Metaphor under the scope of negation

Claire Joyce
B.A. (Mod.) CSLL
Final Year Project, May 07
Supervisor: Dr. Carl Vogel
Declaration

I hereby declare that this dissertation is entirely my own work and that it has not been submitted as an exercise for a degree at any other university.

_____________________________ 04 May 2007
Claire Joyce
Abstract

This paper details a set of two experiments designed to show that metaphor is more difficult to understand under the scope of negation. Negation effects all discourse, so people should be much worse at understanding negative metaphors than negative literals. As metaphor is literally false, negating this will be a kind of double negative, intuitively more difficult to understand. This study gives a complete and thorough set of experiments, covering all areas of negation and established metaphor.
Acknowledgements

I would like to thank Dr. Carl Vogel, CSLL coordinator and supervisor, who has always been there to help me with everything throughout my years in Trinity and abroad, and has made the experience all the more enjoyable.

Next I would like to thank my family for being there when I need them and stepping back when the pressure is on.

Finally I would like to thank my friends, who distracted me at every turn.
List of Figures

Figure 3.1. Example of format of analogy for all experiments.
Figure 4.1: Screenshot, my experiments home page
Figure 4.2: Screenshot, editing Experiment 1
Figure 5.1. Chart showing the affect of sex on performance by individual target
Figure 5.2. Number of participants in each age category
Figure 5.3. Average Strictly Correct answers by age
Figure 5.4. Average Strictly Correct answers by age and by target.
Figure 5.5. Average Strictly Correct answers by handedness and target.
Figure 5.6. Average Strictly Correct answers by linguistic background and target.
Table 5.1. This table shows the number of Strictly Correct answers per target.
Table 5.2. This table shows the total time taken in seconds to give a Strictly Correct answer by target.
Table 5.3. This table shows the number of Strictly Correct answers by source.
Figure 5.7. This figure charts strict correctness against error and rationality by men and women.
Figure 5.8. Average Strictly correct answers by sex and target.
Figure 5.9. This figure charts strict correctness error and rationality by handedness.
Figure 5.10. Average Strictly correct answers by handedness and target.
Table 5.4. This table shows the number of Strictly Correct answers per target.
Table 5.5. This table shows the number of Strictly Correct answers per target ignoring oppositeness.
Figure 5.11 This is a graph of the average number of answers by Rationality, Strictness and Error.
CHAPTER 1

INTRODUCTION
1.1 Introduction

This is an introductory chapter, which expands upon the aim of this study and gives an overview of what to expect from the rest of the paper.

1.2 Aims

The aim of this study is to prove by psycholinguistic experiments that metaphor is more difficult to understand under the scope of negation. Intuitively it seems that it would be as negation makes things more complicated and metaphor is already complicated, so combined it would be doubly complicated. This study expands upon the studies of past students Sheehan (2005) and McGillion (2006). I wanted to build upon their work by making my experimentation more complete and presenting my results in a more meaningful way.

1.3 Structure of paper

This section details the chapters of this study.

Chapter 2- Background Reading

This chapter gives some description of the different views and theories there are about metaphor in the literature.

Chapter 3 – methodology

This chapter explains everything involved in creating my experiments. It details the demographic questions I asked and the two different experiments I conducted.

Chapter 4 – experimentation

This chapter describes the web tool and how I used it to create my experiments.

Chapter 5 – Results and discussion

This chapter shows the results for both experiments. It contains many graphs and tables of the data obtained in the study. All aspects of the results are analysed closely and discussed in detail.

Chapter 6 – Conclusion

This chapter summarises the project and notes the achievements made. It also gives suggestions for future work in the area.

1.4 Conclusion

This chapter has explained the aim of this project and has given an overview of this study contains, with a brief description of the contents of each chapter.
CHAPTER 2

BACKGROUND READING
2.1 Introduction

In this chapter I will present some of the literature written on the topic of metaphor. Firstly, I will present what metaphor has been defined as. Then I will give an explanation of the terms used to identify the different parts of metaphor. Then I will give a description of some theories of interpreting metaphor. Finally I will discuss what is meant by negative metaphor.

2.2 What is Metaphor?

The etymology of the word is as follows “from metapherein "transfer, carry over," from meta- "over, across" + pherein "to carry, bear". "a transfer," especially of the sense of one word to a different word, lit. "a carrying over".” (Harper, 2001). WordNet (2006) defines metaphor as “a figure of speech in which an expression is used to refer to something that it does not literally denote in order to suggest a similarity”. When learning what a metaphor is in school, most people will be familiar with the definition that a metaphor is a simile without using ‘like’ or ‘as’.

But defining what metaphor is and what it is for has long been a problem for writers on this area. There is no one clear definition even though it has been worked upon by many from various different disciplines. Glucksberg (2001) gives two reasons for why this is the case. “First, the term is used in several different … senses. Second … definitions vary to reflect sharply different theoretical agendas and assumptions.” (2001, p.3). So there is not one single definition that all writers agree upon, or a standard they subscribe to. We can see this from the fact that linguist have coined many different terms to describe the parts of metaphor (see section 2.4). Instead there are many different theories as to what metaphor is and to how it works (see section 2.3).

There are two main opposing views when considering metaphor which relate to metaphors importance or lack there of to how we conceptualise. Supporters of it not being important, see it as merely a nice feature of language while those who do see it as incredibly telling about our cognitive and conceptual systems. Lakoff is the main writer in this area. He suggests that metaphor actually structures our thoughts. Lakoff and Johnson (1980) propose metaphor can help us make judgements about the nature of our conceptual system and as this is linked with cognition, metaphor can give insight here too. They talk about structural metaphors e.g. Ideas are objects, and orientational metaphors. These metaphors come from our physical surroundings e.g. the notion that ‘sad’ is ‘down’. Reddy suggests that all language is metaphorical. This is an interesting view as the world is not language, so any label we apply is a metaphor.

2.3 INTERPRETING METAPHOR

In this section I will describe some theories of metaphor interpretation.
2.3.1 The Substitution Theory

The Substitution theory is a view that is not really considered these days, but was the main theory of metaphor in the past. It states that the metaphor is merely a substitute for a literal term and its meaning can be found by substituting the literal term back in. This interpretation quickly falls down under investigation, as Goatly (1997) explains, it does “not allow … that … a metaphorical expression might receive a number of different … interpretations” (1997, p.116). The substitution theory suggests there is only one precise way to literally paraphrase a metaphorical expression but this is obviously not the case.

2.3.2 The Interaction Theory

To explain this theory Goatly uses an example from Black (1962),

A battle is a game of chess.

This theory involves the suppression and emphasis of certain features, which come from the interaction of the Vehicle and the Topic. So in the above example, certain features of the game of chess that can be applicable to battle will be emphasised, for example, the soldiers are pawns to be moved around. Others, for example, that the soldiers are very small and made of wood, are suppressed.

2.3.3 Conceptual Mapping Theory

Lakoff (1993) proposes this theory that metaphor is a mapping from a source domain to a target domain. We can understand this from the example Love is a Journey. What it involves is understanding an abstract concept in terms of a simpler, more concrete concept. The mapping is very structured and there is a systematic correspondence between the entities of each domain.

2.4 Terms

There are three parts to a metaphor, the Vehicle, the Topic and the Grounds. Goatly (1997) describes these terms as follows: “The conventional referent of the unit is the Vehicle. The actual unconventional referent is the Topic. The similarities and/or analogies involved are the Grounds.” (1997, p.9) The ‘unit’ he refers to is a unit of discourse that the metaphor occurs in. The meaning of these terms becomes more apparent given an example.

1. The past is a foreign country; they do things differently there. (Goatly, 1997, p.9)

In this example taken from Goatly (1997) the Vehicle is ‘a foreign country’, the Topic is ‘the past’ and the Grounds is ‘they do things differently there’. Often the Grounds is left unstated, and the similarity between the two terms is inferred by the hearer. For example, in 1 above, the Grounds could have been left out and the hearer would fill it in by finding a feature or features belonging to the second term that they can apply to the first in a meaningful way.
There are numerous different terms for these parts of metaphor that are used by different writers on the subject. Vehicle is also known as frame, conventional referent. Topic can be tenor, subject, focus or unconventional referent.

2.5 What does metaphor do?

Metaphor has been attributed many functions. Goatly (1997) has listed some of these functions. One such function he describes is that of filling lexical gaps. This occurs when no word exists that adequately describes what is meant. We see this a lot in computer science, e.g. mouse. We can liken this to how children use language when they are learning. They form metaphor in the same way adults do when they are overextending words to cope with objects they have not yet learnt. The only reason that their efforts are errors and adults are metaphors, is that adults intentionally used a metaphor, not just out of lack of knowledge of the language, but to make a specific comparison.

Another function Goatly describes is metaphor as an aid to explanation. In the field of science, for example, complex ideas are hard to visualise in their own terms. We can explain these ideas by comparison to concepts we are more comfortable with. An example is light being described as a wave, or Goatly’s example of electricity as water.

Goatly(1997) describes further functions of metaphor, but it is important to remember that these explanations of metaphor occur after the fact. It is not the case that someone comes across a situation in which they are faced with a lexical gap and they think to themselves, “I will now use this function of metaphor to fill the gap”. Nor is it the case that they are using metaphor specifically to complete some function. While these descriptions are interesting to show what metaphor can do, they are not to be focused on too heavily I feel, as they pertain more to descriptions of metaphor in language use rather than how metaphor relates to our conceptual systems, which is much more interesting.

2.6 Negative Metaphor

Negation makes any utterance more difficult to understand. When combined together with metaphor, which is already complicated, the difficulty increases. Unless they are negated from their creation and are established as negative eg no man is an island, negative metaphors will be more complicated to understand, even though they are literally true. We can see metaphor as a kind of negation itself because it is literally false. So when we are asking questions about negative metaphor, we are dealing with a kind of double negation, which is intuitively more difficult to understand than a single negation.

Despite the interest metaphor has enjoyed for many years, the area of negative metaphor is rarely mentioned. Vogel (2001) talks about the truth-values of metaphors. If read as literal sentences, there are simply false. So when they appear under the scope of negation, they are true. Vogel believes that negative metaphor will be hard for people to understand and that novel negative metaphor will be rejected completely. This study aims to lend proof to the first part of this belief.
2.7 Conclusion

In this chapter, I have detailed some of the main ideas in the literature about metaphor. I have given some definitions of metaphor and detailed some of its uses. I have also described some theories of interpretation and an explanation of negative metaphor. This explanation gave Vogel’s (2001) views on the subject of negative metaphor which this study aims to prove.
Chapter 3
Methodology
3.1 Introduction

In this chapter, I will detail how I went about forming my experiments and how I expected them to prove my aim. Firstly, I will detail some previous work in the area and why I took on and rejected some of their methodology. Next, I will describe the layout of each section of my experiments as they appeared to participants. This will include a full explanation of the demographic questions I asked and how they may reflect on my results, an account of the general idea behind the format of the experiment analogies with more detail on each individual experiment, how I found appropriate metaphors to use and how I expected the Users Comments to help in my analysis.

