailed). There was no statistical difference between the correct response rate for false statements and targets (t= -0.193; d.f.= 72; p= 0.847, two-tailed).

The findings from these two analyses are consistent with the hypothesis that the context of the questionnaire affects the correct response rate depending upon which kind of statement is predominantly used for fillers.

10.3 Discussion

As predicted, target sentences embedded in a context of mostly false filler statements were less likely to lead to semantic illusion responses (and conversely to more correct responses to target questions) than the same target statements in a context of mostly true fillers. This indicates that the context of the task has an effect upon how the sentences are processed. If participants find a large number of statements-to-be-verified to be false, they appear to be more likely to assume any sentence to be false, if there is any doubt about its truthfulness, as is potentially the case for semantic illusion sentences. This assumption was also supported by the results of the analysis of the correct response rates for each group. In the 'mostly true' group, participants gave the greatest number of correct responses to true fillers. The combined analysis of correct response rates for both experimental groups provided evidence for the claim that truthful statements are easier to verify than false ones (cf. Horn, 1989; Clark & Chase, 1972).

In terms of semantic illusions and their occurrence, these results imply that there is an effect of the overall context of the specific task in which the target sentences are included on the semantic illusion rate, but this effect is small and the correct response rate to target sentences is not affected by the task context: indeed for both types of context, the correct response rate for semantic illusion sentences is very similar to the correct response rate for the not-favoured type of filler (i.e. in a 'mostly false' context it is similar to the correct response rate for true fillers).

The finding from the 'mostly true' group seems straightforward enough: semantic illusion sentences are strictly speaking false statements, and hence the correct response rate for semantic illusion sentences is the same as that for false fillers. But the finding from the 'mostly false' groups suggests that something else is going on: in a 'mostly false' context the semantic illusion sentences appear to be treated as if they were true statements. This finding is in keeping with the data described in Chapter 7, which showed that semantic illusions are not like other sentences.

It seems as if a functional fixedness explanation as put forward in the introduction of this chapter cannot account for these data. Maybe a better – less obvious – analogy would be to describe the response strategy used by participants as a variant of the matching effect observed in syllogistic reasoning. In the matching effect (Wetherick & Gilhooly, 1990), the premises of a syllogism are processed superficially and the participant makes a response by matching the more conservative of the two premises. In the semantic illusion matching analogy, a statement about which there is any doubt – such as a target statement which may lead to a semantic illusion – is matched to other statements already verified, and the more conservative response is given depending on the context.

10.4 Conclusions

The experiment described in this chapter was concerned with the effect of filler statements on semantic illusion rate. It was assumed that a semantic illusion would be less likely to occur for the same target sentence, when it was encountered in the context of mostly false fillers. The results of this study confirmed this assumption. However, the context did not appear to affect people's ability to spot substitutions in the target sentences, as there was no effect of context on the correct response rate for semantic illusions. The data from this study also re-confirmed the claim that false statements on the whole are harder to verify than true statements as discussed in chapter 5. It also showed that semantic illusions were even harder to verify than false statements providing further evidence that semantic illusion sentences are highly atypical as discussed in Chapter 7.

Once again it seems that semantic illusions are difficult to process successfully even though they do not appear to be difficult to understand. This raises the question of whether semantic illusions occur as a result of the capacity limits of the processing system. The next chapter will explore the processing demands semantic illusion sentences place upon the system by relating semantic illusion to working memory research.

Chapter 11:

Working Memory and Semantic Illusions

From previous research including the experiments described in Chapters 3 to 6, it seems clear that it is not easy to affect semantic illusion rate by anything short of a drastic manipulation such as changing the sentence format into a kind of puzzle (Chapter 9) or by changing the focus of the sentence by a syntactic mechanism such as putting the substituted word into a cleft phrase (Brédart and Modolo, 1988). But even with these relatively extreme measures to draw participants' attention to the substitution, the occurrence of semantic illusions is not entirely eradicated. In fact, none of the experimental manipulations to date have managed successfully to prevent semantic illusions from taking place. In Chapter 7, semantic illusion type sentences were investigated in comparison with sentences taken from sources of written discourse. Participants rated semantic illusion type sentences (without a substituted target word) on a series of scales designed to provide information about the nature of semantic illusion sentences compared to other sentences. It was found that semantic illusion sentences are markedly different from other forms of written discourse. Furthermore, in Chapter 10 it was demonstrated that the task context in which semantic illusion sentences are presented has a significant effect upon semantic illusion rate. Building on these findings, this chapter is concerned with how semantic illusion sentences are handled within the cognitive system and examines the interaction between semantic illusions and working memory.

11.1 Working memory

The term ‘working memory’ refers to that part of our cognitive system which is in charge of ongoing processing. It is also a temporary storage place in which information can be held long enough for a cognitive process to be carried out. Working memory – as originally proposed by Baddeley and Hitch (1974) – consists of a central executive, which is responsible for reasoning, decision making, and for the co-ordination of the operation of the subsidiary ‘slave’ systems, which form the other major components of working memory. The central executive has a limited capacity of three to four ‘chunks’ of data which can be dealt with at any one time (e.g. Baddeley, 1986). The ‘slave’ systems are essentially no more than specialised rehearsal loops in which a limited amount of information can be held until the central executive is ready to deal with it. Baddeley and Hitch suggested two such systems. One is the visuo-spatial sketchpad, which maintains and rehearses visual and spatial inputs – as though sketched down rapidly on a pad of paper (e.g. Logie, 1995). The other is the articulatory loop (now often called the phonological loop) which maintains and rehearses verbal inputs by essentially repeating them over and over until the central executive is free to deal with the input (e.g. Baddeley, 1986).

Since semantic illusions are primarily a verbal phenomenon, even when presented in written format, the phonological loop will be described in greater detail. Since past research (e.g. Gilhooly, Logie, Wetherick and Wynn, 1993) provided only limited evidence for the use of the visuo-spatial sketchpad and no evidence for the use of visualisation in problem solving, the visuo-spatial sketchpad will be ignored.

A number of phenomena, such as the phonological similarity effect which can be observed for both read and heard material (e.g. Baddeley, 1966; Conrad, 1964) provides evidence for the involvement of the phonological loop in reading comprehension.

People generally find it harder to recall words or letters (especially in the order in which they were presented) when these words or letters sound alike, whether such words or letters are presented aurally or visually (as written material). This implies that when participants are asked to remember a verbal sequence they appear to translate visually presented material into a phonologically based code for temporary storage.

The observation that semantic illusion sentences are often misremembered as having been presented in a correct version (e.g. Erickson and Mattson, 1981) could be seen as evidence for one of the major changes that the working memory model has undergone, since its original description by Baddeley and Hitch. Current working memory theory (e.g. Logie, 1995) has one key difference to Baddeley and Hitch's theory: in Logie's model, working memory is not seen as a gateway between perceptual inputs and long-term processing and memory. Instead, working memory is seen as an active structure in which all components can interact with prior knowledge as stored in a knowledge base, and ongoing perceptual inputs are constantly being interpreted via the knowledge base. This point of view is very similar to the view that knowledge is stored in schemata in long-term memory (Bartlett, 1995[1932]; see section 8.2).

11.1.1 Working memory and prior knowledge

The relationship between working memory and long-term knowledge appears to be interactive: material in temporary memory is influenced by information from long-term knowledge as well as by the operation of the components of working memory. The retrieval of information relevant to a current task appears to be a major role for working memory. The central executive activates and integrates representations in long-term memory to facilitate and make sense of the task at hand. There is empirical evidence to suggest that the phonological loop's performance of short-term memory verbal tasks is in fact subject to long-term knowledge influences. For example, the memory span (usually 7±2 items or ‘chunks’; Miller, 1956) of the phonological loop is greater for words in sentences (15-16 words) than for unrelated words (5-6 words) (Baddeley, Vallar & Wilson, 1987). Similarly non-words that approximate English result in better immediate recall than those that do not (e.g. Gathercole and Baddeley, 1989), and familiarising participants with non-words increases immediate serial recall capacity (Hulme, Maughan & Brown, 1991). These findings suggest that previous information from long-term memory allows participants to chunk short-term inputs more efficiently.

A similar situation might arise when semantic illusion sentences are processed. As semantic illusion sentences – by their very definition – deal with old knowledge that has already been stored in long-term memory at a previous point in time, it seems reasonable to assume that this information from long-term memory is activated via the central executive as the sentence is read and is used to deal more efficiently with the contents of the sentence at hand. It is possible that the propositions within a semantic illusion sentence are dealt with as chunks to increase the speed and efficiency of processing, and the chunking process obscures the substitution within the semantic illusion sentence, as the part of the sentence containing it is contracted into a meaningful unit.

11.1.2 Working memory and language comprehension

The idea of working memory involvement in language comprehension can help make sense of familiar experiences as well as laboratory phenomena. For example, most people have experienced the sudden need for reinterpretation of so-called ‘garden path’ sentences such as “I saw that gasoline can explode. And a brand new gasoline can it was too”. Usually “can” would first be interpreted as a verb, but with the added information from the second sentence, it is reinterpreted as a noun. So maybe working memory provides a buffer in which the exact wording of connected discourse can be held briefly, so that one can achieve the kind of rapid re-interpretation required in understanding garden path sentences. However, there is a need for chunking information even in short sentences in order to process these, and this can lead to disinformation as the example above demonstrates.

Two models of language understanding in particular have attempted to relate working memory to language comprehension. Clark and Clark's (1977) 4-step model of language comprehension implies an important role for a hypothetical structure that can easily be identified with the phonological loop. In the first step of Clark and Clark's model, the comprehender constructs a phonological representation of the message in working memory. This is then used to identify the content and function of the separate components of the message. From this information, a hierarchical structure of the sentence is built using the underlying propositions within the message. Then the working memory representation of the original input is wiped and the comprehender retains the meaning of the message, but not its exact wording (e.g. Sachs, 1967). Even though Clark and Clark do not identify the working memory component carrying out the buffering function described in their model, it corresponds directly to the phonological loop, which is specialised for temporary storage of the phonological form of the language material being processed.

But not all understanding is based on the analysis of working memory representations in this 'off-line' fashion: there is empirical evidence that lexical semantic information can be accessed very rapidly within a few hundred milliseconds. For example, spoken words are on average recognised within 200 ms of the onset of the acoustic signal and both sensory and contextual cues influence the word recognition process (e.g. Marslen-Wilson, 1987; Marslen-Wilson and Tyler, 1980). The very speed of the process implies that at least at the single word level of comprehension, lexical semantic information can be accessed without the need of a working memory representation. But in the case of semantically or syntactically complex sentences the message interpretation may well proceed off-line with recourse to working memory representations.

In a second model of text comprehension (Kintsch and Van Dijk, 1978; Van Dijk and Kintsch, 1983) another aspect of working memory processing is stressed which corresponds much more closely to the central executive. Van Dijk and Kintsch's model suggests that the message is processed in cycles, each cycle representing a chunk containing several propositions. In this model, the theoretical component analogous to the central executive holds as many of the propositions as it has 'space' for in its limited capacity, and the propositions form the basis of coherence processing. Text will be accepted as coherent if a certain amount of overlap is found between propositions within any given chunk and an earlier chunk stored in the working memory buffer. If there is no such overlap, inferences have to be made by reference to long-term memory when necessary. The need for inferences places a much heavier load upon general processing resources. Thus, when dealing with particularly complex messages, some of the limited capacity available for processing may be allocated to aspects of processing other than the short-term storage of propositions. According to the model, therefore, working memory requirements for sentence processing are likely to vary as a function of the syntactic and semantic complexity of the sentences being processed.

11.1.2.1 The phonological loop and language comprehension

There are two main theoretical approaches to the relationship between the phonological loop and language comprehension. Both these theories share the assumption that comprehension of simple syntactic/semantic clauses or sentences occurs 'on-line', without reference to a phonological representation of the message held in working memory. This assumption is of interest with regard to the question of working memory involvement in semantic illusion occurrences. As stated in Chapter 7, participants appear to regard semantic illusion sentences as perfectly normal and of relatively low complexity. If this is indeed the case in terms of the processing demands, there should be no effect of articulatory suppression upon the semantic illusion rate compared to a control condition. If there is an effect of articulatory suppression upon semantic illusion rate, however, this would imply that semantic illusion sentences are more complex in processing terms than they seem and thus require the construction of a working memory representation, which according to Clark and Clark (1977) will lead to the purging of the actual wording of the sentence from working memory as the sentence is being processed.

