Creativity and Convention: the Pragmatics of Everyday Figurative Speech

Rosa Elena Vega Moreno

Thesis submitted for the degree of Doctor of Philosophy in Linguistics.

University College London
2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

To my family,
with love and gratitude

Abstract

The aim of this thesis is to provide a unified account of the comprehension of novel and conventional figurative uses of language in everyday conversation. The first two chapters present a view of human cognition as having evolved towards increasing efficiency, characterised by the selection and processing of just those sources of information which are likely to yield cognitive benefits and incur limited costs in processing effort. Pursuing the pragmatic framework of Relevance Theory developed by Dan Sperber and Deirdre Wilson, I take this view of cognition to set the grounds for a relevance-driven, inferential approach to human communication. The rest of the thesis focuses on the role that selective accessing and inference play in the comprehension of a range of everyday figurative uses varying in their degree of familiarity or conventionality.

After reviewing existing cognitive approaches to metaphor in chapter three, I provide my own relevance-theoretic analysis in the next chapter. I argue that metaphor interpretation involves processing assumptions selected from the encyclopaedic entry of the concept encoded by the metaphor vehicle and the inferential construction of a new (broader) concept. In chapter five, I analyse the claim that the meaning of idioms is not completely arbitrary but partly related to the meaning of their parts. In chapter six, I provide an account of the interpretation of idioms (and idiom variants) which captures this relation. In these chapters, it is argued that the selective processing resulting from relevance-driven comprehension often leads to the construction of new concepts for novel metaphorical uses and the development of pragmatic routines for processing familiar figurative uses.

Contents

Acknowledgements

I feel proud and privileged for having had the chance to work under the supervision of Robyn Carston who has been an inexhaustible source of inspiration, support, patience and wisdom. I am grateful to her for the long hours she has dedicated to reading, thinking about and commenting on my research, for our many illuminating discussions and her always constructive criticism. I am particularly thankful to her for respecting and encouraging my ideas and for waiting patiently as I shaped them through the years.

My most sincere gratitude also to my second supervisor, Deirdre Wilson, for our enlightening discussions, her cheerful and dedicated involvement with my work and progress and for making me the recipient of her knowledge, her kindness and her continuous support.

I owe thanks to a great number of people not all of whom I can mention here. I am grateful to Neil Smith and Valerie Hazan for creating a pleasant working atmosphere in the department and for the PhD room where I have spent long hours working and socialising. Thanks to Steve Nevard and Molly Bennet for making my life easier by solving all sorts of technical and administrative troubles. I am also grateful to the members of the Department of English Philology at University of Las Palmas de Gran Canaria, in particular Francisco Alonso, Mercedes Cabrera, Isabel González, María del Mar Pérez, Alicia Rodríguez and Francisco Vizcaíno for encouraging me into research and for being always willing to help.

On the financial front, I thank the Canary Government (Gobierno Autónomo de Canarias) for a four-year scholarship and the UCL Department of Phonetics of Linguistics, the UCL Graduate School, Linguistics Association of Great Britain and EURESCO for their help in financing conference presentations. I am also thankful to UCL and its Department of Phonetics and Linguistics for offering me various jobs, which have helped me improve my CV and finance my PhD studies.

My work has benefited from discussions with a number of people in and outside UCL and from the generosity of those who have helped with references and have shared their own research with me. I am thankful to all those who have helped me in any way and to those I have met in conferences and workshops for their interest in my work and their feedback. Special thanks are due to Nicholas Allott, Michael Ashby, Cristina Cacciari, Barbara Eizaga, Ray Gibbs, Richard Horsey, Corinne Iten, Sergio Maruenda, Paula Rubio, Bjliana Scott, Neil Smith, Begoña Vicente, Tim Wharton, Francisco Yus and the members of the Relevance Reading Group. I am particularly thankful to Billy Clark and Adrian Pilkington for their detailed reading of my thesis and their insightful comments.

I am extremely grateful to each and every one of my friends. To those in England, for looking after me and to those in Spain for not forgetting about me. To my fellow PhD students, for understanding the sacrifices and rewards involved in writing a thesis and to my non-linguist friends for reminding me that there is more to life than linguistics and research. I owe thanks to Nicholas Allott, Javier Ayestarán, Ursula Budnik, Dirk Bury, Ana Carrera, Robyn Carston, Héctor Conde, Vanessa Contreras, Hayley Dauben, Abdel Fiala, Dorota Gowaska, Alison Hall, Vikki Janke, Ann Law, Marika Lekakou, Sergio Maruenda, Eva Naranjo, Sylvia Navarro, Daniel Quintana, Hiruma Ramos, Irene Sanoudaki, Hitoshi Shiraki, Marco Tamburelli, Nina Topintzi, Vina Tsakali, Hiroyuki Uchida, Raquel Vega, Liliana Vega, Olivia Vega, Reiko Vermuelen, Peter Vermuelen, Kim Walker, Hilary Warner, Tim Wharton and Deirdre Wilson for their continuous support, their affection, and the priceless moments we have shared together throughout the years.

Above all, my deepest and most heartfelt gratitude is to my family. I have no words to thank my parents, José Vega Pérez and Rosa Moreno Díaz, and my sisters, Liliana and Olivia Vega Moreno, for the financial and emotional sacrifices they've made for me. I am most grateful to them for their unconditional love, generosity and dedicated patience, for coping with my mood swings, my busy schedules and, particularly, with my absence. I am incredibly lucky for having the family I do and for knowing I can count on their

warmth and their help every day of my life. I dearly remember my grandfathers, José Vega Alemán and Andrés Moreno Torilo, who sadly died in the course of my research in the UK and who cannot see, but from heaven, how I become a doctor. To my grandfathers, grandmothers, sisters and parents in honour to their love, patience and devotion, I dedicate this thesis with all my love, respect and gratitude.

Chapter 1

Human Creative Cognition and Selective Processing

> "I see thinking as bringing ideas together, as ideas flirting with each other, learning to dance and embrace [...] Ideas are constantly swimming around the brain, searching like sperm for the egg they can unite with to produce a new idea. The brain is full of lonely ideas, begging you to make some sense of them [...] The lively brain picks and chooses and creates new works of art out of ideas"
>
> Theodore Zeldin (1998: 56-57)

1. Introduction

The human mind is exceptionally creative. It is this creativity, developed through evolution that has helped the human species to survive and to differentiate from the rest of the species. Creativity is often defined as the ability to construct something new (typically useful) out of existing elements (Sternberg, 1999). A poem, a book, a painting, a plan to escape prison, a story, an utterance, and a thought may all be instances of creative cognition. In fact, although not everyone is capable of creating a masterpiece, whether a book, a sonnet or a film, every human being with a healthy cognitive system can be cognitively creative by constructing, combining and modifying mental representations in thinking or in understanding what others think. The emergence of something novel, of a mental entity that has never been represented before, an idea that has never been expressed before, makes thought and the communication of thoughts beautiful, interesting and useful.

At some level, asserting that the human mind is creative, flexible or generative seems a mere platitude. We know our minds are complex and efficient. We recall memories, events, and faces. We invent stories. We dream. We lie. We form mental images, construct categories, establish similarities, draw analogies, etc. We acquire an incredible amount of simple and complex concepts in infancy and throughout our lives. We combine these concepts to create new ones. We understand people's novel conceptual combinations, their novel uses of words. We produce and understand utterances, whether

literally or figuratively intended. We construct mental representations all the time, when thinking and when working out what people think, to explain people's behaviour and to make sense of things. In short, our minds are always at work: acquiring, organising, classifying, storing, recalling and modifying information. What is needed is a theory to explain the cognitive processes underlying this creativity. This is a serious and complex enterprise.

Different branches of cognitive psychology have approached the topic from different perspectives. Work has been carried out in perception, language comprehension, information processing, problem solving, memory recall, imagery, etc. In their experiments scholars have tested subjects on a wide variety of cognitive tasks, some involving for instance solving a problem, building up a category, working out the meaning of a novel word, imagining creatures from other planets, forming mental images, producing and understanding novel conceptual combinations, etc. (See Barsalou, 1983, 1991; Cacciari et al., 1997; Ward et al., 1997). Regardless of the different natures of the studies and methodologies employed, similar results have been obtained: evidence of people's striking ability to construct new representations (e.g. new meanings, new ideas, new categories, new images, etc.) on the basis of existing knowledge. This common interest across disciplines in understanding the creativity involved in human cognition gave birth to the field known as creative cognition. Creative cognition is a new area of research which aims to investigate both the cognitive processes that lead to the emergence of novel cognitive structures and the role of existing cognitive structures in this emergence (Finke et al., 1992; Smith et al., 1995; Ward et al., 1997).

In line with research on creative cognition and creative thinking, this dissertation studies some aspects of cognitive creativity involved in ostensive verbal communication. The aim of this chapter is to present some cognitive background for this dissertation, as it investigates existing psychological evidence on the role played by selective processing and depth of processing in the construction of new representations and the formation of ad hoc categories. The next chapter will present a cognitive view of the mind (that defended by Sperber and Wilson's Relevance Theory) which is consistent with this psychological evidence and which provides grounds for the construction of a psychologically plausible theory of human ostensive communication (also defended by

Sperber and Wilson's Relevance Theory) which will be outlined in that chapter. The rest of the thesis will show how the pragmatic interpretation of ordinary metaphorical (Chapters Three and Four) and idiomatic uses (Chapters Five and Six) is essentially a creative process even in the most conventional cases.

2. The Selective Mind

A great deal of research has been done on how the mind works and, although the mind is still largely a mystery to humans, we understand it better now than ever before. This section focuses on one particular property of human cognition, its selectivity. Scholars from a wide variety of cognitive disciplines have observed that the human cognitive system seems to be rather selective, at least when it comes to perceiving, processing, memorising and recalling information. Let's consider some evidence.

Experiments on visual perception typically involve tasks like these. A medium-size white square is presented to subjects in the following contexts: a) together with a set of several black squares of roughly the same size, b) together with a set of medium-size white triangles, c) together with several big squares of different colours, d) together with a set of tiny squares. Findings show that subjects perceive the same target item (i.e. the medium-size white square) differently on each occasion. For example, although in c) the item is normally perceived as small, in d) it is perceived as large. Our perceptual system seems to select from all available stimuli that which is most salient and distinctive at the moment of perception. Salience of the properties of the target item (size, colour, shape) varies across contexts and situations, so that priming of one property (e.g. colour) in one context, e.g. (a), does not necessarily presuppose priming of the same property in a different context, e.g. (c) (see Barsalou et al., 1999; Garner, 1974; Glass et al., 1979). Some scholars take this as evidence that subjects form different representations of the target figure in each case.

The context-dependence of perceptual attention has implications for theories of visual perception. Although these implications need not carry over to theories of language comprehension (but see Barsalou, 1999; Barsalou et al., 1993; Barsalou et al., 1999; Smith and Samuelson, 1997), they do show that at least some aspects of cognition seem to be geared to focusing on those bits of information which are potentially relevant in the

context. In fact, they also show that selectivity seems to go hand in hand with the ability to pick out similarities and enhance distinctiveness among available stimuli. At least at an intuitive level, what subjects seem to be doing in the experiment above is perceiving members and non-members of a certain category. They seem to be taking the target item not to belong to any of the following categories: the category of black squares, the category of triangles, the category of big squares, and the category of small squares. In other words, experiments like this seem to suggest the following:

(1) a) Selective attention plays an important role in cognition.
    b) The inputs (e.g. properties) our cognitive systems pay attention to are those which are potentially relevant to the individual at the time.
    c) Salience of a property depends on the context in which the item is processed/ perceived.
    d) Attending to a certain subset of the properties that a particular stimulus gives access to has an effect on how we classify this stimulus.

Evidence for informational selectivity is not unique to perception but plays an important role in performing a wide variety of cognitive tasks. Consider some problem-solving tasks typical, for instance, of IQ tests. IQ tests normally present subjects with 'odd-one-out' tasks. Subjects are typically presented with a set of elements and asked to say which of the elements 'does not fit'. One example of a set would be something like 'milk', 'chicken', 'farm', 'potatoes'. Here, it is the element 'farm' that should be odd in virtue of not being edible. A set like 'milk', 'chicken', 'farm', 'computer', however, may lead subjects to exclude not 'farm' but 'computer'. This is so because all elements except 'computer' belong to the category of 'things one normally finds in an agricultural countryside farm'. This category may be created ad hoc ('ad hoc category') in solving the task and may never be created by the subject again. Similarly, these elements are grouped together in performing this task but may never be grouped together in the mind of the subject under the same category again. A possible explanation of what makes the element 'farm' suitable on one occasion but not on another is that subjects access different bits of knowledge about farms in each context. The property of 'edibility', which is highly

salient in processing the words *milk, chicken and potatoes*, leads subjects to infer the property 'not edible' from their knowledge of farms. The assumptions 'farms have chickens', 'farms have cows', 'vegetables are grown on farms', are possibly retrieved directly from memory, where they have been stored as a result of previous experience with farms.

These types of problem-solving tasks seem to test subjects - however indirectly - on their ability to construct categories on the spot and to identify the members and non-members of these categories. In fact, not only is there a wide array of games and tests that explore our ability to create new categories on the fly, some actually exploit just the opposite ability. In these cases, rather than giving subjects a list of elements and asking them – directly or indirectly – to provide the category to which the elements belong, subjects are provided with the description of an ad hoc category (i.e. a label) and asked to supply members that belong to it. The famous table game Scattergories, for instance presents participants with the following categories: 'things one can be allergic to', 'things one can find on a beach', 'things that bounce', etc. The Spanish TV Program '1, 2, 3' provides not only the ad hoc category but also a member of it, such as the following: 'actresses that have been married to actors, for example, Nicole Kidman', 'main characters of European novels, for example, Tess of the D'Urbervilles', and so on. In each case, those participants who are faster at providing members for the ad hoc categories and give the greatest number of correct responses win the game. Mind games and IQ tests such as the ones presented here may make no predictions about human cognition. They may make no theoretical claims, follow no experimental procedures and hence have no scientific psychological validity. Still, they are intuitively appealing as they illustrate humans' extraordinary ability to construct categories on the spot, to identify the members and non-members of these categories, and to provide members of novel categories. The need is for serious experimental work to test this ability, to make theoretical claims about the processes involved in ad hoc category formation and its implications for a theory of human cognition.

The psychologist Lawrence Barsalou has dedicated much of his work to investigating ad hoc categories (e.g. Barsalou, 1983, 1991). He has studied how and why ad hoc categories are formed, how they behave, and what implications they have for a

theory of memory, in particular for how the mind works in retrieving, classifying and reorganising information in memory. His work shows that people seem to construct ad hoc categories all the time. They may construct, for instance, an ad hoc category such as 'things to eat on a diet' when planning to lose weight, an ad hoc category 'things to do with the neighbour's children on a rainy day' when babysitting, an ad hoc category 'places to visit in Madrid' in planning a holiday, etc. Although ad hoc categories play an important cognitive role at the time when they are formed, they are typically only temporarily represented in the mind. They are formed, used and forgotten. In some instances, when the same category is frequently constructed, it may end up being stored in memory. So for people who diet a lot or baby-sit regularly in rainy places, the above categories may become stabilised in their minds.

Although ad hoc categories differ from common taxonomic categories (e.g. birds, fruit, cars, furniture) in that they are not permanently stored in the mind, they seem to behave in very much the same way as taxonomic categories. A common assumption about such categories is that they exhibit prototype structure. A world without categories would be chaotic, as every new element we perceive would have to be considered in isolation. Our minds have the ability to classify things in the world so that people can recognise examples of a certain category (e.g. examples of birds) when presented with them. In fact, people also seem to be good at recognising which members of a certain category (e.g. birds) are good examples of that category (e.g. robins) and which are not (e.g. penguins) (Rosch, 1978; Smith and Medin, 1981). Prototype theories have argued that the more typical a certain member is of a certain category, the faster it is to be judged as a member of that category. The less typical, the more likely it is that people hesitate in considering it a member of the category (e.g. is a chicken a bird?). Barsalou's experiments show that graded structure is not restricted to common categories but is also exhibited by one-off ad hoc categories. That is, when presented with an ad hoc category (e.g. 'things to eat on a diet'), subjects are able to provide members that belong to the category (e.g. low fat yoghurt), elements that are not members of the category (e.g. chocolate), typical members (e.g. fruit) and less typical members of the category (e.g. meat).

Some of Barsalou's experiments presented subjects with a short text about a person involved in a certain activity (e.g. moving to a new city), a label for an ad hoc category (e.g. 'ways to make friends') and a list of possible candidates for that category (e.g. 'join a card playing club', 'go back to school', 'get convicted for a crime'). Subjects were asked first to circle those members they believed to belong to the category and then to rank them according to typicality, from most typical to least typical (Barsalou, 1983). Although participants had never come across the categories presented in the experiment (e.g. 'things that can fall on your head', 'places to go to escape being killed by the Mafia', etc.) and hence had never thought about them before, they were consistent in deciding both which members belonged to the ad hoc category they were tested for and which members were more typical of that category. How is this possible without any previous experience with the category or exemplars? Barsalou suggests that in performing the experiment subjects may have considered certain typical properties members of the category may have. This includes the property of 'geographical distance' for the category of 'places to go to escape from being killed by the Mafia'. The greater the geographical distance of the place one would go to escape from the Mafia, the better the chance of not being found and hence of not being killed. For the ad hoc category 'things to eat on a diet', typical properties include 'low in fat' or 'low in calories', so that food which is both low in fat and calories (e.g. vegetables, skimmed milk) has a good chance of being considered not only a member of the category, but also a typical member.

Some of Barsalou's experiments explored the opposite ability, that is, the ability of subjects to provide a certain ad hoc category given a set of potential members. A typical instance-to-concept task would provide subjects with a question such as 'what category do coffee, perfume, leather, and skunk all belong to?' (Barsalou, 1983). It was found that subjects had more difficulty both in retrieving the members of an ad hoc category and in providing an ad hoc category for a set of members than they did in retrieving the members of a common taxonomic category (e.g. furniture) and in providing a taxonomic category for a set of members (e.g. what category do 'bed', 'sofa' and 'cupboard' belong to?). The reason seems obvious. A certain item, say 'apple', may belong to an indefinite range of ad hoc categories, for example 'things to eat on a picnic', 'things that can fall on your head', 'things to have for dessert', etc., but to a finite number of common taxonomic

categories, namely 'fruit' and 'food'. Both concept-to-instance and instance-to-concept associations are well-established for common categories but not for ad hoc categories, making retrieval easier for the former. The provision of specific contextual information improved the performance of the tasks considerably for ad hoc category formation and recognition. In both cases, subjects consistently treated ad hoc categories in the same way as common taxonomic categories and assumed they behaved in similar ways.

What Barsalou's experiments seem to show is that we have a striking ability to form new representations by selecting different subsets of knowledge from memory on different occasions. This ability enables us to construct ad hoc categories, that is, to group elements we have never encountered together before and may never encounter together again, as well as to provide members for a certain ad hoc category we have never thought about before. Ad hoc categories behave very much like normal taxonomic categories: they play a fundamental role in cognition, they have typical properties and exhibit prototypical structure. We may conclude that this work on category formation shares with work on visual perception, such as the work described above, the following picture of human cognition (which is slightly more general than that presented in (1)):

(2) a) Selective attention plays an important role in cognition.
    b) The inputs (e.g. properties) our cognitive systems pay attention to are those which are potentially relevant to the individual at the time.
    c) Accessibility of a mental input (e.g. a concept) or saliency of a property depends on the context in which the item is processed.
    d) Attending to a certain subset of the properties that a particular stimulus gives access to has an effect on how we classify this stimulus.

The output of selection and classification involved in these instances of ad hoc category formation is a novel mental representation, namely a new (unlexicalised) ad hoc concept which denotes the category created ad hoc (e.g. it denotes the set of things one may eat on a diet, the set of things one may do to make friends, etc.). In the next chapter and throughout the thesis, we will see how the construction of ad hoc concepts has important implications for a theory of utterance interpretation. A main goal of this thesis is indeed

to study the pragmatic inferential mechanisms responsible for the lexical-pragmatic fine-tuning that takes place in creating ad hoc concepts. Although these mechanisms have not been studied within psychology, work in this field provides interesting experimental evidence which should be accommodated within an adequate pragmatic framework.

3. Human Memory and Information Processing

Research on utterance interpretation benefits from looking at work on human memory, work on information processing and the relation between the two. The close link between memory and information processing is not strange when we consider the following assumptions. On the one hand, the way in which a stimulus is processed affects the information we store for that stimulus. Since the retrieval of a target item from memory depends both on which information we retain and on how we store it, the way we process information has strong implications for theories of memory. On the other hand, the effectiveness of certain cues for retrieving information from memory indicates the presence of a memory trace, and hence gives some insight into the information that played a role during the processing of the stimulus. In this section, I will present a principle of human memory, the Encoding Specificity Principle which, like the Levels/ Depth of Processing Hypothesis to be presented in Section 5, has important implications for a theory of utterance comprehension.

3.1 The Encoding Specificity Principle

Memory tests typically involve two phases. First there is a phase in which participants are presented with a stimulus and a task to perform on the stimulus (e.g. read it). This is known as the 'encoding phase'.[1] At a later stage, subjects are tested for memory of the stimulus presented during the encoding phase. This is known as the 'retrieval phase'. Although much of the early work on memory and information processing focused on the storage of information rather than on how it is encoded and recalled from memory, in the last few decades an increasing number of scholars have been interested in the processes

[1] The term 'encoding' here is different from the term 'encoding' as used in linguistics. It merely refers to the information that is considered in processing a stimulus in the study phase and that determines how the stimulus is represented (temporarily) in the mind.

that take place at encoding and retrieval, and particularly in the close link between the two (see Anderson, 2000; Baddeley, 1997 for good overviews on issues on memory and cognitive psychology).

In proposing the Encoding Specificity Principle for human memory, Tulving and Thompson (1973) argue that the "specific encoding operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored" (Tulving and Thompson, 1973: 369). In other words, the Encoding Specificity Principle postulates that when presented with a stimulus (e.g. a word or set of words), we engage in certain mental processes (i.e. 'encoding') which will determine what information we end up storing in working memory for this stimulus. Only the information that plays a role during encoding leaves a memory trace, so only the presence of a retrieval cue whose relationship with the target item was processed at the moment of encoding can – following this trace - track the target item in memory and retrieve it.

Here is an example. A typical experiment on memory involves giving subjects a list of twenty four words and asking them to read through the list. On some occasions the subjects are explicitly advised to try to memorise the items because they will be asked about them in a later task. Words may appear in isolation (e.g. BLACK), together with a strong cue (i.e. a word in a relation of free association) (e.g. white-BLACK), or together with a weak cue (i.e. a word which does not normally prompt recall of the word in free association) (e.g. train-BLACK). Traditional theories of memory assume that a word can act as a good retrieval cue for another word in virtue of their pre-experimental relations. This implies that a word will always prime the same words whatever the experimental setting. If so, the word *white*, which is a strong associate to the word *black* in free recall, should act as an effective cue in all three cases above. The Encoding Specificity Principle originated as a reaction against traditional views. The prediction of the Encoding Specificity Principle is that processing items in different contexts should affect the information people access during the encoding phase. Different encodings leave different memory traces so only those recall cues that use information present at encoding will be effective retrieval cues.

Scholars have repeatedly tested the hypothesis of context-dependence of processing, finding consistent and conclusive evidence for the encoding specificity view and against traditional models (e.g. Barclay et al., 1974; Brown and Craik, 2000; Hannon and Craik, 2001; Tulving, 1983). Tulving and Thompson (1973), for instance, found that for all three context settings above, the word *white* acted as a highly effective cue in the first two cases, where the word *black* appeared in isolation or together with its strong associate *white*. However, contrary to the prediction of traditional theories, the cue *white* was much less effective in the 'train-BLACK' context. A possible explanation of this finding is that the concept of 'blackness' perceived and stored by subjects in this condition differs from the concept of 'blackness' stored by subjects in the other two conditions. A similar result is found when the word *water* is presented in the pair 'whisky-WATER'. In this case, the cue *lake* does not act as a good recall cue for *water*, because *whisky* and *lake* emphasise different properties of water (e.g. 'something drinkable', 'something one can swim in') (Hannon and Craik, 2001). Although Tulving and Thompson never tested this possibility, it is probable that a cue related to the 'blackness of trains' (e.g. smoke) would act as an effective cue to recall *black* when the word has been presented in the 'train-BLACK' condition, as it would provide information related to properties salient at the moment of encoding.