3.2 Previous Experiments

In determining each different type\(^1\) of metaphor’s ease of understanding the question of awareness has to be dealt with. It is not possible to come to any conclusions from a comparative analysis of a metaphor’s comprehensibility unless one can be sure the participant is aware that what they are presented with is figurative. In McGillion’s work on measuring awareness in metaphor, her intuition was that most people are unaware of metaphors and she found that people are divided into those who are aware and those who are merely sensitive to metaphor (McGillion, 2006). I used the same format for my experimentation as she did, but I interpreted the results I obtained in a different way, as explained below (see section 3.3.3).

In a previous experiment on the subject of negative metaphor (Sheehan, 2005) participants were asked to how long it took them to understand a sentence and whether they would correct a foreign speaker of English were they to use the sentence in question. Sheehan chose this second question as preferred to asking whether they found the sentence grammatical. In this way, the participant was not specifically told what was being tested. In psycholinguistic studies we do not want the participants to know what is being studied as this will affect the results. If participants know what is being tested they answer in a way they think they should given this knowledge, rather than how they would normally answer if they were unaware. While this was a good start, I decided to use a more covert method in my testing as explained below (section about analogy).

3.3 Layout of Experiments

The layout of the experiments is described here. It was partly motivated by the Experiment Tool used to create the experiments (see section 4.4)

3.3.1 Instructions

\(^1\) Type of metaphor here refers to whether the metaphor had either negative or positive polarity.
The participants are first presented with a short list of instructions (see Appendix A). These instructions explain that they are going to be first asked to answer some questions about themselves (see section 3.3.2) and then they will be asked to complete a series of analogies. The term ‘analogy’ is explained with a simple example and the participant is asked to choose whichever term they think fits best stressing that there are no right or wrong answers. The user is warned not to use the ‘Back’ button on their browser (see section 4.3.2), complete all questions and not change their answers. This is to ensure that this choice was their first instinct so I can see that this is how they would normally answer the question.

While analysing the results for the first experiment a problem became apparent with the data obtained from the web tools inbuilt timed response, which times how long each participant spends on each question. The data obtained for individual’s response times was corrupted, with standard deviations being far greater than the means in some places so no meaningful conclusions could be drawn from this (see section 5.6.2). For this reason, included in the instructions for the second experiment was an instruction to spend no longer than 15 minutes completing the experiment and not to use any other browser windows while participating. This was to encourage participants to focus on the task at hand and not browse around other web pages or locations on their computer, which seemed to be the reason for the strange data in the response times.

In addition to these instructions a one-line instruction was included at the top of every slide in the analogies section of each experiment just to make sure the participants understood what was being asked of them.

3.3.2 Demographics, Section 1

While the main aim of this study is to find out more about people’s ease of understanding of metaphor in general, it is important in psycholinguistic studies to see whether there are certain factors which correlate to better or worse performance by participants with a view to finding out if these factors are consistent with the performance of the whole population. In this case it is interesting to see if any factors pertaining to individual participants affect their ease of understanding metaphor. Understanding of metaphor as mentioned above (see section 2.2) is thought to give insight into the level of people’s cognitive skills, as do some of the factors I am including in the demographics section of my experiment so it will be interesting to see what correlations can be found here.

In the first section of my experiments the participants were asked to answer five questions about themselves.

Sex

The first question participants are asked is: “Are you male or female?" There is no study that proves there is a direct correlation between gender and cognition.

---

1 The 15 minutes completion time was decided by summing the mean answer time for each section and rounding it off to the nearest 5-minute mark.

2 In the first experiment the users were not asked what their mother tongue was. I thought I could control this by only sending the experiment link to people whose first language I knew was English but as the link was passed on to others I realised I could not control this factor and it needed to be included in the demographic questions section in the subsequent experiment.
While some studies show there are differences for example in the areas of spatial reasoning with men performing better on average and women performing better in some areas of language skills, these differences are generally negligible and could be explained by a range of other factors at play. It will be interesting to see if my results give any insight into this domain.

**Age**

The second question participants are asked is which age bracket they fall into. The brackets are ‘under 18, 18-24, 25-34, 35-44, 45-60 and over 60’. Language change is a phenomenon that is easily visible between different age groups with children’s language use varying significantly to adults. Different generations speak differently; language has its own age graded patterns. Generally one would imagine that as people age their cognitive skills improve but I feel that there is a cut off point where people plateau or even begin to decline. Getting a good number of participants in each of these age brackets will enable me to show a correlation between age and cognition and see if my instincts are correct.

**Handedness**

The participant is asked whether they are right-handed, left-handed or mixed. There is a lot of interest in this area as it is known that different sides of the human brain control different areas of cognition. Most right-handed people process language with the left side of their brain while left-handed people process language with a mix of left and right, sometimes just one or the other but mainly both. So handedness seems to give us an indirect insight into the workings of the brain and the organization of language in the brain. For example, it is common in the literature from many studies, e.g. (Geschwind & Behan 1982), that there is a correlation between language difficulties, such as dyslexia, and left-handedness. This field is a constantly changing one though, with studies always arising to support each viewpoint and discredit others. It is thought by most, however, that a difference in handedness correlates to a difference in language processing. Left-handed people are thought to be more accepting of weirdness than right handed people, which Schuetze documents in his discussion of grammaticality judgements (Schuetze, 1996). It will be interesting to see if I will find significant differences in the results which support the above theory although I expect it will be difficult to be able to infer much from my sampling. Though official information sources, for example the U.S Census, do not keep track of this kind of information, the figure of lefthanders in the population is generally given at 11-15% so I do not expect to have a large number of left handed participants.

---

1. Gender is a word that can have different meanings for different people. For many their gender identity is a private and complex part of who they are. Although I am of the opinion that gender is a fluid spectrum and not just a two-sided coin clearly demarcated with ‘male’ on one side and ‘female’ on the other, I only included these two options in my experiments. This was to keep the study similar to questionnaires people would be used to filling out and also I thought it might confuse some people to have ‘other’ as an option here, or list out alternatives.

2. When analysing my results, I will count mixed-handedness as instances of left-handedness. ?WHY?

First Language

This question was not included in my first experiment but in subsequent experiments I felt it was necessary to include\(^2\). There is always confusion over which term to use to describe what is meant by first language, eg. native language, mother tongue. There are problems with all these terms\(^1\). What I mean by first language in this study is the language that the participant uses most in their lives and feels most proficient in. If they feel more proficient in a language that they use less than some other language some kind of adjusted scoring of each language they use has to be made judged by where it scores on these two scales. In the case of true bilingualism, I expect participants to enter both languages. I do not feel that any participants will have trouble answering this question, it is more often a case of linguists having trouble in choosing a preferred term and I think that most participants will answer the same regardless what term is used.

The main reason for asking this question is that I only want to include people whose first language is English. Learners of English would be at varying stages of their proficiency and this may lead to noise, which I could not attribute to the affects of negation on metaphor. Also sentences which I few as literal may be viewed as metaphorical in other languages so it was best to avoid speakers of other languages.

Linguistic background

The last question participants are asked is if they have any background in linguistics and to give details of this background if they do. This is an important question to include as responses from participants with a background in linguistics may affect the overall results from the experiment. As Linguists are trained to look closely at language, they are more likely to guess at what is being tested or to be quicker to actually discover what is being tested than others. This may affect how they answer the questions, trying to get them right by having an idea what the study is about rather than answering as they would normally. This question and the comment section (see section 3.3.4) will help me to normalise the results should it occur that the linguists among the group of participants score markedly better than others. I expect that the highest scores in my experiments will be from the linguists but not necessarily that all the linguists will have perfect scores. This does not mean I will disregard the data obtained from linguists, but more that I will have some explanation should results go this way. It will be interesting to see if the linguists do perform better than others.

3.3.3 The Analogies, Section 2

If participants are signalled that the experiments are about metaphors nothing could be deduced about their everyday awareness and ease of comprehension of them. The method used was one that worked well in a previous experiment on the awareness of metaphor (McGillion, 2006). In this study, analogy is used much in the same way as in McGillion’s. The idea is to have some combination of source and target in a guiding analogy, and ask participants to choose between four targets, which are literal and metaphorical, given the literal or metaphorical source being questioned. If the

\(^1\)For example, the participant may not speak the same language as his/her mother and so would not like to use the term mother tongue to describe what I mean by first language.
participant matches the correct target to source following from the guide then they are aware of metaphor and the ease of understanding of each type of metaphor can be gauged from the overall number of correct responses for that metaphor. The example below illustrates this:

The first two sentences about Sharon are the guiding analogy. In this example the matching in the question should be metaphorical source to literal target because the guide is metaphorical source to literal target (“straight shooter” to “tells the truth”). The four choices of are always the same: a literal equivalent, a metaphorical equivalent, a literal opposite and a metaphorical opposite. In the experiment these, as well as the questions themselves, were presented in random order so as to ensure participants didn’t see a pattern. If the order is fixed it may be easier for the participant to deduce what the experiment is about.

<table>
<thead>
<tr>
<th>Sharon is a straight shooter</th>
<th>Guiding source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is to</td>
<td>Guiding target</td>
</tr>
<tr>
<td>Sharon tells the truth</td>
<td></td>
</tr>
<tr>
<td>As</td>
<td></td>
</tr>
</tbody>
</table>

*Bob is a beast*

<table>
<thead>
<tr>
<th>Is to</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob is a dog.</td>
<td>Metaphorical equivalent</td>
<td></td>
</tr>
<tr>
<td>Bob is ugly.</td>
<td>Literal equivalent</td>
<td></td>
</tr>
<tr>
<td>Bob is handsome.</td>
<td>Literal opposite</td>
<td></td>
</tr>
<tr>
<td>Bob is an Adonis.</td>
<td>Metaphorical opposite</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.1. Example of format of analogy for all experiments.

In the example shown in Figure 1, the correct answer is 2, the literally equivalent answer. If participants choose the correct match a summing of these matches over all participants would allow a judgement to be made from results about whether certain types of metaphors are easier to understand than others.

As I mentioned above, this format was used by McGillion in her experimentation on this subject. It may seem on looking at her work that I am merely repeating the same experiments. This is not the case. I am building on her work using the same methods she used as they seemed to be the best way to me to carry out my experiments in a covert way. I wanted also to be more rigorous in my experimentation by trying to include all possible combinations in my experimentation to give a more complete set of results (see discussion of experiments below).