The first theory about the involvement of the phonological loop in sentence comprehension is that phonological working memory is used as a buffer store to be consulted during off-line linguistic analysis (e.g. Saffran and Marin, 1975; Baddeley, Vallar and Wilson, 1987; Martin, 1987). The kinds of sentences that require back-up from such a buffer are usually highly complex and may not have a single pragmatic interpretation based upon the words present; the comprehender has to carry out fairly detailed syntactic and semantic analysis off-line (e.g. Caramazza, Basili, Koller and Berndt, 1981; Saffran and Marin, 1975). In fact, it can be shown that only long and complex sentence structures are affected by the disruptive effect of articulatory suppression (e.g. Baddeley, Eldridge and Lewis, 1981) and that neither speed nor accuracy are affected in a simple verification task of active sentences such as "Canaries have wings" vs. "Canaries have gills" (Baddeley, 1978). A similar requirement for a phonological representation in order for successful sentence comprehension to take place exists also in the case of long sentences, in which word order is critical (e.g. in an anomalous sentence like "One could reasonably claim that sailors are often lived on by ships of various kinds"). It has been argued that for such sentences, a phonological representation of the full word sequence is required for the interpretation of the full sentence form, and thus comprehension of these forms places higher demands upon the limited capacity of working memory (Baddeley, Vallar and Wilson, 1987; Martin, 1990).

Another suggestions has been put forward by Caplan, Waters and colleagues (Caplan, Baker and Dehaut, 1985; Waters, Caplan and Hildebrandt, 1987): the working memory representations held in the phonological loop is used as a post-syntactic checking mechanism for syntactically complex sentences in which there is no straightforward assignment of lexical items to the proposed syntactic structure. For example, Waters, Caplan and Hildebrandt (1987) demonstrated that participants' performance in a semantic acceptability judgement task declines with articulatory suppression, implying phonological loop involvement in the task. But articulatory suppression does not interfere more with syntactically complex sentences than simple ones, while it does interfere more with two-proposition than one-proposition sentences. Waters, Caplan and Hildebrandt interpreted this as evidence that the phonological loop is not involved in the syntactic analysis of a sentence, but in the post-syntactic interpretative processes involved in the acceptability judgement task. In other words, the phonological loop plays a role in interpreting the semantic content extracted from a complex sentence's syntactic structure.

Articulatory suppression thus appears to have an effect upon the reading of sentences for meaning, but only for sentences which are relatively complex and/or lengthy. Based on this finding it seems likely that an articulatory suppression task would interfere with the sentence-verification task often used in semantic illusion research, and thus that articulatory suppression would have an effect upon semantic illusion rate.

11.1.2.2 The central executive and language comprehension

There is considerable evidence linking the general processing resources of the central executive to language processing for meaning. Daneman and Carpenter (1980) used the term working memory to refer to general-purpose cognitive resources upon which the process of understanding connected discourse (written or spoken) must draw to meet processing demands. These resources correspond closely to the central executive of working memory (Baddeley and Hitch, 1974; Logie, 1995). This theory draws upon the notion that language comprehension involves both a processing and a storage aspect. Processing is required for word recognition from surface representations, to access syntactic and semantic information, and to interpret the meaning of sentences. The storage requirement arises from the need to store intermediate representations of linguistic material to provide input for more advanced processing activities. Thus, if there is a common limited pool of cognitive resources, there will be a trade-off between processing and storage whenever a processing task exceeds the limited resources available (cf. attention theory, for example, Broadbent, 1958; Treisman, 1964; Deutsch and Deutsch, 1963).

Daneman and Carpenter (1980) developed the reading span test, which is highly correlated with other language skills, as a measure of an individual's working memory capacity, because individual differences in the capacity of working memory influence the point at which the hypothesised trade-offs between processing and storage occur for any particular person. In the reading span test, participants are required to read a number of sentences for comprehension, while holding the last word of each of the sentences read in memory for later sequential recall. The reading span is defined as the largest number of sentences read for which a participant can perform this task correctly. Daneman and Carpenter suggested that participants' capacities for processing and storing linguistic information (as determined by the reading span) would directly determine the accuracy and efficiency of the participants' ability to process language for meaning.

Once individual working memory capacity is determined using the reading span test, certain theories implicating central executive involvement in the processing of certain comprehension tasks can be tested empirically. For example, Daneman and Carpenter (1983) hypothesised that working memory capacity would have a drastic effect upon the ease with which readers can integrate new material with preceding text. In a text like "…he went and looked among his baseball equipment. He found a bat that was very large and brown and was flying back and forth in the gloomy room…" the initial context would usually lead a reader to interpret "bat" as the object used for baseball, but the following information about the bat flying about forces the reader to re-interpret the word as referring to the animal. Daneman and Carpenter assumed that, in order to achieve this re-interpretation, a reader must recover active representations of the original surface form of the word "bat" (thus drawing on storage requirements). As predicted, the results showed that readers with small reading spans were especially poor at recovering from such textual ambiguities when there was a sentence boundary between the ambiguous word and its disambiguating context. Readers with larger reading spans performed much better in this task.

11.1.3 Working memory and error detection in texts

It has been shown that semantic errors are harder to detect in written texts, than orthographic or typographical errors (e.g. Sommers, 1980). This might be due to the fact that the detection of semantic errors requires text comprehension and thus generates a higher workload in working memory (Hacker, Plumb, Butterfield, Quathamer and Heineken, 1994). If the working memory workload is too great, semantic errors are no longer detected, even if participants are actively looking for semantic errors and know how to correct them when processing demands are lower (e.g. Markman, 1985).

Several factors affect the cognitive demands placed upon working memory by the specific requirements of an error detection task. For example, semantic error detection improves if participants are specifically asked to read a text for comprehension (Beal, Bonitatibus, Garrod, 1990; cf. van Jaarsveld, Dijkstra and Hermans, 1997 on semantic illusions and task demands). Similarly drawing attention to specific portions of text by using paratextual cues, such as bold print, increases the detection of all error types (e.g. Hacker, Plumb, Butterfield, Quathamer and Heineken, 1994). Working memory load also decreases with repetition and with successive re-readings, while the detection of syntactic and semantic errors improves (Levy, Newell, Snyder and Timmins, 1986). One can assume that the load upon the central executive to be greater when processing for meaningful errors (such as semantic, syntactic or spelling errors) than when processing for typographical errors. It also appears likely that the greater the amount of text that has to be processed for an error to be identified, the greater the load upon the central executive. Similarly the phonological loop would also appear to be more involved in the detection of errors requiring processing above the word level, such as semantic errors or errors spanning a greater amount of text.

Larigauderie, Goanac'h and Lacroix (1998) tested these assumptions in a series of experiments using articulatory suppression and random generation as secondary tasks to observe the effects of these upon error detection. With articulatory suppression as a secondary task, participants' overall processing speed did not increase, but performance on semantic or syntactic error detection above the word level (i.e. relating to individual words at a time) declined significantly. On the word level, error detection was as good as in the control group.

When the secondary task was random generation, the overall time taken to complete the detection task increased significantly. Despite the increase in processing time there was still a significant decline in the detection rate for semantic or syntactic and also for orthographic errors, while typographical errors and errors involving only word-level processing remained constant. The random generation task always led to a greater decline in performance than articulatory suppression, and the decline in error detection was always particularly strong where integrative processing was required. Larigauderie, Goanac'h and Lacroix concluded that these results showed that the demands on the central executive were greater when the detection of errors required any meaningful processing and/or memory search processes, such as those required by an orthographic check.

11.1.4 Semantic illusions and working memory

Empirical evidence from working memory studies suggests that both the phonological loop and the central executive are involved in language comprehension. The phonological loop plays an important role in the processing of long and complex sentences, while the central executive is involved in processing information for lexical and semantic content.

Semantic illusion sentences are generally fairly long and complex with a certain amount of redundancy included (e.g. the statement "Two animals of each kind went on the Ark" could be verified and understood easily without reference to Noah at all), although participants subjectively do not perceive the type of sentence that can be turned into an semantic illusion sentence as being particularly hard to interpret (see Chapter 7). This may be due to the fact that the semantic illusion sentences generally deal with familiar material, whereas most forms of natural written discourse serve the purpose of imparting new information.

But how complex are semantic illusion sentences really?

Since there is evidence that complex sentences require the involvement of the phonological loop component of working memory in order to be interpreted for meaning, but syntactically simple sentences do not, an experiment was designed to examine the processing demands that semantic illusion sentences make upon our cognitive resources. If phonological loop involvement in the processing of semantic illusion sentences can be demonstrated, a case can be made for the idea that semantic illusion sentences require high working memory effort to comprehend. As a result of this it could be argued that the target word might fall victim to a trade-off between storage and processing, as hypothesised by Daneman and Carpenter (1980).

The following studies were designed to attempt to find a dissociation between the involvement of the phonological loop and that of the central executive in the processing of semantic illusion sentences. Using the well-documented method of articulatory suppression to test for the involvement of the phonological loop, it was assumed that there would be an increase in semantic illusion rate in an articulatory suppression condition, due to the higher processing demands made by semantic illusion sentences. A further increase in semantic illusion rate was anticipated when participants were asked to perform a secondary task involving the central executive. In the first experiment the task was concurrent mental addition. In the second it was a random generation task. It was thought that even without the added demands of a secondary task, semantic illusion sentences would need a great proportion of all available resources to be comprehended and for the substituted word to be identified successfully.

11.2 Experiment 6a

11.2.1 Method

Participants: 54 Oxford Brookes psychology students took part in this study as a part of a practical class for a course in Human Information Processing.

Stimuli: 24 target sentences were constructed and mixed with 60 fillers. The target sentences took the form of true/false statements of roughly equal length and form as the filler sentences (a list of target sentences used is presented in Appendix 13).

Procedure: 3 experimental groups were tested: the *control group* was presented with each statement individually. A sentence appeared in the middle of a computer screen in 14 point Times New Roman font and was displayed until the participant made a response, by hitting either one of two response keys. (*T* and *F* on a standard keyboard were chosen for "true" and "false" respectively, despite their relatively awkward positioning, as they were considered to be intuitive for the relevant responses. Participants also could respond by hitting the *x*-key if they felt that they really did not know the response to any of the statement-to-be-verified.) Response times to first keystroke and the response made were recorded. After each response the screen went blank for a 100 millisecond inter-stimulus interval. The sentences were presented in 5 blocks of 16. At the end of each block the screen displayed the message "Hit the spacebar when you are ready to continue." Before recording the data, participants were given a 'tryout block' to familiarise them with the program (written in Superlab 2.0) worked. Then the five experimental blocks were run continuously and all data recorded.

The *articulatory suppression group* was required to follow the same basic procedure as the control group, but simultaneously they were asked to say "the-the-the-" out loud throughout the experiment. The experimenter reminded participants as necessary to maintain this task throughout the course of the experiment. Again participants were given the opportunity to ‘try out’ a few sentences to get a feel for the procedure.

The *mental addition group* had to carry out a secondary task of mental addition of double-digit numbers while verifying the statements. A sum of two double-digits was displayed above the sentences in larger and different coloured font. Participants were instructed to verify the statements by hitting the *T* or *F* key as appropriate before entering the result of the addition task using the number pad on the keyboard. Responses and response times for both tasks were recorded. Before participants did the dual-task experiment, baseline performance rates for the mental arithmetic and the statement-verification task were established separately. Then participants were given a chance to practice doing both tasks simultaneously. After participants felt reasonably confident about performing the two tasks concurrently, they performed the double task for 3 blocks of 16 trials.

Knowledge check: After the experiment was completed participants were asked to complete a knowledge check consisting of 24 multiple choice questions based upon the target sentences.

11.2.2 Results

Data from 48 participants was used in the analysis. Data from 5 participants, one from the control group and four from the articulatory suppression group, had to be excluded from the analysis, because no knowledge check data was available for these participants. Data from a further participant from the mental addition group was excluded as this participant had consistently given the response to the mental arithmetic task first, so that there was no record of the statement verification response. For the remaining 48 participants only the data for the target statements was analysed.

11.2.2.1 Semantic illusion rate, correct response rate and "don't know" rate

A record was made of all target statements answered in each group, including correct, semantic illusion and "don't know" responses. A semantic illusion was considered to have occurred when participants had judged the respective statement to be "true" *and* had given the correct response to the corresponding knowledge check question. A response was classed as correct if the participant responded with "false" during the verification task *and* gave the correct answer in the knowledge check. All other responses were classified as "don't know" responses. The mean semantic illusion rate was calculated for each group, as a percentage of all responses to statements which could have led to a semantic illusion (see table 11.1).