The literature on encoding specificity is not concerned with language comprehension per se, so it does not study, for instance, the implications that these type of experiments might have for a theory of word meaning and utterance interpretation. Still, the results seem to shed light on the ability of the human mind to convey different meanings on different occasions, and on the ability of our cognitive systems to pick out from memory only a subset of information, with different subsets being accessed on different occasions. Tulving (1979), in arguing against traditional theories, concludes that "systematic variability in the effectiveness of retrieval cues cannot be attributed [just] to pre-experimentally established relations between cue and target words, it must be *determined by processes occurring in the study episode"* (Tulving, 1979: 204) (my italics). This statement makes sense if we accept the context-dependence of lexical comprehension. If different bits of knowledge are retrieved and used during the

comprehension of a word in context, it seems natural that a certain recall cue may trigger memory for the word on some occasions while exhibiting no priming effect on others.

One of the first experiments to test the relation between semantic encoding specificity and lexical flexibility was carried out in the mid-seventies by Barclay, Bransford, Franks, McCarrell and Nitsch (1974). These scholars presented subjects with a list of sentences, including either 'the man lifted the piano' or 'the man tuned the piano'. After a three-minute break, subjects were given a set of cues (e.g. 'something heavy', 'something with a nice sound', etc.) and asked to write down the noun(s) from the previous sentences that each cue reminded them of (so *piano* in each of these cases). The results showed that 'something heavy' acted as a good recall cue when the word *piano* had been processed in a context in which the piano was lifted, but not when it had been processed in a context in which the piano was tuned. The cue 'something with a nice sound' had the opposite effect. Barclay et al. (1974) took the encoding specificity hypothesis to explain their findings. The idea is that since only those bits of information which are actually considered while processing a word in context are involved in encoding the meaning conveyed by the word (and hence included in the representation we store temporarily in working memory), only these bits of knowledge would leave a memory trace. Since only these bits of knowledge leave a memory trace, only cues which are related to this information should be effective in retrieving the words (or utterances) present at encoding. Other cues, related to other bits of information the words are associated with in long-term memory should not be so effective cues. That is, at the moment of processing ('encoding') an utterance such as 'the man lifted the piano', only certain information associated with the word *piano* in long term memory is accessed, for instance the assumption that pianos are heavy. Considering this piece of knowledge during processing leaves a memory trace, so that only a cue related to this property of pianos (i.e. the property of being heavy) and not cues related to other properties (e.g. the property of emitting nice sounds, of having pedals, etc.) are effective in recalling the word (or utterance). The predictions made by the Encoding Specificity Principle are therefore borne out.

One possible objection to this experiment is that a property such as 'being heavy' may not actually be retrieved from the subject's knowledge of pianos, but may be

activated by use of the word *lift*. To test whether this was the case, Barclay and colleagues set up another experiment. This time subjects were presented with sentences such as 'the man lifted the piano' or 'the man lifted the infant' and with cues such as 'something cuddly' or 'something heavy'. The rationale underlying the experiment was that if the property 'heavy' is accessed from processing the word *lift* rather than *piano*, then it should act as an effective cue in either case. The results suggested otherwise. The property 'heavy' was a good recall cue in the piano context but not in the infant context. The property 'cuddly' was an effective cue in the infant context alone. Taken together, these findings suggest that selective processing takes place naturally and spontaneously during utterance comprehension. Different bits of knowledge associated with a word in memory are accessed in processing that word on different occasions. Only those bits of information which are considered during the encoding process should act as good retrieval cues.

4. The Flexibility of Concepts

In section 2, I presented some experimental work by the psychologist Lawrence Barsalou on the ability of humans to construct ad hoc categories on the fly (e.g. the categories 'things that can fall on your head', 'places to go to escape being killed by the Mafia', etc.) by accessing and assembling different bits of information stored in long-term memory. The instability of category representations is not restricted to the formation of ad hoc categories but is also found in the processing of words denoting common taxonomic categories (e.g. Barsalou, 1981, 1985; Barsalou et al. 1993). Barsalou and colleagues have shown that taxonomic categories are represented differently across individuals and populations, within a single individual and population, and across contexts and times.

In a series of experiments Barsalou (1987) showed that property generation varies both between individuals and within a single individual. He asked different subjects to provide one of the following: properties of ideal birds, properties of an average bird, properties they view as typical of birds, or properties they would include in their definition of birds. The results showed that as little as a third of the descriptions provided

by one subject overlapped with those provided by another subject. In a further experiment, he tested whether subjects would actually produce the same or different properties for a given category (e.g. the category of birds) if they took different points of view. Results showed that the properties subjects produced when asked to take the point of view of a housewife differed from those they produced when taking their own point of view, which differed again from those they produced when asked to take the point of view of a Chinese citizen, a university professor, and so on. One possible explanation for these variations is that people store a wide range of information about a single category. To test this, Barsalou, Sewell and Spindler (ms) (see Barsalou, 1993) pooled together all the properties that subjects had given for a particular category in the generation tasks, and tested another group of subjects, who were asked to judge those properties as true or false for a particular category (e.g. 'it flies', 'it is small', 'it sings', etc. for the category of birds). The agreement between subjects was between 97% and 98%, showing that people have stable knowledge stored for categories and that this knowledge is essentially the same. So if people do not differ in the amount of knowledge they have stored for a certain category, they must differ in the bits of this knowledge they retrieve from memory on different occasions.

Typicality rankings also proved to be unstable. In one experiment, Barsalou and Sewell (1984) asked students from the same university to rank a set of exemplars of a category from least typical to most typical (or to rate their typicality on a scale 1 to 7) while taking either their own point of view or someone else's. For instance, they were asked to rank a list of exemplars of birds from the point of view of the average American citizen or from the point of view of a Chinese citizen. The results showed instability, and even reversibility, in rankings. For instance, when taking the point of view of an average American, subjects judged robins as the most typical bird and swans as the least typical. When taking the point of view of a Chinese citizen, this ranking was reversed. Finally, a great degree of instability was found not only across subjects but within subjects participating in the same experiment at different times (see Barsalou, 1993). For example, the prototypicality ratings for a person who judged the exemplars of a category twice with a few weeks' break only overlapped in 64% of cases.

Barsalou and colleagues took these findings as evidence against the idea that categories are organised around ready-made prototypes. Instead, prototypicality judgements for a given category vary across contexts, times, individuals, populations, tasks, points of view, etc. In fact, the instability of graded structure suggests that people are forming different representations for a single category on different occasions. They seem to assemble different subsets of their knowledge of a category (e.g. birds) when performing different tasks, taking different points of view, processing the category in different contexts, at different times, etc. Different representations result in different typicality judgements and rankings. The different representations of a certain category are what Barsalou refers to as concepts:

> "On a particular occasion, different people retrieve different subsets of features from their knowledge of a category. These varying subsets of features are what I am defining as concepts. Rather than being stable structures in long-term memory and retrieved as needed, concepts are temporary constructions in working memory" (Barsalou, 1993: 34)

Most standard accounts of concepts take a concept to provide access to a relatively stable stock of information – e.g. a stereotype or prototype - in long-term memory. This stock of information is often assumed - explicitly or implicitly - to be retrieved as a whole when the concept is activated. A wide range of experimental research on concept processing, encoding specificity and lexical access, including the work presented here, brings this assumption into question. It suggests instead that our minds select a certain accessible subset from a wide array of information, with different subsets being accessed on different occasions. Barclay et al. (1974) propose that different senses of a single unambiguous word are created on each occasion. That is, if in processing the word *piano* in different contexts, different properties of pianos are accessed, then different representations must be formed. Barsalou's argument is very much the same, he refers to these new representations being formed by a one-off assembly of properties from long-term memory as *concepts*.

According to Barsalou, categories are sets of entities of a certain ontological type (e.g. weddings, newspapers, pianos, etc.) for which people store large amounts of information in long-term memory. Concepts are cognitive representations of a certain

category that people construct in working memory on a particular occasion by selecting different subsets of their information about that category (see Barsalou, 1982, 1989; Barsalou and Bower, 1980; Barsalou and Medin, 1986; Barsalou et al., 1993). The accessibility of this information is seen as varying across times, with information that remains inactive or rarely accessed in most contexts becoming highly accessible on certain occasions. So although the property 'edible' is not typically accessed in processing the word *frog* in isolation, it becomes highly accessible when the word is presented in the context of a discourse about a French restaurant (Barsalou, 1982). Also, although the property 'flammable' is not typically accessed when processing the word *newspaper* in most contexts, it becomes highly accessible in the context of a discourse about making a fire. Evidence for the variability and context-dependence of information accessibility is that a piece of information (e.g. flammable) about a certain category (e.g. newspapers) is verified significantly faster (by 145 milliseconds) in a context where this information is appropriate (e.g. discourse about a fire) than in isolation or in a neutral context (Barsalou, 1982).

4.1 The Notion of Concept

The view of concepts taken by many approaches in psychology differs from the view of concepts within more philosophical accounts. This distinction is often, though not always, linked to a more general distinction between internalism and externalism. In an internalist view of concepts, concepts differ when people form different mental representations even if the represented category remains the same. To externalists, concepts differ when they have different denotations, i.e. pick out different categories. Barsalou, Barclay and colleagues, seem to hold a broadly internalist view. That is, they claim that for every subset of information associated with a category that we select on a certain occasion, a new representation is formed, a new concept is constructed and a new meaning conveyed. In their view, the fact that people select different subsets of their information about pianos in processing the utterances 'the man lifted the piano' and 'the man tuned the piano' is an indication that they are constructing different meanings, and hence an indication that an unambiguous word such as *piano* exhibits semantic flexibility. Similarly, the concept that someone constructs for a category such as cars may include

the feature 'radiator' on one occasion and 'air conditioning' on another (Barsalou and Bower, 1980). In fact, different properties of cars may be incorporated into a one-off representation by different individuals (e.g. a driver, a mechanic) or by the same individual in different contexts (e.g. when taking the car to the mountains, the beach, the motorway, the garage, etc.). Since, in each case, people retrieve different subsets of features from their knowledge of cars, Barsalou argues that different concepts are being created.

An externalist position, the one assumed in this thesis, is rather different. Accessing different subsets of knowledge of a category on different occasions may or may not result in the construction of a different concept. When selective processing results in the construction of a concept whose denotation differs from that of the encoded concept (e.g. it is narrower), we can say a new concept has been constructed. When accessing different subsets of information does not affect the denotation of the encoded concept, we cannot say a new concept has been formed.[2] Processing the word *bird* in the context of a discussion of a person eating in a restaurant involves considering information about edibility which we would not consider when processing the same word in some other contexts. Considering the assumption that 'some birds are edible' as opposed to other information (e.g. 'some birds fly', 'some birds sing', 'birds have feathers') may result in the on-the-spot formation of a concept BIRD*[3] which denotes a subset of birds in the world, namely those which are edible, taste nice and are likely to be served in a restaurant. Considering the assumptions that 'pianos are heavy' or 'pianos make nice sounds', however, may not generally result in the on-the-spot formation of a new concept, because all pianos are typically heavy and made with the purpose of sounding nice. Processing those assumptions in comprehending the utterance would not alter the denotation of the encoded concept.[4]

---

[2] The externalist view defended in Relevance Theory and this thesis is the one proposed by Fodor (1998). According to Fodor, although two conceptual representations that pick out different things in the world are necessarily different concepts, picking out the same thing in the world, does not make two conceptual representations identical. To be identical they must also share what Fodor refers to as their Mode of Presentation. The expressions 'evening star' and 'morning star', for instance, denote the same entity in the world, yet they encode different concepts with different Modes of Presentation.

[3] Henceforth, I will use capital letters to refer to concepts (e.g. BIRD) and add an asterisk when the concept is created ad hoc (e.g. BIRD*).

[4] The next chapter will look more closely at how ad hoc concepts are created and at how the encoded concept and the communicated concept may differ.

5. The Depth of Processing Hypothesis

We have seen that researchers on memory generally agree that the operations we engage in during encoding when presented with a stimulus affect (representation and) memory of that stimulus. Craik and Lockhart (1972) propose, more specifically, that the *depth* to which a stimulus is processed affects retention. Their widely influential proposal is known as the Depth of Processing Hypothesis, or Levels of Processing Hypothesis (see Craik, 2002; Craik and Lockhart, 1972; Craik and Tulving, 1975; Cremark and Craik, 1979). Craik and Lockhart (1972) claim that a stimulus (e.g. a word) can be processed to different levels or depths, and distinguish between 'shallow' and 'deep' processing. Whereas in shallow processing subjects process structural or phonological information (e.g. about case or rhyme), deep processing involves the processing of semantic information (e.g. about word meaning or the encyclopaedic information made accessible by use of a word).[5] In their experiments, they presented subjects with a set of words and a task to be performed on each word. Tasks were of three types: a structural decision task (e.g. 'is the word written in capital letters?'), a phonological decision task (e.g. 'does the word rhyme with....?') or a semantic decision task (e.g. 'does the word fit the sentence ...?'). In the memory phase, subjects were tested for their memory of the words presented in the encoding phase (e.g. they were asked either to recall or to recognise the words). The results showed that memory for words which were processed at a deeper level (paying attention to meaning) was significantly better than memory for words which had been processed at a shallow level (e.g. paying attention to phonological or structural features). They explained these findings by arguing that longer-lasting traces for semantically processed information lead to better retention and memory for items processed in a deep (semantic) condition than for items which are in a shallow condition. In virtue of these and other similar findings, Craik and Lockhart propose the Levels of Processing or Depth of Processing view of memory, which claims that the depth to which

[5] The notion of 'semantics' used by these scholars is much broader than the one which is defended by modern pragmatic theories such as Relevance Theory (to be presented in the next chapter). 'Semantics' here is roughly synonymous only with meaning, and so covers not only aspects of semantics proper (i.e. encoded meaning) but also of pragmatics (e.g. the use of encyclopaedic knowledge).

a certain stimulus is presented affects memory for this stimulus. The deeper the processing, the better the memory.

The phenomena studied by Craik and Lockhart clearly show selective attention. Subjects' attention to the target item is affected by the experimental condition in which the item is presented. In fact, the Encoding Specificity Principle presented above was developed within the Levels of Processing framework. The idea that ties both notions together is the hypothesis that there exists a close link between encoding and retrieval, in that what happens at encoding determines what is stored in memory for the items encoded. In processing a stimulus, subjects pay attention to selected properties of the stimulus. It is the properties one pays attention to at encoding that determine what is stored for the stimulus. In processing the word *train* in a context such as: 'Does the word in capitals rhyme with *Spain*? – TRAIN', subjects engage in shallow processing. As a result, they process merely some phonological features of the word and hence are not likely to consider assumptions about their knowledge of trains (e.g. that they are a mode of transport, they go from station to station, they move on railway tracks, etc.). In processing the same word in a context such as: 'Does the word in capitals fit the sentence: "the ___ arrived at 10:00pm?" – TRAIN', subjects engage in deep processing. They are, therefore, likely to consider some assumptions about trains, but unlikely to pay close attention to the word's phonological features (although they need to construct the phonological representation before arriving at the semantic representation). The claim of the Levels of Processing hypothesis is that memory for the word *train* would be better in the deep processing condition because semantic processing leaves stronger traces than structural processing. The claim of the Encoding Specificity Principle is that the effectiveness of the recall cues would depend on whether the information in the cue was present at encoding. For the case in which the target word *train* was presented in a rhyming context, cues rhyming with it (e.g. *plane, Spain*) are likely to be more effective than cues related to the meaning of the target item (e.g. *station, journey*). For the case in which subjects paid attention to the meaning of the target item during encoding, the opposite effect should be found. In fact, since the Encoding Specificity Principle claims that for an item to be retrieved the recall cues employed should contain information that had been present at encoding not just any cue related to people's knowledge of trains

should be effective in every case. I would predict that if one follows the Encoding Specificity Principle argument, effective recall cues for the word *train* when presented in the semantic context above (e.g. cues related to train schedules) would not necessarily be effective in retrieving the same word when encoded in a context such as: 'Does the word in capitals fit the sentence: "the passenger was run over by the ___?" – TRAIN', and vice versa.

Possibly the most obvious problem with the Levels of Processing hypothesis is the risk of circularity: what is processed deeply is optimally retrieved; what is optimally retrieved must have been processed deeply. Part of the problem lies in the definition of 'depth'. What is depth? Craik and Tulving (1975) tested the hypothesis that depth could be defined in terms of the time it takes to process the stimulus. The longer one takes to process, the more deeply one would process the stimulus, and the better the predicted recall for that stimulus. They tested the hypothesis with an experiment designed so that the structural task would necessarily take longer to perform than the semantic task. The structural task involved deciding whether a certain word fitted a certain CV template (e.g. 'is the word *train* a CCVVC word?'). The results show that although subjects spent longer on the structural decision task, words presented under the semantic condition still exhibited better recall. We can conclude from this that depth cannot be defined in terms of the length of time spent in processing. But if depth is not purely a function of time, what is it? In answer to this, Craik replies that "the concept of depth of processing is not hard to grasp – "deeper" refers to the analysis of meaning, inference, and implications, in contrast to "shallow" analysis such as surface form, colour, loudness, and brightness." (Craik, 2002: 308). It is therefore the engagement in inferential work that seems to have a positive effect on memory retention.

The Levels of Processing hypothesis has been extremely influential, and also widely criticised (see Baddeley, 1978 and papers in *Memory* 2002 vol. 10 (5/6) for more recent reviews). It is worth pointing out that the terminology used by the theory is rather confusing: although the expressions 'levels of processing' and 'depth of processing' are used interchangeably, the notions of 'level' and 'depth' are intuitively rather different. The term 'level' suggests a qualitative distinction, namely between two (or more) different cognitive mechanisms (e.g. one perceptual and one conceptual or inferential).

The term 'depth', on the other hand, suggests that depth of processing is a matter of degree, with the same item processed more or less shallowly or more or less deeply on a certain occasion. Which possibility did Craik and Lockhart have in mind? It seems that the Levels of Processing hypothesis favours the idea that different levels are qualitatively different: "deeper processing is not simply an extension or prolongation of shallow processing" (Craik, 2002). Still, Craik and Tulving agree that it is highly likely that both stimulus-driven bottom up processing and conceptually driven top-down processing interact (Craik and Tulving, 1975; Craik, 2002). One would like to think, in line with the Levels of Processing hypothesis, that our minds have dedicated mechanisms for processing different types of stimuli, e.g., perceptual, phonological and semantic representations. The main concern of the Levels of Processing hypothesis has been to show that in processing the meanings of words, subjects engage in deep-level processing which they do not engage in when paying attention merely to the formal features of words and that this deep-level processing improves memory retention, so that words which are processed conceptually are more likely to be remembered than those processed perceptually.

5.1 Beyond the Depth of Processing Hypothesis

The idea that an input may be processed to different depths is interesting for reasons not considered by the Levels of Processing hypothesis. We have seen that people generally process different bits of information associated in memory with a certain stimulus (e.g. a word) on different occasions. On the basis of this evidence, I would like to propose that it is necessary to distinguish between 'levels' of processing and 'depth' of processing. On the one hand, in line with the Levels of Processing hypothesis, the level of processing of a certain stimulus can be taken to refer to the qualitatively different processes (e.g. perceptual, semantic, linguistic) involved in processing this stimulus. On the other hand, information may be processed to different depths *within* a single level, so that only a subset of the information made available at this level is actually considered. The experimental evidence on the flexibility of concepts and the effect of encoding specificity seems to be consistent with this division. In fact, in showing that in processing a word in context people generally consider only a few assumptions associated in memory with the

words being used (e.g. the assumption that pianos are heavy, the assumption that birds can be edible), these experiments suggest that "semantic" processing is generally relatively shallow.

Further evidence of people's ability to process information to different depths and of their tendency to process information in a relatively shallow manner comes from hearers' failure to spot mistakes, slips of the tongue and incongruencies such as those in (3)-(6):

(3) I saw the accident. It was awful, I was just feeding the penguins in the park when this car came out of nowhere and killed the poor lady. (Pigeons) – Adapted from an example in Wilson (2000).
(4) I was really angry with their decision. It is clear that we are not getting any money from them; research funds for Linguistics are as scarce as pig's teeth. (Hen's teeth) – Adapted from an attested example cited in Glucksberg (2001).
(5) An electric train is travelling at 100 kilometres per hour from London to Oxford. In which direction is the smoke going? (Electric trains do not emit smoke) – Famous logic problem.
(6) There has been a terrible plane crash. The plane has come down right on the border between France and Spain. Where should the survivors be buried? (Survivors are alive) – Famous logic problem.

The slip of the tongue in (3) might have been caused by a mixture of things including the semantic relation between the words (both refer to animals), their syntactic relation (both are nouns) and especially their phonological relation (both have similar pronunciation). The slip of the tongue in (4) seems to be largely semantic, as both pigs and hens belong to the same category of 'farm animals'- also, they are both monosyllabic). The point of the examples is, however, not to explain mistakes in production but to explain why people typically fail to spot these mistakes and incongruencies in normal comprehension. The hypothesis that people process information to different depths sheds light on these examples. We can argue that the hearer of (3) may not spot the misuse of the word *penguin* for *pigeon* because he has processed the linguistic information associated with

the word rather shallowly, merely perceiving that the word contains certain letters or certain sounds. He may also have considered very little of the logical and encyclopaedic information associated with the word in memory (e.g. property of being a bird). In this case, the reason for shallow processing seems to lie in the nature of the surrounding text. The most important information to be recovered in interpreting the utterance is not that pigeons are being fed but that a lady was killed. In (4), shallow processing seems to be linked not just to accompanying textual information but also to familiarity with the meaning of the expression being used (*as scarce as hen's teeth* (roughly, 'very scarce')). The fact that the idiom is frequently processed makes its meaning highly accessible, speeding up the interpretation process and preventing the hearer from spotting the misuse of the word *pig* for *hen*. As a result, the hearer does not access the assumption that 'pigs have plenty of teeth', and (hence), misses the incongruency.

The last two cases are somewhat different, in that they have been deliberately constructed to test whether people can spot the anomalies. In fact, they have been constructed in such a way as to make people miss the anomalies. Once more, people who fail to notice the incongruencies in (5) and (6) seem to do so because, in processing the text, they do not access certain assumptions, namely the assumption that electric trains do not emit smoke in (5), and the assumption that survivors are alive (intrinsic to the meaning of the word *survivor*) in (6). Other examples of illusion often considered in the literature include questions, such as the following:

(7) Can a man marry his widow's sister?
(8) How many animals of each kind did Moses put in the ark?

Subjects' typical answers to these questions are 'yes', in (7) and 'two', in (8). Not noticing the anomaly is an indication that in processing the question, subjects did not access certain essential information, namely that a widow's husband is dead and that it was Noah (and not Moses) who put animals in the ark.[6]

[6] Semantic analysis of examples like these have been proposed (e.g. Erickson and Mattson, 1981; Sanford, 2002; Sanford and Sturt, 2002) but see Allott and Rubio Fernández (2002) for an interesting pragmatic (relevance-theoretic) approach.

Examples (3)-(8) fit well with the picture of cognition developed in this chapter, for they seem to indicate that utterance comprehension is also a rather selective process. A word is not generally processed fully and deeply before moving on to the next. Instead, it gives access in memory to a range of information, some of which is actually considered and used on-line. What we need is a theory of utterance interpretation capable of integrating with this view of human cognition, which is also capable of explaining what determines the depth to which an utterance or a word is processed on a particular occasion.

6. Conclusion

In this chapter, I have presented evidence in favour of the view that our minds are creative and capable of constructing new mental representations on the fly (e.g. representations denoting categories we construct on the spot) by accessing different subsets of (encyclopaedic) information in memory, while skipping others. Although the study of selective processing, depth of processing and of their relation to creative cognition is essential to the understanding of how our mind works, little attempt has been made within either psychology or linguistic theory to integrate this view on human cognition with work on ostensive communication. Little attention has been given, for instance, to explaining the mechanisms underlying the selection of information during on-line (utterance) processing, or the principles that determine how much or how little information is to be considered on a particular occasion. In the next chapter, I will present a pragmatic framework grounded in an account of human cognition that suggests some solutions to these problems.