Creating the questions

I thought the best source from which to build my questions would be The Irish Times Newspaper. I read the paper and tried to find naturally occurring metaphors that would be suitable as sources for my study. I then tried to approximate the matching and the opposite meaning of the source as closely as possible with two literal and two metaphorical target statements. To do this, I would use a dictionary and a thesaurus. I did not always keep the exact wording of the text I found in the
newspaper. I left out names and specifics that might limit the usability of the experiment due to essential knowledge of local events required in understanding the context. Often, if I found a metaphor in the newspaper and I required the source to be literal, I tried to turn the metaphor into a literal statement and used that for the source and kept the metaphor to use as a target in that question. For the guiding analogy, I constructed a simple source to target combination using some idea related to the context chunk of text. For example in Experiment 1, when my context was as follows: “**I was not involved in the meeting**, in the terms or anything else. I was simply asked to go along with him to the office,” I used the following guiding analogy as the context contained references to the business world:

“He is a successful business man is to he is a business giant.”

Finding metaphors with the same and opposite meaning was that hardest part of making the experiments. It was of the utmost importance to me to make the meanings of targets match as closely as possible to their literal counterparts. Having for example a literal target that fit better than the metaphorical target when the guiding analogy indicated a metaphorical target would severely affect my results. It is very difficult to find two metaphors with the same meaning and judging whether something can be viewed as a metaphor or not can be tricky. Idioms and hyperbole were avoided as much as possible as they are not to be confused with metaphor in the sense of this study.

**Experiment 1**

The questions asked in Experiment 1 and 2 are in Appendix B.

In the first experiment, participants are tested on established metaphors with mixed polarity. I felt it was very important to include all possible combinations of negative and positive with literal and metaphorical sources and targets to have a more complete set of results than obtained in previous experiments on this subject (Sheehan, 2005; McGillion, 2006). There were 16 questions overall covering the following combinations: Metaphorical positive (source) to Literal positive (target), Metaphorical positive to Literal Negative, Metaphorical negative to Literal positive, Metaphorical negative to Literal Negative and so on for Metaphorical to Metaphorical, Literal to Metaphorical and Metaphorical to Metaphorical (see appendix for full text). For questions where both the source and the correct choice of target are negative, both the metaphorical and literal equivalent in the choice of targets contains negation. This is so that the participant will not choose their answer by picking the only negative target from the selection available. In questions where only the desired target is negative, there is only one negated answer in the selection of targets.

In Experiment 1, the correct answer in the analogy is always of an equivalent meaning as in the example in fig.1 above, and never one of an opposite meaning. This is in keeping with McGillion’s experiment. The reason for only matching equivalents at this stage was two-fold. Firstly I wanted to give myself an idea of what was involved in creating the experiment and seeing what kind of results I would obtain before deciding how to incorporate opposite matching into the project. Secondly, including opposite matching at this stage with all the possible combinations of source, target and polarity I thought necessary would lead to a large number of questions.
This could lead to participants getting tired and not giving as much attention to the questions as they would given a shorter number of questions, giving up halfway through, or not passing the experiment on to others which I requested in my email asking for participants (see section 4.4). Leaving it out at this early stage enabled me to create a relatively short experiment that was easier for participants to complete, and therefore they would be more likely to give more attention to the task and in turn give me a good set of results. Trying to include all possibilities of making the sources of the analogies map to the semantically equivalent AND opposite targets, while also having all combinations of the sources and targets being literal or metaphorical and positive or negative, would have doubled the size of the experiment. In this way, Experiment 1, while still important and included in the overall results of the study, could be viewed as a pilot experiment.

This may seem to make the inclusion of opposites in this experiment unnecessary but they play a vital role in two areas. Firstly, the inclusion of opposites allows me to assess the sanity of participants. If participants persistently choose the target that has the correct meaning semantically but not necessarily the strictly correct target in terms of metaphorical or literal, then I know that they understand how the analogy works and the meaning of the source and target. Secondly, if participants choose opposites for questions that involve negation, I may be able to say that this is an affect of the negation thus providing evidence to prove the aim of this study, that negation has an adverse affect on the understanding of metaphor.

Experiment 2

In Experiment 2 participants were again tested on established metaphor with mixed polarity but this time opposite matching was included. The experiment contained 16 questions with the following combinations: Metaphorical to Literal, Metaphorical to Literal Negative, Metaphorical to Literal Opposite and Metaphorical to Literal Negative Opposite, and so on for Metaphorical to Metaphorical, Literal to Metaphorical and Metaphorical to Metaphorical (see appendix for full text). This time the sources do not have mixed polarity as results showed source variation to have little significant impact on the results. In this experiment, negation is expected to affect ease of understanding most in the questions where both negation and opposite matching is included with more noise (i.e. choosing the wrong meaning) than correct answers occurring in these areas. In this experiment there are three forms of negation: explicit negation, oppositeness and metaphor. Questions which include a combination of any of these things are intuitively going to be hard to understand as they constitute a double negative.

3.3.4 User Comments, Section 3

Upon completion of the experiment, the participants are asked if they would like to leave a comment and if not to enter ‘no comment’ here. This may be useful as

---

1 This issue was dealt with in Experiment 2 as findings from Experiments 1 showed that the choice of source (metaphorical/literal and positive/negative) was not significant (see section Results 1) which effectively cut the experiment in half.

2 The entry of ‘no comment’ here was because the experiment tool would not let the participant submit their answers unless something was typed in this box.
participants may comment that they knew what the experiment was about. People may comment that they found the experiment difficult or did not understand what they were supposed to do. They may comment on whether they found the experiment boring or not. This feedback helps me make judgements about their results, for example if someone scores very well and they knew what the experiment was about I know that this will affect the generality of the results.

3.4 Conclusion

This chapter detailed the methodology of the experiment. I began by summarising some previous work in the area and justifying my choice for creating the experiments in the way I did. I gave a detailed layout of what the experiments would look like. Firstly, I explained the instructions the participant would see when participating and I explained why I included the instructions I did here. I detailed all the categorical questions that form the demographic section by given information on what correlations are thought to be present between these and language and cognitive skills and suggesting how my experiments may lend proof to these theories. I explained the format of the analogies in the main body of each experiment. Then, I went through each in detail saying what results I expected. Finally I explained the user comments section of the experiment, and how this would be beneficial to judging the coveryness of this study.
Chapter 4

Experimentation
4.1 Introduction

In this chapter I will explain the process of experimentation. I will begin by giving my reason for conducting the experiment online. Then I will give an in depth explanation of the web tool used to create the experiments. I will list some problems I encountered using the web tool and finally I will give details of how I obtained participants.

4.2 Why Online?

The experiments were conducted using an online experiment tool. I chose to conduct this experiment online for many reasons. I thought it was the best way to reach as many people as quickly as possible as the majority of people have regular access to the Internet. The participants could complete the experiment from anywhere they had access to the internet and did not have to set aside time to meet for a face to face experiment which would have been more time consuming and probably resulted in a fewer number of respondents. Their anonymity was assured so would not feel self conscious about their performance which would affect how they answered. A further issue I had with a face-to-face approach was the affect the interviewer can have on participant’s performance. As documented by Holmes (2001) in An Introduction to Sociolinguistics, characteristics of the interviewer's language use and even their gender can affect how the participants perform in face-to-face interviews. While her explanation of this phenomenon is specific the study of sociolinguistics¹, I feel it has some resonance for all face-to-face interviews. Holmes (2001) states that the majority of interviewers “are middle-class, well-educated academics” (p.162) and interviewees tend to change their speech to match that of their interviewer when they are trying to cooperate. She says women will do this more than men as women are “more cooperative conversationalists than men” (p.162). Men on the other hand might do the opposite, purposely making their speech even less like the interviewers using more colloquial expressions as a reaction against the speech of a middle-class academic from the university” (p.162). The gender of the interviewer also affects interviewees’ language behaviour. Holmes writes that as most interviewers in the studies she is referencing are males, “it is likely that the interview context would be considerably more comfortable for men than for women”(p.163). This will cause them to use a less formal register with the interviewer and so affected the results of the particular studies. While this is obviously in the context of a different field of linguistics this shows that the interviewer can skew the results of the experiment and this is of course undesirable. For these reasons I felt an online experiment was the best way to conduct my study.

¹ In particular Holmes was writing about the affects observed in interviews that aim to find particularities common to a social dialect and explain the differences in language use between men and women.
4.3 Web-Based Experiment Tool

To create my experiments I used web based experiment tool that was created as a final year project by former CSLL student Sarah Kenny (Kenny, 1998). This tool was built upon in subsequent final year projects\(^1\). I found this preferable to using one of the variety of other survey tools available online due to the fact that it is free and also that it was built specifically for the purpose of computational linguistics experimentation and has been used in many studies in the CSLL department before so it is nice to carry on the tradition within the course of using a tool built for us.

4.3.1 Building Experiments

Logging into the tool presents you with your personal homepage (shown in fig 1), which shows you what experiments are currently active. Selecting these with the mouse shows the following information specific to each one: the number of participants, the username and password, the URL for participating, the URL for previewing and the description you provided of the experiment. At the bottom of this page there are six buttons used for creation, editing and analysis. These are: ‘Preview Experiment’, ‘Analyse Experiment’, ‘Create Copy’, ‘Edit Experiment’, ‘Delete Experiment’ and ‘Create Experiment’. Before an experiment has been created only the ‘Create Experiment’ button is clickable.

![Experiment System](http://web.cs.tcd.ie/Experimenter/web/fig4.1/fig4.1.html)

**Figure 4.1: Screenshot, my experiments home page**

\(^1\) (McGowan, 1999; Guennouni, 2000; Ryan, 2001; Hourihane, 2002; Graham, 2005)
When the Create Experiment button is clicked, you are brought to a new page where you must choose a title for the experiment\(^1\), which is included in the URL for previewing and participating. You must also choose a username and password that participants will need to access the experiment. Then you have the option of entering a brief description of the experiment, which is very useful if you have a number of similar experiments and need to keep track of the differences in them. The final text box is for entering instructions that the participants will see after they enter the username and password (see section instructions). There is also a checkbox you can tick to display the sections of the experiment in a random order. When this is completed, clicking the ‘Continue’ button at the bottom of the page brings you to a new page where you begin to create the body of the experiment.

This page is labelled ‘Section 1, Slide 1’. It contains seven buttons: ‘Add Passage’, ‘Add Graphic’, ‘Add Question’, ‘Add Answer’\(^2\), ‘Next Slide’, ‘Next Section’, and ‘End Experiment’. There is also a checkbox you can tick on this (and all other slides) to make your questions appear in a random order\(^3\). To explain how these buttons work, I will briefly go through some of how I created my experiments. I

---

\(^1\) The name of the experiment should have as little to do with the area of the study as possible. This is to ensure participants do not know what is being tested. This was a reason why I gave my experiments the names of foods I liked, the other reason being I was hungry at the time of creating them.