Table 11.1: Mean rates (percent) of correct, semantic illusion and "don't know" responses for each group.

Group	Number of participants	SI rate	Correct rate	Don't know rate
Control	17	39.3	50.7	9.9
Articulatory suppression	15	41.3	38.8	20.0
Mental addition	16	36.5	42.2	21.4

A one-way analysis of variance was carried out on the semantic illusion rate data using participants as subjects. The results of the analysis of variance were found to be insignificant ($F= 0.43$; $d.f.= 2, 45$; $p= 0.653$) indicating that there was no effect of the type of secondary task upon the semantic illusion rate.

Similar analyses were also carried out for the correct response rate and the "don't know" response rate, to investigate if secondary task type affected performance. For the correct response data, the one-way analysis of variance was found to be insignificant ($F= 2.134$; $d.f.= 2, 45$; $p= 0.130$), and no effect of secondary task type upon the rate of correct responses was observed. Neither was secondary task type found to affect the rate of "don't know" responses ($F= 2.613$; $d.f.= 2, 45$; $p= 0.084$).

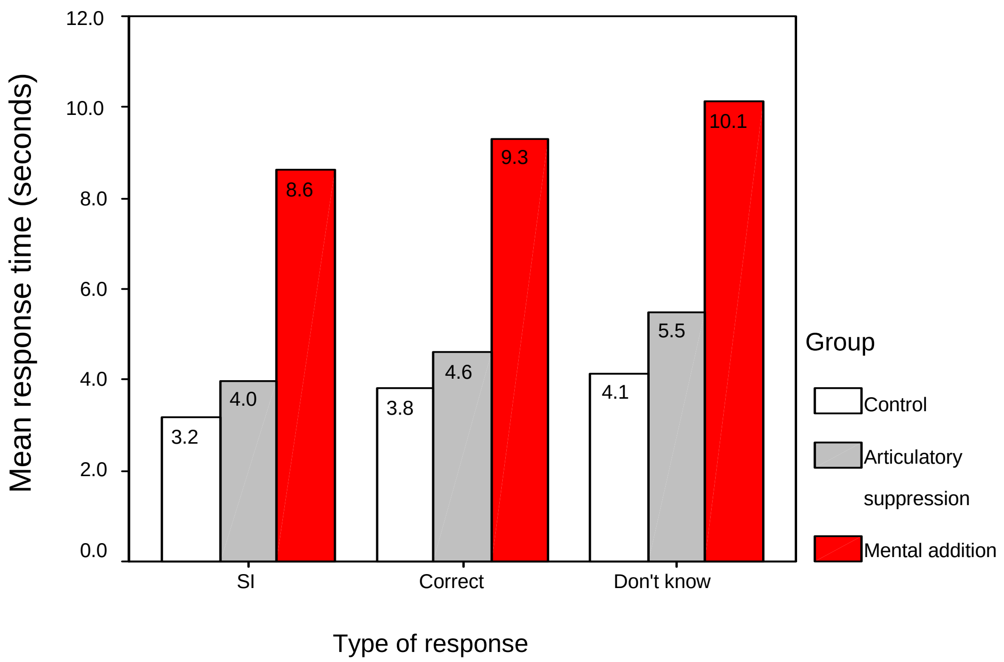
11.2.2.2 Response times

The mean response times were recorded for each group and each type of response (semantic illusion, correct and "don't know"; see table 11.2 and Figure 11.1).

Table 11.2: Mean response times (seconds) for semantic illusion responses compared with correct and "don't know" responses in each experimental group.

Group	SIs	Corrects	Don't Knows
Control	3.2	3.8	4.1
Articulatory suppression	4.0	4.6	5.5
Mental addition	8.6	9.3	10.1

Figure 11.1: Mean response times (seconds) for semantic illusion responses compared with correct and "don't know" responses for each experimental group.



A repeated measures analysis of variance was carried out upon the response time data, using participants as subjects. Mauchly's test of sphericity was found to be significant for type of response (W= 0.738; d.f. = 2; $p<0.001$) and Pillai's Trace was used to determine the significance of this effect. Response times were significantly different for type of response ($F= 9.137$; d.f.= 2, 44; $p<0.001$) with semantic illusion responses being made most rapidly for each group, followed by correct responses and then "don't know" responses for all three secondary task groups. Response times were also significantly different for each group ($F= 706.724$; d.f.= 2, 45; $p<0.001$) with participants in the control group responding most rapidly for all three types of responses, followed by the articulatory suppression group. Response times were longest for the mental addition group for each response type. There was no significant interaction between group and response type ($F= 0.167$; d.f.= 4, 90; $p= 0.955$).

To investigate the observed effects further, a series of post-hoc tests was carried out. For the between-participants factor of secondary task type, a Student-Newman-Keuls test was carried out. It was found that the response times for the mental addition group were significantly longer than those for either of the other two groups. There was no significant different between response times in the control group and the articulatory suppression group (see Table 11.3).

Table 11.3: Student-Newman-Keuls post-hoc test showing homogeneous subsets for the effect of secondary task type upon response times (p= 0.05).

Secondary task type	**N**	**Subset**	
		1	**2**
Control	17	3.7	
Articulatory Suppression	15	4.7	
Mental Addition	16		9.5
Significance		0.076	1.0

The within-participants factor of response type was further investigated by using a series of paired samples t-tests with a Bonferroni correction ($p= 0.0167$ for a significance level of 0.05). The results indicated that semantic illusion responses were made significantly more rapidly than correct responses ($t= -3.226$; d.f.= 47; $p= 0.002$) or “don’t know” responses ($t= -3.956$; d.f.= 47; $p<0.001$). However – allowing for the Bonferroni correction – correct responses were not made significantly more rapidly than “don’t know” responses ($t= -2.072$; d.f.= 47; $p= 0.044$), although there is a definite trend in that direction.

For the mental addition group, individual response times for the mental arithmetic only and the sentences only tasks were calculated to provide a baseline against which to compare the performance for the joint task. Since the processing times for the mental addition group are extremely long compared to the other two experimental groups, mean response times for the combined task were worked out and then compared to the sum of the two baseline response times for each participant. The paired samples proved to be highly correlated (correlation coefficient = 0.714; d.f.= 14; p= 0.002) and a paired samples t-test was carried out upon the mean response times. The overall time taken for participants to perform the combined task was found to be significantly shorter than the time taken to perform one task after the other (t= -2.148; d.f.= 15; p= 0.048, two-tailed); see also table 11.4 for means and standard deviations).

Table 11.4: Mean time taken (seconds) for the performance of both maths and sentences simultaneously and for maths and sentences consecutively in the mental addition group.

	N	Mean	Std. Dev.
Simultaneously	16	9.5	2.0
Consecutively	16	10.4	2.3

11.2.3 Discussion

In contrast to the hypothesis that secondary tasks requiring working memory involvement would affect the rate at which semantic illusion responses would occur, the results of this study found that neither the semantic illusion rate, the correct response rate nor the “don’t know” response rate varied significantly between the experimental groups. It was thought that this finding could be an unwanted side effect of the way in which the experiment was conducted. There was, for example, no way of monitoring whether participants in the articulatory suppression group maintained a constant stream of verbal output and due to restrictions of the Superlab Program used the accuracy of the mental arithmetic task could not be checked.

The response time data, however, did indicate that there is an effect of the secondary task upon processing. The control group verified the sentences most rapidly, while the articulatory suppression group's response times were slightly (but not significantly) slower than those of the control group. The mental addition group took significantly longer to respond to the sentence verification task than either of the other two groups. This finding suggests that the unexpected absence of a difference in semantic illusion response rates for the three experimental groups could be explained by a trade-off between accuracy and processing time: in order to achieve the same level of accuracy in the sentence verification task, participants had to spend more time on each sentence to compensate for the secondary task requirements upon working memory. It seems possible that in the mental addition group, participants performed the two tasks consecutively rather than simultaneously as the instructions required, because the response times were so much greater than those for the other two groups. However, compared to the performance when the mental addition and the sentence verification tasks were performed separately, the joint maths-and-sentences task appears to have been performed more rapidly than if the tasks had been done one after the other. Nevertheless, it is still highly likely that the tasks were performed with participants effectively switching between maths and sentence verification.

Response times also varied with the type of response given to each sentence. Semantic illusion responses were made most rapidly within each group. Correct responses were made slightly more slowly, suggesting that unmasking the substitution required more processing time than simply dealing with its gist (c.f. Reder and Kusbit, 1991 for similar response time findings; and van Jaarsveld, Dijkstra and Hermans, 1997 for a discussion of the trade-offs between speed and accuracy). "Don't know" responses took the longest time to be made, as would be expected based upon theories of memory search when no match can be found (cf. serial exhaustive scanning model, Sternberg, 1966). However, this difference did not reach significance in a post-hoc check, and only cautious conclusion can be drawn. The finding that semantic illusion responses were made more rapidly than other incorrect responses (i.e. "don't know" responses) is in keeping with the idea that semantic illusion responses are a type of processing error and not related to retrieval failures (cf. Reder and Kusbit, 1991).

Overall the results from this experiment clearly showed that the load upon working memory affects the speed at which sentences are processed for verification. It seems extremely unlikely that the semantic illusion rate is truly independent of working memory load, as suggested by the results of this study, rather it is likely that inadequate control over performance of the secondary tasks was the explanation of the lack of effect upon the semantic illusion rate. As a consequence it was decided to repeat this study with more carefully monitored secondary tasks.

11.3 Experiment 6b

11.3.1 Method

Participants: 48 participants from a number of different educational backgrounds were invited to take part in this study. As an incentive a £10 cash prize was raffled off amongst the participants.

Materials: 96 general knowledge statements were created using material from general knowledge and trivia quizzes found on the world-wide web, as well as previous statements used by other semantic illusion researchers in the past. 24 statements were target statements, designed to lead to potential semantic illusions (see Appendix 13). Three target statements were used in the trial phase of the experiment. One additional statement proved to be considered "false" by almost all participants for reasons other than the substituted target word. This sentence was later excluded from the analysis as were the three sentences used during the trial phase. 72 filler statements were constructed along the same format as the target sentences, and varied in topic as well as in overall length.

The statements were presented individually one after the other in the centre of a computer screen using Cedrus Superlab Pro Version 2.0 for Windows. The order of presentation of the sentences was arranged so that no two target sentences ever followed on from each other directly, and so that they were relatively evenly spread through the experimental blocks. Each block consisted of 16 or 17 statements presented consecutively with a 100 millisecond inter-stimulus interval after each response made by a participant during which the computer screen was left blank. There were five experimental blocks and one initial trial block of statements-to-be-verified.

Design: There were three experimental groups in this experiment, and participants were randomly assigned to these groups, so that at the end each group included data from 16 participants. The three groups were referred to as *Control*, *Articulatory Suppression* and *Random Generation*.

The control group: In this group the participants were required to verify the statements presented on the computer screen by hitting either a “true” or a “false” key on a six button response box (RB-600), on which two adjacent keys had been labelled as “true” (key 3) and “false” (key 4) respectively. The response box was arranged in such a way that participants could place it either to the left or the right of the computer screen depending upon their hand-preference. Participants were required to read each sentence as it came up on the screen and to decide as quickly as possible if they thought it was true or false. If they did not know an answer, they were encouraged to make a guess.

When a choice had been made by key press, the screen was cleared for 100 milliseconds until the next sentence appeared. In the background a cassette-tape player played back the sound of a metronome clicking once per second. Participants were instructed to ignore the sound. Participants were given the trial block of 16 statements to try out the task, at the end of which there was an opportunity for a brief break and for any questions that participants might have about the procedure. Once participants felt ready to tackle the experiment proper, the five experimental blocks were run with the opportunity for short pauses between blocks at participants' discretion. After all statements had been verified, each participant completed a knowledge check consisting of 24 multiple choice questions testing for correct knowledge of the facts pertaining to the target statements used in the experiment. Data from 16 participants was gathered and used for this group.

The articulatory suppression group: In this group the participants were required to complete a secondary task of articulatory suppression (e.g. Levy, 1971, 1975; Murray, 1967): in time with the tape-recorded metronome clicks, participants were asked to repeat the digits one to four (i.e. "1, 2, 3, 4, 1, 2, 3, 4, 1, 2 …" and so on). First a baseline articulation rate was measured for each participant, as they performed the counting task on its own for two minutes. The experimenter measured the time for the baseline articulation, while a voice-key attached to the secondary computer recorded the timing of each participant's articulation once per second. Inter-response intervals were calculated for each participant.