Chapter 2

Relevance Theory: Communication and Cognition

> "As a result of constant selection pressure towards increasing efficiency, the human cognitive system has developed in such a way that our perceptual mechanisms tend automatically to pick out potentially relevant stimuli, our memory retrieval mechanisms tend automatically to activate potentially relevant assumptions, and our inferential mechanisms tend spontaneously to process them in the most productive way."
>
> Deirdre Wilson and Dan Sperber (2002: 254)

1. Introduction

In this chapter, I present an outline of Sperber and Wilson's Relevance Theory (Sperber and Wilson, 1986/1995), an approach to the study of human communication which is grounded in a view of human cognition that sheds interesting light on the psychological evidence presented so far. Sperber and Wilson propose that the human mind has evolved in the direction of increasing efficiency and is now set up in such a way that its attention and cognitive resources tend to be automatically directed to the processing of information which seems relevant at the time. This relevance-driven processing of stimuli in general is exploited in human communication and comprehension where the hearer's investment of effort, attention and cognitive resources is oriented to deriving the interpretation that the speaker intended to convey. The relevance-driven comprehension process generally involves the selective processing of available information, often resulting in the construction of new representations (e.g. new conceptual representations). The theoretical assumptions about human cognition and communication which are presented here in outlining the relevance-theoretic framework provide background to the arguments of this dissertation.

2. Basic Notions of Relevance-Theoretical Pragmatics

2.1 The Cognitive Principle of Relevance and the Definition of Relevance

Relevance Theory (Sperber and Wilson, 1986/1995) is grounded on the assumption that our minds have evolved in the direction of increasing efficiency and are now set up so that they tend automatically to maximise relevance. According to Relevance Theory, any new information, whether derived from external stimuli (utterances, sounds) or from internal representations (thoughts, memory, inference) will be relevant to an individual if it yields some positive cognitive effects when processed in context (e.g. by answering a question, confirming a hypothesis, correcting a mistake, etc.). Positive cognitive effects can be achieved by the warranted contradiction and elimination of an existing assumption, by the warranted strengthening of an existing assumption or by combining with existing assumptions to yield a true contextual implication. Imagine I have some existing assumptions (1):

(1) a. Jane will probably visit me.
    b. She will probably come this summer.
    c. If Jane comes to visit me, I'll take her to see The Mousetrap.

My friend Jane phones and says (truthfully) that she is definitely coming to visit me in January. This information is relevant in the context (1) because it yields at least the following positive cognitive effects. It provides evidence which strengthens the assumption that she is coming to visit me, contradicts and eliminates the assumption that she is coming during the summer and combines with my existing assumption (1c) to yield the warranted contextual implication that I will take her to see The Mousetrap. Other things being equal, the greater the positive cognitive effects a new piece of information has for a certain individual, the more relevant that information will be to that individual at the time.

The definition of relevance in terms of positive cognitive effects alone is, however, incomplete. Deriving cognitive effects, like engaging in any mental process, involves the expenditure of mental effort, and stimuli which yield the same positive cognitive effects may differ in the amount of processing effort they require. Intuitively,

the amount of processing effort required makes a difference to relevance. Relevance Theory claims that, other things being equal, the less processing effort is needed to derive a given set of cognitive effects, the greater the relevance. Consider the following scenario. Mary Jennings is planning to go to the cinema with her husband and wants to know when he is coming home. She calls him at his office and his secretary tells her (truly) one of the following:

(2) a. Mr Jennings's meeting finishes at 4pm.
b. Mr Jennings's meeting finishes sometime after 2pm.
c. Mr Jennings's meeting finishes 1800 seconds after 3:30.

According to this definition of relevance, (2a) would be more relevant than (2b) for reasons of cognitive effect, as (2a) entails (2b) and hence carries all the implications (2b) would have, plus some more. Also, (2a) is more relevant than (2c), this time for reasons of processing effort. Although both utterances express the same proposition and would therefore carry the same positive cognitive effects when processed in the same context, (2c) is logically and linguistically more complex than (2a), and would require the expenditure of extra processing effort.

Processing effort is the effort of perception, memory and inference required to represent an input, access contextual information and derive positive cognitive effects. Processing effort may thus be affected by a number of factors as varied as legibility, syntactic complexity, audibility, familiarity with particular dialect, register, style or construction, the accessibility of contextual assumptions, the effort of imagination involved in constructing a context, etc. On this approach, the relevance of an input to cognitive processes is a positive function of the positive cognitive effects achieved in processing this input and a negative function of the processing effort required to achieve these effects:

Relevance of an input to an individual

a. Other things being equal, the greater the positive cognitive effects (of an input to an individual who processes it) the greater the relevance (to that individual at that time).

b. Other things being equal, the smaller the processing effort required to derive those effects, the greater the relevance (of the input to the individual at that time).

The most basic theoretical claim of Relevance Theory is that human cognition, having evolved towards increasing cognitive efficiency, is now geared to the maximisation of relevance. That is, it is geared to striking the best balance between costs and effects and so to achieving the greatest cognitive effects for the least processing effort. This tendency of the human mind to allocate the available attention and cognitive resources to the processing of potentially the most relevant information is an outcome of biological evolution, and of the need to provide solutions to the problems and challenges in the environment in which early humans evolved. The result is what Sperber and Wilson call the First or Cognitive Principle of Relevance:

Cognitive Principle of Relevance

Human cognition tends to be geared to the maximisation of relevance

Although humans are capable of monitoring a wide array of features in the surrounding environment, paying simultaneous attention to every single aspect of these stimuli would result in a cognitive explosion. One consequence of the Cognitive Principle of Relevance is that the processing of available stimuli is quite selective. From all the information present in our environment at a particular moment, our cognitive system automatically picks out only those bits of information which are potentially (most) relevant to us, and processes them in a way that tends to maximise overall relevance. For instance, our perceptual systems tend to pick out sounds and sights that are potentially relevant to us (e.g. smoke in our house rather than in someone else's). Our memory retrieval mechanisms tend to activate assumptions which are potentially relevance-enhancing additions to the context in which inputs are processed (e.g. in the

above example, assumptions about fighting a fire rather than about the effects of smoke on the environment). Our inferential systems tend to derive the most relevance-enhancing cognitive effects (e.g. plans to rescue the children) from the combination of the new information and context, and so on.

This theoretical assumption about human cognition sheds light on much of the experimental work presented in the last chapter. We have seen how in processing an input, people generally focus their attention on a selected sub-set of the information available at the time. Mental effort and cognitive resources are directed towards those bits of information which are potentially relevant, or which enhance the relevance of stimuli currently being processed. We are now in a position to claim that this pattern is a by-product of a biologically evolved tendency of the mind to maximise relevance. Although psychologists often assume that what makes a certain input likely to be processed at a certain time is that it is contextually appropriate or relevant to the task being performed (e.g. Barsalou, 1993), they make no attempt to define the notion of relevance or to explain how the context in which a certain input is processed is constructed. Relevance Theory, by contrast, is built around a theoretically adequate and cognitively plausible definition of relevance. In this dissertation, I will try to show how this tendency of the human mind to maximise relevance has important bearing on human communication and the comprehension of ostensive stimuli, including utterances. Unlike existing cognitive and pragmatic models, Relevance Theory has adequate cognitive and communicative tools to account for the selective processing that clearly operates in the comprehension of utterances.

2.2 The Communicative Principle and the Comprehension Procedure

An important consequence of the Cognitive Principle of Relevance is that people can predict and manipulate, at least to a certain extent, the mental states of others. People may be able to predict, for instance, which stimulus is likely to attract a person's attention at a given time, which background assumptions he would be likely to consider in processing this stimulus, which inferences he would be likely to derive, and so on. Imagine I come home covered in mud from head to toes. I can correctly predict that my parents will notice this and that they will infer that there must be a reason why I am

covered in mud, and will ask me about this. There are at the time a number of other stimuli around them that they could notice and pay attention to, and lots of other things they could ask me about when I get home; yet I can predict, to a certain extent, which inputs they will be paying attention to, which set of assumptions they will consider in processing these inputs, and which inferences they may draw.

An essential assumption in Relevance Theory is that as a result of cognitive pressures towards the efficient processing of potentially relevant information, the human mind has developed in evolution an ability to recognise people's intentions, and more generally an ability to read the minds of others (see Sperber, 2000; Wilson, 2000; Wilson, 2002). We see a person in the street looking at a map and also around him, and infer that he is trying to locate himself; we observe a person throwing something from a distance to land near a bin and infer he intended the object to fall inside the bin, etc. The ability to recognise what other people are thinking about or paying attention to, and the ability to predict how their thoughts will develop in response to a stimulus, plays an essential role in communication.

For Sperber and Wilson, as for Grice, human communication is essentially a mind-reading activity in which an audience infers what the speaker intended to communicate on the basis of the evidence she has provided. For these authors, the most important type of human communication is overt, or ostensive, communication. The crucial feature of ostensive communication, as opposed, say, to accidental information transmission, is that the communicator not only has an informative intention but also intends the hearer to recognise that she has this intention (in Sperber and Wilson's terms, she has not only an informative but also a communicative intention). Consider, for instance, the following situation. We are at a party and I ostentatiously yawn at you, point to my watch and close my eyes. Alternatively, I say: 'I am tired'. You see me doing or saying this, and infer that I mean that I am tired and that I want to go home. In this example, my informative intention would be to inform you that I am tired and want to go home and my communicative intention to inform you of my informative intention (i.e. the communicative intention is to have the informative intention recognised). Imagine again the scenario above; except that on this occasion you infer that I am tired not by seeing me produce an ostensive stimulus but merely by looking at the redness of

my eyes. Since no ostensive stimulus was produced (I did not intend to inform you of anything), this inference is derived on your own responsibility with no encouragement from me. This second scenario would not count as an act of ostensive (or overt) communication.

Relevance Theory claims that in evolving towards increasing efficiency, the human mind has developed a dedicated mechanism for recognising the intentions underlying ostensive stimuli. Among all the available stimuli in our environment, ostensive stimuli (e.g. utterances) have a special property, which is that they inevitably raise an expectation (or presumption) of relevance in an addressee. This expectation is not raised by non-ostensive stimuli, which cannot be treated as evidence of a person's intention to communicate:

Communicative Principle of Relevance

Every utterance communicates a presumption of its own optimal relevance.

It is because our minds have a dedicated mechanism for recognising ostensive stimuli as evidence of someone's intentions to communicate something that when an ostensive stimulus (e.g. an utterance) is presented to us, it automatically triggers a more or less specific expectation about how this information will be relevant to us on this occasion. In producing an ostensive stimulus, say an utterance, the speaker requests the hearer's attention. In return for the demand on his attention and processing resources, the hearer is entitled to presume that the speaker's utterance will be optimally relevant to him:

Optimal Relevance

An utterance is optimally relevant iff

a. It is relevant enough to be worth the hearer's processing effort.
b. It is the most relevant one compatible with the speaker's abilities and preferences.

Although, as posited by the Cognitive Principle of Relevance, our minds tend to be geared to the maximisation of relevance, and so aim to derive the greatest cognitive effects for the minimum processing effort, in interpreting an ostensive stimulus (e.g. an utterance), the hearer is only entitled to expect this stimulus to be optimally relevant, as defined above. The reason for this is that the choice of an ostensive stimulus in communication may be affected by the speaker's own abilities and preferences. Consider again the situation in (2b). The speaker of (2b) may not know exactly when Mr Jennings's meeting finishes, or she may be unwilling (or not allowed) to disclose that information. In this situation, (2b) may be the most relevant utterance the speaker is willing and able to produce in the circumstances. Since, given the presumption of optimal relevance, the hearer of (2b) is entitled to expect the speaker's utterance to be the most relevant one she could have produced based on her abilities and preferences, and in any case, at least relevant enough to be worth processing, he might accept the interpretation derived from (2b) as relevant enough to satisfy the particular expectations raised by the utterance.

As noted above, it follows from the Cognitive and Communicative Principles of Relevance that a speaker should be able to predict, to some extent at least, what kind of information a hearer is likely to pay attention to when she speaks, what background assumptions he is likely to use, and what inferences he is likely to draw. As a result, she should be able to formulate an utterance (depending on her own abilities and preferences) which will enable the hearer to derive the cognitive effects that she intends him to derive for a minimal amount of processing effort. On the other hand, on the assumption that the speaker is aiming at optimal relevance and so putting him to no unnecessary processing effort in deriving the intended cognitive effects, the hearer is entitled to follow a path of least effort in deriving cognitive effects, and treat the first interpretation that satisfies his expectations of (optimal) relevance as the one the speaker intended to convey.

As noted above, Relevance Theory claims that the success of ostensive communication does not simply depend on general mind-reading abilities. According to this theory, continuous pressure towards increasing efficiency has resulted in the development of a dedicated cognitive mechanism (a module) that recognises and

processes ostensive stimuli (Sperber, 2000; Wilson, 2000, 2002, 2003; Wilson and Sperber, 2002). The relevance-theoretic comprehension procedure is given below:

Relevance-theoretic Comprehension Procedure

a. Follow a path of least effort in computing cognitive effects: test interpretive hypotheses (disambiguations, reference resolutions, enrichments, implicatures, etc.) in order of accessibility.
b. Stop once your expectations of relevance are satisfied.

This comprehension procedure is seen as having biologically developed in evolution in such a way that it is automatically triggered by an ostensive stimulus, and guides the recovery, by inference, of the hypotheses about the intended interpretation. The relevance-theoretic comprehension procedure, together with the notion of optimal relevance and the communicative principle of relevance, are the key components of relevance-theoretic pragmatics.

3. Relevance Theory and Utterance Interpretation

Work on linguistic communication has long accepted two assumptions: the assumption that human cognition is flexible and creative enough to enable us to construct an indefinite number of thoughts in our lifetime, and that natural languages are productive enough to enable us to communicate any of these thoughts. Natural languages provide us with a limited lexicon and a finite set of rules, out of which we can create an indefinite number of utterances, each capable of being used to convey a different set of thoughts. Unlike classical code theorists, modern pragmatists share the view that identifying the thought(s) a speaker intends to convey in using an utterance involves a mixture of decoding and inference (Bach, 1994, 1997; Carston, 2002a, 2002b; Grice, 1975/1989; Recanati, 1989, 2003; Sperber and Wilson, 1986/1995, 1987; Wilson and Sperber, 1993a, 1993b; Wilson, 1998; 2002). Utterances are automatically decoded by the language module into a semantic representation or logical form (a structured set of encoded concepts). This semantic representation or logical form generally

underdetermines the thoughts that the speaker intended to convey. On the one hand, an utterance typically underdetermines what is explicitly communicated by the speaker. On the other hand, it may underdetermine also what is implicitly communicated, as in (3):

(3) a. *Tom:* Has John left yet?
b. *Mary:* His car is parked over there.

The semantic representation obtained from the linguistic decoding of (3b) is sub-propositional and needs to be enriched pragmatically at the explicit level so as to yield a complete thought capable of being true or false. This involves, for instance, assigning reference to the pronoun 'his' and the adverb 'over there', as in (4):

(4) JOHN(x)'S CAR IS PARKED OUTSIDE TOM(y)'S HOUSE ON THE 7TH MARCH 2005.

Presumably, the speaker of (3b) does not only want to convey the thought in (4); she also wants to communicate the thought in (6), which can be derived by adding an extra premise to the context, as in (5), and combining it with the proposition explicitly expressed in (4):

(5) JOHN(x) IS LEAVING BY CAR.
(6) JOHN(x) HAS NOT LEFT YET.

The distinction between explicit and implicit content has generated considerable discussion within linguistics, philosophy and psychology (e.g. Bach, 1994, 1997; Carston, 1998a, 1998b, 2000, 2002a, 2002b; Gibbs, 1999a, 1999b, 2002a; 2002b; Gibbs and Moise, 1997; Levinson, 1983, 2000; Recanati, 1989, 2003). Virtually every scholar working on pragmatics nowadays accepts what has become known as the 'underdeterminacy thesis'; the assumption that the sentence (or other linguistic expression) uttered does not fully determine the proposition a speaker explicitly expresses. They disagree, however, on the role that pragmatic processes play in fleshing out the encoded logical form into the proposition explicitly expressed. In line with

Grice, modern scholars generally agree that to arrive at the explicit content of an utterance, the hearer needs to disambiguate ambiguous terms and assign reference to indexicals, as in (7)-(8) and (9)-(10):

(7) Peter went to the *bank.*
(8) The inspector was looking for *bugs* in the room.
(9) *It* is *there.* [The TV magazine is on the bedside table]
(10) *She* is pretty. [Mary Evans]

For some pragmatists (e.g. Grice, 1975/1989), the fleshing out of the logical form stops here. Many modern scholars agree, however, that some further pragmatic enrichment of the logical form, in which extra conceptual material is built into the explicit content of the utterance, is often needed to arrive at a proposition capable of being true or false (e.g. Bach, 1994; Recanati, 2003). Relevance theorists have also claimed that the pragmatic enrichment of a logical form does not stop at the point where a full proposition is obtained but may go further in order to satisfy the hearer's expectations of relevance. Examples such as (11) and (12) illustrate this point:

(11) *Kath:* Would you like to have some lunch?
Mary: No, Thanks, I haven't eaten.
(12) a. Mary Smith has eaten some food in her lifetime.
b. Mary Smith has eaten lunch today.

Unlike some modern pragmatists, (e.g. Bach, 1994), relevance theorists do not claim that the logical form of (11B) is enriched simply to the point where the minimal proposition in (12a) is obtained but treat it as enriched into the more specific proposition in (12b), where the period of time involved and type of food eaten are further narrowed down in context. According to Relevance Theory, then, what is accepted by the hearer as explicitly communicated is the first inferential enrichment of the encoded sentence meaning that yields enough cognitive effects to satisfy the hearer's expectations of

relevance, as in (4) and (12b) (Carston, 1988, 1998a, 2002a; Sperber and Wilson 1986/1995; Wilson and Sperber, 2002).[1]

In relevance-theoretic terms, a proposition which is explicitly communicated is an 'explicature', and a proposition which is not explicitly but implicitly communicated is an 'implicature'. In Relevance Theory, an implicature may be either an implicated premise (intended contextual assumption), as in (5), or an implicated conclusion (intended contextual implication), as in (6). Implicated premises may be more or less accessible to the hearer at the time of processing and implicated conclusions may be derived with different degrees of strength (the weaker the implicature, the more the responsibility of the hearer in its derivation). Constructing a hypothesis about the interpretation the speaker intended to convey on a particular occasion involves thus the following sub-tasks:

(13) a. Constructing an appropriate hypothesis about explicit content (explicatures), as in (4).
b. Constructing an appropriate hypothesis about the intended contextual assumptions (implicated premises), as in (5).
c. Constructing an appropriate hypothesis about the intended contextual implications (implicated conclusions), as in (6).

A common assumption in pragmatic theory is that an implicature is derived AFTER the explicature has been determined (e.g. Grice, 1975/1989; Levinson, 2000). A crucial claim of Relevance Theory is that explicatures and implicatures are derived in parallel, via a process of mutual adjustment regulated by the relevance-theoretic comprehension procedure given above, and guided by more or less precise expectations of relevance raised by the utterance (e.g. Wilson, 2000; Wilson and Sperber, 2000; 2002). According to this view, the hypotheses about explicit content, context and implicatures in (13) are not seen as sequentially ordered. That is, the hearer does not first decode the complete logical form of the utterance, then construct an explicature and select an appropriate context, and then derive a range of implicated conclusions (or

[1] For excellent detailed insight into the contribution of pragmatics to explicit communication see

other positive cognitive effects). Instead, following a path of least effort in deriving cognitive effects, the hearer considers hypotheses about explicit context, content and implicatures in order of accessibility, and stops when he has enough implications to satisfy his expectations of relevance.

To illustrate this mutual adjustment process, let's analyse the interpretation of (14) in the following situation: Jill and Larry are lovers who have killed Larry's depressive wife Kate to inherit her fortune and made it look like suicide. The day after Kate's death, a police inspector arrives at the house to search for clues about Kate's death:

(14) *Larry:* I saw you were flirting with the inspector. What was that about?
Jill: He was looking for bugs in the house.

The interpretation of Jill's utterance requires Larry to construct a hypothesis about the set of thoughts she may have intended to communicate, both explicitly and implicitly. This may be done as follows:

(15)[2]

(a) Jill has said to Larry, "He_x was looking for $BUGS_1$/ $BUGS_2$ in Kate's house."	*Embedding of the decoded (incomplete) logical form of Jill's utterance into a description of Jill's ostensive behaviour.*
b) Jill's utterance will be optimally relevant to Larry.	*Expectation raised by recognition of Jill's ostensive behaviour and acceptance of the presumption of relevance it conveys.*
(c) Jill's utterance will achieve relevance by explaining why Jill had to flirt with the inspector.	*Expectation raised by (b), together with the fact that such an explanation would be most relevant to Larry at this point.*

Carston (2002a).

[2] I am not claiming here that this is necessarily the sequence in which comprehension occurs. According to Relevance Theory, the mutual adjustment takes place in parallel, rather than in sequence.

(d) If a policeman is looking for $BUGS_2$, it means he has some suspicions.	*First assumption to occur to Larry which, together with other appropriate premises, might satisfy expectation (c). Accepted as an implicit premise of Jill's utterance.*
(e) For a policeman to find $BUGS_2$ in a house would confirm his suspicions.	*Second assumption to occur to Larry which, together with other premises such as (d), might satisfy expectation (c). Accepted as an implicit premise of Jill's utterance.*
(f) The best way of distracting someone is to flirt with him.	*Third assumption to occur to Larry which, together with other premises such as (d) and (e), might satisfy expectation (c). Accepted as an implicit premise of Jill's utterance.*
(g) The inspector was looking for $BUGS_2$ in Kate's house.	*First enrichment of the logical form of Jill's utterance to occur to Larry which might combine with (d)-(f) to lead to the satisfaction of (c). Accepted as an explicature of Jill's utterance.*
(h) Jill was flirting with the inspector to distract him from finding the bugs.	*Inferred from (g) and (d)-(f), satisfying (c) and accepted as an implicit conclusion of Jill's utterance.*
(i) Larry and Jill are in danger of being discovered.	*From (g) plus background knowledge. One of several implicatures of Jill's utterance which, together with (h), may satisfy expectation (b).*

Larry assumes that Jill's utterance will be optimally relevant to him by answering his question and thus providing a reason for her inappropriate behaviour with the inspector. Following a path of least effort in computing cognitive effects, he starts considering hypotheses about explicit context, context and implicatures in their order of accessibility. When potential implicated premises are considered, as in (d)-(f), the hearer enriches the explicit content on this basis, as in (g). When a contextual assumption is added as premise to the context, it combines with explicit content to derive a potential implicature of the utterance, as in (h) and (i). If the combination of explicit content, context and implicatures derived at this point satisfies Larry's expectations of relevance, he stops processing. If not, he will continue considering the next most accessible contextual assumption, the next most accessible enrichment, the next most accessible implicature and so on until the process of mutual adjustment provides a combination which satisfies his expectations of relevance, at which point he will stop.

4. Accessibility of Contextual Assumptions

I have tried to show how the relevance-driven comprehension of utterances is a considerably selective process which does not aim to consider all the possible contextual assumptions or derive all the plausible contextual implications (that would result in cognitive explosion). Instead, the comprehension process follows a path of least effort till it arrives at the right combination of explicit content, context and implicatures. Although the process is one of non-demonstrative inference, the expectations of relevance and the relevance-theoretic comprehension procedure impose strong pragmatic constraints which guide the hearer towards the set of assumptions the speaker might have intended to convey. Since the relevance-theoretic comprehension procedure starts by considering only highly accessible contextual assumptions, to understand the utterance comprehension process we need to look at the factors which play a role in ordering or altering the accessibility of the assumptions a hearer may consider in interpreting an utterance. I look here at cognitive and communicative factors which may affect accessibility in interpretation.

In line with much research in psychology, Relevance Theory assumes some version of a spreading activation model of memory, a view of the model which is consistent with the Cognitive Principle of Relevance assumed in Relevance Theory. Since our minds are almost always processing new information, a subset of our stored encyclopaedic assumptions and concepts is likely to be active at any given point, with different subsets being activated at different points. On this approach, our memory is seen as organised in such a way that the activation of a certain concept (e.g. DOCTOR) immediately activates semantically related concepts (e.g. NURSE, HOSPITAL), which themselves activate related concepts to different degrees. Evidence for this spreading activation is found across studies of memory (e.g. Collins and Loftus, 1975; Meyer and Schvaneveldt, 1971; Ratcliff and Mcknoon, 1981). Free-recall experiments and word decision tasks are clear cases. For instance, subjects have been reported to take considerably less time (85 ms less) to decide whether a word (e.g. *butter*) is an English word when it is presented after a semantically related word (e.g. *bread*) than when it is presented after a semantically unrelated word (e.g. *nurse*) (Meyer and Schvaneveldt, 1971).