\(^2\) The ‘Add Answer’ button is only clickable after you have added a question with the ‘Add Question’ button.

\(^3\) The box to tick for displaying the questions on the slide in a random order cannot be used until more than one question has been added.
clicked the ‘Add Passage’ button, which placed an editable text box on the slide that I used to add either brief instructions or a heading. Then I clicked the ‘Add Question’ button, which placed another editable text box below the first one in which I entered a question, e.g. ‘What is your first language?’ When looking at the complete experiment, participants will see this as a box with the question in it followed by a text box for them to enter their answer. When I wanted them to choose from a set selection of answers, a choice of targets in the Analogies section for example, then as before I created a question with the ‘Add Question’ button and enter a question, e.g. ‘Are you male or female?’, and then click the ‘Add Answer’ button the same number of times as answers I want them to be able to choose from. This creates text boxes below the question box for me to type the choice of answers in. When participating in the experiment, these will appear underneath the question as the text of the answer preceded by radio buttons which the user can click. You have the option of allowing the users to enter multiple answers for the same question (by ticking the ‘Allow Multiple Answers’ checkbox which appears after you have added more than one answer) and randomising the order in which the answers appear to the participants (by ticking the ‘Answers Randomised’ checkbox). For my experiments I did not allow multiple answers, as I wanted participants to choose one answer among the options that they thought fit the analogy the best. I did however choose to have the answers appear in a random order to keep the experiment as covert as possible. There was no need to randomise the order in which the questions appeared to participants as the only section of the experiment that had more than one question was the Demographics section and there was no benefit in randomising the order of these questions. A further note on the questions is that during the experiment, the participant cannot proceed to the next slide without answering all the questions on the current slide, which is a further good point about the web tool.

When all the questions and answers were added to one slide, I clicked the ‘Next Slide’ button and repeated the process. When all the slides in one section were complete (the experiment was spit up into three sections, Demographics, Analogies and User comment which are detailed above see method sections 1? 2? 3?) I clicked on the ‘Next Section’ button. If the section I had just completed contained more than one slide, this brought up a page giving me the option to display the slides of that section in a random order by ticking a checkbox. I randomised the order of the slides in the Analogies section. Randomising is important here as I inputted the analogies in a certain order e.g. all four metaphorical to literal questions followed by all four literal to literal questions etc, so that it would be easier to work with the results. The order might have an affect on how participants answer the questions, e.g. questions at the end being answered worse due to less attention paid, or it might be easier to figure out what the experiment is about if they notice the pattern of the questions. After the final section was added, the Comments Section, where participants could leave a comment if they wished, I clicked on the ‘End Experiment’ button. This brought up a text box where I could put a closing comment the participants would see when they had finished the experiment. This was a note of thanks for participating. You can also give the participant the option of not submitting their answers at the end of the experiment by ticking the appropriate checkbox. I did not choose to do this as it could have led to people not submitting their answers and then going back and participating at a later time having thought more about the analogies when what I wanted was their first

---

1 The text boxes placed on the screen after clicking the ‘Add Passage’, ‘Add Question’ and ‘Add Answer’ buttons are only editable by the experimenter when creating or editing the experiment, when the experiment is being participated in, this text is fixed and participants cannot edit it.
reactions. Then I clicked the ‘End Experiment’ button on this page and was given the
option to save before exiting and being brought back to the home page.
When experiments are completed they are editable from the home page by clicking
the ‘Edit Experiment’ Button. This brings up the slides you have created with buttons
and checkboxes much like those used when creating the experiment (see figure edit).
You can edit most things about the experiment, including all the text and whether or
not things appear in a random order, and you can add and delete passages, questions
and answers. When adding these things you are presented with a drop down menu to
specify the position you wish to add them, e.g. after question 2, after question 3 etc.
When you click on the ‘Exit’ button on the top right hand corner of the screen you are
given the option of saving or disregarding any changes you made to the experiment.
From the home page you can also delete, copy or analyse the results of an experiment.
Creating a copy is a useful tool as both my experiments are in the same format and
have the same number of slides so all I needed to do was edit the text.

4.3.2 Problems

The tool is being steadily improved each time it is worked upon and has many
good qualities. It is user friendly, secure enough for the purposes of this study,
experiments created are easily editable for the most part, and the analysis tools were
perfect for my needs (see section 5.3). However, there are some problems with the
tool.

The biggest problem I had with the tool is that it does not work on some
computers. This could be due to the browser being used not having java enabled or
the security settings being too high but it is not a problem I was able to fix on any
computer that it did not work on. One of the main reasons I used the web tool was to
reach as many respondents as possible so it was very annoying when so many people
contacted me saying they could not get the experiment to work. It worked on Trinity
College computers and many friends’ personal computers but did not work in some
other college campus computer rooms. This was the only really major problem I had
with the tool that I felt affected my experimentation process. The other problems I had
with it were just awkward and time consuming for me to get around.

Another problem I had was that, once you have clicked ‘Next Slide’ in the
creation stage, you cannot delete this slide. This was extremely annoying at first when
I was learning how to use the tool and wasn’t sure what my layout should be like. If I
had added too many or too few slides by mistake I had to go back to the beginning
and create a new experiment, which was very time consuming.

Furthermore, there was no text formatting tools and some symbols in my text
appeared as question marks in the finished experiments if I had used copy and paste
from Microsoft® Word 2000 to input my text. This was easily solved by deleting the
character and inputting it in the textbox instead of pasting it. The formatting tools
were not essential to the working of the experiment and while I had initial wanted to
highlight the source sentence of my analogy in the context section by putting it in
italics, I got around this problem in the same way as McGillion by marking the text
with double asterisks instead.

Finally, using the back button in the browser did not take you to the previous
slide but to the web page you were viewing before logging into the experiment system
which effectively cancels your answers if you were participating or makes you lose
what you have done if you were creating an experiment. This is annoying at first but
once you have gotten used to it, it has little impact on creating the experiments.
Participants are warned of this in the instructions at the beginning (see section 3.3.1) and I did not receive any complaints about this feature so I presume it did not hinder any participants and so did not affect my experimentation.

4.4 Participants

To gain participants for my experiments, I composed an email, which explained that I was conducting an experiment for my final year project, and included the URL, username and password of the experiment. I asked the recipients to forward the email on to anyone who they thought would be interested in participating and encourage any questions to be emailed to me. I also encouraged anyone who wished to learn more about the experiment and its results to contact me and I would let them know once the study was finished. I felt it was important to give people this option as I am always interested in finding out what an experiment was for after I have participated. For the full text of this email, see Appendix C.

I then sent this email to my Course Director who forwarded it to all the years of the CSLL degree course and postgraduate students. I also sent this email to a list of friends and family. To reach more people, I posted the email on Bebo.com, which is “a social networking website, designed to allow friends to communicate in various ways.” (Wikipedia, 2007). On this site, I am connected to many friends and acquaintances. Participants can easily reply to this posting a to me personally or to the whole group, as can I, which is useful when trying to solve problems participants may have. It is easy to send an email to everyone you are connected with so I found it a very useful tool in gaining respondents.

4.5 Conclusion

In this chapter I detailed the experimentation process. Firstly, I gave reasons why I chose to conduct the experiment online. Then I gave a very detailed account of the web-tool and how I used it to create my experiments. I detailed some problems I had with the web-tool when I was creating my experiments and finally I described how I contacted my participants.

---

1 For my second experiment the email was also sent to Dr Dianna Laurent of Southeastern Louisiana University who offered her students extra credit of ten bonus points upon participation of the experiment.
Chapter 5

Results and Discussion
5.1 Introduction

In this chapter, I will detail the results of my experiments and discuss what these results mean with regard to my overall aim. I will go into detail about the results from each experiment separately, first showing the results of the demographic questions and how these factors had an effect, if any, on the results, then a more detailed description of how the respondents fared overall and how that data is meaningful. The aim of this chapter is to show that the theory I am working towards is correct, namely that negation hinders metaphor.

5.2 Respondents

The overall number of respondents for the two experiments was 53, 24 for the first experiment and 29 for the second. I was very pleased with this response rate. In the first experiment, it took some time to get participants and these were mainly my family, friends and people from course. For the second experiment, progress was much quicker, taking only a few days to reach the final number, and mainly consisted of students from the university of Louisiana. The reason for this is no doubt the extra credit offered for completing the experiment. The URL was distributed as explained above (section participants).

5.3 WebTool Analysis

The webtool, which I used to create both experiments (see section 4.3.1), came with analysis features. The results could be viewed by participant, by question or as a summary of all participants answers to all questions. It also gave the option of viewing a Response Times Summary, which measured, in seconds, the time taken by participants to respond to each question. These two summaries were formatted in such a way\(^1\) that they could be opened in Microsoft® Excel as a workbook in order by question. Excel has many attractive features for data manipulation and graph creation\(^2\) and many built in functions for statistical analysis. All the analysis was carried out using this tool.

5.4 Chi-square Analysis

Chi-square was used to test for statistical significance in the results, as shown below. As Prof. Jeff Connor-Linton (2003) explains it is formally defined as a non-parametric test of statistical significance for bivariate tabular analysis. It is usually used to test whether two samples are different enough in some respect that one can generalise that the population these samples come from is also different in that respect. It lets you know the degree of confidence you can have in your hypothesis.

---
\(^1\) The structure of columns and rows were maintained by commas and new lines respectively which separated the data. Excel has the option to recognise these as delimeters and so everything stays in the right place.
\(^2\) All the graphs in this experiment were created using Microsoft® Excel 2000
I have used it many times in this study and will give the relevant figures obtained in the discussion below.

5.5 Strictly Correct, Rational and Error

I will explain some terms I will be using to discuss my results. When picking a target in the analogies, the Strictly Correct target is the correct choice of target by meaning and metaphoricity or literality. This is as opposed to the Rational target, which is correct in terms of meaning but is not correct in terms of being metaphorical or literal. An example of Rational would be if the Strictly Correct target was the metaphorical opposite and the participant chose the literal opposite. We can see that they are rational in that they understood the meaning match but not strictly correct as they failed to achieve the metaphor match. An Error answer occurs when the participant fails to choose the meaning match and metaphor or literal match. I will use the Strictly Correct/Rational distinction throughout this discussion of results and it becomes very meaningful in the discussion of Experiment 2. I will not mention much about Strictly Correct as opposed to Rational in my discussion of the first experiment except to say that if people were not Strictly Correct, they were Rational, this tells us that they are sane. Participants are judged as sane if they choose Rational or Strictly Correct answers instead of making errors. On average, only 2% of answers were true errors in this experiment. Distinguishing results by Rational and Strictly Correct becomes much more interesting in Experiment 2 where there were more errors. Participants are said to be sane when they answer either rationally or strictly correct.