Participants were then asked to perform the statement-verification task in the same fashion as the control group while at the same time continuing with the articulatory suppression task. The joint task of statement-verification and articulatory suppression began with a trial block as for the control group. Then five experimental blocks were run, with potential for breaks between the blocks. After the end of the verification task, the participants were asked to complete the knowledge check for the target statements as before. Again, inter-response intervals were calculated for the articulatory suppression secondary task. Data from 16 participants was gathered and analysed for this group.

The random generation group: The last experimental group was the random generation group. The secondary task for this group was random number generation (e.g. Baddeley, Lewis and Vallar, 1984; Farmer, Berman and Fletcher, 1986; Logie, Baddeley, Mane, Donchin and Sheptak, 1989; Saariluoma, 1991; Gilhooly, Logie, Wetherick and Wynn, 1993; Larigauderie, Goanac'h and Lacroix, 1998) in time with the pre-recorded metronome click. Participants were asked to randomly generate numbers between zero and nine (e.g. " 4, 2, 3, 3, 7, 2, 0, 1…") A performance baseline for the random generation task was established, by each participant performing the random generation task on its own for about two minutes. The experimenter recorded in writing each number generated, while the voice-key recorded the timing of the responses. Inter-response intervals and randomness scores were calculated for each participant.

Then participants were required to perform the random generation task simultaneously with the statement-verification task. Trial phase and the experimental blocks were run as for the other two experimental groups, as was the knowledge check after the end of the verification task. Data from 16 participants was gathered for this group, but data from 1 participant was excluded from the analysis as they had completely abandoned any attempt at random number generation, repeating the same digit over and over while they verified a statement, and only then changing to a new digit which was then repeated until the next statement had been verified.

Apparatus: The statements-to-be-verified were displayed using a PC running Superlab Pro 2.0 for Windows. A 6-button response box (RB-600) from Cedrus was attached to accept input for the verification task. The middle two buttons (3 and 4) were labelled as "true" and "false", while the other buttons (1, 2 and 5, 6) were covered over with cardboard sleeves to avoid confusion.

On a second computer (Apple Macintosh Classic II) data from the two number generation tasks was recorded using a small hand-held microphone, which was set up as a voice-key using Superlab 1.34 for Macintosh. The voice-key was set up to take a reading once every second, so that inter-response intervals could be calculated. To facilitate the generation of numbers at a rate of one per second, a metronome click was recorded onto a cassette-tape and played back using a basic tape deck. Participants were encouraged to generate in time with the clicks, but told not to worry unduly if they felt they were out of synch with it, as long as the speed of their number generation remained as constant as possible.

11.3.2 Results

11.3.2.1 Semantic illusion rate and correct response rate

The rate of semantic illusion occurrence was calculated for each target sentence for each group. 20 sentences were analysed. The three sentences that had been used during the trial phase of the experiment were excluded from the analysis, as was one sentence used in the main part of the experiment, as a large proportion of the participants pointed out that they had not recognised the semantic illusion in question but had thought the sentence wrong for other reasons ("spider"; see Appendix 13). As previously a semantic illusion was considered to have occurred when a participant had responded to the relevant statement with “true” *and* had given the correct response in the knowledge check (see table 11.5 for semantic illusion rates).

Table 11.5: Mean semantic illusion rate (percent) and standard deviations in each experimental group.

Group	N	Mean SI rate	Std. Dev.
Control	20	24.1	17.0
Articulatory suppression	20	31.6	15.0
Random generation	20	35.0	15.4

A repeated measures analysis of variance was carried out on the remaining data using sentences as subjects. The sphericity assumption was not violated for any of the effects measured. There was a significant effect of secondary task type upon mean semantic illusion rate (F= 5.293; d.f.= 2, 38; p= 0.009). In order to determine if all three groups differed significantly from each other, a post-hoc Bonferroni t-test was carried out with the significance criterion adjusted to p<0.0167. Only the difference between the control group and the random generation group proved to be significant (t= -2.970; d.f.= 19; p= 0.004, one-tailed) in the post-hoc analysis, with the semantic illusion rate being greater in the random generation group than in the control group. The articulatory suppression group was not significantly different from either the control group (t= -1.914; d.f.= 19; p= 0.036, one-tailed) or from the random generation group (t= -1.342; d.f.= 19; p= 0.098, one-tailed).

Correct response rates were similarly analysed for each experimental group, using sentences as subjects. A correct response was considered to have occurred when a participant had responded correctly to a statement *and* the corresponding knowledge check question (see table 11.6 for mean correct response rates).

Table 11.6: Mean correct response rate (percent) and standard deviations in each experimental group.

Group	N	Mean correct rate	Std. Dev.
Control	20	64.1	19.4
Articulatory suppression	20	57.5	17.6
Random generation	20	52.7	16.6

A repeated measures analysis of variance was carried out on the correct response data. The sphericity assumption was not violated. As with the analysis of the semantic illusion rate data, there was a significant effect of secondary task type upon mean correct response rate (F= 4.980; d.f.= 2, 38; p= 0.012). The effect was further investigated in a series of post-hoc t-tests using a Bonferroni correction (significance criterion adjusted to p<0.0167). The only difference to prove significant in the post-hoc analysis was that between the correct rates in the control condition and the random generation condition (t= 2.713; d.f.= 19; p= 0.007, one-tailed), with the correct response rate being significantly greater in the control group than in the random generation group. As with the semantic illusion rate analysis, the articulatory suppression group was not significantly different from either the control group (t= 1.926; d.f.= 19; p= 0.035, one-tailed), or from the random generation group (t= 1.517; d.f.= 19; p= 0.075, one-tailed).

11.3.2.2 Processing times

11.3.2.2.1 Total time taken

The total processing time taken to complete Blocks 1 to 5 (excluding pauses at the 'between-blocks' display on the computer screen) was recorded for each participant, and mean processing times and standard deviations were calculated for each group. (see Table 11.7).

Table 11.7: Mean processing times (seconds) for all 5 blocks in total:

Group	N	Mean time	Std. Dev.	fastest time	slowest time
Control	16	379	147	143	672
Articulatory suppression	16	279	112	127	468
Random generation	15	285	95	169	489

A one-way analysis of variance was carried out. There was a significant effect of secondary task type upon mean processing times (F= 3.454; d.f.= 2, 44; p= 0.040) and a Student-Newman- Keuls post-hoc test was carried out. The results of the post-hoc test suggested that the observed difference in processing times was unlikely to reflect a real difference, suggesting that secondary task type did not have an effect upon the total time taken to complete the task (see table 11.8 for the results of the Student-Newman-Keuls test).

Table 11.8: Student-Newman-Keuls post-hoc test showing homogeneous subsets for the effect of secondary task type upon total processing times (p= 0.05).

Secondary task type	N	Subset
		1
Control	16	279
Articulatory Suppression	15	285
Random generation	16	379
Significance		0.063

11.3.2.2.2 Mean response times for correct responses and semantic illusion responses by sentence

The mean response times for each semantic illusion sentence were recorded for both correct responses and for semantic illusion responses and a repeated measures analysis of variance was completed comparing performance for correct and semantic illusion responses on each of the 3 experimental groups. Mauchly's test of sphericity was used to test the sphericity assumption and found to be significant for the group effect (W= 0.619; d.f.= 2; p= 0.021) and Pillai's Trace was used to determine if the secondary task requirements affected the response times for semantic illusion sentences. There was a significant effect of secondary task upon response times (F= 9.182; d.f.= 2, 16; p=

0.002). Whether the participants gave a correct or a semantic illusion response did not affect the response times (F= 0.004; d.f= 1, 17; p= 0.953), although there was a significant interaction between group and response given (F= 17.341; d.f.= 2, 34 ; p= 0.001).

To investigate the actual times measured further, three a priori comparisons were made to compare the response times for correct and semantic illusion responses within each group (the significance criterion was adjusted to $p<0.0167$; see Table 11.9 for mean response times and standard deviations).

Table 11.9: Mean response times (seconds) and standard deviations for type of response in each group.

Group	Response	N of participants	Mean	Std. Dev.
control	C SI	18 18	4.6 3.6	0.7 1.2
articulatory suppression	C SI	20 20	3.2 3.3	0.5 0.9
random generation	C SI	20 20	3.0 3.9	0.6 1.3

While there was a significant difference between the time taken to give a correct response to an semantic illusion statement and the time taken to make the incorrect semantic illusion response in two of the 3 groups: in the control group (t= 4.031; d.f.= 17; p= 0.001, two-tailed) and in the random generation group (t= -2.758; d.f.= 17; p= 0.013, two-tailed), the direction of that difference is opposite for the two experimental groups. In the control condition, correct responses were given more slowly than semantic illusion responses. Whereas in the random generation group, semantic illusion responses were the ones given more slowly. However, there was no overall significant effect of the type of response given across all 3 groups (t= -0.219; d.f.= 57; p= 0.827, two-tailed).

11.3.2.3 Secondary task analysis

Mean inter-response intervals were calculated for each participant, both for base rate number generation and for number generation as secondary task during the course of the main part of the experiment (the two measurements are referred to as BR and IRI/Snt respectively). For the articulatory suppression group, there was no significant difference between BR and IRI/Snt ($t = 0.350$; d.f.= 15; $p=0.731$, see table 11.10 for means).

Table 11.10: Mean inter-response intervals (milliseconds) and standard deviations for articulatory suppression task on its own (BR) and as secondary task (IRI/Snt).

	N	Mean	Std. Dev.
BR	16	1007	249
IRI/Snt	16	1006	243

For the random generation group, sentence processing did not affect the rate of random number generation ($t = -0.520$; d.f.= 14; $p = 0.611$, see table 11.11 for means). The randomness of numbers generated was also determined using a runs-test with a cut-off point at 4.5 (since participants had been instructed to randomly produce digits between 0 and 9), and the number of runs was recorded for each participant. Randomness was defined as:

$$R = \frac{(r - d/2)^2}{d/2}$$

Where **R** = randomness, **r** = number of runs, and **d** = number of digits generated. With this definition of randomness, an R-value of 0 indicates truly random production of digits; the greater the R-value, the less random the production of digits. When the randomness of numbers generated was compared for single-task random generation and dual-task random generation, it was found that the numbers generated were significantly more random when the random generation task was not combined with the sentence processing task ($t = -2.340$; d.f.= 14; $p = 0.035$, see table 11.11).

Table 11.11: Mean inter-response intervals (milliseconds) and standard deviations for random generation task on its own (BR) and as secondary task (IRI/Snt) including mean measure of randomness of numbers produced.

		IRIs (msec)		Randomness	
	N	Mean	Std. Dev.	Mean	Std. Dev.
BR	15	1060	324	0.828	0.897
IRI/Snt	15	1080	356	4.364	5.840

11.3.3 Discussion

In Experiment 6b, there was a significant effect of secondary task upon the semantic illusion rate, at least between the control group and the random generation group, and the semantic illusion rate for the articulatory suppression group was larger than that for the control group and smaller than that for the random generation group. Though the difference was not found to be significant, the trend was in keeping with the hypothesised effect. The results of the analysis of the correct response rates mirrored these effects. Even so, the differences in semantic illusion rate were smaller than expected, suggesting that the central executive of working memory is involved relatively peripherally in the occurrence of semantic illusions. This could be due to the labour-saving partial matching process that underlies semantic illusions (Reder and Kusbit, 1991) and which seems to be an automatic aspect of human processing. Since much of our previous knowledge of the world is stored in active schemata (Bartlett, 1995[1932]), recognising the contents of a new sentence to verify them requires a matching process of some sort to take place and since linguistic information is not normally presented in exactly the same form twice, the form that a linguistic message such as a sentence takes is not remembered. This can be demonstrated empirically (Sachs, 1967) as can the fact that linguistic information is very rapidly integrated once we start processing it. This was demonstrated by experiments on the abstraction of linguistic ideas (e.g. Bransford and Franks, 1971) in which participants were shown to integrate information rapidly without being able to distinguish what they had seen and what was new but in keeping with the ideas they had previously learned (see Section 8.5.1).