The activation of a concept (or conceptual address) gives access in memory to a set of encyclopaedic assumptions about the entities the concept denotes which will themselves be activated. The activation of the concept DOCTOR, for instance, activates encyclopaedic assumptions such as 'doctors work long shifts', 'they wear white coats', 'they cure patients', 'they work in hospitals', etc. The spreading activation view of memory assumes that our mind is organised in such a way that the degree of activation of these encyclopaedic assumptions would be affected by the prior activation of other concepts and their presence in working memory. For instance, processing the concept DOCTOR after processing the concepts NURSE and HOSPITAL would enhance the activation of a subset of assumptions about doctors (i.e. those connected with nurses and hospitals) rather than others. Since different sets of concepts and assumptions are present in working memory at each point, different subsets of encyclopaedic assumptions receive different degrees of activation at different times. Consider the following examples:

(16) *A:* I always like dressing up during carnival but I am tired of uncomfortable heavy costumes. I prefer something simple. What do you recommend?
B: Why don't you go as a doctor?

(17) *A:* My father is having problems with his knee again. He can hardly go upstairs.
B: Why don't you take him to see a doctor?

(18) *A:* All the boyfriends I have had have been vulgar and broke. I really deserve something better, someone a little more sophisticated.
B: I'll introduce you to my cousin. He is about your age, single and a doctor.

In (16), activation of the concepts encoded by the words DRESS, COSTUME and CARNIVAL would activate a further range of concepts related to fashion and festivities. Processing the concept DOCTOR in this context may lead to the increased activation of some of our encyclopaedic assumptions about doctors (e.g. that they dress in long white coats). Processing the same concept DOCTOR in the context of another set of concepts such as PAIN, KNEE in (17) may result in other sub-sets of assumptions about doctors being more highly activated (e.g. the assumption that doctors heal pain) and so on.

Because the associative spreading of activation through memory and recency of processing have an impact in the accessibility of information, they should affect utterance interpretation. It is important to bear in mind, though, that the process of spreading activation is an automatic one which reflects the way human memory is organised and operates regardless of people's intentions or their assumptions about other people's intentions. Spreading activation is therefore at work not just in the processing of utterances or other ostensible stimuli, but in the processing of information more generally. Therefore it operates independently of the relevance-theoretic comprehension procedure, though it may affect the premises used in the interpretation process.

When processing an ostensive stimulus, for example an utterance, the accessibility of assumptions in memory may be crucially affected by the expectations of optimal relevance raised by that stimulus (e.g. that utterance). That is, it is affected by more or

less specific expectations about the type and level of cognitive effects the utterance may yield in the circumstances. In interpreting (16), for instance, A would normally expect B's utterance to achieve relevance by answering her question, and more specifically, by recommending a type of costume. Similarly, in interpreting (17) and (18), A would expect B's utterance to achieve relevance by commenting on his problem or by providing a piece of advice. The expectations of relevance raised by the utterance play a major role in interpretation. Despite all the information potentially available at the time of the utterance, the hearer's expectations of relevance are powerful enough to narrow the search space and add an extra layer of activation to some of the information available at the time. The expectations raised by an utterance lead the hearer to construct a hypothesis about the speaker's meaning by enhancing the accessibility of assumptions that might help to satisfy those expectations of relevance (e.g. in (18), the assumption that doctors are of high social status, that they earn good money, that they are intelligent, etc.). This subset of assumptions would then be considered in the regular way by following a path of least effort as laid down in the relevance-theoretic comprehension procedure. The assumptions which are not likely to help satisfying the hearer's expectations of relevance are less likely to be considered, and if considered, they are likely to be rejected for not satisfying those expectations and helping to make the utterance relevant in the expected way.

Not only do expectations of relevance contribute to altering the accessibility of existing assumptions, they also provide access to assumptions which wouldn't be available otherwise. Consider (19) and (20):

(19) *A:* Look at those costumes, they are great!

B: I love carnival, Pity I have to miss the parade this year. We doctors have the most unfair timetables.

(20) *A:* All the boyfriends I have had have been vulgar and broke. I really deserve something better, someone a little more sophisticated.

B: I think they were all great guys. You should see a doctor. There may be some insecurity issues you need to deal with.

As in the example (16) above, processing the concept DOCTOR after having processed concepts such as PARADE, CARNIVAL, etc. may increase the activation of assumptions about doctors which are related to culture, fashion, carnival, etc. The expectation that B will provide a reason for why she has to miss the parade should also increase the accessibility of encyclopaedic assumptions about doctor's shifts, schedules, response to emergencies, etc. These assumptions should be added to the context in order of accessibility and used as premises to derive the implications that the speaker may have intended to convey. (20) is a similar example. In that the accessibility of the assumptions considered during the interpretation process ultimately depends on the specific expectations of relevance raised by that utterance in that situation. In this case, the expectations of relevance are themselves likely to be revised as the utterance proceeds, eliminating some of the assumptions automatically activated and increasing the activation of others.

The way that expectations of relevance raised by an ostensive stimulus guide the interpretation process is an aspect of our cognitive make-up that plays a major role in utterance comprehension, but one which is hardly acknowledged by accounts of utterance interpretation outside relevance-theoretic pragmatics. Understanding an utterance is a creative process, as it involves the selection of a particular set of contextual assumptions, and the construction of a particular combination of explicit content, context and implicatures, all constrained by considerations of relevance. As we have seen in the first chapter, selective processing and the one-off assembly of different sub-sets of information is generally linked to creativity, and particularly to the construction of ad hoc categories and new representations to denote those categories. In the next section, I will argue that utterance interpretation is also a creative process in that it often involves the construction of ad hoc concepts and that these ad hoc concepts are also a by-product of the relevance-driven mechanisms at work in interpreting utterances.

5. Lexical Pragmatics

Although most modern pragmatic approaches acknowledge the mismatch between the logical form encoded by an utterance and the thoughts conveyed in producing that utterance, very little attention has been paid to the pragmatic mechanisms involved in constructing the concepts which are taken to be the constituents of those thoughts. The reason for this is that it is generally assumed that the concept expressed in using a word is the very same concept encoded by that word. The flexible view of human cognition presented in the last chapter suggests, however, that, in many cases at least, people often use a single unambiguous word (e.g. *bird*) to express a wide range of different concepts, each one with its own denotation and graded structure, as in (21)-(24):

(21) a. The *fish* attacked the swimmer. (Narrowed to shark-like fish)
b. The *fish* was nice but the potatoes were cold. (Fish typically served in restaurants)
c. Please feed the *fish* in my room twice a day. (Goldfish-like fish)

(22) a. I like listening to the *birds* in the morning. (Robins, canaries)
b. The *birds* flew above the waves. (Seagulls)
c. She was feeding the *birds* in the square. (Pigeons)

(23) *Red* hair, *red* car, *red* apple, *red* tomato, *red* eyes, etc. (Understood to denote different shades of red covering different aspects of an object)

(24) *Cut* the grass, *cut* my hair, *cut* the cake, etc. (Understood to denote different manners of cutting involving different instruments)

Solving this divergence between psychological reality and theories of utterance interpretation is imperative. This is because a theory which is capable to account for how the thoughts a speaker intended on a certain occasion are derived but not for how the constituents of those thoughts are derived cannot account for utterance interpretation successfully.

I argued in the last chapter that two concepts differ in content if they differ in denotation, and so pick out different sets of entities in the world. These differences in denotation should affect the truth-conditions of utterances in which they occur. What is

important about the examples above is that a single word is used to express a range of different concepts, each making a different contribution to the truth-conditional content of the utterance in which they appear. We can assume, for instance, in line with Searle (1980), that if I ask you to cut my grass and you take the kitchen scissors and cut the grass in the way you would cut a someone's hair, you stab it with a knife or you cut out a whole piece of grass and leave an empty square behind, you will not be interpreting my utterance in the intended way. To test the different contributions that the concepts conveyed by using a single word *cut* on different occasions may make to the truth-conditional utterance in which they appear, consider (25)-(27):

(25) If you cut the grass, I'll give you some money.
(26) If you cut my hair, I'll give you some money.
(27) If you cut the cake, I'll give you some money.

If (25), (26) and (27) were statements in a contract, I would only be obliged to pay the sums agreed on the basis that you had CUT* my grass, CUT** my hair and CUT*** the cake, where each of the concepts conveyed by the word *cut* denotes a different manner of cutting, involving different instruments. If you CUT* my cake with a lawnmower, CUT** a slice of my hair with a knife or CUT** my grass by cutting the tips with small scissors, you would not be entitled to payment.[3] We can conclude that an adequate pragmatic theory should incorporate some account of how hearers construct the concepts they understood the speaker to have expressed in using a word, on the basis of the concept encoded by that word, plus contextual assumptions and more general pragmatic principles.

A crucial aspect of current lines of research in Relevance Theory is that the distinction between the two processes of decoding and inference (which in this theory corresponds to the distinction between linguistic semantics and pragmatics) holds not only at the level of the sentence but also at the level of the word (Carston, 1996, 2002a, 2002c; Sperber and Wilson 1986/1995, 1998; Wilson, 1998, 2004; Wilson and Sperber,

[3] For semantic and pragmatic analysis of this phenomena as well as reviews on the existent literature see also Pustejovsky (1995) and Recanati (2003).

2002). The approach to lexical pragmatics defended in this framework abandons the code-like assumption that the concept expressed by the use of a word is always the very same concept encoded by that word; instead, the encoded concept is seen as acting merely as clue to the intended concept. A more general idea underlying this hypothesis is that, contrary to what has been traditionally assumed in linguistics and philosophy, the stock of concepts we can represent in our minds, and are therefore capable of communicating, is much greater than the stock of words available in a given language to encode those concepts.

Sperber and Wilson describe a concept as consisting of a conceptual address, a constituent of conceptual representations which may give access in memory to three different types of information: lexical, logical and encyclopedic. The linguistic entry of a concept gives access to linguistic information about grammatical category, phonological representation, etc. The logical entry specifies the (one-way) logical relations the concept has with other concepts. Finally, the encyclopaedic entry contains a range of assumptions the person has stored about the denotation of the concept (e.g. the category of birds in the world):

(28) *Conceptual address:* BIRD
Linguistic entry: Noun; [bɜ:d]
Logical entry: one-way inferential links to other concepts ANIMAL
Encyclopaedic entry: information about the denotation (e.g. it typically flies, it has feathers, it typically sings, it can be of different colours, etc.)

Not every concept we are capable of forming in our minds will, however, have all three entries. Some concepts do not have a logical entry, as may be the case with proper names (LONDON, NOAM CHOMSKY). Some concepts do not have an encyclopaedic entry: for instance, a logical concept such as AND. Crucially, some concepts also lack a lexical entry. In other words, many of the concepts we can represent in our minds and use as constituents of our thoughts have no corresponding linguistic form. These unlexicalised concepts may be of two types. On the one hand, there are concepts which

have a stable and permanent entry in our minds, yet remain unlexicalised in our public language. We might be able to recognise a particular state of mind, type of pain, degree of happiness, mood, etc., have a stable representation for it (e.g. PAIN*, HAPPY*) and be able to derive inferences from that representation each time we encounter it, but still have no word for it. On the other hand, there are concepts which do not have a stable conceptual address in memory but are constructed in our minds at a moment's notice using stable concepts as templates. We might represent, for instance, different ad hoc concepts for pain (e.g. PAIN**, PAIN***) on different occasions, different ad hoc concepts for states of tiredness (e.g. TIRED*, TIRED**) and so on. On this approach, the mind may be extremely rich in unlexicalised conceptual resources, and these may play a role in communication and comprehension, as I will argue below.

5.1 Stable Concepts and Ad hoc Concepts

There are at least three ways of viewing the relation between our stock of concepts and our stock of words (see Sperber and Wilson, 1998). According to the Classical View, the relation seems to be one-to-many: that is, few basic concepts and many words. On this view, humans have a very limited range of primitive concepts (e.g. ADULT, MALE, FEMALE, MARRIED) which combine to provide the meanings of a great many words (e.g. *woman, man, wife, husband*, etc.). The view that the relation between our stock of concepts and our stock of words is roughly one-to-one is possibly the most popular in current research on linguistics and philosophy. It is the one defended by Fodor (Fodor, 1998) and perhaps the one that is held by most pragmatic approaches outside Relevance Theory. A consequence of this view is that the concept expressed by use of a word is the same as the concept encoded. Relevance Theory suggests a third possibility which is confirmed by some of the psychological evidence presented in the last chapter: the number of concepts we can construct and represent in our mind is much greater than the number of concepts which are lexicalised in our languages. It is because of this mismatch that a single word (e.g. *bird, cut, red*) may be used to convey a range of different concepts, with different denotations, on different occasions.

Unlike many existing pragmatic approaches, Relevance Theory not only acknowledges the gap between the concept encoded by a word and the concept

expressed by a speaker in using that word on a particular occasion, but also aims to provide an explanation of how the hearer bridges the gap between the concept encoded and the concept expressed. This gap may arise in at least two ways. In the first place, the concept encoded by a word may be more general than the concept the speaker intends to convey by using the word on that particular occasion, as in (22)-(25). In this case, the encoded concept needs to be narrowed down in context to pick out only a subset of its linguistically-specified denotation (lexical narrowing). In the second place, the encoded concept may be narrower or more specific than the concept the speaker intends to convey on that occasion. In this case, the hearer may need to broaden the denotation of the encoded concept so that it includes a greater range of entities. Cases of lexical broadening include those in (29)-(30):

(29) I was born with a *square* mark on my foot. (Roughly square)
(30) There is a *round/oval/square/rectangular* stain on the kitchen floor. (Roughly round, oval, square, and rectangular)
(31) The salad was delicious but the lamb was *raw*. (Not cooked enough)
(32) You are a *genius*! (Very clever)
(33) I have *thousands* of things to do. (A lot)
(34) I am *starving*. (Very hungry)
(35) My daughter is an *angel*. She sleeps all night through.
(36) My mother is a *witch*, she always knows what I am up to.
(37) My husband has *married* his computer.
(38) Getting married and settling down will kill her. She is a *butterfly*.

(29)-(31) would generally be classified as instances of approximation. In these cases, a word has a strict definition (e.g. exact number, geometric figure, etc.) but is used on a specific occasion to denote a broader set of entities, some of which (strictly speaking) fall outside that definition. In (29) for instance, the word *square* is used to denote not only perfect geometric figures but also things which are roughly square. Similarly, in (31), the word *raw* is used to denote not only uncooked food but also food that, although not totally uncooked, is not cooked enough for the speaker's taste. More

radical cases of concept broadening include category extension, in which the denotation of the encoded concept is broadened to include not a just a few cases which almost fall under the definition, but a range of items which may fall well beyond the scope of this definition. Examples include hyperbolically or metaphorically intended words, such as those in (33)-(34). The denotation of the concept expressed in (33), for instance, should be broad enough to include not only real geniuses but people who are very clever; the denotation of the concept expressed in (34) should be broad enough to include states in which someone is quite hungry; the denotation of the concept expressed in (35) should include people who are extremely nice and easy to deal with; the denotation of the concept expressed in (38) should denote people who are lively and enjoy freedom, and so on.

The treatments given in the existing literature to narrowing, approximation and category extension have been rather different, and each of these cases has typically been studied in isolation (see Wilson, 2004; Wilson and Sperber, 2000 for reviews). For instance, narrowing has often been seen as involving the application of default rules (Blutner, 1998, 2002; Levinson, 2000). Approximation is typically seen as a type of lexical vagueness governed by contextually-determined standards of precision (Lasersohn, 1999; Lewis, 1979). Hyperbole and metaphor are typically seen within linguistics and pragmatics as involving a blatant violation of conversational maxims (Grice, 1975/1989). Relevance Theory is a pioneer in treating narrowing and broadening as two different instantiations of a single process of pragmatic fine-tuning of the linguistically-specified meaning of a word (Carston, 1996, 2002a; Sperber and Wilson, 1998; Wilson and Sperber, 2000, 2002; Wilson, 2004; but see also Recanati, 1995, 2003).[4] This pragmatic adjustment of encoded concepts is itself a by-product of a more general process of mutual adjustment of explicit context, context and cognitive effects which, guided by the relevance-theoretic comprehension procedure and the

[4] It is worth noticing that Recanati has defended a similar view (see Recanati, 1995, 2003). He proposes the idea that local pragmatic processes may operate modulating word meanings at explicit level as part of arriving at the explicit content of the utterance (see Recanati, 1995, 2003). Modulation of senses may result in the derivation of a more specific interpretation (e.g. 'he was wearing *rabbit*' – rabbit fur) or a looser interpretation (e.g. 'the machine *swallowed* the card', *'the ham sandwich* is getting restless'). The modulation of word meanings is believed to be guided by a search for coherence and by the activation of stereotypical scenarios or schemata.

expectations of relevance raised by the utterance, helps to make the utterance relevant in the expected way.

According to Relevance Theory, when an utterance is decoded, the concepts encoded by the constituent words are activated, giving access to a range of logical implications and encyclopaedic assumptions. Following a path of least effort, the hearer starts considering these assumptions in their order of accessibility, adding them to the context in order to derive contextual implications (and other positive cognitive effects). In interpreting (21a), for instance, the hearer may access from the encoded concept FISH the assumption that fish can be dangerous; in interpreting (38), he may access the assumption that butterflies are delicate, and so on. Following a path of least effort, he may add this highly accessible assumption to the context and use it as implicit premise in deriving implications (e.g. the implication that the fish that attacked the mariner was dangerous in (21a) and the implication that Mary is delicate in (38)). When he has enough implications to make the utterance relevant in the expected way, he stops, and assumes that this was the speaker's intended meaning.

The expectations of relevance raised by an utterance may make a certain hypothesis about the intended implications highly accessible to the hearer (e.g. the implication that Mary does not like staying in the same place for too long in (38)) before a full explicature is derived. Considering this assumption as a possible implicated conclusion may help the hearer to flesh out the explicature by backwards inference and select the context in such a way that it will warrant the expected conclusion. This process of backwards inference may therefore involve looking into the encyclopaedic entry of the encoded concept BUTTERFLY and accessing the assumptions that butterflies fly, that they hardly stay put, etc. A consequence of this mutual adjustment process is the selective processing of the encoded concept BUTTERFLY and the resulting broadening of this concept into an ad hoc concept BUTTERFLY*, which denotes not just butterflies but delicate creatures that enjoy freedom and can't stay put.

What current lines of research in Relevance Theory propose is thus that it is the ad hoc concept constructed on-line (in the way just described), and not the concept encoded by the word, which the hearer treats as a constituent of the explicature. It is

therefore this ad hoc concept that contributes to the truth-conditional content of the utterance and which warrants the derivation of the intended implications, as in (39):

(39) She is a *butterfly.*

Explicature: MARY(x) IS A BUTTERFLY*

Implicit premises: BUTTERFLIES* LIKE MOVING FREELY, BUTTERFLIES* DO NOT SETTLE ANYWHERE FOR LONG.

Implicatures: MARY LIKES MOVING FREELY, MARY DOES NOT SETTLE ANYWHERE FOR LONG.

It follows from this account that although the concepts encoded by an utterance give access in memory to a wide array of assumptions, only a subset of these assumptions is actually processed in context (e.g. the assumptions which were most accessible at the moment of processing). This selective processing often results in the denotation of the encoded concept being modified, either narrowing or broadening it, as in (21)-(24) and (29)-(38), respectively. The encoded concept is pragmatically adjusted in every case to the point where it can warrant the set of expected implications. The weaker the strength to which these implications are derived, the weaker the ad hoc concept being constructed and the weaker the explicature of which it is a constituent (e.g. for rather creative loose uses).

Since lexical narrowing and broadening are not distinct processes but merely two instantiations of a single process of lexical pragmatic adjustment that fine-tunes the interpretation of virtually every word in context, it is not strange that they often take place simultaneously (see Carston, 1996, 2002a, 2002c), as in (40):

(40) My daughter, my princess.

In (40), arriving at an optimally relevant interpretation after processing a subset of encyclopaedic information about princesses (e.g. the assumption that princesses are beautiful, adorable, lively, etc.) results in the formation of a new ad hoc concept PRINCESS*, which is narrower than the encoded concept in that it denotes a subset of real princesses (e.g. princesses who are lively, beautiful and loveable), and broader than

the encoded concept in that it denotes a set of entities which typically fall outside its definition (e.g. a set of young women who, although they are not princesses, are lively, beautiful and loveable). It follows from the relevance-theoretic account that a single word may be used to convey a range of different concepts which may be narrower or broader than the encoded one. Each narrowing or broadening of the encoded concept, or each combination of the two, would make a different contribution to the truth-conditional content of the utterance, as in (41):

(41) a. Her name is *written* in her email. (Narrowing - typed) WRITTEN*

b. Her name is *written* on the wall. (Narrowing - painted) WRITTEN**

c. Her name is *written* in my heart. (Broadening) WRITTEN***

d. Her name is written in English history. (Broadening and narrowing) WRITTEN****

6. Conclusion

Relevance Theory provides an approach to human communication which is grounded on the assumption that the human mind has developed a tendency to direct its cognitive resources and mental effort to the processing of information which is potentially most relevant at the time. This tendency towards cognitive efficiency has important consequences for human communication. Relevance Theory claims that our minds have developed dedicated mechanisms for the recognition of ostensive stimuli and interpretation of the intentions underlying their use. The processing of such a stimulus (e.g. an utterance) triggers both a more or less specific expectation of relevance about how it may achieve optimal relevance, and a comprehension procedure geared to following a path of least effort in deriving an appropriate set of cognitive effects. Relevance-driven comprehension is generally a selective process which picks out those assumptions, enrichments and implications which are most likely to yield an interpretation which makes the utterance relevant in the expected way. Selective processing of potentially relevant information during comprehension generally results in the construction of new conceptual representations, with the same mutual adjustment

and pragmatic fine-tuning process at work whether the resulting interpretation is literal, approximate, hyperbolic or metaphoric.

Chapter 3

Metaphor, Interaction and Property Attribution

> "When we use a metaphor, we have two thoughts of different things active together and supported by a single word, or phrase, whose meaning is a resultant of their interaction."
>
> I.A. Richards (1936: 93)

> "It would be more illuminating [...] to say that metaphor creates the similarity than to say that it formulates some similarity antecedently existing"
>
> Max Black (1979: 37)

1. Introduction

Metaphor is the trope per excellence, and an adequate theory of linguistic communication should have something to say about how people communicate with metaphorical expressions. This chapter presents an overview of the literature on metaphor, showing that modern (psycholinguistic) approaches have systematically moved away from a view of metaphor as a purely linguistic device and towards a more cognitive approach. The traditional view that what is conveyed by the use of a metaphor can be successfully paraphrased, and that metaphorical meaning is based on pre-existing similarities between metaphor topic and metaphor vehicle, has been gradually abandoned in favour of the idea that the use and understanding of a metaphor is essentially a creative process. Rather than depending on pre-existing similarities, metaphor comprehension is seen as depending on the emergence of similarities as a result of interaction between mental representations. Although modern research has opened up the interesting possibility that metaphors can communicate new meanings, this chapter is designed to raise important questions about whether existing models can actually account for how these new meanings are constructed in the course of the interpretation process. I suggest that the reason for this failure is that models of metaphor often lack adequate pragmatic principles and inferential procedures which might guide and constrain the interpretation process.

2. Traditional Views on Metaphor

Interest in the study of metaphor goes back a long way. The first serious concern with the use of metaphor we know dates back two thousands years ago, when teaching the art of Rhetoric was common practice. Aristotle, to our knowledge the first philosopher to show a positive interest in metaphor, saw in it a tool that could be used by the politician to persuade and by the poet to delight (Aristotle, *Poetics; Rhetoric*). However, this classical treatment of metaphor was rather paradoxical (Sperber and Wilson, 1990). On the one hand, the main concern of rhetoricians and philosophers, Aristotle included, was the search for truth and the portrayal of reality. Literal language was seen as the most appropriate tool for communication, as it was taken to describe reality in an objective, accurate and direct way. The directness and lack of ornamentation of literal language was seen as reducing the risk of ambiguity, imprecision and consequent misunderstanding typical of figurative speech. On the other hand, classical rhetoricians encouraged the teaching of tropes, which they described as obscure and misleading. Indeed, it was by virtue of being vague and mysterious that tropes could be used to amuse and persuade.

These two views of metaphor as something attractive that deserves close attention, an art one should aim to master - "the greatest thing by far is to be a master of metaphor" (Aristotle, *Poetics*) - and as a deviation from principles of literalness and truthfulness would persist throughout the centuries. With few exceptions (e.g. the Romantics' adoration of the trope), the approach to metaphor that has dominated scholarly thinking sees metaphor as very much like the song of Ulysses' sirens: beautiful, enchanting, delightful, yet dangerous and to be avoided. Perhaps one of the most radical versions of this position is reflected in the thinking of the British empiricist John Locke:

> "Figurative applications of words [...] insinuate wrong ideas, move the passions, and thereby mislead the judgement; and so indeed are perfect cheats [...] They are certainly, in all discourses that pretend to inform or instruct, wholly to be avoided; and where truth and knowledge are concerned cannot but be thought a great fault." (Locke, 1690: III.x.34)[1]

[1] (III. X.34) stands for Book III, Chapter X, Section 34.