5.6 Experiment 1

First I will show and discuss the results for the demographic section of this experiment. The results obtained from the web tool were processed and organised in such a way that it was possible to show if there were any correlations between membership of a specific category and a higher success rate, both overall and for particular targets.

5.6.1 Demographics

Sex

This experiment consisted of 10 males (42%) and 14 females (58%). While it may have been preferable to have equal numbers of each in this category, this is difficult given the anonymity aspect of the experiment and I feel that this is a reasonable spread. On average, men performed slightly better overall in choosing the Strictly Correct target. Men gave an average of 9.2 Strictly Correct answers (52%) and women gave an average of 8.64 (48%). This means being male or female did not have an effect on performance. This does not support the claim mentioned earlier (see section 3.3.2) that women are better at language tasks than men. Figure 5.1 shows a further breakdown of these results into Strictly Correct answers for each target.

The distribution is not Chi-square significant, so does not let us deduce anything about the larger population. Women show as very slightly better with metaphorical and literal negative targets and the biggest difference is for metaphorical negative target where women do worse, but these differences are very slight.
Figure 5.1. Chart showing the affect of sex on performance by individual target

**Age**

The spread of participants with respect to age was not as varied as it could have been as the chart in Figure 2 shows. There was only one participant in each of the categories ‘Under 18’, ‘45-60’ and ‘Over 60’. There were no participants in the ‘35-44’ category. The most participants were in the ‘18-25’ category, most students who the experiment was sent to and my friends being in that age group. As there was not an even enough spread of participants across categories, I felt this would not really help to show a correlation between age and ease of understanding. Figures 5.3 and 5.4 are included for completeness sake but not with a view to making any meaningful connection between age bracket and understanding.

Figure 5.2. Number of participants in each age category
Of all participants, 1 was left-handed (4%) and 3 were mixed (13%). Left-handed alone was quite lower than the figure of 11% of the population given above (see method handedness). Both left-handed and mixed were counted as one giving a figure of 17%, which is higher than the average is supposed to be. However, I feel this is quite representative of the population as there is a large difference between the number of right handers (83%) and others.

Handedness does not have a much of an effect on Strict Correctness. The figures here are right handed on average getting 9.1 correct answers (54%) and left and mixed handers getting 7.75 correct answers (46%). This is not a significant distribution. The differences separated into targets in Figure 5.5 are very slight and also not significant.
Figure 5.5. Average Strictly Correct answers by handedness and target.

**Linguistic Background**

There were 12 participants in each category, with and without linguistic background. This was ideal as it would show if my intuition that having linguistic training would cause participants to realise what was being questioned and perform better. Those with linguistic background did perform better with an average of 9.58 Strictly Correct answers, (56%) and those with no linguistic background with an average of 7.6 (44%). This difference is not very great and I expected it to be higher. An interesting point however was that, of the participants that scored above average (over 9 Strictly Correct responses), 71% (5 out of 7) of them had linguistic training. This shows that while not all linguists are better than the general population at the task, the majority of the people who are best at it are linguists.
5.6.2 Main Results

Now I will discuss the results over all participants, comparing scores for each choice of target and source using Chi-square analysis and explaining what the results mean for this study. The timed response data will also be analysed and explained.

The main result of the first experiment supports the hypothesis this study is trying to prove, that metaphor is harder to understand under the scope of negation. This was deduced from Table 5.1.

<table>
<thead>
<tr>
<th>TARGET</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>METAPHOR</td>
<td>57</td>
<td>34</td>
</tr>
<tr>
<td>LITERAL</td>
<td>51</td>
<td>72</td>
</tr>
</tbody>
</table>

Table 5.1. This table shows the number of Strictly Correct answers per target.

These figures were obtained by summing each participant’s total Strictly Correct answers for each target choice. These figures were analysed with Chi-square analysis and found to be significant. (Chi-square 9.38, p <= 0.01) As the figures show, negative metaphor is the worst answered question with positive metaphor being answered significantly better. While I expected positive literal to score the best, it was literal negative which scored the highest, followed by positive metaphor, positive literal and then negative metaphor. I expected the order to be literal positive, literal negative, metaphorical positive and metaphorical negative. But this aside, this distribution supports the claim that negation makes metaphor more difficult to understand.

The response times of each target in Table 5.1 above were summed and are shown in Table 5.2 below. The figures show a similarity in the negative column with the figures in Table 5.1. The difference between metaphor positive, negative and literal positive is not so great as it was in table 1 but this distribution is still significant (Chi-square 17.72, p<= 0.001). This suggests that participants took the longest to
answer the literal negative questions, which seemed to be the easiest from the data in table 1, and spent the least amount of time on metaphor negative, the hardest question. There seems to be a correlation then between scoring worse on Strictly Correct metaphorical negative targets and the less time taken. I expected more time to be spent on the harder questions and felt this would help support the hypothesis of the study by equating time spent with increasing difficulty of understanding. The reason these values do not seem to support my hypothesis comes from the timed response data. When means and standard deviations were taken of each participants total response times it was found that the standard deviation in a lot of cases was higher than the mean (e.g. one participant had a mean of 57 seconds and a standard deviation of 116 seconds, that is double the mean). This is very unusual and it became apparent from looking at the data that participants had taken long pauses while completing the experiment. As this was a web-based experiment, participants may have been looking at other web pages at the time or using different programs on their computer. The downside to conducting the experiment online is that you cannot control the participants’ attention to the task. For this reason, I do not think much can be deduced from the timed response data for this experiment. As explained in the instructions above (section method), this prompted me to add to the instructions for the second experiment in an effort to stop participants doing other things while participating in the experiment so that the timed response data could be meaningful.

<table>
<thead>
<tr>
<th>TARGET</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>METAPHOR</td>
<td>885.77</td>
<td>814.72</td>
</tr>
<tr>
<td>LITERAL</td>
<td>967.81</td>
<td>1170.01</td>
</tr>
</tbody>
</table>

Table 5.2. This table shows the total time taken in seconds to give a Strictly Correct answer by target.

Data was also tabulated in the same way as Table 5.1 for the different sources for all participants. The results of this are in Table 5.3. This distribution is not significant (Chi-square 0.48, p<=1). This means that the source has no effect on the successful choice of a target. This enabled me to keep Experiment 2 the same size as Experiment 1 by not including polarity variation on the source. This way I could keep Experiment 2 short and not worry about losing meaningful results by not varying the source as much.

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>METAPHOR</td>
<td>56</td>
<td>48</td>
</tr>
<tr>
<td>LITERAL</td>
<td>54</td>
<td>56</td>
</tr>
</tbody>
</table>

Table 5.3. This table shows the number of Strictly Correct answers by source.
5.6.3 User Comments

Most users entered ‘No Comment’ in the comment section at the end of the experiment. Some interesting comments were noted however. One participant commented that there were always two obviously incorrect answers and two very similar answers. This shows that this particular person was probably not aware of the difference between the similar answers, one being metaphorical and the other being literal. This also shows that this person was not aware of what was being tested which was my aim in creating the experiment. They scored an overall of 6 Strictly Correct answers (37.5%). On the other hand, another participant commented that they wondered how much the experiment related to metaphor versus literal meaning so it was obvious they knew what was being tested. This did not, however, cause them to score very high above average with 9 Strictly Correct answers (56.3%).

5.7 Experiment 2

As with Experiment 1, I will first show and discuss the results for the demographic section to see if any correlations exist between category membership and scoring.

5.7.1 Demographics

Sex

Again, there were more female participants than males; this time there were 19 females (66%) and 10 males (34%). It would have been preferred that this was more even but the results were considered nonetheless. Figure 5.7 charts strictly correct answers, Errors and Rational answers against each other for each sex. This shows that men fare worse than women at strict correctness, unlike in the previous experiment where they did slightly better. There is only slight difference in rationality between the two. Again this distribution is not significant so it does not let us generalise about either sexes performance. One thing to notice is that error is much higher in this experiment than in the previous one no doubt due to the inclusion of opposite matching.
The figures were broken down further into each of the target categories. This is what is shown in Figure 5.8. This graph shows greater differences between the two sexes in particular areas. The greatest of these differences is for the metaphorical negative target, though it seems women performed better when the target was metaphorical, literal and literal negative opposite. We cannot narrow this effect down to a single particular factor and say women are better at matching when negation or opposites or metaphor are involved for example as the data seems to be mixed. While women score better with opposite metaphor, for example, we cannot make a judgement about oppositeness because men perform better on both negative metaphor opposite and literal opposite.
Figure 5.8. Average Strictly correct answers by sex and target. ‘LN’ stands for literal negative, ‘LO’ stands for literal opposite, ‘LNO’ stands for literal negative opposite and so on for metaphor.

**Age**

The spread of participants by age was disappointing. The vast majority (90%) fell into the 18-24 bracket, which was also the bracket with the highest number in Experiment 1. There were only 2 participants in the over 60 bracket and 1 in the 25-34 bracket. The reasons for these results are given in the discussion of age in Experiment 1 (see section). Due to this very uneven category membership, I do not feel I can deduce anything meaningful by examining these results in the same way as I have above, as the sample does not reflect the population. At a glance the difference between results across categories is very small and would not be significant.

**Handedness**

In this experiment there were no left-handed participants, 3 mixed handed (10%) and 26 (90%) right handed. Figure 5.9 shows the scoring separated into handedness. This shows mixed handed participants performed slightly better on strict correctness and there was little difference between the two for rationality. Looking at the data, the only participant scoring no errors was mixed handed. Mixed and left handed people are supposed to be more accepting of weirdness so maybe this is why this particular participant scored so well on rationality. However, another mixed handed participant scored the most errors in the experiment, which is against the hypothesis. It is clear that these results do not lend much support to a hypothesis about handedness.
Figure 5.9. This figure charts strict correctness error and rationality by handedness.

Figure 5.10 shows the category broken down into target choice. We can see a big difference for the literal negative opposite target and the metaphorical opposite target and right handed participants performing slightly better on literal negative targets, and much better on metaphorical negative targets where there were no strictly correct answers by the mixed group. Again this may not be very indicative of mixed being particularly better at literal negative opposite targets and is probably due to the high performance of one individual. It would be interesting to see if this result was maintained if a more even spread of participants across categories could be observed.

Figure 5.10. Average Strictly correct answers by handedness and target.
First Language

All participants who took part in the study spoke English as their first language. As anticipated, there was no confusion from participants over the choice of term used to describe what I meant by first language. This was ideal, if there had been any other choices for first language, I would have had to leave them out of the study (see section 3.3.2)

Linguistic Background

Unfortunately, there were no participants with linguistic background in the second experiment so no further investigation could be made in this area. This was due to the fact that I had ‘used up’ all of my contacts with linguistic training in the first experiment.