Another factor that could have affected the overall semantic illusion rates in this study is the possibility of a trade-off between speed and accuracy (cf. van Jaarsveld, Dijkstra and Hermans, 1997; see also Section 11.2.2.2). However, the analysis of the processing times was found not to support this hypothesis, and participants took the same overall time to complete the task in all three groups. If anything, participants in the control group appeared to take longest to perform the sentence verification task, with the longest processing times in the control group being noticeably longer than those in either of the two groups which also performed secondary tasks. One possible explanation for this could be that participants in the control group were wrong-footed by the apparent ease of the task, and so spent more time looking for any catches or trick questions. This could be regarded as a processing luxury which the other groups could not afford as they also had a secondary task to perform. Despite this hypothesised search for trick questions, however, the semantic illusion rate for the control group was still well within the usual parameters for semantic illusion experiments, and even apparent caution at responding to the statements-to-be-verified did not help participants in the detection of substitutions (c.f. Erickson and Mattson, 1981; Reder and Kusbit's co-operative hypothesis (1991); van Jaarsveld, Dijkstra and Hermans, 1997).

In an attempt to replicate the response time findings for type of response from Experiment 6a (see section 11.2.2.2), the mean response times for correct and semantic illusion responses were compared. Contrary to expectation, however, there was no significant effect of the type of response upon response time. There was an effect of group and a significant interaction between group and type of response. For the control group, the findings were in keeping with previous studies looking at response times (e.g. Reder and Kusbit, 1991): semantic illusion responses were given significantly more rapidly than correct responses. In the articulatory suppression group the type of response did not affect response times in either direction. But for the random generation group, semantic illusion responses were made significantly more slowly than correct ones. This finding could indicate that when the central executive meets heavy processing demands, the integrative processing that leads to semantic illusion responses takes longer and is more arduous, than when that process is the only task that needs to be performed.

Since secondary task demands did not seem to have much effect upon primary task (sentence verification) performance, performance on secondary tasks was examined. It was found that articulatory suppression did not suffer at all when participants verified sentences simultaneously – the inter-response intervals remained constant. Similarly, inter-response intervals for the random generation task were not affected by the sentence verification task, but the numbers produced were significantly less random when participants verified sentences compared to when they generated the numbers on their own. This provides some evidence that a trade-off takes place between the sentence verification task and random number generation, and most participants in this study chose to concentrate on the sentence task over the number generation task.

11.4 Conclusions

Both the experiments described in this chapter showed evidence of the involvement of working memory in the processing of semantic illusions. While the results were far from identical on the surface, they appeared to support the same basic finding.

In Experiment 6a, the secondary tasks, which participants were required to perform in order to allow for an examination of which components of working memory were involved with semantic illusion processing, did not lead to any change in overall semantic illusion rate and participants in each of the three experimental groups made the same number of semantic illusion responses on average. However, the overall response times increased with secondary task: participants in the articulatory suppression group responded more slowly than those in the control group, while participants in the mental addition group responded much more slowly than participants in either of the other two groups. The results suggested that working memory is required to process semantic illusions, and there is evidence for a trade-off between accuracy and processing times, as had been previously demonstrated by van Jaarsveld, Dijkstra and Hermans (1997), for example. The processing times also varied with the type of response given. Semantic illusion responses were made most rapidly, followed by correct responses, and finally, “don’t know” responses took the longest to be made. It was thought that this last finding might reflect a serial search process of the knowledge relevant to the sentence at hand.

In Experiment 6b, which had been designed as a better controlled version of Experiment 6a, the processing time findings were not replicated: for the control group, the mean response times closely resembled those described by Reder and Kusbit (1991), and semantic illusion responses were made most rapidly, followed by correct responses. However, in the articulatory suppression group there was no difference in response times for semantic illusion responses and correct responses, and in the random generation group semantic illusion responses were the ones that took longest to be made, while correct responses were made more rapidly. The finding, that semantic illusion responses in this group required more time to be processed, might be due to the fact that in the random generation group the cognitive resources, that would normally be assigned to completing the integrative process that allows partial matching requirements to be met by a substituted target word, are not available. This would make the integration of the substituted word a slightly harder task than usual, which requires more time for successful processing.

Even though the response time data were not conclusive, evidence from the semantic illusion rates and from an analysis of secondary task performance again showed that at trade-off between speed and accuracy seems likely. In this experiment speed seems to have been considered more important, and as a consequence the semantic illusion rate in the random generation group was significantly higher than in the articulatory suppression group, which in turn was slightly higher than that in the control group. Data from the secondary task analysis showed that participants did not need to slow the average rate of number generation when responding to the sentences in the primary task compared to number generation on its own. But for the random generation group, the randomness of the numbers generated was significantly better (more random) when participants only generated numbers and did not verify statements at the same time.

The results from both experiments indicate that the central executive is certainly involved in dealing with semantic illusion sentences, as was expected, and that either more semantic illusion responses are likely to occur, or it takes longer to process the relevant statement if the central executive is otherwise engaged. However, the phonological loop does not appear to play an important role in the occurrence of semantic illusions: the semantic illusion rate does not appear to be significantly affected by a secondary task involving the phonological loop such as articulatory suppression. While Experiment 6a seemed to show that statements take longer to be processed when participants are involved in articulatory suppression, the time difference was not great. This suggest that semantic illusion sentences are not only easy to process because they deal with familiar ideas, but also because they were phrased in a straightforward-to-process fashion, which did not require phonological loop involvement to disentangle for meaning. It would appear the participants' subjective opinion of semantic illusion sentences (as shown in Chapter 7) that they are not difficult to process is corroborated by the results of this study, which used a more objective measure of processing difficulty.

Chapter 12:

General Discussion

This chapter provides an overview of all that is known about semantic illusions from previous studies (as reviewed in Chapter 2) and from the studies described in this thesis. These findings are related to the aims of the thesis as presented in Chapter 1 and implications are discussed, and finally directions for future research are proposed, building on the findings of experiments described in previous chapters.

The main aim of the thesis was to explore semantic illusions as a phenomenon, to relate these findings to theories of cognitive processing and to examine what semantic illusions can reveal about ordinary sentence processing. Semantic illusions are not only important because of the implications that their existence has for theories of human information processing, but also because they imply that great care has to be taken to ascertain that routine communications do not contain unexpected details, as it appears very hard to pay attention to anomalous elements of an otherwise coherent context. For example, questions in examinations have to be phrased in accordance with examinees' expectations of the format a problem on a certain topic should take, if the examinees' knowledge is to be assessed fairly. If for example, a question in a multiple choice exam were to be phrased in a similar way to semantic illusion sentences, it would be highly likely to lead to roughly 30 percent of the individuals taking the exam and who possess the relevant knowledge failing to apply their knowledge correctly in response to such a 'trick' question.

Previous research on semantic illusions has been remarkably unsystematic, and there is little coherence in what is revealed about semantic illusions in the literature. Some researchers used names only, others used a mix of names and other words; some reported correct response rates, some reported semantic illusion rates; some used questions, some used statements-to-be-verified. Individual researchers appeared to follow particular interests in a rather sporadic fashion, dealing with one issue or another, but with little systematic follow-through. In this thesis, an attempt was made to organise these findings and to establish a profile for semantic illusions, including – where necessary – further empirical investigation.

12.1 On the surface: A description of semantic illusions

In this section, the profile of semantic illusions that emerges from previous research and from the studies in this thesis is reviewed.

1. Semantic illusions are a sentence processing phenomenon in which an individual mis-processes a substituted word within a sentence context dealing with a familiar topic, as if the substitution was identical in meaning to the word it replaces, without awareness of this on the part of the individual.

2. Semantic illusions are remarkably robust. They can withstand – to some degree – all experimental manipulations that have been carried out upon semantic illusions to date, implying that the mechanism underlying semantic illusions is very fundamental to sentence processing. Semantic illusions occur not only for names, but also for words from different form classes (Chapter 3, Pilot study). They also occur in a number of different languages (see Chapter 6).

3. Semantic illusions are not under the conscious or deliberate control of the reader/listener (Erickson and Mattson, 1981; Reder and Cleeremans, 1990; Reder and Kusbit, 1991), except to a minor degree when given specific instructions stressing the importance of giving accurate responses to the task at hand (van Jaarsveld, Dijkstra and Hermans, 1997). The effect of task demands upon semantic illusion rate is fairly small. Unless explicit warning about semantic illusion sentences is given and participants are specifically instructed to respond as accurately as they can, effects of task demands are not manifest (e.g. Chapter 9, experiment 4a).

4. The position of sentence focus has an effect upon semantic illusion rate, and substitutions that coincide with the sentence focus are more likely to be detected. If a manipulation is performed to ascertain that the point of focus is on the substituted word, semantic illusion rates are generally lower than for matched control sentences. However, many of the manipulations used to draw attention to the substitution are problematic. Using cleft-phrases (“It was X, who…”) causes both a syntactic and a semantic change (such a cleft would generally indicate a requirement to confirm the assertion in the initial phrase, or it would indicate an enhanced focus). Nonetheless, this manipulation does not entirely eradicate semantic illusion responses (Brédart and Modolo, 1988). Using typographical cues, such as capitalisation or underlining – either in the target sentence itself, or in a prime preceding the presentation of the target sentence – leads to both an increase in detection rate (Brédart and Docquier, 1989) and to a response bias causing participants to be inclined to make many false detections (Kamas, Reder and Ayers, 1996). Even sentences which naturally have the substitution in the position that is perceived to be the focus of the sentence lead to semantic illusions (Erickson and Mattson, 1981; Chapter 7, focus determination task). The fact that the physical position of the substitution within a target sentence

has no effect upon semantic illusion rate (Chapters 5 and 6, Experiments 2 and 3) suggests that the target word position cannot be used as a guideline to determining the sentence focus as had been suggested by Erickson and Mattson.

5. It is clear that target word and substitution have to be similar to each other for a semantic illusion to occur. Both phonological and semantic similarity contribute to semantic illusion rate (Erickson and Mattson, van Oostendorp and de Mul, 1990, van Oostendorp and Kok, 1990; Shafto and MacKay, 2000). Semantic similarity plays the more important role, but what exactly is required for terms to be considered similar is not clear – as long as a substitution provides a satisfactory match for the role-slot it occupies, the exact nature of the similarity does not appear to matter. More or less similar terms have different effects upon 'literal' and 'gist' tasks (see section 2.2). More similar terms facilitate processing in a 'gist' task, as it is easier to process a sentence as if there was no substitution if the substitution is very similar to the word it replaces. More similar terms also lead to a greater semantic illusion rate in a 'literal' task, as it is harder to detect a substitution that is very similar to the word it replaces. Less similar terms have the opposite effect – they are easier to detect and harder to ignore (van Oostendorp and de Mul, 1990; van Oostendorp and Kok, 1990; van Jaarsveld, Dijkstra and Hermans, 1997). The substitutions that are most likely to lead to the greatest semantic illusion rate are those that are both phonologically and semantically similar to the target word (Shafto and MacKay, 2000).

6. Of the two tasks used in semantic illusion research, statement-verification and question-answering, the latter leads to a far greater semantic illusion rate in a straightforward comparison (Chapter 4, Experiment 1). However, there are also many more "don't know" responses for questions than for statements-to-be-verified.

Apart from the different in overall semantic illusion rate, statement-verification and question-answering are affected differently by various experimental manipulations. The similarity effect (see section 2.4) differs for statements and question depending on the position of the substitution within the sentence. For questions, a substitution closer to the start of the sentence leads to a greater similarity effect than a substitution nearer the end of the sentence. For statements, there is no difference in the size of the effect between the two positions (van Jaarsveld, Dijkstra and Hermans, 1997). Similarly, extra information contained in the target sentence (i.e. sentence length) has a different effect for the two tasks. In the question-answering task, the more information relevant to the answer is contained in the sentence, the higher the likelihood of a semantic illusion response (Reder and Cleeremans, 1990; Reder and Kusbit, 1991). But in the statement-verification task, there is no effect of sentence length on semantic illusion rate (Chapter 5, Experiment 2).

7. When compared with other examples of written discourse, semantic illusions are seen to be part of a different group than control sentences taken from day-to-day sources of written materials. They are more alike to each other than to other sentences, are seen to be more stand-alone, to deal with more familiar topics, and they appear less formal or complex and as such are easy and straightforward to process, at least when rated on a series of rating scales (Chapter 7).

12.2 Beneath the surface: Mechanisms that might underlie semantic illusions

In this section various theoretical suggestions that have been made to explain semantic illusions are reviewed.

1. Semantic illusions cannot be explained by a failure to encode the substitution. Reading the sentences aloud, thus ascertaining that the word is processed at least at

a phonemic level, has no effect upon semantic illusions. There is also no evidence from studying participants’ reading times for individual words that there is particularly shallow processing of substituted words when a semantic illusion occurs. Nor can semantic illusions be explained by participants’ failure to retrieve the relevant knowledge from long term memory, as even when the facts relevant to semantic illusions are studied in advance of the semantic illusion task, the semantic illusion rate is not substantially affected (Reder and Kusbit, 1991).