It is not surprising that the privileged status that literal language has enjoyed almost uninterruptedly for two thousand years – from the time of classical rhetoric until about thirty years ago – has affected the way that scholars have approached the study of figurative language in general, and metaphor in particular. Standard pragmatic approaches to metaphor comprehension have been directly influenced by the classical view that metaphor is a deviation from a norm of literal truthfulness. Grice's account of tropes is a good illustration (Grice, 1975/1989). The approach to human communication proposed by Grice is based on the assumption that people generally follow a Co-operative Principle and certain maxims of conversation whose purpose is to promote successful communication. According to Grice, in producing a figurative utterance (e.g. a metaphor), the speaker says something blatantly false, thus violating one of these maxims, the first maxim of Quality (or Truthfulness) given below:

(1) Supermaxim of Quality: Try to make your contribution one that is true.
Do not say what you believe to be false. (Maxim of Quality/Truthfulness)
Do not say that for which you lack adequate evidence. (Grice, 1975/1989: 27)

Grice rationally reconstructs the figurative interpretation process as follows. In interpreting an utterance, the hearer first considers the proposition literally expressed by the utterance (i.e. what is said) which in the case of a figurative (e.g. metaphorical) utterance is blatantly false. Recognising that the first maxim of Quality has been overtly violated (flouted), and on the assumption that the speaker is still obeying the Co-operative Principle and supermaxim of Quality, he looks for a related, true proposition which the speaker might have intended to convey.

It is important to notice that this account of tropes introduces an inconsistency in Grice's model of communication, since it forces him to abandon the most convincing rationale for the existence of implicatures. Grice's framework is based on the assumption that a rational speaker should in general obey the Co-operative Principle and conversational maxims. An implicature is a belief imputed to the speaker so as to preserve the assumption that he has obeyed the Co-operative Principle maxims. However, in the case of figurative utterances, the speaker is taken to say (or 'make as if

to say') something blatantly false. The conversational implicature in this case replaces 'what is said', thus confirming the hearer's suspicion that the maxim of Quality has been v*iolated* (rather than obeyed). That is, though in Grice's framework speakers normally convey both what is said and what is implied, in the case of tropes they only convey what is implied. Since implicatures have to be calculated on the basis of what is said, it is not clear how these implicatures are calculated. We are therefore left with the possibility that a speaker can blatantly violate the maxim of truthfulness – that is, he can say anything at all – as long as an implicature (whether or not calculable) is triggered (for discussion on these issues see Wilson and Sperber, 1981, 2000; Wilson, 1995). This lack of constraints is the price Grice has to pay for trying to accommodate figurative uses of language in his framework.

3. A Challenge to the Literal Priority Claim

Grice's account of metaphor has been brought into question by modern psycholinguistic research. The Gricean account and the classical approach which inspires it, have been criticised in particular for what is seen as their commitment to the Literal Priority Claim given below:[2]:

(2) *The Literal Priority Claim:* Metaphorical meanings are derived after literal meanings.

Models based on this claim take metaphor comprehension to involve at least the following steps:

(3) a. Derivation of literal meaning.
b. Rejection of literal meaning.
c. Derivation of figurative (e.g. metaphorical) meanings.

[2] We may notice, as said above, that what Grice was doing was providing a 'rational reconstruction' of the comprehension process – he wasn't claiming that that is how it works in actual on-line processing. It is really the psychologists who have made this jump.

This multi-step comprehension process, typical of pragmatic accounts (e.g. Grice, 1975/1989; Searle, 1979b) but also assumed by early psycholinguistic approaches (Bobrow and Bell, 1973), is commonly known as the 'literalist' model or the 'serial' model. The serial model has been challenged theoretically (Carston, 2002a; Pilkington, 2000; Recanati, 1995, 2003; Rumelhart, 1979; Sperber and Wilson, 1986/1995, 2002; Wilson, 1995; Wilson and Sperber, 2000) and experimentally (e.g. Gibbs, 1980, 1994a; Glucksberg, Gildea and Bookin, 1982; Inhoff, Lima and Carroll, 1984; Ortony, Schallert, Reynold and Antos, 1978), not only with regard to the comprehension of metaphors but with regard to other non-literal uses too (Gibbs, 1986; McGlone, Glucksberg and Cacciari, 1994).

The basis for the serial model is the assumption that the failure of a literal interpretation is the starting point for a metaphorical interpretation. A problem with this is that literal falsehood is neither a necessary nor a sufficient condition for triggering a metaphorical interpretation. First, there are utterances which can be literally true and still have metaphorical readings. These include John Donne's famous example 'no man is an island' and more mundane cases such as 'my daughter is no angel' or 'my brother is an animal'. Second, there is an indefinite number of utterances which are literally false, semantically incongruous or defective but are not metaphorically (or figuratively) intended. An utterance such as 'cats don't have tails', is literally false, an utterance such as Chomsky's famous 'colourless green ideas sleep furiously' is semantically incongruous. However, producing those utterances would not always prompt a figurative interpretation. Third, there are utterances which may be interpreted literally, metaphorically or both literally and metaphorically depending on the context in which they are processed. Consider (4) and (5), for instance:

(4) *A:* How is the weather there now? Has it improved at all since I left?
B: A little bit of sun is finally shining through my window.

(5) *A:* How is the paper going? I hear you have had quite a rough time writing it.

B: A little bit of sun is finally shining through my window.

In (4), the speaker in B may be just expressing the thought that there is a ray of sun shining through her window at that very moment, which would lead the hearer to infer that the weather is good and that it has improved since he left. In (5), the speaker in B may be speaking metaphorically and intending to communicate that she is finally being inspired, enlightened and hence making some progress on her paper (case 5a). Alternatively, she may be delaying answering A's question for a moment out of the excitement caused by her seeing a ray of sun finally shining through her window (after weeks of rain) (case 5b). In fact, there is at least another possible scenario. It may be that she is only intending to convey that she is inspired, but a ray of sun shines through her window as she speaks - both speaker and hearer may be aware of this fact (case 5c). The literalist model predicts that the same steps should occur in processing each of these scenarios: the literal meaning is derived, if it is false, the literal meaning is rejected, and the metaphorical meaning is derived. The rigidity of this model forces it to stipulate a set of rules, or principles that will explain why a literal interpretation is rejected by the hearer in (case 5a) and (case 5c) – even though it is true in (5c) - while it is accepted in (4) and (5b). It is also worth noticing that in many cases, the metaphorical meaning is the first to come to mind, as in (6) and (7). This is clearly inconsistent with the literal priority claim:

(6) *A:* My heart is aching.
B: I know sweetie, breaking up with someone is painful. But it'll get better.
A: No, I mean my heart is really aching. Please take me to hospital.

(7) *Student A:* (after shopping) I have a hole in my pocket.
Student B: I know, I shouldn't be spending this much either.
Student A: No, I mean. I have a real hole in my pocket. I just noticed when I put my hand on it.

A consequence of the serial model is that since metaphorical meanings are only inferred after literal meanings have been derived and rejected, processing metaphorical language should require additional inferential work, and hence more time and effort than processing literal uses. A number of psycholinguistic experiments have challenged this view, showing how, given an adequate context, people do not take longer to understand metaphorical than literal uses. This has been shown for novel as well as for conventional metaphors, and for both individual words and whole sentences (e.g. Gibbs, 1994a; Inhoff, Lima and Carroll, 1984; Ortony, Schallert, Reynold and Antos, 1978).

By taking metaphorical meanings to be derived only after literal meanings are rejected, the serial model assumes that metaphorical meanings, unlike literal meanings, are not automatically constructed. In a series of interesting experiments, Glucksberg Gildea and Bookin (1982) have challenged this assumption showing how hearers automatically derive metaphorical meanings whenever they are accessible. Their experiments are inspired by the Stroop task. Stroop's famous experiment (Stroop, 1935) presented subjects with a list of names of colours which were printed in different, non-matching colours (e.g. the word *yellow* was printed in red, the word *blue* in green, etc.). Students were asked to name the colour of the ink and not to read the word. The results showed that subjects had great difficulty in performing the task, because of an interference from reading which, despite the instructions given seemed to happen automatically and involuntarily.

The assumption underlying Glucksberg et al. (1982) was that, just as people cannot avoid reading words when presented to them, so they cannot ignore metaphorical meanings when processing utterances that can be metaphorically interpreted, even when a literal interpretation is possible. To test this, they presented subjects with one sentence at a time on a screen and asked them to judge whether the sentence was literally true or literally false. The material used included literally true sentences (e.g. 'some birds are robins'), literally false sentences (e.g. 'some birds are apples'), literally false but metaphorically true sentences (e.g. 'some jobs are jails') and scrambled sentences (e.g. 'some flutes are jails') which were both literally and metaphorically false. They found that subjects took significantly longer to reject metaphorical sentences as literally false

(1239ms) than they did to reject literally false sentences (1118ms) and scrambled metaphors (1162ms).

The authors interpreted these findings as evidence that people cannot help but derive metaphorical interpretations whenever they are accessible. That is, people take longer to reject sentences of the sort 'some jobs are jails' than to reject other literally false sentences because of what the authors refer to as a 'metaphor interference effect'. Because the sentences were metaphorically true, subjects found it difficult to reject them as false. In order to show that these results were not caused by the particular combination of topic and vehicle used in those sentences, Glucksberg et al (1982) set up another experiment.[3] This time, they presented subjects with sentences whose topics were modified by one of two quantifiers 'some' or 'all'. The prediction was that subjects should have less difficulty in rejecting as false a sentence such as 'all jobs are jails' than a sentence such as 'some jobs are jails'. The reason is that although both use the same combination of topic and vehicle, only the 'all' sentences are literally and metaphorically false while the 'some' sentences are literally false but metaphorically true. Being metaphorically true, the 'some' sentences, but not the 'all' sentences, should be expected to produce a metaphor interference effect. This prediction was borne out. The authors concluded from these experiments that, contrary to common belief, some metaphorical meanings are computed automatically and non-optionally.

Dascal (1987) criticises Glucksberg et al. (1982) for using literally false sentences on the ground that they are difficult to interpret. He argues that people might generate metaphorical meanings in these cases (e.g. 'some books are tables') due to the difficulty of generating a literal interpretation that they can accept. He suggests that a delay in rejecting metaphorical utterances such as 'some jobs are jails' may be seen as providing evidence in favour of the serial model. That is, people may take longer to reject metaphorical sentences as literally false not because they interpret them as metaphorically true but because they are going through the process of rejecting the literal meaning and reinterpreting the sentence metaphorically. In an attempt to rescue

[3] Research on metaphor generally assumes a metaphor consists of two main elements which following Richards (1936) are generally referred to as the metaphor topic and the metaphor vehicle. The metaphor topic of a metaphor such as 'my flatmate is a pig' corresponds to the subject 'my flatmate' and the metaphor vehicle corresponds to the object 'pig'. A core assumption in metaphor research is that in a metaphor the topic is characterised by means of the vehicle.

the hypothesis considered by Glucksberg and colleagues (1982) hypothesis, Keysar (1989) set up another experiment. Because of the possibility that literally false sentences may trigger an alternative (e.g. metaphorical) interpretation, he tested the metaphorical interference effect with examples including both literally false and literally true sentences. In fact, these examples included sentences which, depending on the context in which they were presented, could be judged as a) literally true and metaphorically false (L+/M-), b) literally false and metaphorically true (L-/M+), c) literally and metaphorically true (L+/M+), or d) literally and metaphorically false (L-/M-). For example, a sentence such as 'Bob Jones is a magician' was presented in one of the following contexts:

(8) a. A context in which Bob Jones is a magician by profession but has serious troubles trying to administrate his earnings. (L+/M-)

b. A context in which Bob Jones is not a magician by profession but has the ability to magically stretch the little money he earns. (L-/M+)

c. A context in which Bob Jones is a magician by profession and has the ability to magically stretch the little money he earns. (L+/M+)

d. A context in which Bob Jones is not a magician by profession and is a terrible finance administrator. (L-/M-)

The experiment involved presenting a short story about a character, in this case Bob Jones, followed by the target utterance (e.g. 'Bob Jones is a magician') which was to be judged as literally true or literally false. The key cases were those context-stories biasing an L-/M+ or an L+/M- interpretation:

(9)

L-/M+: Bob Jones is maestro and manager of a famous orchestra. They are known for their drama and style. He earns his living travelling around the world, but the expenses of a major orchestra are not minor. Sometimes it seems as if Bob's money is made of rubber because he stretches it so far. How does he create such a healthy profit despite these expenses?

'Bob Jones is a magician'

L+/M-: Bob Jones is an expert at such stunts as sawing a woman in half and pulling rabbits out of hats. He earns his living travelling around the world with an expensive entourage of equipment and assistants. Although Bob tries to budget carefully, it seems to him that money just disappears into thin air. With such huge audiences, why doesn't he ever break even?

'Bob Jones is a magician'

According to Dascal, if metaphorical interpretations are triggered by literally false sentences, an interference effect should only be found in the L-/M+ context. Since in L+/M- the literal interpretation is adequate, the metaphorical meaning should, according to the serial model, not be derived and so no metaphor interference should be expected. Keysar predicted a different result: since both literal and metaphorical meanings are derived automatically whenever they are accessible, one should expect an interference effect whenever the literal and metaphorical meanings are incongruent: that is, in either the L-/M+ or the L+/M- context. As in Glucksberg et al. (1982), the results showed that subjects took longer to reject a sentence as literally false when it was metaphorically true. Furthermore, they also showed that subjects took longer to judge as literally true a sentence which, although literally true, was metaphorically false, suggesting that the metaphorical falsity of the sentence was interfering with the judgement. Keysar took these findings to support the hypothesis that both literal and metaphorical meanings are derived in parallel automatically and non-optionally, whenever they are accessible.

We may conclude that with regard to the literal priority claim it is possible to distinguish two contrasting positions in the cognitive science literature. On the one hand, we may group under the label 'literalists' all scholars who, following the traditional line of thought, see metaphors as deviations from a norm of literalness. Literalist models would include all those theories of metaphor interpretation which take the derivation of metaphorical meanings to be dependent on the failure of literal interpretations. On the other hand, we may group under the label 'non-literalists' or 'constructivists' all scholars who, reacting against the literalist position, see metaphorical and literal meanings as equally natural and worth considering, with metaphorical meanings being automatically derived whenever they are accessible. In fact, given appropriate contexts, literal and metaphorical interpretations have equal chances of being derived. Many modern psycholinguists and some philosophers can be

seen as following a constructivist approach (e.g. Glucksberg and Keysar, 1990; Gentner, 1983; Tourangeau and Rips, 1995; Tourangeau and Sternberg, 1981). A constructivist position is also proposed by Relevance Theory, and will be defended in this dissertation.

4. From Property Matching to Property Attribution

The model of metaphor comprehension based on the Literal Priority Claim which dominated research since classical times is known as the Comparison View. The Comparison View, first proposed by Aristotle (*Poetics xxi; Rhetoric III, ii*), treats metaphors as elliptical similes. On this approach, understanding a metaphor of the form *X is Y* (e.g. 'John is a lion') involves transforming this (literally false) statement into a (literally true) comparison of the form *X is like Y* (e.g. 'John is like a lion') and identifying the shared property which provides a ground for the comparison (e.g. bravery). The assumption that metaphors are understood as implicit similes, like the related assumption that metaphorical meanings are parasitic on literal meanings, has been deeply rooted in research on metaphors for centuries. The classical comparison view has influenced psycholinguistic models which take identification of the ground for the comparison to be the basis for metaphor interpretation. (e.g. Malgady and Johnson, 1976; Ortony, 1979; Tversky, 1977). A standard assumption is that the ground for the comparison is a set of features common to topic and vehicle. Thus, interpreting a metaphor such as 'Peter is a pig' would involve taking the features associated with 'Peter', on the one hand, and 'pig', on the other, and matching them against each other so as to arrive at the subset of features which they both share (e.g. 'dirtiness', 'filthiness' and 'gluttony') (e.g. Tversky, 1977).

In an attempt to refine this 'feature matching' view, Ortony (1979) proposed the Salience Imbalance Hypothesis. This suggests that in order to understand a metaphor, it is not enough to find a set of features common to topic and vehicle: it is also crucially necessary to identify which of those common features are highly salient for the vehicle (here 'pig') but low in salience for the topic (here 'Peter'). It is this subset of features which is taken to form the grounds for comparison and hence for the interpretation of the metaphor. Take for example the utterance 'cigarettes are time bombs'. According to

Ortony, the reason why this utterance is judged as metaphorical is that the feature 'deadly' is highly salient for the vehicle ('time bombs') but only low in salience for the topic ('cigarettes'). He argues that this salience imbalance does not carry over to literal comparisons, in which the relevant features should be high in salience for both terms: so, for instance, in 'copper is like tin', both terms in the comparison share a number of highly salient features. It is because of this difference, Ortony argues, that comparing cigarettes to time bombs results in a metaphorical interpretation, while comparing copper to tin does not.

Glucksberg and colleagues raise an important problem for this view (see Cacciari and Glucksberg, 1994; Glucksberg and Keysar, 1990). They argue that salience imbalance cannot be a valid cue for metaphoricity because it occurs in all comparisons and not only in metaphorical ones. The reason is that, according to the 'given-new convention', which they see as applying to all assertions, for a comparison (or, more generally, an assertion) to be informative, it must attribute salient properties of the predicate to the subject. If the intended properties are already highly salient for both subject and predicate, the comparison is not informative (and hence pragmatically unacceptable). For instance, what makes a literal statement such as 'limes are like lemons' informative (e.g. to someone who does not know what limes are) is that it helps the hearer to incorporate a set of highly salient properties of lemons into his knowledge of limes.

In fact, models of metaphor comprehension based on the view that metaphor interpretation involves the identification of similarities between topic and vehicle are generally quite problematic. On the one hand, as Goodman (1972) pointed out (see also Hahn and Chater, 1997), everything is like everything else to a certain extent. The lamp on my desk, the pen I write with, a tissue on the floor resemble each other in an indefinite number of ways: they belong to me, they are in my room, they are less than a metre from me, they are more than a metre from the building next door, they are more than a mile from France, more than a thousand miles from the moon, etc. In fact, if we have the capacity to construct ad hoc categories, these three objects are similar in that they belong to an indefinite number of possible categories (e.g. the category of not being red, the category of being more than a centimetre long, the category of things I am not

taking to college with me on Thursday, etc.). The instability of the similarity relation raises questions about its suitability as the basic mechanism governing metaphorical interpretation. This is not to say that the exploitation of resemblances and perception of similarities plays no role in interpretation, but it does suggest that metaphor comprehension cannot be based solely on the identification of common properties unconstrained by general cognitive or pragmatic mechanisms.

Even if we restrict ourselves to common taxonomic categories as the basis for comparisons, it does not take us very far. Consider the set of common taxonomic categories to which both Peter and pigs belong: the category of living kinds, the category of animals and the category of mammals. It is by virtue of being mammals (and hence animals and living kinds) that Peter and pigs share a set of properties: they are born, reproduce and die, they breathe, have internal organs, etc. Although it is true that these are properties common to both Peter and pigs, they play no role in processing the metaphor 'Peter is a pig'. In fact, even if what the speaker of the metaphor intends to convey is, roughly, that Peter is dirty, filthy and gluttonous, the assumption(s) that Peter is dirty, filthy and gluttonous may not have been part of the hearer's encyclopaedic information about Peter before the utterance took place. The hearer may know a bunch of things about Peter (e.g. that he is the speaker's flatmate, that he studies law, that he was born in Paris, that he likes hiking, etc.) but may not have known that he is dirty, messy and difficult to live with. It is this new information that helps to make the utterance informative.

More generally, property matching models have often been criticised for their inability to account for cases in which the hearer is not familiar with the metaphor topic at all. These include utterances such as 'the lecture was a sermon', 'Mr Smith is an ogre' and 'my room is a pigsty'; uttered to someone who is not familiar with the lecture in question and has never heard about Mr Smith or seen the speaker's room. The argument against matching models is that, with no stored information about these entities, the hearer cannot match the properties of the vehicle with those of the subject.[4] In fact, being dirty may not be a permanent feature of the speaker's room but a temporary state,

[4] One way out of this problem might be to argue that people do have knowledge about rooms, lectures and people in general which they can use to understand these expressions. Still, a matching of features does not seem the right way to go. I'll come back to this issue shortly.

and so is unlikely to be part of the information the speaker or anyone else stores as part of their knowledge about that room. In other words, the assumption that the speaker's room is dirty does not have to be part of the hearer's knowledge before the interpretation process, but rather arises as a result of it.

We can conclude from these arguments that based on the matching of pre-existing features, feature-matching models face two main problems: a) how can a hearer match properties of topic and vehicle when they are simply not stored? And b) how can a hearer base his interpretation on the matching of common properties when there may be indefinitely many of these, and when the set of common properties does not include the set of properties which are actually intended? Some scholars have therefore abandoned the idea that a feature-matching process takes place as part of the metaphor comprehension process, and argued instead that the process involved is one of property attribution, in which properties of the vehicle are attributed to the topic (e.g. Gentner, 1983; Glucksberg and Keysar, 1990; Tourangeau and Rips, 1995; Tourangeau and Sternberg, 1981).[5] According to this position, even if the hearer of the utterances above did not attend the lecture the speaker is referring to, has never met Mr Smith, and knows nothing about the speaker's room, he can still assign a subset of the stored properties of sermons, ogres and pigsties to the subjects of these utterances. In doing so, he would take the speaker to communicate, among other things, that the lecture was tedious and difficult to follow, that Mr Smith is bad and to be feared, and that the speaker's room is dirty and messy. In other words, whereas in matching models, the ground of the metaphor is discovered by identifying the set of features common to topic and vehicle, in attribution theories, the ground of the metaphor is the set of properties of the vehicle which can be understood as attributed to the topic – whether or not they formed a prior part of the hearer's knowledge of the topic.

Experimental research seems to support this position. The results show that although common properties of topic and vehicle figure in people's reported interpretations of metaphorical expressions, vehicle-based properties are judged as

[5] These theories are influenced by Black's Interaction Theory (1962) and therefore see attribution as a result of interaction or alignment of topics and vehicles. The Structure Mapping framework proposed by Gentner and colleagues may be better seen as involving a mixture of matching and attribution processes (see Bowdle and Gentner, 1999; Gentner, 1983, 1989; Gentner et al.., 2001; Gentner and Wolff, 1997; Wolff and Gentner, 2000).

considerably more important to interpretation than topic-based properties (Becker, 1997). While the main concern of matching models has been to determine how to spot similarities between topic and vehicle, the main concern of feature-attribution models is to explain how the hearer identifies which properties of the vehicle can be appropriately attributed to the topic. The next section discusses in some detail one popular attribution view of metaphor known as the Class-Inclusion Theory.

5. The Class-Inclusion Theory: Attribution, Interaction and Categorisation

The Property Attribution or Class-Inclusion theory of metaphor proposed by Glucksberg and Keysar (1990) and developed by them and colleagues (e.g. Camac and Glucksberg, 1984; Glucksberg, 2001; Glucksberg and Keysar, 1990, 1993; Glucksberg, Manfredi and McGlone, 1997; Glucksberg, McGlone and Manfredi, 1997; McGlone and Manfredi, 2001) suggests that we should abandon not only the idea that metaphor interpretation is based on the matching of common features but also, crucially, the idea that this matching is motivated by the prior derivation of an implicit comparison. Metaphors, according to Glucksberg and Keysar, are not elliptical similes but class-inclusion assertions, and identifying the properties of the vehicle which are to be attributed to the topic depends primarily on the category which the metaphor vehicle is taken to represent.

Theories of metaphor have generally assumed that metaphors typically involve category violations. Saying 'Tom is a rat' is asserting, or at least appearing to assert, that Tom belongs to the category of rats, that he is a member of the rat species. When Tom is indeed a rodent of this type, this is just a case of class inclusion assertion and no category violation is involved. However, when the utterance is metaphorically intended, for instance if Tom is the speaker's ex-boyfriend, a category violation is clearly involved as people cannot be rats. As a result, hearers are forced to look for a plausible interpretation on another level. As noted above, the most common view is that the metaphorical utterance is an elliptical simile, or that it implies a related simile. In either case, what is intentionally conveyed is not the (false) assertion *X is Y* but rather the (true) comparison *X is like Y*.