5.7.2 Main Results

As in the above discussion of Experiment 1, I will now go through the analysis over all participants by target using Chi-square analysis on the numbers of answers in particular categories. In this experiment, the source was not tested for significance as results from Experiment 1 showed it to be insignificant. Following the same analysis performed in Experiment 1, Table 5.4 was created by summing each participant’s total Strictly Correct answers for each target choice.

<table>
<thead>
<tr>
<th>TARGET</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>METAPHOR</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>METAPHOR OPPOSITE</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>LITERAL</td>
<td>36</td>
<td>22</td>
</tr>
<tr>
<td>LITERAL OPPOSITE</td>
<td>19</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 5.4. This table shows the number of Strictly Correct answers per target.

Participants answer the least number of Strictly Correct times when presented with negative metaphor. While the scores are all lower for negated targets, there is not a great difference for metaphorical opposite targets. When we collapse this table further, not distinguishing between equivalent and opposite matching for a moment, we get Table 5.5.

---

1 Unfortunately, the same problems occurred with the timed response data in Experiment 2 as Experiment 1, so it was abandoned as a way to judge ease of understanding.
2 By equivalent in this discussion I mean having the same meaning as the Source.
This distribution is not significant (Chi-square 0.11, p <= 3.84). While negation affects choosing the correct answer for metaphorical targets, the affect is not enough to support the aim of this paper. The results were tabulated in the same way as above for Rational answers and Errors (see Appendix D). These distributions were also not significant.

Figure 5.11 gives us a graphical representation of all these results together. It is interesting to see that participants are more rational overall when not matching to opposites, this form of negation really seems to affect performance, note that it has the highest error rate. Literal positive matches were the easiest as expected. Participants were Strictly Correct over three times more often than they were rational or wrong. Explicitly negated metaphor scored lowest for strictly correct but it did not affect participants so much that they were not rational. Negative metaphor scored highest on rationality. In total, results for metaphorical targets were much worse than for literal targets as is expected.

<table>
<thead>
<tr>
<th>TARGET</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>METAPHOR</td>
<td>31</td>
<td>24</td>
</tr>
<tr>
<td>LITERAL</td>
<td>55</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 5.5. This table shows the number of Strictly Correct answers per target ignoring oppositeness.
Figure 5.11 This is a graph of the average number of answers by Rationality, Strictness and Error.

So far it is apparent in this sampling that all three kinds of negation have an effect on participants performance. But as the figures are not significant, the data needs to be looked at in a different way. Questions had to be asked about the three different kinds of negation present in this study. Different groupings were examined in order to answer these questions.

**Explicit Negation**

**Metaphor**

The first part of the first question asked about the data was, when the target is explicitly negated, is there more error when it is metaphorical? To calculate this, the results were tabulated and summed by looking at all the negated targets by rational, strict and error results.

<table>
<thead>
<tr>
<th></th>
<th>MN</th>
<th>MNO</th>
<th>LN</th>
<th>LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>32</td>
<td>13</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td>16</td>
<td>22</td>
<td>16</td>
</tr>
<tr>
<td>E</td>
<td>15</td>
<td>27</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>

Adding the first two columns together and the last two columns together gave the following table, in which we do not distinguish between opposite and same matching.

<table>
<thead>
<tr>
<th></th>
<th>MN&amp;MNO</th>
<th>LN&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>S</td>
<td>25</td>
<td>38</td>
</tr>
<tr>
<td>E</td>
<td>42</td>
<td>43</td>
</tr>
</tbody>
</table>

Chi-square = 5.74, for significance chi-square should be greater than or equal to 5.99.

We notice in the strict row that there is a difference comparing the two columns. Participants score worse on strict correctness for metaphor under the scope of explicit negation. This distribution is almost significant.

Collapsing the first two rows gives the following table.

<table>
<thead>
<tr>
<th></th>
<th>MN&amp;MNO</th>
<th>LN&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;S</td>
<td>69</td>
<td>69</td>
</tr>
<tr>
<td>E</td>
<td>43</td>
<td>43</td>
</tr>
</tbody>
</table>

This table treats Rationality as a type of correct answer. This table is obviously not significant. It tells us that under the scope of negation, participants are as likely to be sane whether the target is metaphorical or not. Collapsing the Error row with the Rational row gives the following table.
This distribution is significant. Here we are treating rationality as a type of error. This table shows us that, under the scope of negation, participants are worse at getting the answer strictly correct if its metaphorical than they are when it is literal. We can see that, while the scores are not particularly good for literal either, there is less difference between being strictly correct and getting its wrong or rational than there is in the metaphorical case. We can view negative metaphor as a kind of double negation, as metaphor is literally false. Intuitively double negatives are more difficult to process. This is the case here as we have shown that, under the scope of explicit negation, while people are still sane, they are less likely to get metaphor strictly correct and hence that metaphor is more difficult to understand, which is our goal.

**Opposite matching**

The next part of the first question is examining whether under the scope of negation opposite matching fares worse than equivalent matching. We use the same targets as above but this time we add the opposites together and the equivalents together. Collapsing this down we get the following:

<table>
<thead>
<tr>
<th></th>
<th>MN&amp;LN</th>
<th>MNO&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>54</td>
<td>22</td>
</tr>
<tr>
<td>S</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>E</td>
<td>27</td>
<td>58</td>
</tr>
</tbody>
</table>

Chi-square = 24.79, \( p \) is less than or equal to 0.001.

This distribution is significant and shows a strong effect of opposite matching on the data. This table shows that participants are more rational and make fewer errors when dealing with equivalent matching under the scope of negation than they are with opposite matching. This is expected, as opposite matching is a form of negation, which makes the right hand column a column of double negatives, which are intuitively more difficult to understand, hence the larger number for errors. Collapsing the table down again we get the following:

<table>
<thead>
<tr>
<th></th>
<th>MN&amp;LN</th>
<th>MNO&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;S</td>
<td>85</td>
<td>54</td>
</tr>
<tr>
<td>E</td>
<td>27</td>
<td>58</td>
</tr>
</tbody>
</table>

Chi-square = 18.2195514176894, \( p \) is less than or equal to 0.001.

This distribution is significant and the effect of opposite matching is still strong. This tells us participants are just as likely to get the wrong target as they are to be rational or correct when they are matching to opposites, but they are far more likely to be sane when they are matching equivalent targets than they are to be wrong.

Collapsing this table by combining rationality and error as we did above we get the following:
This distribution is clearly not significant. It tells us that people score the same for strict correctness under the scope of explicit negation whether they are matching opposites or not.

We can see in this question that, under explicit negation, opposite matching has a strong effect on how sanely people answer. This is a stronger result than was seen for the case of metaphor. Above we saw that metaphor just affected people’s performance on getting it strictly correct, while oppositeness here makes people answer much less sanely than equivalent matching. As we see above, people are as wrong as they are sane in the case of opposites, but their sanity greatly improves in the case of same matching. So, while there is a significant effect of both metaphor and oppositeness under explicit negation, the effect of oppositeness is greater. This suggests to me that oppositeness is a stronger form of negation than metaphor under the scope of explicit negation.

**Opposite Matching**

Metaphor

The second question about negation was when the target is of opposite meaning, firstly, what effect does metaphoricity have? As above, tables were created to determine the significance with Chi-square. I will display the tables and their significance below, and then discuss what they mean.

<table>
<thead>
<tr>
<th></th>
<th>MO&amp;MNO</th>
<th>LO&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>S</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>E</td>
<td>59</td>
<td>55</td>
</tr>
</tbody>
</table>

This distribution is not significant.

<table>
<thead>
<tr>
<th></th>
<th>MO&amp;MNO</th>
<th>LO&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;S</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>E</td>
<td>59</td>
<td>55</td>
</tr>
</tbody>
</table>

This distribution is not significant.

<table>
<thead>
<tr>
<th></th>
<th>MN&amp;LN</th>
<th>MNO&amp;LNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>R&amp;E</td>
<td>79</td>
<td>76</td>
</tr>
</tbody>
</table>

This distribution is not significant.
This distribution is not significant. These results show that under the scope of oppositeness, metaphor does not have a significant affect on performance. The results are similar between the two columns for all tables. This is in keeping with earlier results from which I proposed that metaphor does not have a strong effect at negating things.

Explicit Negation

Still under the scope of oppositeness, we will now see if explicit negation has an effect.

<table>
<thead>
<tr>
<th></th>
<th>MNO &amp; LNO</th>
<th>MO &amp; LO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>S</td>
<td>32</td>
<td>37</td>
</tr>
<tr>
<td>E</td>
<td>58</td>
<td>56</td>
</tr>
</tbody>
</table>

This distribution is not significant.

<table>
<thead>
<tr>
<th></th>
<th>MNO &amp; LNO</th>
<th>MO &amp; LO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;S</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td>E</td>
<td>58</td>
<td>56</td>
</tr>
</tbody>
</table>

This distribution is not significant.

<table>
<thead>
<tr>
<th></th>
<th>MNO &amp; LNO</th>
<th>MO &amp; LO</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>32</td>
<td>37</td>
</tr>
<tr>
<td>R&amp;E</td>
<td>80</td>
<td>75</td>
</tr>
</tbody>
</table>

This distribution is not significant.

These tables show us that, while the number of errors is higher for explicit negation under oppositeness, it is not significantly high to make any generalisations about the population.

What the two parts of this question indicate is that under the scope of oppositeness, metaphor and explicit negation do not have much of an effect.

Metaphor

Explicit Negation

The last question was when the target is metaphorical, firstly, what effect does explicit negation have?

<table>
<thead>
<tr>
<th></th>
<th>M&amp;MO</th>
<th>MN&amp;MNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>36</td>
<td>45</td>
</tr>
<tr>
<td>S</td>
<td>31</td>
<td>25</td>
</tr>
<tr>
<td>E</td>
<td>45</td>
<td>42</td>
</tr>
</tbody>
</table>

This distribution is not significant.
What these tables show is that under the scope of metaphor, explicit negation does not seem to have a significant effect, which is disappointing. We can see from the tables participants are more Strictly Correct when it comes to positive metaphor, but unfortunately this difference is not big enough to let us generalise about the population. I feel the data from the rest of Experiment 2 and Experiment 1 is enough to prove the aim of this study.

Opposite Matching

The second part of this question is what effect does oppositeness have under the scope of metaphor.

<table>
<thead>
<tr>
<th></th>
<th>MO&amp;MNO</th>
<th>M&amp;MN</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>20</td>
<td>61</td>
</tr>
<tr>
<td>S</td>
<td>33</td>
<td>23</td>
</tr>
<tr>
<td>E</td>
<td>59</td>
<td>28</td>
</tr>
</tbody>
</table>

Chi-square = 34.29, $p$ is less than or equal to 0.001.