2. It seems likely that participants adopt a partial matching approach to processing in general, which allows rapid and efficient information processing (Reder and Kusbit, 1991). Such a strategy is in keeping with a schema theory of how knowledge is stored in long term memory (Bartlett, 1995[1932]). The partial match is likely to be carried out on a basis of the global goodness-of-fit (Barton and Sanford, 1993). These strategies, which can lead to an error in processing semantic illusion sentences, are extremely useful in real life situations, where the rapid extraction of meaning from any number of possible situations takes precedence over the need to pay attention to small details. There appear to be two mechanisms, which contribute to partial matching: one is related to integrating information from the sentence context, the other is related to connecting the information from the sentence with previous knowledge (Hannon and Daneman, 2001).

3. There is evidence that semantic illusions are subject to interference from the context in which they are presented (Chapter 10; Experiment 5), so that the fact that many semantic illusions are presented in a context of largely correct statements-to-be-verified or properly phrased questions is likely to discourage participants from noticing anything untoward with semantic illusion sentences. If the context includes a majority of false fillers, more semantic illusions are detected. However, this effect

is small, and the context is less likely to affect people's actual detection rates, and more likely to be the result of a processing bias not unlike the bias that is observed when the target word is underlined (Brédart and Docquier, 1989; Kamas, Reder and Ayers, 1996). When target words are underlined, participants appear to detect more semantic illusions, but they also make many more 'false-alarm' responses. When target statements are surrounded by many false fillers, participants appear to detect more semantic illusions, but they also give fewer correct responses to true fillers. Both these observations indicate a response bias, rather than a change in sensitivity to substitutions.

4. There appears to be a strongly 'automatic' element to sentence processing, not unlike the Stroop effect (1935), in which an incongruous colour name interferes with the naming of the colour in which the stimulus is printed. Evidence for this claim comes from the observation that far fewer semantic illusion responses are made when the sentences containing semantic illusions are not presented as straightforward sentences, but as 'sentence-puzzles' made up of jumbled-up sentence chunks (Chapter 9, Experiment 4b). It was argued that this manipulation removed the automatic component of sentence processing, which uses a partial matching process to achieve comprehension as quickly as possible, since with the puzzles, participants had to process every chunk of the sentence in more depth.

5. Depending on the processing load in working memory, semantic illusions are more or less likely to occur (Chapter 11, Experiments 6a and 6b). When the central executive is busy, there is an increased number of semantic illusion responses, suggesting that semantic illusion sentences place a fairly large load on working memory while being processed. It is likely that in the course of processing a semantic illusion sentence, a 'chunking' process takes place, during which the

substituted word is incorporated into a larger processing unit relevant to comprehension. If working memory is not busy, this ‘chunking process may not lead directly to a semantic illusion, as the system may still have a record of the actual phrasing of the sentence, permitting a belated correction of the processing at hand. But since most semantic illusions are fairly long and complex, a heavy load is placed upon working memory and it seems unlikely that these kinds of sentences are processed ‘on-line’.

12.3 A conspiracy amongst minnows: Semantic illusions as a combination of effects?

Even though none of the explanations for semantic illusions tested to date is on its own sufficient to fully account for the frequency of occurrence of semantic illusions, it seems possible that the phenomenon is due to a combination of some or all of the effects observed. This theory can be crudely tested by summing the percentages of semantic illusion responses that each variable alone can account for. The size of each effect was estimated by subtracting the semantic illusion rate in each experimental condition in which a significant reduction in the semantic illusion rate was observed, from the corresponding control or comparison condition. The resulting value representing the reduction in semantic illusion rate was then transformed into a percentage of the control/comparison semantic illusion rate. This value is referred to as ‘*percentage of semantic illusion accounted for*’ (see Table 12.1 for details). Where more than one study has explored a particular effect, as for example in the case of sentence focus (Brédart and Modolo, 1988, Brédart and Docquier, 1989, Reder, Kamas and Ayers, 1996), an average effect size was estimated. All concerns regarding possible biases were disregarded for the purpose of this estimate.

In the studies concerned with the effect of working memory load upon semantic illusions (see Chapter 11), experimental manipulations were found to increase semantic illusion rate. In these cases, the effect size was estimated by calculating the percentage by which the manipulation had increased in experimental conditions compared to the control condition.

Table 12.1: Breakdown of percentage of semantic illusions accounted for by each effect studied (indicative only).

Effect	*Mean SI rate in control/comparison condition as a percentage of all responses made (SI_C)*	*Mean SI rate in experimental condition as a percentage of all responses made (SI_E)*	*Effect size as a percentage of all responses made ($E=\|SI_C - SI_E\|$)*	*Percentage of SI_C accounted for by E.* **(E / SI_C x 100)**
Focus	*34.4*	*21.1*	*13.3*	*38.7*
Task demands	*31.3*	*20.3*	*11.0*	*35.1*
Similarity	*34.3*	*21.7*	*12.6*	*36.7*
Word position	*30.7*	*24.7*	*6.0*	*19.5*
Automatic processing	29.6	11.5	18.1	61.1
Filler atmosphere	*24.3*	*18.3*	*6.0*	*24.7*
Working memory load	24.7	33.3	9.2	37.2

The calculation described above demonstrated rather dramatically that the sum of all the effects of variables implicated in the Moses Illusion more than accounted for the overall observed semantic illusion rate. In fact, the percentage of semantic illusions that can be explained by a combination of all effects studied is no less than 253 percent. It is impossible to avoid the conclusion, therefore that the observed semantic illusion frequency effect observed in particular studies probably results from a combination of the factors known to affect semantic illusion rate, each contributing a relatively small proportion of the total effect. Under specific experimental conditions, presumably, only a subset of these factors work together, and if each factor could be simultaneously controlled, it should be possible to entirely eradicate semantic illusion responses. Perhaps a little disappointingly, it seems that there is no one ‘big effect’. Instead, it seems likely that the question “How many animals of each kind did Moses take on the Ark?”, for which Erickson and Mattson (1981) observed a semantic illusion rate of more than eighty percent, constitutes an example in which many of the factor associated with semantic illusion rate just happen to act in conjunction. No mysterious sea monster, but a conspiracy of minnows.

12.4 Out in the world: How semantic illusions relate to real life

In this section, the relevance of semantic illusions to research and some possible applications of the knowledge gained through the study of semantic illusions are discussed.

1. As mentioned previously (Chapter 1), semantic illusions are of interest to theories of sentence processing in much the same way that visual illusions help evaluate theories of visual processing. The processes that lead to illusions are likely to be the same processes that lead to successful processing under different circumstances, and hence the nature of the errors made in an illusion situation allows the researcher to gain insight into the mechanisms underlying human information processing. Being a part of sentence processing, semantic illusions can provide evidence supporting, or indeed, calling into question previous theoretical suggestions about how sentences are comprehended. Semantic illusions confirm the importance of the role of context upon comprehension; they provide evidence for the ‘effort-after-meaning’ that people make in processing linguistic material (Bartlett, 1995[1932]); and they suggest that – comprehension being the essential goal of most sentence processing activity – processing is geared primarily towards ‘making sense’ of an input, while detail of wording becomes quickly irrelevant and is forgotten (Sachs, 1967). In this way, semantic illusions can be interpreted as evidence for a model of sentence

processing which requires a mental representation to enable comprehension (e.g. Clark and Clark, 1977).

2. The existence of semantic illusions draws attention to people's lack of ability to pay attention to details in the process of comprehension. As people are constantly bombarded with diverse linguistic inputs, a filtering strategy which can ignore discrepancies such as those encountered in semantic illusions is generally beneficial, allowing for rapid and efficient processing. However, the existence of semantic illusions and the robustness of the phenomenon demonstrates that the approximate nature of sentence processing is not under conscious control, and hence cannot be adjusted easily in situations where attention to detail is of importance. This implies that special care has to be taken when this kind of attention is required to design appropriate equipment to assist monitoring, or in the case of printed materials for which monitoring equipment cannot be provided, it is important to phrase text in a way that meets processing expectations, and is transparent and easy to follow. Such a manipulation would ensure that the processing load on the central executive of working memory is minimised, allowing more resources to be allocated to the understanding of details.

3. The fact that expectation-based contributions play a large part in the processing of language materials, as demonstrated by semantic illusions, may be a direct result of the way in which people in Western countries are usually educated. Generally, teaching and learning conform to a 'didactic contract' (e.g. Schubauer-Leoni, 1986) in which specific bodies of knowledge are transferred according to certain sets of expectations. For example, in the teaching of mathematics, it is usual that only a certain type of problem is dealt with at any one time, in which the formal structure of the problem conforms to the expectation, that all elements presented in the

problem are required for a successful solution. In the case of semantic illusions, the expectations about the task at hand, raised by its familiar structure as a general knowledge test, lead the participant astray with regard to the actual task. This line of reasoning raises the question of how well standard school teaching helps to prepare people for the problem solving in the real world, and whether different didactic methods, encouraging and requiring a variety of reanalyses of teaching problems, could better prepare children for the processing requirements of the real world.

12.5 In future: Suggestions for further research

One of the most intriguing findings of this thesis was the fact that semantic illusion type sentences are viewed as different from control sentences taken from real sources of printed material. However, this finding was based entirely on subjective judgements of the participants involved in the rating scale study (Chapter 7). In order to explore the nature of the different between semantic illusion type sentences and sentences from ordinary discourse further, it might be enlightening to examine functional magnetic resonance imaging (fMRI) scans of brains engaged in the processing of sentences. Will there be different patterns of activation for the two types of material? Further, a comparison of fMRI scans of cases in which semantic illusions occurred to cases in which the substitution is detected might be of theoretical interest, potentially revealing if participants really fail to detect the substitution, or if the only fail to detect it consciously (cf. inattentional blindness, Mack and Rock, 1998).

Another issue that might benefit from further investigation is the question of familiarity and mental flexibility in the processing of semantic illusion sentences. In all the studies described in this thesis and in those described in the literature, the participants were educated normal adults. These participants were assumed to have a certain level of ‘general knowledge’, from which certain familiar ideas were drawn to design experimental stimuli. These participants are also all likely to have a degree of mental flexibility which allows processing to proceed smoothly, even in the face of problems such as substituted words.

It has also been argued (Reder and Cleeremans, 1990) that not everybody is equally likely to experience semantic illusions. People who are experts on given topics were thought not to fall for semantic illusions (e.g. Christian or Jewish clergy would not be expected to confuse Noah and Moses!). Similarly, someone to whom a topic was of special importance would not usually experience semantic illusions. It might be possible, for example, that a child would not confuse Moses and Noah either, because the story of the Ark might be much more important to them than to an adult.

Another reason why children might not be as likely to experience semantic illusions could be that their mental flexibility is still developing and hence not as adaptable to processing problems associated with semantic illusions. In order to investigate whether mental flexibility has an effect upon semantic illusions, a questionnaire could be handed to children and to sufferers of Parkinson's Disease, which has mental rigidity as one of its symptoms. Comparing semantic illusion rates from these two groups of participants to the semantic illusion rates for a group of normal adults might provide an insight into the role of mental flexibility in the processing of semantic illusion sentences.

12.6 A last word

On the whole, it appears as if the phenomenon of semantic illusions is far from exhausted, either in terms of theoretical implications, or in terms of directions for future research. While this thesis has succeeded in providing a more detailed picture of what semantic illusions are like, there has been little progress towards a more complete explanation for the occurrence of semantic illusions. So far, partial matching on the basis of global good fit is still the best candidate theory, but exactly how it works is still not fully understood – it seems at least in part related to processing habits and experiences and in part to working memory limitations. At the end of this thesis, it looks as if – maybe – the wrong question has been asked all along. If semantic illusions are a direct result of ordinary sentence processing strategies, why they occur is not nearly as interesting a question as its converse: Why do semantic illusions not occur a lot more often than they do?

References

ADAMSON, R.E. (1952). Functional fixedness as related to problem solving: a repetition of three experiments. *Journal of Experimental Psychology*, **44**, 288-291.

AKIYAMA, M., TAKEI, S., & SAITO, K. (1982). Statement verification and question answering in children. *Japanese Journal of Educational Psychology*, **30**, 265-272.