The Class-Inclusion theory proposes instead that what is asserted by use of a nominal metaphor is not that the metaphor topic belongs to the category represented by the vehicle taken literally but that it belongs to a superordinate category of which the vehicle is a salient member. That is, in the example above, what is asserted is not that Tom is a rat but that he belongs to a superordinate category that rats exemplify, namely the category of 'entities which are particularly disgusting and repulsive'. Since this category has no name of its own, the speaker borrows the name of a prototypical member (e.g. rats). The dual reference of the vehicle allows for a double interpretation; one in which the vehicle stands for its literal referent (rats) and one in which it stands for this broader category (call it rats*).[6] A metaphorical interpretation results when the vehicle is taken to stand for the latter category. Since this new category constructed ad hoc during interpretation is said to provide a set of properties that can be attributed to the topic, it is often described in the theory as an 'attributive category'. It is in this sense that scholars such as Glucksberg and Keysar treat metaphors as class inclusion assertions: they assert that the topic belongs to the superordinate category exemplified by the vehicle (i.e. of which the vehicle is a salient member).

5.1 Metaphor and Ad hoc Categories

The Class-Inclusion theory seems to have been put forward on both philosophical and psychological grounds. On the psychological front, Glucksberg and colleagues adopt ideas put forward by Lawrence Barsalou about the ability of subjects to construct ad hoc categories on the fly by selecting a subset of information from long term memory, a different subset each time (e.g. Barsalou, 1983, 1991). As discussed in previous chapters, Barsalou's experiments show that people construct ad hoc categories all the time when making plans or attempting to fulfil a goal. They create novel categories when babysitting ('things to do with an eight-year-old on a rainy day'), when planning holidays ('places to visit in Madrid on a holiday weekend'), etc. He also shows that the processing of a certain common taxonomic category (e.g. birds) in different contexts, on different occasions, and taking different points of view, etc., results in the assembly of

[6] In using the terms reference and dual reference here I am merely using the terminology used by Glucksberg and colleagues. It is, however, important to bear in mind that strictly speaking, metaphor vehicles are not referential expressions.

different bits of information associated with the category in long term memory. This leads to the formation of different ad hoc categories (e.g. the category of 'edible birds', 'birds that live in China', 'birds that don't fly'), each one exhibiting different prototypicality effects (e.g. prototypical member may be chickens, swans or penguins, respectively).

The Class-Inclusion theory of metaphor assumes that the ability to construct different ad hoc categories by selecting different bits of information from long term memory is exploited during metaphorical interpretation. More specifically, it takes metaphor interpretation to involve selecting some properties of the vehicle and creating an ad hoc attributive category on the basis of this selection. However, the view of ad hoc category construction defended by Glucksberg, Keysar and colleagues supports a set of claims not directly suggested by Barsalou's experiments. For example, Glucksberg and colleagues take the prototypicality of the vehicle as a motivation for category formation rather than an outcome of it. This is so even though it was the instability of graded structure that Barsalou's experiments were designed to illustrate. The Class-Inclusion theory claims, for instance, that a metaphor such as 'my lawyer is a shark' is successful because sharks are prototypical members of the category of predatory creatures which the metaphor vehicle instantiates.

5.2 Interaction in Interpretation

Like most modern models of metaphor, the Class-Inclusion view also seems to have been influenced by Black's Interaction Theory of metaphor (Black, 1962, 1979). A disciple of Richards (1936), who was himself a disciple of Coleridge, Black's aim was to rescue metaphor from comparison and substitution theories, which treated metaphor as a stylistic device easily paraphrased in literal terms, in order to propose a more romantic view of metaphor as cognitively significant and not paraphraseable without loss. The main idea underlying Black's approach is that metaphor encourages an interaction between topic and vehicle which creates similarities between them. Instead of being based on the identification of pre-existing similarities, metaphor interpretation is a creative process from which something new emerges, such as a new perception of an object.

According to Black, a metaphor such as 'man is a wolf' consists of a primary subject 'man' (the metaphor topic), a secondary subject 'wolf' (the metaphor vehicle), each of which is associated with a system of commonplaces corresponding roughly to the set of encyclopaedic assumptions about the entities they denote. This system of commonplaces includes assumptions which are actually true or folk assumptions which although false, are held as true (e.g. the assumption that wolves are dangerous and aggressive creatures). Metaphor interpretation, he argues, results from an interaction of commonplaces which he describes figuratively:

> "Suppose I look at the night sky through a piece of heavily smoked glass on which certain lines have been left clear. Then I shall see only the stars that can be made to lie on the lines previously prepared upon the screen, and the stars I do see will be organised by the screen's structure." (Black, 1962: 41)

In understanding the metaphor 'man is a wolf', the metaphor topic 'man' acts as a frame highlighting commonplaces associated with the vehicle 'wolf' (the smoked glass), and the vehicle 'wolf' projects back these selected assumptions (the smoked glass with lines on it) which act as a grid to select a set of commonplaces associated with the topic 'man' (the set of stars visible through the glass). Looking at the topic through this grid results in the enhancement of some commonplaces associated with it (visible stars) (e.g. assumptions about man's basic instincts, aggressiveness, competitiveness, etc.), and the suppression of other assumptions (stars which cannot be seen). This reorganisation of assumptions in the topic results in the creation of something new, namely a new way of looking at men, who are somehow dehumanised.[7]

Black's figurative description of the interpretation process is suggestive, but makes no precise predictions about which assumptions will be retrieved and why. However, Black's and Richards' assumption that metaphor interpretation is essentially an interactive process between two concepts or domains has inspired a wide range of psycholinguistic research which does aim to make predictions about this process and to test the assumptions underlying the interaction view. The Class-Inclusion theory is one

[7] In early work, Black (1962) argued that both topic and vehicle change as a result of the interaction: men are dehumanised and wolves humanised. In a later account, he concentrates on the modification of the topic (Black, 1979).

of these approaches, although not the only one; the interactive hypothesis is the background to virtually every modern theory of metaphor, including Blending Theory or Conceptual Integration Theory (Fauconnier and Turner, 1998, 2002), Domain Interaction Theory (Tourangeau and Sternberg, 1981) and the Structure-Mapping Theory (Gentner, 1983). Some of the arguments I will use here against the Class-Inclusion theory, and more particularly against its ability to account for the construction of new representations via an interaction process should also apply to these models.

Class Inclusion theorists see metaphor interpretation as an interactive process in which metaphor topic and metaphor vehicle play different but interactive roles. On the one hand, the metaphor topic provides a context for the interpretation of the vehicle in that it specifies a set of dimensions for the attribution of properties (e.g. for the topic 'lawyer' these include dimensions such as 'skill', 'character', 'cost', etc.). On the other hand, the metaphor vehicle provides access to properties a subset of which can be used to assign (positive or negative) values to some of those dimensions. In metaphor interpretation, topic and vehicle are seen as aligned in such a way that the properties of the vehicle which can assign values to the dimensions in the topic and the dimensions in the topic which can be characterised by the properties of the vehicle are selected (Glucksberg, 2001; Glucksberg, McGlone and Manfredi, 1997).

On Glucksberg's account, metaphor interpretation essentially involves identification of the first higher-order category to which the vehicle can be assigned, on the basis of the properties selected during interaction, which can also include the topic as a member. For the metaphor 'my lawyer is a snake', this category is claimed to be that of 'people and animals which are devious and malevolent'. Here the interaction of topic and vehicle results in the selection of the dimension of 'character' from those made available by the topic, and the properties of 'being malevolent' and 'devious' from those made available by our knowledge of snakes. According to this model, then, the metaphor 'my lawyer is a snake' is understood as asserting that the speaker's lawyer belongs to the ad hoc category of 'entities which are devious and malevolent'. Since he belongs to this category, the hearer can assign to him the set of properties typical of that category (e.g. the properties of being devious and malevolent), thus assigning a (negative) value to the dimension of 'character'.

One implication of this view is that varying the topic of the utterance should involve a change in the dimensions of attribution provided by the topic. This would affect the set of properties of the metaphor vehicle which are selected in interpretation. Take, for instance, the metaphor 'some roads are snakes'. Although roads can be characterised in a variety of ways (e.g. based on shape, surface, safety, etc.), the presence of the vehicle 'snake' highlights only some of these dimensions, and in particular the dimension of 'shape'. At the same time, the presence of this dimension in the topic acts as a filter selecting the set of properties of the vehicle which can assign values to that dimension (e.g. the property of having a twisting shape). The hearer understands the metaphor as conveying that topic and vehicle belong to a single category which the vehicle exemplifies in virtue of those properties and which can assign values to the topic in virtue of that dimension (e.g. the category of 'things with a twisting shape'). In other words, the Class Inclusion theory claims that what determines the construction of a distinct ad hoc category (e.g. snake*, snake**') and so the attribution of different properties of the vehicle (e.g. of snakes) to the topic is a combination of both a) the relevant constraints imposed by the topic and b) the categories the vehicle can be taken to exemplify (e.g. 'things with a twisting shape', 'things which are devious and malevolent') (Glucksberg, 2001: 55). The aptness of a metaphor depends partly on how typical the literal referent of the metaphor vehicle is of the superordinate category it represents.

5.3 Problems with the Class Inclusion View

As we have seen, the Class Inclusion approach to metaphor interpretation makes two main claims. On the one hand, it claims that metaphor interpretation is an interactive property-attribution process in which topic and vehicle play different but interactive roles: one providing dimensions for attribution, the other providing values to those dimensions. On the other hand, it claims that the metaphor vehicle has 'dual reference' and that to be interpreted metaphorically, the vehicle needs to be taken to refer to a superordinate category which the vehicle, taken literally, exemplifies. The interactive property-attribution hypothesis and the ad hoc category construction hypothesis may seem fine when one looks at them separately. However, trying to unify them in the way

proposed by the Class-Inclusion account brings a certain risk of circularity and confusion:

> "The properties of a metaphor vehicle that are attributed to the topic are thus determined by two criteria: a) the higher-order category (or categories) that the vehicle may exemplify, and b) whether the prototypical properties of that category characterise the metaphor topic in a meaningful way. For example, consider again the metaphors *some roads are snakes* and *some lawyers are snakes*. Different properties of snakes are attributed to the topics *some roads* and *some lawyers*. The attribution of properties in these metaphors is a joint function of the categories that the vehicle can exemplify (e.g., "things with a twisting shape" and/or "things that are devious and malevolent") and the relevance constraints imposed by the respective topics (e.g. shape for roads, character for lawyers)" (Glucksberg, 2001: 55).[8]

It follows from what Glucksberg says here that for a certain property to be attributed to the metaphor topic, it needs a) to be a property of a certain superordinate category the vehicle exemplifies, and b) to be able to assign value to a dimension in the topic. In other words, property attribution presupposes the existence of a higher-order category. The question is: how has this category been formed? The theory at this point seems to lead to the conclusion that the ad hoc category is formed by inheriting some subset of properties of the subordinate category represented by the vehicle, taken literally. How is this subset selected and inherited? According to the theory, the selection of properties of the vehicle goes hand in hand with the selection of dimensions of the topic so that only properties which are capable of assigning values to the dimensions provided by the topic are selected.[9]

All this leaves us with a relatively circular-seeming picture. What determines the properties of a metaphor vehicle that are attributed to the topic is the higher-order category (or categories) that the vehicle may exemplify (together with the constraints imposed by the topic). What is used in the construction of this higher-order category is a set of properties we select from the metaphor vehicle (those which can assign values to the topic dimensions). What determines the selection of properties of the vehicle is the set of topic dimensions it can assign values to. Or, to put it in another way, the metaphor

[8] The relevance constraints that Glucksberg is referring to here are the dimensions provided by the topic.

topic provides a set of dimensions which act as a filter for selecting the set of properties of the vehicle which are capable of assigning values to those dimensions. The selection of this subset of properties results in the construction of a new category which is exemplified by the vehicle in its literal sense and capable of including the topic as a member. Since what the metaphor does is to assert that the topic belongs to this category, the hearer is entitled to attribute to it the set of properties of the newly created category - the very same subset of properties which the topic had selected from the vehicle in the first place! In other words, metaphor interpretation begins and ends at the same point: with the metaphor topic and the dimensions provided by it. This leads to the conclusion that the processes of property selection, category construction and property attribution which form the basis of the Class-Inclusion model are ultimately dependent on the constraints imposed by the topic: that is, on the dimensions it provides and how relevant those dimensions are to interpretation.

Founding a theory of metaphor interpretation on the constraints imposed by metaphor topics is rather problematic. The property attribution theory, as we have seen, was proposed as an alternative to matching models, which were not able to account for the interpretation of metaphors where the hearer is not familiar with the metaphor topic. Still, the account of property attribution proposed by the Class-Inclusion Theory highlights the need for the hearer to know enough about both topics and vehicles for an interaction to take place. "Understanding a metaphor thus requires two kinds of semantic and world knowledge. First, one must know enough about the topic [...]. Second, one must know enough about the metaphor vehicle" (Glucksberg, 2001: 55). Unlike matching models, the Class-Inclusion theory may be able to ignore the problem of unfamiliar topics by arguing that hearers always know something about the dimensions a topic provides. For example in the utterances 'that lecture was a sermon', 'Mr Smith is an ogre' and 'my room is a pigsty' discussed above, the hearer may not be familiar with that particular lecture, person or room; but, still, he is familiar with the categories they belong to (e.g. lectures, people and rooms), and it is because of his familiarity with these categories that he may be able to provide a set of dimensions for attribution. The problem with this is that lectures, people and rooms can be characterised in an indefinite

[9] See also Ritchie (2003) on criticisms on the circularity of the Class-Inclusion approach.

numbers of ways. Rooms may be characterised by their size, colour, level of humidity, height, number of windows, people who have lived in them, etc. Lectures can be characterised by their length, the number of students attending, the quality of the audio system, the number of pauses taken, etc. People can be characterised by their height, their mood, their hair colour, their sleeping patterns, etc. A 'dimension' is a very vague term which can be used to refer to general aspects (e.g. character, physical appearance) or specific traits (e.g. experience, shape); the theory makes no distinctions between them.

Although it may solve the problem of unfamiliar topics, assuming that the dimensions provided by the topic depend on the dimensions provided by the category the topic belongs to is problematic. One reason for this is that topics which belong to the same category (e.g. the category of people, the category of professional workers) should provide roughly the same set of dimensions, making it difficult to see how they can be used to pick out different properties of a single vehicle or contribute to the formation of different ad hoc categories on different occasions. Consider the examples in (10):

(10) a. (Of a surgeon who has been negligent) That surgeon is a butcher.
b. (Of a pianist who has played terribly badly) The pianist butchered the sonatas.
c. (Of a teacher who fails most of the class) That teacher is a butcher.
d. (On a gruesome crime scene) This man is a butcher!

The Class-Inclusion model typically uses examples like (10a) and (10b) to show that a vehicle can be used to modify a single dimension in different ways: the way a surgeon's skills are characterised is different from the way a pianist's skills are characterised. The theory does not say, however, how this different characterisation takes place, e.g. how the hearer derives different implications and hence different interpretations in each case. If the theory provides the hearer with the basic ingredients: topic dimensions ('skill'), properties of the vehicle (e.g. 'lack of precision'), and ad hoc categories constructed (e.g. 'people who lack precision'), all of which are the same, how does a hearer derive different interpretations in (10a) and (10b)?

It is also worth noticing that the topic in (10c) provides the same range of dimensions as provided by the topics in (10a) and (10b), the dimension of skill included, and it is combined with the same metaphor vehicle ('butcher'). However, unlike (10a) and (10b), the interaction between topic and vehicle in this case does not result in selection of the dimension of 'skill' in the topic. Why should this be so? After all, teachers, like surgeons and pianists, are people characterised by the profession they belong to and the skills involved in exercising that profession. However, a natural interpretation of (10c) is not that the teacher is incompetent in his job, but maybe that he is just far too strict and inflexible, and, from the point of view of the students, someone to be feared. What these examples suggest is that selection of an appropriate dimension for attribution may be not so much a prerequisite to comprehension, as assumed by the Class-Inclusion view, but, if anything, a by-product of the comprehension process.

Finally, processing (10d) involves the identification of yet another subset of features and the construction of yet another ad hoc category still. This category is not derived from the assumption that murderers lack skill (which they may not), or that they are strict (which they might not be), but on another range of assumptions that the vehicle 'butcher' gives access to. These include assumptions having to do with the use of sharp instruments, the spilling of blood, dismembering, etc. It is not clear what dimensions are selected on this occasion or how the Class-Inclusion theory can capture the fine-grained interpretation and range of implications the hearer derives from these assumptions.

I agree with the 'interactive' idea that the presence of the metaphor topic has an effect on the set of attributes or assumptions which we access from the metaphor vehicle on a given occasion (e.g. the activation of a certain concept in memory may have an effect on how we process incoming information). However, I don't agree with the assumption that by putting a certain topic and a certain vehicle in the same sentence, the right combination of dimension and attribution will emerge, by magic, providing an adequate basis for interpretation. Further evidence for this point can be found in the analysis of examples such as (11) and (12):

(11) That lawyer is a shark.

(12) John is an iron bar.

A metaphor, for example a nominal metaphor of the form *X is Y,* may be used to convey a wide range of different meanings, and involve the formation of a wide range of different ad hoc categories (and ad hoc concepts denoting those categories). (11), for instance, can be used to express the thought that that particular lawyer belongs to the category of sharks*, which denotes 'people who are extremely energetic and hard working, who give everything for their job, who would fight until the end, etc.' Alternatively, the utterance may be taken to express the thought that the lawyer belongs to the category of sharks**, which denotes the kind of 'people who want to win and achieve success at all costs, no matter how many people they victimise on their way'. Being a shark*, one would hope, is preferable to being a shark**: the former are responsible, hard-working people; the latter are wicked and to be feared.

The question is: what determines the formation of the different ad hoc categories shark*, shark** (and so the ad hoc concepts SHARK*, SHARK** that denote those categories) on each occasion? The Class-Inclusion Theory provides no answer to this question. According to this theory, aligning a metaphor topic and a metaphor vehicle should result in the emergence of a combination of topic dimensions and vehicle properties which should form the basis for the construction of the ad hoc category to which topic and vehicle belong, and so the basis for the interpretation of the utterance. If this is all there is to metaphor interpretation, aligning the same topic and vehicle should result in the emergence of the same combination of dimension and property, the construction of the same attributive category and in the derivation of the same interpretation across contexts. This is clearly not the case.[10]

The metaphor 'John is an iron bar' in (12) presents another interesting case, which illustrates how a single metaphor may be used to convey a number of different assumptions and to assert that the topic belongs to a number of different categories, such as those in (13):

[10] This criticism is not unique to the Class-Inclusion theory but applies to interactive views more generally.

(13) a. John belongs to the set of people who are insensitive, have no feelings, wouldn't feel compassion for someone else's suffering, etc.

b. John belongs to the set of people who are not easily upset, who would deal with difficult situations with courage, who are brave and experienced in life's upsets, who are capable of facing bad news, etc.

c. John belongs to the category of people who are very reserved and do not like sharing feelings with others, etc.

d. John belongs to the category of people who are capable of lifting heavy weights, who are strong and muscular, etc.

e. John belongs to the category of people who are difficult to defeat in their area of expertise, etc. (e.g. sport, chess, computer games, etc.).

f. John belongs to the set of people who are inflexible, live by the rules, etc. (e.g. some judges, teachers, parents).

g. John belongs to the category of people who are difficult to convince, persuade, induce to change their mind, etc.

Etc.

Although the utterance in (12) can be used to convey a wide range of different meanings, each involving the construction of a different ad hoc category, such as those in (13), the Class-Inclusion theory does not offer an adequate explanation for how this takes place. Saying that the vehicle is an exemplar (or prototypical member) of the superordinate category does not guide the hearer to a single interpretation for a number of reasons. First, 'iron bars' can potentially be members, and even typical members, of an indefinite number of ad hoc categories (e.g. 'hard things', 'things to use as weapons', etc.). Second, according to Barsalou's experiments, prototypicality is an unstable notion which varies across contexts, points of view, individuals, etc. with the typicality of a given member arising as a by-product of constructing an ad hoc category rather than as a prerequisite to the construction of that category. Third, even if we take prototypicality to be a stable notion, and assume that metaphor vehicle can exemplify only a limited number of ad hoc categories (e.g. the category of heavy things, hard things, etc.), none of these categories may be the one intended by the speaker on a certain occasion, as in

(13g), where the intended category denotes the set of people who are difficult to convince or persuade.

Crucially, saying that metaphor interpretation (and category construction) depends on an interaction of topic dimensions and vehicle properties cannot explain how an utterance can have an indefinite number of possible interpretations, or how the hearer chooses or constructs a hypothesis about the one intended by the speaker. Not only can a single dimension-property combination open the way to a range of possible interpretations (as in (10a) and (10b)), in many cases a good number of properties of the vehicle can be used to characterise a good number of topic dimensions. Since every combination offers a potential ad hoc category to which both topic and vehicle can be said to belong, how does a hearer know which one was intended? The Class-Inclusion Theory lacks adequate interpretive tools to answer this question.

6. The Emergence Problem

I have shown that much of current research on metaphor has moved away from 'feature matching' models of metaphor (e.g. Tversky, 1977; Ortony, 1979) and so from the idea, inherent to these models, that metaphor comprehension involves the matching of properties between topic and vehicle. They have argued instead that metaphor interpretation is very much a matter of attributing a subset of properties of the metaphor vehicle to the metaphor topic (Gentner, Bowdle, Wolff and Boronat, 2001; Glucksberg, 2001; Glucksberg and Keysar, 1990; Tourangeau and Rips, 1991; Tourangeau and Sternberg, 1981). In other words, it is because pigs are dirty and smell bad that in interpreting the utterance 'my flatmate is a pig', the hearer is entitled to assign the properties of dirtiness and bad smell (of the metaphor vehicle) to the speaker's flatmate (i.e. the metaphor topic). A very serious problem for both matching models and attribution models is that sometimes the set of properties which are attributed to the topic are not stored as part of our representation of the vehicle, as in (14)-(15):

(14) *Doctor:* I am afraid the surgeon who performed a caesarean on your wife perforated both ovaries. I had no choice but to remove them.

Husband: I want that surgeon out of the hospital. That surgeon is a butcher!

(15) *Jane:* I know I have to speak to my boss but I am afraid of him. He is such a bulldozer!

The speaker in (14) may be expressing the thought(s) that his wife's surgeon is highly incompetent, dangerous, careless, etc. The speaker in (15) may be expressing the thought(s) that her boss is stubborn, difficult to deal with, that he is not respectful to her, that he undermines her needs, her thoughts, etc. The problem raised by these examples is that our knowledge of butchers does not include the assumption that butchers are negligent and careless and our knowledge of bulldozers does not include the assumption that they are disrespectful or stubborn. Since the set of intended properties are not stored as part of our representation of the vehicle, they can be neither matched with the properties of the topic nor attributed to it. Both matching and attribution models therefore fail to explain how these properties are derived.

Properties which are not part of the hearer's representation for the metaphor vehicle or the metaphor topic, but which seem to emerge in interpreting a metaphor, are often referred to in the literature as 'emergent properties' or 'emergent features'. Examples (14) and (15) show how emergent features play a crucial role in arriving at the meaning the speaker intended to communicate in uttering a metaphor. It follows from this that any adequate account of metaphor interpretation should aim to provide an explanation of how these emergent features are derived. I shall refer to this as the 'emergence problem' of metaphor interpretation.

6.1 Experimental Work on Emergence

Experimental research has shown that 'emergent features' play a fundamental role in metaphor interpretation. Tourangeau and Rips (1991), for instance, found that in providing interpretations for a list of metaphors, subjects produced more emergent features than common features. Furthermore, they judged emergent features as more relevant to interpretation than either topic-based, vehicle-based or common features. So, for a metaphor such as 'the eagle is the lion among birds', a feature such as 'is

respected', which is associated neither to the topic nor the vehicle individually, was found to figure prominently in subjects' reported interpretations as well as being judged as highly relevant to those interpretations. Findings like these are repeated across the literature. Gineste and colleagues, for instance, show that over 60% of the properties produced during the processing of poetic metaphors emerge during interpretation (Gineste, Indurkya and Scart, 2000). So, for a metaphor such as 'the kiss is a fruit', subjects produced properties such as 'intense' or 'reward', which are not normally used to characterise either the topic or the vehicle individually. In a series of experiments, Becker (1997) also found that significantly more emergent features and vehicle-based features appear in subjects' interpretations of metaphors than topic-based or common features. Finally, rather than asking subjects to report interpretations, Tourangeau and Rips (1991) provided subjects with two possible interpretations for a set of poetic metaphors, one based on features common to topic and vehicle, the other based on features which were not commonly associated with either but were nevertheless relevant to interpretation. They found subjects systematically preferred the interpretations based on emergent features.