<table>
<thead>
<tr>
<th></th>
<th>MO&amp;MNO</th>
<th>M&amp;MN</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;S</td>
<td>53</td>
<td>84</td>
</tr>
<tr>
<td>E</td>
<td>59</td>
<td>28</td>
</tr>
</tbody>
</table>

Chi-square = 17.98, $p$ is less than or equal to 0.001.

<table>
<thead>
<tr>
<th></th>
<th>MO&amp;MNO</th>
<th>M&amp;MN</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>33</td>
<td>23</td>
</tr>
<tr>
<td>R&amp;E</td>
<td>79</td>
<td>89</td>
</tr>
</tbody>
</table>

The distribution is not significant.

This set of tables gives us the most significant result for Chi-square yet. This shows that oppositeness has a huge effect under the scope of metaphor. From the first and second table we see that there is much more chance of being rational and sane than there is of making an error.

What I deduce from the results of this question is that again, oppositeness has the strongest effect on making people less sane. When choosing the target, negation made no difference to participant’s ability while oppositeness had a dramatic affect.
From these questions, we see that under the scope of explicit negation both oppositeness and metaphoricity had an effect on how well people answered. Under the scope of oppositeness neither explicit negation nor oppositeness had an effect. Final under the scope of metaphor, only oppositeness had an effect. While the aim of this project is supported by the results of this experiment (see section expli neg metaphor) and the results of the first experiment, it also seems shows that metaphor is a weaker form of negation than oppositeness. Perhaps this indicates something about the way in which people process metaphor, but that is work for further study. The results given here prove that negative metaphor is harder for people to get right.

5.7.3 User Comments

While again most participants left no comment, one participant echoed what another participant commented in the first experiment, that it seemed like two answers fit for every question and that they were not sure what the point of the experiment was. I am quite confident then that this experiment was successful in being covert.

5.8 Conclusion

In this chapter I presented my results and analysis and showed that the data proves the theory that under the scope of negation, metaphor is harder to understand. I dealt with each experiment separately. I detailed all the results for the demographic sections and discussed if any correlations between category membership and performance could be found there, both overall and then split into targets. Then I examined the main results of each experiment with a view to proving my goal. For the first experiment, I found a significant distribution summing the answers by target. This allowed me to generalise that negative metaphor is harder to get right which proved my aim. In the second experiment, there was no significance straight away. Questions were asked about the data with a view to finding the effects of metaphoricity, explicit negation and opposite matching under the scope of explicit negation, opposite matching and metaphoricity. These findings gave further strong results supporting the theory. Furthermore they lead me to conclude that metaphor is quite weak as a form of negation compared to explicit negation and oppositeness.
Chapter 6

Conclusion
6.1 Introduction

In this final chapter of my study I will first summarise the project. Then I will state what achievements I have made in this study. Finally, I will discuss some issues which occurred to me during the completion of this project which would benefit from future work.

6.2 Summary

The aim of this study was to prove that, under the scope of negation, established metaphor is more difficult to understand. This was a claim proposed by Vogel (2001). In my experimentation, I built upon previous work by Sheehan (2005) and McGillion (2006). I conducted two online experiments consisting of 16 analogies each. These were set up to include a complete list of combinations of source and target varied by metaphoricity, polarity and oppositeness. The process of creating the experiment was explained in detail. The results of the experiments were shown by tables and charts and they were discussed in detail. No correlation was found between any of the demographic categories and performance. In both experiments, results proved the aim of the experiment. Further, the data suggested metaphor is a weak form of negation, with oppositeness being the strongest and then explicit negation.

6.3 Achievements

I feel I added to the results found by Sheehan (2005) whose study also proved this hypothesis. I felt I added to this by creating more detailed experiments that encompassed all possible combinations that needed to be explored. I felt that my experiments were also more covert and so more telling about people’s responses. My results are very detailed and I feel I have analysed them very meticulously.

I feel that this study lends much to the domain of negative metaphor as there is very little experimental data out there to assist in making claims. Hopefully the data provided here will be useful in further experimentation of negation and metaphor. While I feel that my results and what they have shown are concrete and well reasoned, I accept that the implications of these results could possibly be challenged.

6.4 Future Work

The main issue for further work in this area is to deal with novel metaphor in a similar way as established metaphor was dealt with in this text to prove Vogel’s (2001) hypothesis, that novel metaphor under the scope of negation would be invalid. A complete study conducted in the same way as this one for novel metaphor in addition to this study would give very solid data to prove Vogel’s entire hypothesis. It would also be very interesting to see a study which includes the timed response data and finds if there is a correlation between performance and time spent. I had hoped
this would be a great aid to my proving my aim so I would be very interested to see if there is a correlation here.

Another issue for future work is the experiment tool. I detailed problems I had when using it in section 4.3.2. These problems seem very fixable and the tool would be much better without these problems.

Finally, any further work into the domain of negation and metaphor would be very interesting to see as there is so little written on the topic. Hopefully, studies like mine and Sheehan’s (2005) will encourage further work in this area.

6.5 Conclusion

The theory I aimed to prove was that metaphor was more difficult to understand under the scope of negation. I proved this theory by experimentation that gave very significant results. I further feel that my study has shown that oppositeness will effect people more than metaphoricity over all. I am satisfied that I proved what I set out to prove and also have shown metaphors weakness as a form of negation which was not my goal but seems to be apparent from the results. I look forward to any criticism or extension on this work.
Bibliography

Retrieved 30 April 2007 from

Ithaca: Cornell University Press.

http://www.georgetown.edu/faculty/ballc/webtools/web_chi_tut.html


http://www.etymonline.com/


Appendix A

Instructions for Experiment 1

Thank you for participating in this experiment. First, you will be asked some general questions about yourself. Then you will be given a number of analogies to consider. An example of an analogy is:

1
is to
2
as
A
is to ...

1.B
2.Q
3.C
4.Z

where most people would respond with answer number 1 since B is the second letter of the alphabet. The questions in the experiment will consist of a short piece of context followed by an analogy in the above format. Be aware that there are no right or wrong answers, simply choose which answer you think is best.

At the end of the experiment you will be given an opportunity to leave any comments you may have. If you have any problems with the experiment or wish to be informed of the results, you may email me at joycece@tcd.ie

Note: Please answer all questions. Please do not change your answers. Please do not use the back button on your browser during this experiment as this will result in exiting the experiment.
Instructions for Experiment 2

Thank you for participating in this experiment. Please spend no longer then 15 minutes doing the following:
First, you will be asked some general questions about yourself.
Then you will be given 16 analogies to consider. An example of an analogy is:

1 is to 2 as A is to which of the following?

Most people would respond with answer number 1 since B is the second letter of the alphabet. The questions in the experiment will consist of a short piece of context followed by an analogy in the above format. There are no right or wrong answers, simply choose which answer you think is best.
At the end of the experiment you will be given an opportunity to leave any comments you may have. If you have any problems with the experiment or wish to be informed of the results, you may email me at joycece@tcd.ie.

NOTE: Please do not use other windows while participating as this will affect the working of the experiment. Please answer all questions. Please do not change your answers. Please do not use the back button on your browser during this experiment as this will result in exiting the experiment.
Appendix B Experiment Questions

Experiment 1

A government official had organised the visit after hearing about one of the speaker’s presentations to a group of students in a local school. The speaker said she accepted that her approach was controversial, but said **it grabbed the attention of every student.**

Every student is a rare diamond
   Is to
Every student is special

As

**It grabbed the attention of every student**
   Is to

1. It interested every student and they listened intently.
2. It seized every student’s mind and they listened intently.
3. It bored every student and they talked amongst themselves.
4. It numbed every student’s mind and they talked amongst themselves.

A new website aimed at promoting responsible online practices among young people was launched yesterday by the Minister for Education, who said **“social networking websites are not taking the bull by the horns in verifying users’ ages.”**

The Internet is a wealth of knowledge.
   Is to
The Internet has a lot of information.

As

**Social networking websites are not taking the bull by the horns in verifying users’ ages**
   Is to

1. Social networking websites are being active about verifying users’ ages.
2. Social networking websites are doing little about verifying users’ ages.
3. Social networking websites are knuckling down on verifying users’ ages.
4. Social networking websites are sluggish in their approach to verifying users’ ages.
**Fashion experts were a skeleton crew** at city hall last night for the spring/summer show. Critics say prominent members of the group were boycotting the show because a lot of designers included fur clothing in their collections.

The fashion industry is a goldmine
Is to
The fashion industry is a profitable business
As

**Fashion experts were a skeleton crew.**
Is to
1. There weren’t many fashion experts present.
2. Fashion experts were out in force.
3. A lot of fashion experts were there.
4. Fashion experts were nearly invisible.

“The bull moved to Enfield in 2005 on the establishment of the National Cattle Breeding Centre. Although in his 15th year, **the veterinarians have not had him put down**, " said a spokesperson for the centre.

The farmer was in hot water.
Is to
The farmer had found himself in a difficult situation
As

**The veterinarians have not had him put down.**
Is to
1. The veterinarians have not ended his life.
2. The veterinarians have had him put to sleep.
3. The veterinarians have killed him.
4. The veterinarians have stopped him living out his golden years.
“**I don’t get very nervous during auditions**,” said one of stars of the film and new six part TV drama series.

Stress can be bad for you
Is to
Stress can cause you harm

As

**I don’t get very nervous during auditions**
Is to

1. I get anxious over auditions.
2. I get butterflies in my stomach at auditions.
3. I am relaxed when I have auditions.
4. I am keep a level head in auditions.

The man said he has had discussions with one of the builders over the development of the proposed hospital, but must have the support of the government for the project. He said **he was disappointed with the response of the minister** which he said was not based on knowledge of the project.

Governments are necessary
Is to
Governments are required

As

**He was disappointed with the response of the minister**
Is to

1. The ministers response was insufficient.
2. The ministers response fell short of the mark.
3. The ministers response was satisfactory.
4. The ministers remark was at par for the course.
He said **the election organisers made little effort to inform the public** about the work record of judicial candidates and so many people simply voted according to ethnic background.

I vote very often  
Is to  
I vote frequently

As

**The election organisers made little effort to inform the public**  
Is to

1. The election organisers did not make much of an effort.  
2. The election organisers made a Herculean effort.  
3. The election organisers made a substantial effort.  
4. The election organisers made a microscopic effort.

The Italian authorities plan to ban spectators from stadia not considered to have met required security standards, and this is among them. **They believe the required work cannot be completed** until the start of next season.

Pasta is included in most Italians diets  
Is to  
Many Italians eat pasta

As

**They believe the required work cannot be completed**  
Is to

1. They think the work will not be finished.  
2. They think the work will get done.  
3. They think the work will not be wrapped up.  
4. They think the work will reach maturity.
He is happy his side will succeed in the qualification match tonight. *He has told his side they will have an easy game* from the opposition who are failing to recover from a disappointing European campaign.