ANDERSON, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

ANDERSON, J.R. (1984). 'Spreading activation'. In J.R. Anderson and S.M. Kosslyn (eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower.* New York: W.H. Freeman.

ANDERSON, R.C., & PICHERT, J.W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, **17**, 1-12.

BADDELEY, A.D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 362-365.

BADDELEY, A.D. (1978). Working memory and reading. In P.A. Kolers, M.E. Wrolstad, and H. Bouma (Eds.), *Processing of visible language,* vol. 1 (pp.355-370). New York: Plenum Press.

BADDELEY, A.D. (1986). *Working memory*. Oxford: Oxford University Press.

BADDELEY, A.D., & HITCH, G. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation.* New York: Academic Press.

BADDELEY, A.D., ELDRIDGE,M., & LEWIS,V. (1981). The role of subvocalisation in reading. *Quarterly Journal of Experimental Psychology,* **33A**, 439-454.

BADDELEY, A.D., LEWIS, V., & VALLAR, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology:Human Experimental Psychology,* **36A**, 233-252.

BADDELEY, A.D., VALLAR, G., & WILSON, B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), *The psychology of reading; Attention and Performance XII* (pp. 509-529). Hillsdale: Lawrence Erlbaum Associates.

BAKER, L., & WAGNER, J.L. (1987). Evaluating information for truthfulness: the effects of logical subordination. *Memory and Cognition*, **15**, 247-255.

BARTLETT, F.C. (1995 [1932]). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

BARTON, S.B., & SANFORD, A.J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition*, **21**, 477-487.

BATES, E., DEVESCOVI, A., & D'AMICO, S. (1999). Processing complex sentences: a cross-linguistic study. *Language and Cognitive Processes,* **14**, 69-123.

BATES, E., MCNEW, S, MACWHINNEY, B, DEVESCOVI, A, & SMITH, S. (1982). Functional constraints on sentence processing: a cross-linguistic study. *Cognition*, **11**, 245-299.

BEAL, C.R., BONITATIBUS, G.J., & GARROD, A.C. (1990). Fostering children's revision skills through training in comprehension monitoring. *Journal of Educational Psychology*, **82**, 275-280.

BEGG, I., & DENNY, J.P. (1969). Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning errors. *Journal of Experimental Psychology*, **81**, 351-354.

BIEDERMAN, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, **94**, 115-147.

BIRCH, H.G., & RABINOWITZ, H.S. (1951). The negative effect of previous experience on productive thinking. *Journal of Experimental Psychology*, **41**, 121-125.

BOOLE, G. (1847; REPR. 1948). *The Mathematical Analysis of Logic: being an essay towards a calculus of deductive reasoning.* Oxford.

BRANSFORD, J.D., & FRANKS, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, **3**, 331-350.

BRANSFORD, J.D., & JOHNSON, M.K. (1972). Contextual prerequisites for understanding: some ivestigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, **11**, 717-726.

BRANSFORD, J.D., BARCLAY, J.R., & FRANKS, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, **3**, 193-209.

BRANSFORD, J.D., MCCARRELL, N.S., & NITSCH,K.E. (1976). Context, comprehension, and semantic flexibility: Theoretical and methodological implications. *Bulletin de Psychologie,* Special Annual, 335-345.

BRÉDART, S., & DOCQUIER, M. (1989). The Moses illusion: a follow-up on the focalization effect. *Cahiers de Psychologie Cognitive/ European Bulletin of Cognitive Psychology*, **9**, 357-362.

BRÉDART, S., & MODOLO, K. (1988). Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica*, **67**, 135-144.

BROADBENT, D.E. (1958). *Perception and communication.* Oxford: Pergamon.

BROWN R., & GILMAN, A. (1961). The pronouns of power and solidarity. In Thomas A. Sebeok (Ed.), *Style in Language.* (pp. 253-276) Cambridge, Mass. : MIT Press. .

CAPLAN, D., BAKER, C., & DEHAUT, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, **21**, 117-175.

CARAMAZZA, A., BASILI, A.G., KOLLER, J.J., & BERNDT, R.S. (1981). An investigation of repetition and language processing in a case of conduction aphasia. *Brain and Language*, **14**, 235-271.

CARPENTER, P.A., & JUST, M.A. (1983). What your eyes do when your mind is reading. In K.Rayner (ed.), *Eyemovements in reading: Perceptual and language processes* (pp.275-307). New York: Academic Press.

CHAPMAN, L.J., & CHAPMAN, J.P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, **58**, 220-226.

CHOMSKY, N. (1957). *Syntactic Structures*. The Hague: Mouton.

CHOMSKY, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

CLARK, H.H., & CHASE, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, **3**, 472-517.

CLARK, H.H., & CLARK, E.V. (1977). *Psychology and Language*. New York: Harcourt Brace Jovanovich.

COLTHEART, M., DAVELAAR, C., JONASSON, J.T., & BESNER, D. (1977). Access to the internal lexicon. In S. Dornic (ed.), *Attention and Performance VI*. New York: Academic Press.

CONRAD, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, **55**, 75-84.

CRAIN, S., & STEEDMAN, M.J. (1985). On not being led up the garden path: the use of context by the psychological parser. In D. Dowty, L. Karttunen and A. Zwicky (Eds.) *Natural Language Parsing.* Cambridge: Cambridge University Press.

CRONK, B.C., & SCHWEIGERT, W.A. (1992). The comprehension of idioms: the effects of familiarity, literalness, and usage. *Applied Psycholinguistics*, **13**, 131-146.

CRONK, B.C., LIMA, S.D., & SCHWEIGERT, W.A. (1993). Idioms in sentences: effects of frequency, literalness and familiarity. *Journal of Psycholinguistic Research*, **22**, 59-82.

DANEMAN, M., & CARPENTER, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, **19**, 450-466.

DANEMAN, M., & CARPENTER, P.A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology:Learning, Memory &Cognition*, **9**, 561-584.

DANEMAN, M., & MERIKLE, P.M. (1996). Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin and Review*, **3**, 422-433.

DEUTSCH, J.A., & DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychological Review*, **70**, 80-90.

DICKSTEIN, L.S. (1978). The effect of figure on syllogistic reasoning. *Memory and Cognition*, **6**, 76-83.

DOOLING, D.J., & LACHMAN, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, **88**, 216-222.

DUNCKER, K. (1945). On Problem-Solving. *Psychological Monographs*, **5** (Whole no. 270).

ENGELKAMP, J., ZIMMER, H.D., & MOHR, G. (1990). Differential memory effects of concrete nouns and action verbs. *Zeitschrift für Psychologie*, **198**, 189-216.

ERICKSON, T.D., & MATTSON, M.E. (1981). From words to meaning: a semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, **20**, 540-551.

FARMER, E.W., BERMAN, J.V., & FLETCHER, Y.L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *Quarterly Journal of Experimental Psychology:Human Experimental Psychology*, **38A**, 675-688.

FIEBACH, C.J., SCHLESEWSKY, M., & FRIEDERICI, A.D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, **30**, 321-338.

FILLENBAUM, S. (1970). On the use of memorial techniques to assess syntactic structures. *Psychological Bulletin*, **73**, 231-237.

FODOR, J. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

FODOR, J. (1985). Precis of The modularity of mind. *Behavioral and Brain Sciences*, **8**, 1-5.

FODOR, J. (1986). The modularity of mind. In Z.Pylyshyn and W. Demopoulos (eds.), *Meaning and cognitive structure* (pp. 3-18 and pp. 129-135). Norwood, NJ: Ablex.

FOERTSCH, J., & GERNSBACHER, M.A. (1994). In search of complete comprehension: Getting “Minimalists” to work. *Discourse Processes*, **18**, 271-296.

GARNHAM, A. (1985). *Psycholinguistics: central topics.* New York: Routledge.

GATHERCOLE, S.E., & BADDELEY, A.D. (1989). Development of vocabulary in children and short-term phonological memory. *Journal of Memory and Language*, **28**, 200-213.

GELUYKENS, R. (1992). *From Discourse Process to Grammatical Construction: On Left-dislocation in English.* Amsterdam: John Benjamins.

GERNSBACHER, M.A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, **17**, 324-363.

GILHOOLY, K.J., LOGIE, R.H., WETHERICK, N.E., & WYNN, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory and Cognition*, **21**, 115-124.

GLANZER, M., & CUNITZ, A.R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, **5**, 351-360.

GLENBERG, A.M., ROBERTSON, D.A., JANSEN, J.L., & JOHNSON-GLENBERG, M.C. (1999). Not propositions. *Journal of Cognitive Systems Research*, **1**, 19-33.

GOLDTHWAITE, D. (1997). Knowledge of pragmatic conversational structure. *Journal of Psycholinguistic Research*, **26**, 497-508.

GRAESSER, A.C., MCMAHEN, C.L., & JOHNSON, B.K. (1994). Question asking and answering. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 517-538). New York: Academic Press.

GRICE, H.P. (1975). Logic and Conversation. In Cole, P., Morgan, J.L. (eds) *Syntax and Semantics 3: Speech acts.* New York: Academic Press.

GRICE, H.P. (1978). Further notes on logic and conversation. In Cole, P. (ed) *Syntax and Semantics 9*. New York: Academic Press.

HACKER, D.J., PLUMB, C., BUTTERFIELD, E.C., QUATHAMER, D., & HEINEKEN, E. (1994). Text revision: detection and correction of errors. *Journal of Educational Psychology*, **86**, 65-78.

HALLIDAY, M.A.K. (1967). Notes on transitive and theme in English (Part 2). *Journal of Linguistics*, **3**, 177-274.

HALLIDAY, M.A.K. (1978). *Language as Social Semiotic*. London: Edward Arnold.

HANNON, B, & DANEMAN, M. (2001). Susceptibility to semantic illusions: An individual-differences perspective. *Memory and Cognition*, **29**, 449-461.

HARRIS, D.M., & KAY, J. (1995). I recognize your face but can't remember your name: Is it because names are unique? *British Journal of Psychology*, **86**, 345-358.

HEALY, A.F. (1981). The effects of visual similarity on proofreading for misspellings. *Memory and Cognition*, **9**, 453-460.

HESS, D.J., FOSS, D.J., & CARROLL, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, **124**, 62-68.

HINTON, G.E., MCCLELLAND, J.L., & RUMELHART, D.E. (1986). Distributed representations. In D.E. Rumelhart, J.L. McClelland, and the PDP research group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* (Vol.1). Cambridge, MA: MIT Press.

HORN, L.R. (1989). *A natural history of negation.* Chicago: University of Chicago Press.

HULME, C., MAUGHAN, S., & BROWN, G.D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, **30**, 685-701.

JACKENDOFF, R. (1972). *Semantic interpretation in generative grammar.* Cambridge, Mass.: MIT Press.

JACKENDOFF, R. (1993). *Patterns in the mind: language and human nature.* New York: Harvester Wheatsheaf.

JOHNSON, M.K., BRANSFORD, J.D., & SOLOMON, S. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology*, **98**, 203-205 .

JOHNSON-LAIRD, P.N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness.* Cambridge: Cambridge University Press.

JOHNSON-LAIRD, P.N., & BARA, B.G. (1984). Syllogistic inference. *Cognition*, **17**, 1-61.

JÖRESKOG, K., & SÖRBOM, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Hillsdale, NJ: Erlbaum.

KAMAS, E.N., REDER, L.M., & AYERS, M. S. (1996). Partial matching in the Moses illusion: response bias not sensitivity. *Memory & Cognition*, **24**, 687-699.

KATZ, J.J. (1977). *Propositional Structure and Illocutionary Force.* New York: Crowell.

KINTSCH, W. (1988). The use of knowledge in discourse processing: a construction-integration model. *Psychological Review*, **95**, 163-182.

KINTSCH, W., & VAN DIJK, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, **85**, 363-394.

KORIAT, A. (1993). How do we know that we know? The accessibility model of feeling of knowing. *Psychological Review*, **100**, 609-639.

LANGENMAYR, A. (1997). *Sprachpsychologie: Ein Lehrbuch.* Göttingen: Hogrefe.

LARIGAUDERIE, P., GOANAC'H, D., & LACROIX, N. (1998). Working memory and error detection in texts: What are the roles of the central executive and the phonological loop? *Applied Cognitive Psychology*, **12**, 505-527.

LEVIN, H., & GARRETT, P. (1990). Sentence structure and formality. *Language in Society*, **19**, 511-520.

LEVINSON, S.C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

LEVY, B.A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **10**, 123-132.