Scholars generally agree that the existence of 'emergent properties' fits nicely within the interaction view of metaphor (e.g. Gineste et al., 2000). However, saying that features emerge from interaction is not explanatory: it is necessary to spell out how it is that they are derived. One should then expect the cognitive models inspired by Black's ideas to provide a detailed account of the pragmatic or cognitive steps involved in the derivation of new mental structures and the emergence of new properties. Unfortunately, although a substantial amount of experimental research has been stimulated by the romantic idea of metaphor as powerful and creative, very little work has been done to explain how emergent properties are derived. In fact, experimental work which deals explicitly with the issue, such as that presented above, has mostly been concerned with presenting evidence for the existence of emergent features rather than explanation of the cognitive processes involved in their derivation.

The lack of work on accounting for the derivation of emergent properties in metaphor interpretation is surprising not only because solving the 'emergence problem' is essential for understanding how metaphors are understood but also because most

modern approaches to metaphor are based on the assumption that something new is created in interpreting a metaphor. The issue of emergent properties is a thus a problem for all theories which aim to account for how hearers arrive at the interpretation intended by the speaker's use of a metaphor, particularly for interactive views which aim to show how metaphor interpretation does not rely on existing similarities but constructs new similarities. It is outside the scope of my work to discuss each and every theory of metaphor, not even each and every interactive approach. So I have chosen to comment on some of the approaches to metaphor which I have considered more representative.

6.2 Emergence and the Class-Inclusion Theory

It follows from the Class-Inclusion approach that the ad hoc category constructed in interpreting the metaphor and which the vehicle is taken to exemplify is constructed by selecting a subset of properties from the metaphor vehicle, those properties which are capable of assigning a value to a dimension in the topic. Glucksberg and colleagues often illustrate their ideas with the example 'my surgeon is a butcher'. They argue that in understanding this metaphor, the hearer aligns vehicle properties and topic dimensions, thus constructing an attributive category 'people who are incompetent and who grossly botch their jobs', which the vehicle typifies and which can assign a negative value to the dimension of 'skill' provided by the topic (Glucksberg, 2001: 43-55).

There is an important problem inherent in this well-known example which has, surprisingly, not being spotted in the literature (to my knowledge): how can people construct the ad hoc attributive category 'people who are incompetent and who grossly botch their jobs' by selecting a subset of properties from the metaphor vehicle if the property of 'botching their jobs' is not part of our representation of butchers? Our knowledge of real butchers may include the assumptions that they cut and sell meat, that they use sharp knives, etc. It does not, however, include the assumptions that butchers are incompetent, negligent, careless or people who botch their jobs. If we thought butchers were generally incompetent, we would not trust them and would never buy food from them. Since these properties are not associated with the metaphor vehicle, and since the Class-Inclusion view takes the ad hoc attributive category to be formed by selecting properties from the vehicle, it is not clear how this category is ever formed.

Lacking adequate machinery to construct the ad hoc category the speaker intended to convey in producing the metaphor, the Class-Inclusion theory cannot account for how emergent properties are derived.

6.3 Emergence and Blending Theory

Supporters of Blending Theory or Conceptual Integration Theory (Fauconnier and Turner, 1998) claim to be able to account for the emergence of properties during utterance processing. I analyse this claim here. According to Blending Theory, an important process in human cognition is that which takes mental spaces as input to blend them into another space. As a result of this blend, a new structure and a set of new properties arise. Blending is argued to be not unique to metaphor processing but a common feature of thinking and processing information more generally.

Pursuing Blending Theory, Grady, Oakley, and Coulson (2000) have provided an analysis of the metaphor 'my surgeon is a butcher' and of the emergence resulting in interpreting this metaphor. According to these authors, understanding this metaphor involves taking the mental space corresponding to the topic (a scenario in which there is a surgeon doing his job) and the mental space corresponding to the vehicle (a scenario in which a butcher is doing his job) and blending them into a single space. During this process, some elements of each scenario are projected to the blended space. Information projected from the butcher scenario may include for instance the role of the agent performing the action and the set of instruments used while information projected from the surgeon scenario may include the role of the patient and place (operating table). This projection, it is claimed, results in the construction of a blended space in which a butcher is operating on a patient in an operating table. According to these authors the projection of elements from each of these conceptual spaces would be altered if the metaphor is reversed, as in 'this butcher is a surgeon' or 'he is not a butcher, he is a surgeon' said of an extremely careful butcher. In these cases, the structure will be reversed so that the blended space will have a surgeon cutting a piece of meat in a butchery.

According to Blending Theory, comprehension involves three main steps: composition, completion and elaboration. The process just described above in which some elements of the input spaces are selected and projected into the blended space

corresponds to the composition process. Once the information is in the blended space, it is completed with information obtained from long term memory. The idea is that in order to make sense of the scene constructed in the blended space, we may introduce other information into that space. In the case of the metaphor above, this may include adding the notions of 'incompetence' or 'malice'. According to Grady et al. (2000) this process may in fact "continue indefinitely" reaching what they refer to as 'the elaboration process' whereby the blended space is enriched even further. The elaboration process may lead hearers to imagine a range of different things, they may imagine, for instance, that the butcher is hanging the patient from a hook or is throwing human tissue into the bin after operating, etc.

We can see from this picture that in Blending Theory, processing a metaphor (in fact, processing information more generally) is essentially a creative process in which a new mental structure (a blended space) is constructed. It is this blended space that is said to account for the emergence of new properties. It is important to notice, however, that the blended space provides us with a certain representation which cannot be the one the speaker intended the hearer to derive. The speaker of the metaphor above for instance does not intend to communicate that there is a butcher operating on a patient but that there is a certain surgeon who does not do his job properly. The blended space provides information which is indeed consistent with a literal interpretation of the utterance, the interpretation that my surgeon is a real butcher! Attempting to explain how one gets from this interpretation to the intended one implies a variant of the standard serial model of metaphor interpretation so widely criticised among psychologists. Maybe the hearer is simply supposed to take the blended space metaphorically so as to derive the set of thoughts the speaker intended to convey. If this is true then forming the blended space does not account for how metaphors are understood and just takes us into needless circularity.

Scholars pursuing Blending Theory argue that emergent properties arise naturally from the construction of the blended space. But if a blended space is constructed by projecting information from different sources, namely input spaces and encyclopaedic information, how can anything 'emerge'? With regard to the example being discussed, it is not clear how the properties of 'being negligent', 'liable to be sued for medical

malpractice', etc. which arise in interpreting the utterance 'my surgeon is a butcher' can be obtained by enriching the blended space proposed by Grady et al. (2000) as in this space there is a butcher operating on a patient and butchers cannot be sued for medical malpractice. Saying that emergent features arise by adding information from long term memory does not provide a solution to the problem as this completion process may take different directions not intended by the speaker.

An important reason why the construction of a blended space alone cannot provide an explanation for metaphor comprehension or property emergence is that the same blended space (or a mental image) can be formed in processing a range of situations, utterances and texts (e.g. in watching a fictional movie in which there is a butcher cutting a man in pieces, in hearing the news that a psychiatric patient made his way to the operating room with a butcher's knife, etc.). In fact, different utterances of a single sentence uttered on different occasions (e.g. 'this butcher is a surgeon', 'he is not a butcher, he is a surgeon', etc. may lead to the formation of the same blended space even if the speaker intended to communicate different sets of implications on different occasions. It is possible that although the construction of a mental space, scenario or image in which a butcher is cutting a patient does not account for the derivation of the interpretation the speaker intended to convey, it may, nevertheless, help direct the hearer's attention towards the intended set of implications (e.g. implications about incompetence, negligence, carelessness) playing a role in interpretation.

One important problem with Blending Theory, and with many psycholinguistic approaches to metaphor, is that it does not take seriously into account the speaker's communicative intentions. I have shown earlier how a single metaphor 'John is an iron bar' or 'my lawyer is a shark' can be used to convey a number of different meanings on different occasions. In order to explain this in terms of Blending Theory, one would have to say the hearer forms a different blend in every occasion. It is not clear how this can be done. Since the projection from input spaces to the blended space is taken to be based on structural similarities between spaces and not in the search for the recognition of speaker's intentions, there is no apparent reason why different elements from an input space would be projected into the blended space on different occasions. In fact, even if the explanation of different interpretations were to be given in terms of different types of

completions of the blend, the theory cannot explain what determines these different completions.

7. The Transformation Problem

That metaphor interpretation is much more than the selection and attribution of features can be illustrated by a type of emergence problem which I refer to as the 'transformation problem', as in (16)-(17):

(16) *Julie:* I am afraid about the divorce. My husband's lawyer is a shark.

(17) *Mary:* Are you sure your husband does not mind looking after the children the whole weekend?
Jane: Yes, don't worry about it. He is a teddy bear![11]

The speaker in (16) may want to communicate that her husband's lawyer is strong and aggressive, that he will attack her in court and persist until he achieves his goals. A case like this presents no apparent problem for attribution theories which take some properties of the metaphor vehicle to be attributed to the topic. In this case, a subset of our knowledge of sharks (e.g. that they are aggressive, persistent, strong, etc.) is selected in context and attributed to the topic of the metaphor. The Class-Inclusion view often uses the (related) metaphor 'my lawyer is a shark' to claim that the hearer takes this metaphor to convey not the assertion that the speaker's lawyer is an animal which lives in deep waters, but rather the assertion that her lawyer belongs to the category of 'people and animals who are aggressive and obstinate' (e.g. Glucksberg, 2001). This category, they claim, is constructed ad hoc by selecting a subset of properties of the vehicle which can assign values to a set of dimensions in the topic. In this case, this would involve selecting the properties of aggressiveness and persistence which can be used to assign a (negative) value to the dimension of 'character' provided by the metaphor topic.

One important reason why this account is problematic is that although lawyers and sharks are both aggressive, obstinate and persistent, they are so in very different ways.

The property of 'aggressiveness' which is attributed to the topic is not the property associated with the encyclopaedic entry of the metaphor vehicle, but a related property which denotes a different kind and degree of aggressiveness. This property, call it aggressiveness*, seems to 'emerge' in interpreting the utterance from this particular subject-predicate combination. Interpreting the metaphor in (16), then, cannot be reduced to the selection of vehicle properties and attribution of these properties; some transformation needs to take place.

(17) presents a clearer case. The speaker of (17) intends to convey some of a range of assumptions such as that her husband is nice, easy going, always willing to help, easy to please, good with children, etc. If metaphor interpretation involves the attribution of vehicle properties to metaphor topic, the hearer of (17) may access the assumption that teddy bears are soft and cuddly and attribute these properties to Jane's husband. Unlike (16), in which one can literally say that lawyers can be aggressive, the way in which Jane's husband is soft is only metaphorical. Thus, understanding the metaphor in (17) cannot be reduced to selecting the physical property of softness common to teddy bears and attributing it to the metaphor topic; some transformation needs to take place.

Although the need to transform the properties of the vehicle into properties that can be appropriately attributed to the topic has been widely acknowledged, very little has been done to provide a solution to the problem. Black himself admits that his model cannot account for it:

> "A fairly obvious objection to the foregoing sketch of the "interaction view" is that it has to hold that some of the 'associated commonplaces' themselves suffer metaphorical change of meaning in the process of transfer from the subsidiary to the principal subject. And these changes, if they occur, can hardly be explained by the account given." (Black, 1962: 42)

> "[B]ecause features are specific to a domain, they must be transformed, i.e. seen in a new way, if we are to find correspondences across domains." (Tourangeau and Sternberg, 1981: 217)

> "The way in which men prey on women is different from the way wolves prey on animals." (Gentner and Bowdle, 2001: 227)

---

[11] The example 'my husband is a teddy bear' has been borrowed from Ortony's work (see Ortony, 1993).

> "The way in which wolves are predators is different from the way men are predators, which in turn is different from the way sharks are predators and lawyers are predators." (Glucksberg, 2001: 36)

One theory which seems to take the issue very seriously is the Domain Interaction theory, proposed by Tourangeau and Sternberg (1981). Tourangeau and Sternberg propose that in metaphor interpretation we align topic and vehicle so that they occupy parallel positions within their own domains. Metaphor invites us to see the topic in terms of the vehicle, for which we construe features and dimensions that apply to the topic (e.g. 'men' in 'men are wolves') that are parallel to those applying to the vehicle ('wolves') (for a similar approach see Gentner, 1983). Since metaphor interpretation involves the mapping of features that apply within one domain (e.g. the domain of wolves) to an object from a different domain (e.g. men) and since a given feature may not apply outside its domain, the authors claim that topic and vehicle need not share any features - they may well all be emergent. So for the metaphor 'men are wolves', the feature 'being a predator' which applies literally to wolves can only apply metaphorically to men.

According to the authors, knowledge about the domains specifies both the features and dimensions which are important to interpret the metaphor on a particular occasion, as well as the nature and degree of the parallel that is constructed between topic and vehicle:

> "Because we know what sorts of things social relations are, we can interpret 'men are wolves'; we know not to apply the characteristics of wolves literally to men; and know how these characteristics must be transformed, i.e., interpreted in a new way, to apply to people. Thus, domains sometimes tell us which characteristics of tenor and vehicle are likely to matter in interpreting the metaphor. And they also tell us how to map the features applying within the one domain onto those applying within the other" (Tourangeau and Sternberg, 1981: 216)

The problem with this is that the hearer's knowledge of social relations or the domains represented by topic and vehicle cannot by itself enable him to derive the intended interpretation of the metaphor 'men are wolves', or of any other metaphor or utterance for that matter; stronger (pragmatic) constraints are clearly needed. That is, although the

knowledge the hearer has stored about the entities denoted by the concepts encoded by an utterance plays an important role in comprehension, it cannot, in itself, either specify the properties which would be relevant to that interpretation nor determine how properties of the vehicle are transformed during the comprehension process so that they can be attributed to the topic.

The main contribution of the Domain Interaction view of metaphor theory has been to suggest that metaphor vehicle and metaphor topic do not need to share any (exact) property. The main disadvantage with this model is that, in order to explain how this takes place, it makes the naïve claim that the notion of domain, together with contextual cues, can determine property emergence and transformation. The need to transform the properties of the vehicle into properties that can be appropriately attributed to the topic, which Tourangeau and Sternberg mention, is what I refer to as the 'transformation problem'.

Although the transformation problem has been widely recognised across metaphor models, very little has been done to solve it. Two main proposals have been considered and rejected. First, the proposal that one may attribute features of the vehicle to the topic on the basis of similarities rather than identity has been rejected on the ground that it would lead to an indefinite regress (Gentner, 1983). Second, the proposal that one may take the assumptions associated with the vehicle as metaphorical has been rejected on the same ground. As Carston points out, considering the encyclopaedic assumptions the concepts encoded by a metaphor give access to as metaphorical "does not break through the metaphorical web" (Carston, 2002c: 87) and so does not allow us to provide an explicit account of how metaphor interpretation takes place. Although I agree that an approach based on similarity of properties cannot adequately account for metaphor interpretation and the transformation of properties, I do think the 'metaphor within metaphor' idea is worth exploring and I do so in the next chapter.

8. Conclusion

Modern cognitive approaches to metaphor have gradually moved away from a view of metaphor interpretation as parasitic on literal meaning and the derivation of comparison

statements. These approaches are generally based on two main assumptions. On the one hand, they argue that metaphor interpretation involves some kind of interactive process where some features (or structural relations) of metaphor topic and, particularly, of metaphor vehicle are (mutually) selected. On the other hand, they claim that metaphor interpretation is essentially a creative process which is not dependent on pre-existing similarities but which results in the emergence of similarities. A problem common to all these approaches is that they lack the pragmatic inferential mechanisms necessary to guide the comprehension process and to account for the attribution of properties and the derivation of emergent properties taking place in interpreting a metaphor.

The greatest advantage of the Class-Inclusion theory over a range of existing (interactive) cognitive models of metaphor is that it treats metaphor as involving not comparison but categorisation. The greatest disadvantage is that it does not provide a precise account of how the new categories being conveyed are formed during interpretation. Glucksberg and colleagues acknowledge these gaps in their theory and the need for a pragmatic theory capable of filling them:

> "How are new categories created by metaphors? We may not be able to answer this question in a complete or definite way, but we are certain about how new categories are not created. They are not created by simply using some subset of the properties in common" (Glucksberg, Manfredi and McGlone, 1997: 346)

> "Our account of metaphor categorizations [...] recasts the problem of how people come to understand metaphors. It does not solve that problem, but does outline what an adequate psychological model might look like. Such a model of metaphor comprehension will have to include general principles of discourse comprehension, such as Grice's co-operative principle and the given-new convention, as well as the more specific principles of conversational interaction and inference discussed by Searle" (Glucksberg and Keysar, 1993: 424)

In the next chapter, I will show how the pragmatic approach to metaphor suggested by current lines of research in Relevance Theory offers solutions to many of the problems raised here. Because it is developed within a general pragmatic framework designed to deal with utterance interpretation and ostensive communication, the relevance-theoretic account of metaphor suggests precise inferential mechanisms which can explain how

hypotheses about the intended meanings are derived on every occasion, how ad hoc concepts (denoting ad hoc categories) are constructed as part of the overall interpretation process and how emergent properties are derived during on-line comprehension. Unlike standard pragmatic approaches, Relevance Theory provides a theoretically adequate and cognitively plausible pragmatic approach to metaphor comprehension which complements nicely the experimental evidence and arguments developed in this chapter.

Chapter 4

Relevance Theory and Metaphor Interpretation

> "If we are right, metaphors are based on fundamental and universal psychological mechanisms. They are in no sense departures from a norm, or, as modern pragmatists would have it, breaches of a rule or maxim of communication [...] Words and sentences have a literal meaning, but that meaning is an instrument of communication rather than its content. What hearers expect is that the literal meaning of an utterance will help them infer with a minimum of effort the thought that the speaker intends to convey. This expectation itself derives from, and is warranted by, a more basic expectation of relevance, which is automatically encouraged by any act of communication."
>
> Dan Sperber and Deirdre Wilson (1990: 149)

1. Introduction

In the previous chapter, I argued that although modern cognitive approaches have moved towards a view of metaphor and metaphor interpretation as a creative process in which new meanings are constructed and similarities emerge, they do not offer the inferential machinery necessary to explain how these new meanings and new representations are derived during the interpretation process. In this chapter, I present the approach to metaphor defended in Relevance Theory and show how it provides the appropriate inferential tools to account for metaphor interpretation successfully, and to do this without resorting to the Literal Priority Claim. I develop the relevance-theoretic approach to metaphor further, showing how it can deal with a number of problems not accounted for by existing cognitive approaches, including such problems as the 'emergence' and 'transformation' problems introduced in the previous chapter.

2. Relevance, Literalness and Metaphor Interpretation.

Standard pragmatic approaches to metaphor, as we have seen, are well-known for pursuing a version of the literalist approach to comprehension (Grice, 1975/1989; Searle, 1979a). This is an approach which looks at metaphor as dependent and parasitic on, as

well as deviant from, literal meaning. The literalist approach to metaphor interpretation has been strongly challenged in the last twenty-five years by the results of psycholinguistic experiments such as those outlined in the previous chapter. The inability of the standard pragmatic position to account for metaphor comprehension adequately has led scholars to approach the issue in 'non-pragmatic' ways, by resorting, mostly, to psycholinguistics. This movement away from pragmatics seems to have been partly based on the misconception that any pragmatic model of metaphor would necessarily favour literalism: "thinking of figurative language as a strictly pragmatic phenomenon perpetuates the traditional view that such speech is deviant or, at best, ornamental" (Gibbs, 1994a: 5). Once this belief is abandoned, there is no reason why pragmatics should not be considered an adequate and, in fact, essential approach to the study of metaphor.

With the birth of the pragmatic framework of Relevance Theory, some psychologists, including Gibbs himself, have accepted the possibility that a pragmatic analysis of metaphor can be provided without committing to the Literal Priority Claim. Relevance Theory, Gibbs and Tendahl acknowledge, "is consistent with many of the findings in psycholinguistics on metaphor understanding, and can account for aspects of metaphor understanding that no other theory can explain" (Gibbs and Tendahl, forthcoming). The relevance-theoretic approach to metaphor is not liable to the criticisms made by psycholinguists against standard pragmatic models because the RT pragmatic framework is not based on the assumption that speakers and hearers aim at literalness or truthfulness but on the assumption that they aim at optimal relevance (Carston, 2002a; Sperber and Wilson, 1986/1995, 1987, 1991; Wilson, 1995; Wilson and Sperber, 2000).

Recall that, according to the Communicative Principle of Relevance, every utterance (or other ostensive act) conveys a presumption of its own optimal relevance. That is, it conveys a presumption that it is the most relevant utterance the speaker could have produced which is compatible with her abilities and preferences at the time and that it is at least relevant enough to be worth the hearer's attention and processing effort. On the one hand, a speaker aiming at optimal relevance would not necessarily be expected to produce a literally true utterance but rather an utterance which she thinks will be (or

will seem to be) optimally relevant to the hearer at the time. On the other hand, a hearer looking for optimal relevance would not necessarily expect to find that the proposition literally expressed by the speaker is true but rather that it meets his expectations of relevance. These expectations are often satisfied by a loose interpretation, as in (1a), (2a) and (3a):

(1) *Peter:* How much did you say you earn in your new job?
   a. *Jane:* 800 pounds a month.
   b. *Jane:* 789.30 pounds a month.

(2) *Jeremy:* Your mother said you are spending some time abroad.
   a. *Ruth:* Yes, I live in Madrid.
   b. *Ruth:* Yes, I live in Alcalá. (On the outskirts of Madrid)

(3) *Max:* Why do I have to clean my room?
   a. *Max's mother:* Because it is a pigsty.
   b. *Max's mother:* Because your room is very dirty, untidy, smells bad and makes me feel sick.

Imagine that the second speaker in (1)-(3) wants to communicate a set of implications I_1, I_2, I_3...I_n. In (1) she would like to convey some implications about her finances, expenses and lifestyle; in (2), about her contact with Spanish culture, her absence from home, her access to new experiences, etc.; in (3), about the filthiness and mess in the hearer's room, etc. These implications are derivable as logical and contextual implications either from a strictly literal utterance P, as in (1b), (2b) and (3b), or from a related but strictly false utterance Q, as in (1a), (2a) and (3a). Deriving these implications from the literally true utterance P would involve more processing effort than deriving them from the strictly false utterance Q. The utterance in (1b) is logically more complex than (1a); (2b) increases the risk of being misunderstood by someone not familiar with the area, and the name Alcalá, being less frequently used, which will cost more processing effort; and (3b) is longer and more linguistically complex than (3a). In processing Q, however, a

hearer who takes it literally could derive not only all the implications he can derive from processing P, but also a range of other implications the speaker does not intend to endorse (e.g. the implication that the speaker earns exactly 800 pounds a month, that there are pigs in the hearer's room, etc.). So which of these two utterances should the speaker choose to express?

Relevance Theory claims that in most circumstances, speakers would choose to produce the more economical but literally false utterance Q. Aiming at optimal relevance, she would choose the utterance which yields the intended implications (and other cognitive effects) for the investment of the least processing effort. Since the hearer can derive the set of intended implications (and other positive cognitive effects) from the more economical utterance Q, the speaker should be expected to produce this utterance, and leave it to the hearer's expectations of relevance and the relevance-theoretic comprehension procedure to guide him towards the intended interpretation.

On the assumption that the speaker of (1a), (2a) and (3a) is being optimally relevant, the hearer is entitled to follow a path of least effort in processing her utterance and to take the first interpretation which satisfies his expectations of relevance as the one she intended to convey. Guided by the relevance-theoretic comprehension procedure and the expectations of relevance raised by the speaker's utterance, he would start considering hypotheses about the speaker's meaning in their order of accessibility. In processing (1a), for instance, this may involve deriving implications about the speaker's finances and her lifestyle. In processing (2a), it may involve deriving implications about the speaker's contact with Spanish culture and her absence from home; and in processing (3a), it may involve deriving implications about the filthy state of the hearer's room and the need to tidy it up. This process continues with the hearer considering possible contextual assumptions and implications until he has enough cognitive effects to satisfy his expectations of relevance, at which point he stops. A consequence of following the path of least effort is that the hearer often arrives at a satisfactory interpretation after processing only a subset of the implications that would have been derivable if the utterance had been literally understood. In processing (1a), (2a) and, (3a), for instance, he may arrive at a satisfactory interpretation without even considering that the speaker earns exactly eight hundred pounds, or that the bedroom is

literally fit for a pig. In fact, in most circumstances, even if these assumptions are considered at some point during the interpretation process, the strong pragmatic constraints at work would lead to their quick rejection for not providing the expected sort of cognitive effects.