Teams work well together

Is to

Teams are well-oiled machines

As

*He has told his side they will have an easy game*

Is to

1. He has told his side to expect no brick walls.
2. He has told his side to expect a simple match.
3. He has told his side to expect a difficult game.
4. He has told his side to expect a kick in the tail.

Her game is not as good as it should be. But *there is still a chance she could play well in tournaments* this year and get herself back into the top world rankings.

Life involves many competitions

Is to

Life is a tournament

As

*There is still a chance she could play well in tournaments*

Is to

1. There is still a chance she could get some good tournaments under her belt.
2. There is still a chance she could perform successfully in tournaments.
3. There is still a chance she could get poor results in tournaments.
4. There is still a chance she could perform miserably in tournaments.
"I was not involved in the meeting", in the terms or anything else. I was simply asked to go along with him to the office.

He is a successful businessman
Is to
He is a business giant

As

"I was not involved in the meeting"
Is to

1. I played a part in the meeting.
2. I steered clear of the meeting.
3. I was associated with the meeting.
4. I avoided the meeting.

The prime minister said it was their duty to report back to the parliamentary party on representations from interest groups. He added: "They had no criticisms whatsoever of the Minister for Health, of me or the Minister for Finance. "Our stated position has been the same throughout the negotiations on this issue."

Negotiations were almost over.
Is to
The end of negotiations were in sight.

As

"They had no criticisms whatsoever of the Minister"
Is to

1. They tore into the minister
2. They didn’t attack the minister
3. They didn’t reproach the minister
4. They reprimanded the minister
The jockey overcame the conditions and **he beat his stable companion after a gruelling battle**. Not unnaturally, he was delighted.

The stables were stuffed to the gills

Is to

The stables were a sardine can

As

**He beat his stable companion after a gruelling battle**

Is to

1. He beat his stable companion after a harrowing duel.
2. He beat his stable companion after a difficult race.
3. He beat his stable companion after an effortless contest.
4. He beat his stable companion after a smooth sailing competition.

An actor and retired school principal was named person of the year yesterday at a ceremony in the town’s newly opened resort and spa. Some felt that the choice of winner no longer reflected the effort that nominees put into the community and that the whole process had turned into a popularity contest. A local businessman who was also nominated for the award was quoted as saying **he wasn’t over the moon** about the outcome.

The principal is a Rottweiler

Is to

The principal is a demon

As

**He wasn’t over the moon**

Is to

1. He was pleased as punch.
2. He was overjoyed.
3. He was upset.
4. He was down in the dumps.
**The 17-year-old amateur thundered out of the blocks** at the national golf championship to move into the top ten.

Life is a voyage
  Is to
Life is a journey
  As

**The 17 year old amateur thundered out of the blocks**
  Is to
  1. The 17-year-old amateur was slow off the mark.
  2. The 17-year-old amateur started off badly.
  3. The 17-year-old amateur began well.
  4. The 17-year-old amateur didn’t start asleep.

The manager has deferred naming his team to play in Saturday's second round of the National Football League. Injury concerns over players coupled with the involvement of several in other matches prompted his decision - and **it's clear he won't be gambling with any players** following their narrow round-one defeat.

The team is in hot water.
  Is to
The team is in a jam
  As

**It’s clear he won’t be gambling with any players.**
  Is to
  1. It’s obvious he won’t be putting players on the block
  2. It’s obvious he won’t be risking players.
  3. It’s obvious he will be careless with his players.
  4. It’s obvious he will be chancing players.
Experiment 2

He added that with the amount of criticism of mental health services recently, staff were becoming demoralised. “I think that what we need to do now is to get some sort of healing…**we need people putting their shoulders behind the wheel**.”

The criticisms made his blood boil
Is to
He was angered by the criticisms
As

**We need people putting their shoulders behind the wheel**
Is to:

- We need people to work at the task.
- We need people to avoid the task.
- We need people to marry the task.
- We need people to bury the task.

A prominent bishop said yesterday he had no difficulty with the State providing civil divorce if **the couple’s marriage had shattered**.

The sermon was lifeless
Is to
The sermon was boring
As

**The couple’s marriage had shattered**
is to

- The couple's marriage was not going well.
- The couple’s marriage was getting better.
- The couple's marriage hit an iceberg.
- The couple's marriage was on cloud nine.
A county councillor was found guilty of theft and fraud involving the misappropriation of over €7,000 of council funding. The judge said **he had lined his pockets with the public money** and that he owed the council the money they claimed he had taken.

Politicians blow smoke
Is to
Politicians speak honestly
As

**He had lined his pockets with the public money**
Is to:

He had unlawfully taken the money.
He had been given the money legally.
He had pinched the money.
He had been awarded the money for the sweat of his labour.

The four are serving sentences ranging from 11 to 14 years in connection with an attempted robbery which resulted in the paralysing of a policeman. **They are seeking an order killing the refusals of temporary release**.

He would rob you blind
Is to
He would give you everything
As

**They are seeking an order killing the refusals of temporary release**
Is to:

They are seeking an order that will make the refusals of temporary release invalid.
They are seeking an order that will not annul the refusals of temporary release.
They are seeking an order extinguishing the refusals of temporary release.
They are seeking an order rubber-stamping the refusals of temporary release.
**The runner recovered after 18 months of injury** to win the cross-country running title for the second time in three years yesterday.

Athletes are healthy
Is to
Athletes are fit
As

**The runner recovered after 18 months of injury**
Is to:

The runner recuperated after 18 months of injury.
The runner deteriorated after 18 months of injury.
The runner bounced back after 18 months of injury.
The runner faded after 18 months of injury.

Extraordinary allegations of foul play were made by the manager against the opposing team.**The allegations diminished the quality of the match**.

The allegations were false
Is to
The allegations were fallacious
As

**The allegations diminished the quality of the match**
Is to:

The allegations didn’t stop the match being ruined.
The allegations made the match better.
The allegations cast a shadow over the match.
The allegations healed the match.
They are understood to be the party most interested in signing the best player. If successful, it will see him playing with the squad for the tournament this summer.

Summer is my favourite season
Is to
I hate summer
As

They are understood to be the party most interested in signing the best player
Is to:

They are understood to be more interested than any other party.
They are understood to be the least interested party.
They are understood to be most willing to trade their souls to get the deal.
They are understood to be last in line in terms of being interested.

A local woman challenged her in no uncertain manner but the former silver medallist pulled clear and won by an extremely comfortable margin.

Cricket is not a popular sport
Is to
Cricket is well-liked by many
As

A local woman challenged her
Is to:

A local woman demanded a contest.
A local woman didn’t compete against her.
A local woman threw down the gauntlet to her.
A local woman threw in the towel.
The victim told the court how her life has changed. “I don’t go anywhere on my own. I should feel safe in the city I grew up in.**The attack has affected me very badly**.”

The city is dangerous at night
Is to
The city is a jungle at night
As

**The attack has affected me very badly**
Is to:

The attack made me very upset.
I was apathetic about the attack.
The attack knocked me for six.
I was thick-skinned about the attack.

Looking down the line to the World Cup, **this performance will have reassured other captains**, who may have been unsettled following Ireland’s great win over England.

Ireland won the match against England by many points
Is to
Ireland slew England
As

**This performance will have reassured other captains**
Is to:

This performance will have relieved other captains.
This performance will have worried other captains.
This performance will not have rattled other captains.
This performance will have thrown other captains.
The man was attempting to stop a fight between two of his sisters when the incident took place.

Sisters argue often
Is to
Sisters argue once in a blue moon
As

**The man was attempting to stop a fight**
Is to:

The man was attempting to end the argument.
The man was attempting to encourage the argument.
The man was attempting to douse the argument.
The man was attempting to fuel the argument.

A 55-year-old woman was claiming damages of €38,000 against a supermarket following an accident. **The woman left the Civil Court without being compensated** yesterday.

The supermarket was nearby
It so
The supermarket was a dot in the distance
As

**The woman left the Civil Court without being compensated**
Is to:

The woman left the Civil Court without being given any money.
The woman left the Civil Court fully recompensed.
The woman left the Civil Court in the same boat as she went in.
The woman didn’t leave the Civil Court with only sand in her wallet.
A new agency was appointed by the Government to advise on the situation. The first full report was published yesterday following a major study. However, **some said it was long-winded.**

The situation was savage
Is to
The situation was fiery
As

**Some said it was long-winded**
Is to:

It was very lengthy.
It was concise.
It was longer than the Nile.
It was boiled-down.

Workers at the factory have been called to a meeting this morning at which managers will disclose the outcome of a review of the operation. **A large number of jobs are likely to be binned,** but the company indicated that the plant will not close altogether.

The plant is an eyesore
Is to
The plant is a blemish
As

**A large number of jobs are likely to be binned**
Is to:

Many people will become unemployed.
Many people will get to keep their jobs.
Many people will no longer take home bacon
Many people will stay glued to the factory.
**The men disappeared into the ether** after they set out to find the Northwest Passage to America through the Arctic in 1845.

The minister said: “I am very disappointed with the guilty verdict. **He was an exemplary drone of our nation** performing many years of public service.”

**The men went missing.**
The men were suddenly present.
The men vanished into thin air.
The men precipitated into view.

**He was an exemplary drone of our nation**
Is to:

1. His work benefited our country
2. He made few contributions to our country.
3. He was a good slave to our country.
4. He didn’t lift a finger for our country.
Appendix C

Email to find participants

Hi,

I am in the final year of a course in Computational Linguistics and would very much appreciate your assistance in participating in an online experiment. The link to the experiment is:


where you will be asked to enter the following:

user name: user3
password: w33n1
(note the password '1' as in one, not lowercase 'L').

The instructions will be on the first page. I would ask you to complete the experiment all in one without pausing between slides.

If you have any questions regarding the experiment, please contact me at joycece@tcd.ie.

Don't hesitate to forward the experiment on to others who may be interested.

Thank you in advance for your time,

Claire Joyce
### Appendix D

**Tables from Experiment 2**

#### RATIONALITY

<table>
<thead>
<tr>
<th>TARGET</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>29</td>
<td>32</td>
</tr>
<tr>
<td>MO</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>L</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>LO</td>
<td>12</td>
<td>9</td>
</tr>
</tbody>
</table>

This table shows the number of Rational answers per target.

#### ERROR

<table>
<thead>
<tr>
<th>TARGET</th>
<th>POSITIVE</th>
<th>NEGATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>MO</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>L</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>LO</td>
<td>24</td>
<td>31</td>
</tr>
</tbody>
</table>

This table shows the number of Errors per target.