LEVY, B.A. (1975). Vocalization and suppression effects in sentence memory. *Journal of Verbal Learning and Verbal Behavior,* **14**, 304-316.

LEVY, B.A., NEWELL, S., SNYDER, J., & TIMMINS, K. (1986). Processing changes across reading encounters. *Journal of Experimental Psychology:Learning, Memory and Cognition*, **12**, 467-478.

LIBERMAN, A.M. (1995). The relation of speech to reading and writing. In B. de Gelder and J. Morais (Eds.), *Speech and Reading: A Comparative Approach.* (pp. 17-31). Hove: Lawrence Erlbaum.

LOGIE, R.H. (1995). *Visuo-spatial working memory.* Hove: Lawrence Erlbaum.

LOGIE, R.H., BADDELEY, A.D., MANE, A., DONCHIN, E., & SHEPTAK, R. (1989). Working memory in the acquisition of complex cognitive skills. *Acta Psychologica*, **71**, 53-87.

LUCHINS, A.S. (1942). Mechanization in problem solving: the effect of Einstellung. *Psychological Monographs*, **54**(1), (entire).

LUCHINS, A.S., & LUCHINS, E.H. (1950). New experimental attempts at preventing mechanization in problem solving. *Journal of General Psychology*, **42**, 279-297.

LUCHINS, A.S., & LUCHINS, E.H. (1959). *Rigidity of behavior: A variational approach to the effects of Einstellung.* Eugene: University of Oregon Books.

LYONS, J. (1977). *Semantics* (2 vols.) Cambridge: Cambridge University Press.

MACK, A., & ROCK, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

MACWHINNEY, B., BATES, E., & KLIEGL, R. (1984). Cue Validity and sentence interpretation in English, German and Italian. *Journal of Verbal Learning and Verbal Behavior*, **23**, 127-150.

MAIER, N.R.F. (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, **12**, 181-194.

MANDLER, G., PEARLSTONE, Z., & KOOPMANS, H.S. (1969). Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, **8**, 410-423.

MARKMAN, E.M. (1985).Comprehension monitoring: Developmental and educational issues. In S.F. Chipman, J.W. Segal and R. Glaser (Eds.) *Thinking and Learning Skills*, Vol. 2: *Research and open questions.* (pp.275-291) Hillsdale, NJ: Lawrence Erlbaum Associates.

MARSLEN-WILSON, W.D. (1987). Function and process in spoken word recognition. *Cognition*, **25**, 71-102.

MARSLEN-WILSON, W.D., & TYLER, L.K. (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1-71.

MARTIN, R.C. (1987). Articulatory and phonological deficits in short-term memory and their relation to syntactic processing. *Brain and Language*, **32**, 159-192.

MARTIN, R.C. (1990). Neuropsychological evidence on the role of short-term memory in sentence processing. In G. Vallar and T. Shallice (Eds.), *Neuropsychological impairments of short-term memory.* (pp. 390-427). Cambridge: Cambridge University Press.

MCCLELLAND, J.L., ST. JOHN, M., & TARABAN, R. (1989). Sentence comprehension: a parallel distributed processing approach. *Language and Cognitive Processes*, **4**, 287.

MCELREE, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, **29**, 111-123.

MCKOON, G., & RATCLIFF, R. (1992). Inference during reading. *Psychological Review*, **99**, 440-446.

MCWEENY, K.H., YOUNG, A.W., HAY, D.C., & ELLIS, A.W. (1987). Putting names to faces. *British Journal of Psychology*, **78**, 143-149.

MEY, J.L. (2001). *Pragmatics: an introduction* (2nd ed). Oxford: Blackwell Publishers.

MEYER, M., FRIEDERICI, A.D., & VON CRAMON, D.Y. (2000). Neurocognition of auditory sentence comprehension: event-related fMRI reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research*, **9**, 19-33.

MILLER, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.

MILNE, R.W. (1982). Predicting garden path sentences. *Cognitive Science*, **6**, 349-373.

MURRAY, D.J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, **21**, 263-276.

NI, W., CRAIN, S., & SHANKWEILER, D. (1996). Sidestepping garden paths: assessing the contributions in syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, **11**, 283-334.

POSNER, M., & SNYDER, C. (1975). Facilitation and inhibition in the processing of signals. In P. Rabbitt and S. Dornic (Eds.), *Attention and performance V.* New York: Academic Press.

POSTMA, A. (1999). The influence of decision criteria upon remembering and knowing in recognition memory. *Acta Psychologica*, **103**, 65-76.

RAEBURN, V.P. (1979). The role of the verb in sentence memory. *Memory and Cognition*, **7**, 133-140.

REASON, J. (1990). *Human Error.* Cambridge: Cambridge University Press.

REDER, L.M. (1982). Plausibility judgements vs. fact retrieval: alternative strategies for sentence verification. *Psychological Review*, **89**, 250-280.

REDER, L.M. (1987). Strategy selection in question answering. *Cognitive Psychology*, **19**, 90-138.

REDER, L.M., & CLEREMANS, A. (1990). The role of partial matches in comprehension: the Moses illusion revisted. In A.C. Graesser and G.H. Bower (eds.), *Inferences and Text Comprehension.* The Pyschology of Learning and Motivation (Vol. 25, pp 233-258). New York: Academic Press.

REDER, L.M., & KUSBIT, G.W. (1991). Locus of the Moses illusion: imperfect encoding, retrieval or match? *Journal of Memory and Language*, **30**, 385-406.

REVLIN, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior*, **14**, 180-195.

REVLIN, R., LEIRER, V., YOPP, H., & YOPP, R. (1980). The belief bias effect in formal reasoning: the influence of knowledge on logic. *Memory and Cognition*, **8**, 584-592.

REYNOLDS, A.G., & FLAGG, P.W. (1976). Recognition memory for elements of sentences. *Memory and Cognition*, **4**, 422-432.

ROBERTSON, S.P., ULLMAN, J., & MEHTA, A. (1992). Simultaneous question comprehension and answer retrieval. *Proceedings of the Fourteenth Annual meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

ROESLER, F., PECHMANN, T., STREB, J., ROEDER, B., & HENNIGHAUSEN, E. (1998). Parsing of sentences in a language with varying word order: word-by-word variations of processing demands are revealed by event-related braing potentials. *Journal of Memory and Language*, **38**, 150-176.

ROSS, J.R. (1967). *Constraints on variables in syntax.* Doctoral Dissertation. Massachusetts Institute of Technology, Cambridge [Published 1986 as *Infinite Syntax!* Norwood, NJ: Ablex.].

RUBENSTEIN, H., GARFIELD, L., & MILLIKAN, J.A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, **9**, 487-492.

RUMELHART, D.E., HINTON, G.E., & WILLIAMS, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart, J.L. McClelland, and the PDP research group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* (Vol.1). Cambridge, MA: MIT Press.

SAARILUOMA, P. (1991). Visuo-spatial interference and apperception in chess. In R.H. Logie and M. Denis (Eds.) *Mental images in human cognition*. (pp. 83-94). Amsterdam: North-Holland.

SACHS, J.S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, **2**, 437-442.

SAFFRAN, E.M., & MARIN, O.S. (1975). Immediate memory for wordlists and sentences in a patient with deficient auditory short-term memory. *Brain and Language*, **2**, 420-433.

SANFORD, A. J., & GARROD, S.C. (1989). What, when and how?: Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes*, **4**, 235-262.

SCHANK, R.C., & ABELSON, R.P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

SCHLESINGER, I.M. (1968). *Sentence Structure and the Reading Process.* The Hague: Mouton.

SCHRIEFERS, H., TERUEL, E., & MEINHAUSEN, R.M. (1998). Producing simple sentences: results from picture-word interference experiments. *Journal of Memory and Language*, **39**, 609-632.

SCHUBAUER-LEONI, M.L. (1986). Le Contrat Didactique: un Cadre Interpretatif pour Comprendre les Savoirs Manifestes par les Eleves en Mathematique. / The didactic contract: An interpretative framework for understanding the mathematical knowledge shown by students. *European Journal of Psychology of Education*, **1**, 139-153.

SEARLE, J.R. (1975). Indirect speech acts. In Peter Cole and Jerry Morgan (Eds.), *Syntax and Semantics*, vol. 3: *Speech Acts.* (pp. 59-82). New York: Academic Press.

SELLS, S.B. (1936). The atmosphere effect: An experimental study of reasoning. *Arch. Psychol., N.Y.*, **29**, 3-72.

SHAFTO, M., & MACKAY, D.G. (2000). The moses, mega-Moses and Armstrong iluusions: integrating language comprehension and semantic memory. *Psychological Science*, **11**, 372-378.

SHIFFRIN, R.M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual Learning, automatic attending, and a general theory. *Psychological Review*, **84**, 127-190.

SMITH, E.E. , SHOBEN, E.J., & RIPS, L.V. (1974). Structure and process in semantic memory: a featural model for semantic decision. *Psychological Review*, **81**, 214-241.

SOMMERS, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition and Communication*, **31**, 378-388.

STALNAKER, R.C. (1977). Pragmatic presuppositions. In Andy Rodgers, Bob Wall and John P. Murphy (Eds.), *Proceedings of the Texas Conference on Perfomatives, Presuppositions, and Implicatures* (pp. 135-147). Washington, D.C.: Center for Applied Linguistics. [1974]..

STANNERS, R.F., & FORBACH, G.B. (1973). Analysis of letter strings in word recognition. *Journal of Experimental Psychology*, **98**, 31-5.

STERNBERG, S. (1966). High speed scanning in human memory. *Science*, **153**, 652-654.

STROOP, J. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, **18**, 643-662 .

THOMAS, J. (1995). *Meaning in interaction: an introduction to pragmatics.* London: Longman.

TREISMAN, A.M. (1964). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology*, **77**, 206-219.

VALENTINE, T., HOLLIS, J., & MOORE, V. (1998). On the relationship between reading, listening and speaking: It's different for people's names. *Memory and Cognition*, **26**, 740-753.

VAN DIJK, T.A., & KINTSCH, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

VAN JAARSVELD, H.J., DIJKSTRA, T., & HERMANS, D. (1997). The detection of semantic illusions: Task-specific effects for similarity and position of distorted terms. *Psychological Research/Psychologische Forschung*, **59**, 219-230.

VAN OOSTENDORP, H., & DE MUL, S. (1990). Moses beats Adam: a semantic relatedness effect on a semantic illusion. *Acta Psychologica*, **74**, 35-46.

VAN OOSTENDORP, H., & DEN UYL, M.J. (1984, June). *Semantic relatedness effects in text processing.* Paper presented to the joint meeting of the Experimental Psychology Society and the Netherlands Psychonomic Foundation, University of Amsterdam.

VAN OOSTENDORP, H., & KOK, I. (1990). Failing to notice errors in sentences. *Language and Cognitive Processes*, **5**, 105-113.

VICENTE, K.J. & BREWER, W.F. (1993). Reconstructive remembering of the scientific literature. *Cognition*, **46**, 101-128.

VOS, S.H., GUNTER, T.C., KOLK, H.H.J., & MULDER, G. (2001). Working memory constraints on syntactic processing: an electrophysiological investigation. *Psychophysiology*, **38**, 41-63.

WALLAS, G.C. (1926). *The art of thought.* New York: Harcourt, Brace.

WASON, P.C. & JOHNSON-LAIRD, P.N. (EDS.) (1968). *Thinking and reasoning: selected readings.* Harmondsworth: Peguin Books Ltd.

WATERS, G., CAPLAN, D., & HILDEBRANDT, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *The psychology of reading; Attention and Performance XII* (pp. 531-555). Hillsdale: Lawrence Erlbaum Associates.

WETHERICK, N.E., & GILHOOLY, K.J. (1990). Syllogistic reasoning: effects of premis order. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie and G. Erdos (Eds.), *Lines of thinking* (Vol.1, pp. 99-108). Chichester: Wiley.

WILLIAMS, J.N. (1992). Processing of polysemous words in context: evidence for interrelated meanings. *Journal of Psycholinguistic Research*, **21**, 193-218.

WILSON, S.G., RINCK, M., MCNAMARA, T.P., BOWER, G.H., & MORROW, D.G. (1993). Mental models and narrative comprehension: Some qualifications. *Journal of Memory and Language*, **32**, 141-154.

WITTGENSTEIN, L. (1958). *Philosophical Investigations.* Oxford Blackwell.

WOODWORTH, R.S., & SELLS, S.B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, **18**, 451-460.

ZWAAN, R.A. (1994). Effect of genre on text comprehension.*Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**, 920-933.