This does not mean, however, that hearers never derive literal interpretations, but only that they are often satisfied by an interpretation which falls short of being strictly literal. The expectations of relevance raised in the hearer by the speaker's utterance play a major role in interpretation, and whether these expectations can be satisfied by a loose interpretation or a literal interpretation varies from utterance to utterance. Imagine that the speaker in (1)-(3) is not talking to a friend or a relative but to a tax inspector or a council officer. In this case, unlike in most other circumstances, the hearer's expectations of relevance would be very precise and would not be satisfied by anything short of a literal interpretation. Since the set of intended cognitive effects the speaker intends to cannot be derived by processing the more economical but literally false utterance (Q), as in (1a), (2a) and (3a), the speaker would be expected to produce the literally true utterances (P), as in (1b), (2b) and (3b).

The examples in (1)-(3) illustrate an essential assumption of Relevance Theory: the assumption that there is no clear-cut division between cases of approximation, hyperbole and metaphor, which are all types of loose uses of language, nor between these loose and literal uses. An utterance, whether intended literally or loosely, is approached with more or less precise expectations of optimal relevance and processed following the same comprehension procedure until those expectations are satisfied. Standard pragmatic models, which look at verbal communication as governed by a maxim, norm or convention of truthfulness or literalness, are not capable of explaining the pervasive use of approximation in communication (as in (1a) and (2a)), or the relation between different loose uses of language including metaphorical speech and approximation (e.g. the relation between (1a), (2a) and (3a)). Also, in taking the view that a literal interpretation is necessarily derived before a non-literal interpretation is considered these models cannot account for the ease with which metaphorical interpretations are generally derived.

2.1 Lexical Pragmatics and Loose Use

The treatment of loose uses in Relevance Theory has undergone some changes through the years (see Carston, 1996, 2002a, 2002c; Sperber and Wilson, 2000; Wilson, 2004). In their book *Relevance* and several other papers, Sperber and Wilson argued that the speaker of a loosely used utterance, such as those in (1a), (2a) and (3a), does not endorse the proposition her utterance literally expresses (Q above), which is, therefore, not an explicature, but merely uses it as an effective means by which to communicate a set of implications (I_1, I_2, I_3...I_n.) which she does endorse (e.g. Sperber and Wilson, 1985; 1986/1995, 1987, 1990). On current views, both the set of implications (I_1, I_2, I_3...I_n) and the (looser) proposition expressed by the speaker's utterance are communicated (Carston, 1996, 2002a, 2002c; Sperber and Wilson, 1998, 2000; Wilson, 2004; Wilson and Sperber, 2002).

Departing from earlier code model approaches to communication and developing the inferential approach to communication of Paul Grice, we have seen that Sperber and Wilson's Relevance Theory makes the assumption that decoded meaning is not directly accepted as the speaker's intended meaning, but merely taken as evidence from which to infer the meaning she intends to communicate. A common assumption in Relevance Theory, which affects both the early and new accounts equally, is the idea that the concepts encoded by the words in an utterance provide access to a range of assumptions in memory which the hearer will consider, in their order of accessibility, constructing hypotheses about the implications the speaker might have intended to convey. Decoding the concept encoded by the word *pigsty* in (3a), for instance, gives the hearer access to assumptions about pigsties (e.g. that they smell badly, that they are messy, full of rubbish, etc.) which he considers, in their order of accessibility, in the interpretation process. According to the early view, the proposition expressed by the speaker's utterance was seen as containing as constituents the concepts resulting from linguistic decoding. Being literally false, the propositions expressed by (1a), (2a) and (3a) are not communicated by the speaker but merely used as a way to derive a range of true implications which she is taken to endorse. Under the current view, processing assumptions made accessible by the encoded concepts generally results in adjustment of the denotation of these concepts so that new concepts arise ad hoc (ad hoc concepts). In

(3a), for instance, considering a subset of assumptions about pigsties in deriving the set of intended implications (e.g. that the speaker's room is dirty, filthy, untidy and smells bad) may alter the denotation of the encoded concept PIGSTY, which is broadened so as to include in its denotation places which are not real pigsties but which are dirty, filthy, untidy and smell bad. It is this new, broader ad hoc concept, PIGSTY*, constructed during the comprehension process, and not the encoded concept that, according to the current approach, is taken to be a constituent of the proposition expressed by the utterance. The speaker is thus taken to endorse not only the set of implications derived but also the proposition expressed, which is therefore, in relevance-theoretic terms, an explicature, as in (4):

(4) MAX(x)'S ROOM IS A PIGSTY*.

The process of pragmatic adjustment of lexically encoded concepts was presented in Chapter Two and will be discussed further in this chapter. We saw in Chapter Two how not all of the assumptions a concept gives access to in memory are equally accessible at a given moment, and how the expectations of relevance raised by an utterance generally add an extra level of activation to a certain subset of assumptions which become more accessible to the hearer as a result. It follows from the relevance-theoretic comprehension procedure that the hearer should consider the most accessible assumptions first, and proceed through the accessibility hierarchy until he reaches an interpretation that satisfies his expectations of relevance. Since the accessibility of encyclopaedic assumptions in memory varies from utterance to utterance and occasion to occasion, a different subset of assumptions may be considered on virtually every occasion on which the concept is processed. The selected subset of encyclopaedic assumptions affects the implications derived and so has a bearing on whether the utterance is understood literally, approximately, hyperbolically or metaphorically, as in (1)-(3). Consider the following example, discussed in Sperber and Wilson (2002: 19-20):

(5) *Peter:* Can we trust John to do as we tell him and defend the interests of the Linguistics Department in the University Council?

Mary: John is a soldier!

(6) *Peter:* What does John do for a living?
Mary: John is a soldier!

The concept encoded by the word *soldier* gives access in memory to a wide array of encyclopaedic assumptions, which are activated to different degrees on hearing the word. Some of which contextual assumptions might be those in (7) which Sperber and Wilson (2002) suggest might combine with Mary's utterance to yield the implications in (8):

(7) a. A soldier is devoted to his duty.
b. A soldier willingly follows orders.
c. A soldier does not question authority.
d. A soldier identifies with the goals of his team.
e. A soldier is a patriot.
f. A soldier earns a soldier's pay.
g. A soldier is a member of the military.

(8) a. John is devoted to his duty.
b. John willingly follows orders.
c. John does not question authority.
d. John identifies with the goals of his team.
e. John is a patriot.
f. John earns a soldier's pay.
g. John is a member of the military.

In processing an utterance containing the word *soldier*, such as those in (5) and (6), some subset of the encyclopaedic assumptions in (7) would receive a higher degree of activation, and so be more accessible, than others. In processing (5), for instance, given the context provided, by Peter's question, the order of accessibility may be as shown in

(7) with (7a) as the most accessible, whereas in processing (6) it may be just the reverse. In either case, the hearer, guided by the relevance-theoretic comprehension procedure, would follow a path of least effort, considering the encyclopaedic assumptions in (7) in their order of accessibility, and start deriving the linked implications in (8) in context. The process stops once the hearer has arrived at a combination of explicit content, context and implicatures which satisfies his expectations of relevance. In processing (5), Peter may satisfy his expectations of relevance by only the implications in (8a)-(8d). It follows that, contrary to the prediction of standard pragmatic models, he may not even consider the assumption in (7g), and so may never derive the implication in (8g) or derive a literal interpretation of Mary's utterance. By contrast, in interpreting (6), Peter may satisfy his expectations of relevance after having considered only the encyclopaedic assumptions in (7g) and (7f) and so deriving only the implication in (8g) and (8f). The result of this process would be a literal interpretation. Thus, although the same comprehension procedure operates in interpreting Mary's identical utterances in (5) and (6), the consideration of a different subset of the encyclopaedic assumptions associated with the concept SOLDIER results in different implications being derived, and so in different interpretations being constructed. On some occasions, the interpretation resulting from this relevance-driven comprehension, may be classified as metaphorical, as in (5), and on other occasions, as literal, as in (6).

Relevance-driven comprehension, then, may involve relatively shallow processing of the encoded concept, which results in the (one-off) assemblage of a subset of the encyclopaedic assumptions made accessible. Processing only selective bits of encyclopaedic information may result in the construction of a new ad hoc concept, with a different denotation and implications. This is not the case with (6), as selecting the assumptions in (7g) and (7h) and deriving the implications in (8g) and (8h) does not result in any modification of the encoded concept. In (5), however, deriving a certain set of implications about John, such as those in (7a)-(7d), results in a modification of the encoded concept, which is broadened on-line so that the denotation of the newly constructed concept includes not only actual soldiers but also people who are not members of the military but who are team workers, do as they are told, and defend other people's interests. It is this new broadened concept SOLDIER*, constructed on the fly in

interpreting the speaker's utterance, which is taken to be a constituent of the explicature of the utterance:

(9) Explicature of (5): JOHN(x) IS A SOLDIER*
(10) Explicature of (6): JOHN(x) IS A SOLDIER

We have seen in this thesis how, according to current lines of research in Relevance Theory, virtually every encoded concept may be pragmatically adjusted in the course of interpretation in this way. The output of this fine-tuning may be a concept which has roughly the same denotation as the encoded concept (a literal use, as in (6)); a narrower denotation (a case of lexical narrowing, as in 'the *fish* attacked the swimmer', which denotes shark-like fish, or 'I fed the *fish* in the tank', which denotes goldfish-like fish); a denotation which is slightly broader than that of the encoded concept (a case of approximation, as in (1a)); or a denotation which is considerably broader than that of the encoded concept, and which may result from category extension (as in the comprehension of metaphor and hyperbole in (3a) and (5)).

Thus, unlike standard approaches to metaphor, the pragmatic account defended in Relevance Theory sees the comprehension of metaphor, like the comprehension of any utterance, as a constructive process which does not depend on the prior derivation of a literal interpretation. Utterance comprehension is not guided by a presumption of literalness (or of truthfulness) but of optimal relevance (Sperber and Wilson, 1986/1995, 1987, 1991; Wilson, 1995; Wilson and Sperber, 2000). Whether the resulting interpretation is characterised (post-comprehension) as literal, approximate, hyperbolic, metaphorical, etc. is, to a certain extent, an arbitrary folk-linguistic matter. The relevance-theoretic approach to metaphor is therefore not liable to the criticisms properly levelled at pragmatic views which adopt the Literal Priority Claim.

Relevance Theory, with its unified approach to lexical pragmatics, not only argues (in line with psycholinguistie research) that metaphor interpretation is a spontaneous, non-optional constructive process, it also, and crucially, offers a single comprehension procedure to account for the inferential derivation of both literal and metaphorical interpretations. A common feature of other cognitive research on metaphor, such as that

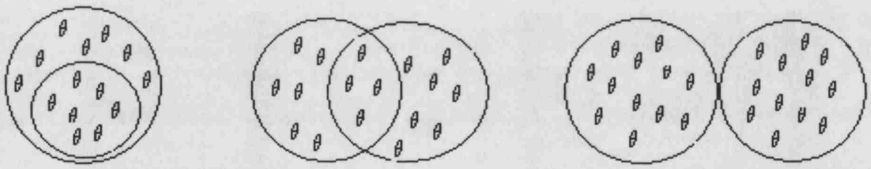
presented in the previous chapter, is that it treats the phenomenon of metaphor interpretation in isolation, not attempting to incorporate it within a broader framework in which approximation, hyperbole and literal interpretation are all approached in the same way. In approaches influenced by cognitive linguists, for instance, the comprehension of metaphor is often seen as involving interaction, mapping or alignment between two distinct domains of knowledge, a process which is not generally seen as occurring in the processing of literal speech. In Relevance Theory, the comprehension of metaphor is not seen as depending on the alignment or interaction of two domains or concepts but involves a process of mutual adjustment (of explicit content, context and implicatures) and lexical pragmatic fine-tuning which is at work in the interpretation of every utterance (whether literally or figuratively intended).

Although experimental research on metaphor provides invaluable insight into the processing of metaphorical utterances the models proposed on the basis of those findings are far from accounting for utterance comprehension, or for the place of metaphor within linguistic theory. For instance, although the experimental findings surveyed in the previous chapter provide convincing evidence against serial model of metaphor comprehension, saying that the interpretation of a metaphor does not need to involve the derivation and rejection of a literal interpretation does not, in itself, explain how utterances are comprehended and what mechanisms are involved in deriving the intended interpretation whether literal, metaphorical or otherwise. The study of figurative language in general, and metaphor in particular, desperately needs a pragmatic framework to complement experimental research. The relevance-theoretic framework provides cognitive and communicative tools which might help to fill this gap. The relevance-theoretic approach to metaphor is not only cognitively plausible and consistent with recent experimental research in psychology, but also sheds light on many basic notions in linguistic theory (e.g. notions such as explicit content, implicatures and truth conditions) which psycholinguists are not generally concerned with. In this chapter, I want to look more closely at the relevance-driven pragmatic processes at work in deriving interpretations for ordinary metaphorical utterances.

2.2 Pragmatic Adjustment and Metaphor Interpretation

We have seen how in Relevance Theory metaphor interpretation involves a pragmatic broadening of encoded concepts which results in category extension (Carston, 1996, 2002a; Sperber and Wilson, 1998; Wilson and Sperber, 2002; Wilson, 2004). I want to show, following Carston (1997; 2002: 353), that there are, in fact, at least three possible ways in which the encoded concept can be adjusted in processing a metaphor. The pragmatic adjustment process may result in the construction of an ad hoc concept which (a) denotes all the entities denoted by the encoded concept plus a range of other cases; (b) denotes only some of the entities denoted by the encoded concept plus a range of other entities or (c) denotes none of the entities denoted by the encoded concept but only a range of other entities. Providing the circles stand for sets denoted by concepts, these three possibilities can be illustrated by (11):

(11) A B C



These figures can be exemplified by the metaphorical uses in (12)-(13), (14)-(15) and (16)-(17), respectively:

(12) *A.* Why does your boyfriend want you to go with him everywhere?

B. Because he is a baby.

BABY*: denotes actual babies and also people who cannot be independent, cannot look after themselves, can't do things alone, et., including some adults such as the speaker's boyfriend.

(13) My love, my treasure.

TREASURE*: denotes extremely valuable things. These include all physical treasures and the speaker's lover.

(14) Being the only boy, Dave has always been the prince of the house.

PRINCE*: denotes that subset of actual princes who are spoilt and do as they please, as well as young boys who are not princes but are spoilt and do as they please.

(15) I am getting divorced because my husband turned out to be an eternal bachelor.

BACHELOR*: denotes a subset of unmarried adult men who party a lot with friends, flirt with women, avoid responsibilities, etc. while excluding others (e.g. the Pope, catholic bishops). It also denotes men who are not bachelors but behave as if they were (e.g. the speaker's husband).

(16) My boss is a bulldozer.

BULLDOZER*: denotes people who are disrespectful, obstinate, undermine other people's feelings and thoughts, etc. (e.g. the speaker's boss).

(14) I tried to persuade Mr Smith to change the essay topic but there was no way. He is an iron bar.

IRON BAR*: denotes people who are difficult to convince, persuade, make change their minds, etc. (e.g. the speaker's teacher).

The relevance-theoretic argument, as we have seen, is that although the concept encoded by a word gives access in memory to a range of encyclopaedic assumptions, a hearer, following a path of least effort in deriving the speaker's meaning, would only consider

and process a subset of these assumptions in interpreting an utterance. The point here is that arriving at an optimally relevant interpretation by selecting only a subset of these assumptions may result in one of the three types of pragmatic broadening in (11). In (12), for example, processing assumptions about a baby's need of care and its inability to look after itself may result in the construction of a concept BABY* which denotes people who need a lot of care and cannot look after themselves. Similarly, in (14), processing assumptions about a certain kind of bachelor's inability to accept responsibilities, tendency to promiscuity, etc. may result in the construction of a concept BACHELOR* which denotes men (single or married) who do not accept responsibilities and cannot be faithful. Finally, the pragmatic fine-tuning in the comprehension of (15) and (16) results in the construction of a concept whose denotation does not overlap at all with the denotation of the encoded concept. So the concept expressed in (16) (BULLDOZER*) denotes a range of people who behave in stubborn and insensitive ways, and not the kind of machinery denoted by the encoded concept BULLDOZER.[1]

So the set of assumptions considered in interpreting the utterance and the set of implications derived as a result may affect the pragmatic adjustment of the encoded concept and the denotation of the resulting ad hoc concept. One can see how at least some of the implications conveyed by using the metaphorical examples in (12)-(15) might be derived from combining the encyclopaedic assumptions associated with the encoded concepts (e.g. the assumption that babies need to be looked after, in (12); that treasures are valuable things in (13); that princes are spoilt, in (14); that certain bachelors do not like responsibilities, in (15), etc.) with other contextual assumptions being considered in processing the utterance. Arriving at a satisfactory interpretation by processing these contextual assumptions may involve the construction of a concept

[1] According to Relevance Theory, pragmatic adjustment processes are responsible for the comprehension of neologisms of the sort discussed by Clark and colleagues (e.g. Clark, 1983; Clark and Clark, 1979; Clark and Gerrig, 1983), such as 'the boy *porched* the newspaper', 'the rabbit *houdinied* his way out of the cage' or 'I *snaked* my way out of the classroom'. This ability to use old words with novel meanings seems to be constraint by what Clark (1987) refers to as 'the principle of constrast' which posits that no two words in a language can have exactly the same meaning. The reason why utterances such as 'the man *netted* the fish' are possible while utterances such as 'the man *ovened* the cake' are not is that whereas the English language has no word for the meaning intended by the speaker's use of '*netted*' in the first sentence, the English language does have a word to convey the meaning the speaker intended in using '*ovened*' in the second sentence, the verb 'bake'. This idea is consistent with that proposed in Relevance Theory that the stock of concepts we can create and communicate with (and even store) is much greater than the stock of words in our languages.

which is broader than the encoded concept, as in (12) and (13), or both narrower and broader, as in (14) and (15). However, examples such as those in (16)-(17) seem to be more problematic. The speaker in (16) may intend to convey the implications that his boss is stubborn, disrespectful and obstinate and the speaker in (17) may intend to convey the implications that his teacher is difficult to convince, that he has fixed ideas, etc. The problem here is what I referred to as the 'emergence problem' in the previous chapter. Our knowledge of bulldozers does not include the assumption that they are disrespectful and obstinate, and our knowledge of iron bars does not include assumptions about their lack of a flexible mind. Being disrespectful, stubborn or hard to persuade are psychological traits, which cannot be stored as encyclopaedic assumptions about inanimate objects. But if they are not encyclopaedic assumptions associated with the encoded concepts, how does the hearer access this information and construct the concepts BULLDOZER* and IRON BAR* as intended by the speaker of (16)-(17)?

In discussing this example ('Robert is a bulldozer'), Robyn Carston makes just this observation (2002a: 350, 2003: 86). She suggests that examples in which none of the properties of the metaphor vehicle can apply literally to the metaphor topic (generally falling into the category illustrated in (11C)) are highly problematic for virtually every approach to metaphor, including the relevance-theoretic approach just outlined. She points out that, although the encoded concept BULLDOZER gives access in memory to a set of assumptions about the physical properties of an inanimate object (e.g. the assumption that a bulldozer is a tractor-like machine, that it is used for moving earth, rocks, etc.), it is not clear how adding these assumptions to the context can result in the derivation of the set of intended implications (e.g. the implications that Robert behaves ruthlessly towards others, is insensitive and obstinate, etc.) and hence in the formation of the ad hoc concept the speaker is taken to endorse, namely BULLDOZER*, which denotes the set of obstinate and disrespectful people.

If Relevance Theory claims that metaphor interpretation involves constructing an ad hoc concept just by selecting a subset of encyclopaedic assumptions already stored in the encyclopaedic entry of the encoded concept, then it will certainly suffer from the same shortcomings I highlighted for attribution theories of metaphor (e.g. the Class Inclusion Theory) in the previous chapter and, like those approaches, it would not be

able to deal with the 'emergence problem'. This would mean it will not be able to account for the examples in (16) and (17), or for any metaphorical use involving a certain degree of emergence, such as the 'butcher' example discussed in the previous chapter and repeated here in (18):

(18) *Doctor:* I am afraid the surgeon who performed a caesarean on your wife perforated both ovaries. I had no choice but to remove them.
Husband: I want that surgeon out of the hospital. That surgeon is a butcher!

The problem with this example is as the one just described for 'bulldozer': how can a hearer construct a concept BUTCHER* that denotes people who are negligent and botch their jobs if 'being negligent' and 'botching their jobs' are not encyclopaedic assumptions we have stored in our representation of butchers?

In my view, the problem here is not just whether theories of metaphor, including the one proposed in Relevance Theory, can deal with these particular cases where the encoded concept and the communicated concept do not overlap (or overlap very minimally, as with (18) where the denotation of the new concept would include those butchers who are negligent and careless). I want to argue that an inability to deal with the emergence of new features (assumptions) in interpretation makes it difficult (if not impossible) to deal with even relatively simple cases, such as those in (12)-(15). Even in (11)-(14), the speaker predicates of the subject a number of properties which are unlikely to have been stored ready-made in the encyclopaedic entry of the encoded concepts. In interpreting (12), for instance, the hearer may take the speaker to convey that her boyfriend does not behave in a way suitable for someone of his age. Since babies do indeed behave in a way suitable for people of their age, this is not an assumption the speaker could have retrieved ready-made from his knowledge of babies. In fact, it is an assumption which is stored as part of our representation neither of the speaker's boyfriend nor of babies, but which seems to emerge during the comprehension process. The hearer of (a) may indeed take the speaker to be conveying a number of implications which are emergent in just this way (e.g. the implication that the speaker's boyfriend is spoilt, incapable of having a grown-up romantic relationship, etc.).

Similarly, in interpreting (15), the hearer may derive the implication that the speaker's husband is not a good husband, yet the property of 'being a good husband' could not have been retrieved ready-made from the encyclopaedic entry of the concept BACHELOR, as bachelors are not married. Other properties may be taken to be predicated of the metaphor topic (e.g. 'neglecting his family', 'not behaving as it would be expected', 'upsetting his wife', 'risking his marriage', etc.) even if they are not properties stored as part of our representation of the encoded concepts. My point here is thus that virtually every metaphor enjoys some degree of emergence which needs to be accounted for. Virtually every ad hoc concept constructed in interpreting a metaphor may end up giving raise to a range of implications not based on assumptions which were directly stored as part of our representation of the encoded concept. These 'emergent properties' may be just one or two, or they may be so predominant that the denotation of the encoded concept and that of the ad hoc concept do not overlap at all, as in (11C).

An adequate account of metaphor should aim to provide an approach to metaphor comprehension capable of accounting for all these degrees of emergence, and for all the cases of broadening falling somewhere in the continuum represented by the figures in (11). In this chapter, I will try to show how a relevance-theoretic approach to metaphor can adequately account for the construction of different metaphorical interpretations, involving different degrees of pragmatic adjustment and for the derivation of emergent features during that process. An important reason why psycholinguistic approaches cannot deal with emergent features is indeed because they lack pragmatic inferential mechanisms and a distinction between contextual assumptions and contextual implications. Unlike most experimental approaches to metaphor, Relevance Theory needs not take metaphor interpretation, or ad hoc concept construction, to reduce simply to the selection of a set of assumptions associated with the metaphor vehicle in long term memory (and the attribution of these assumptions to the metaphor topic). Instead, the relevance-theoretic approach to metaphor is grounded on the assumption that metaphor interpretation is essentially and above all an inferential process (guided by more or less precise expectations of optimal relevance). Emergent properties can be seen simply as contextual implications and so may rise from processing the utterance as a whole (e.g. the implications 'not been mature enough for his age' in (12) or 'not being a good

husband' in (15)). In other words, there is no reason in Relevance Theory why the hearer of an utterance (e.g. a metaphorical utterance) should not be entitled to take the concept encoded by the speaker's words, and the set of assumptions this concept gives access to in memory, as a mere starting point to infer the speaker's meaning. The gap between the set of existing assumptions associated with the encoded concepts and the full set of contextual assumptions and implications which end up being used during the comprehension process would thus be bridged by pragmatic inference, and in this way we may find a solution to the emergence problem.

3. Relevance Theory and Emergence

The emergence of features does not seem to be exclusive to the comprehension of metaphorical examples but has been also found in the comprehension of literal language, more particularly in the comprehension of conceptual combinations (Estes and Glucksberg, 1999, 2000a, 2000b; Franks, 1995; Hampton, 1997; Rips, 1995). Hampton (1997), for instance, observes that in interpreting the combination 'Oxford graduate factory worker' or 'rugby player who knits', people typically produce properties such as 'failure' and 'confused', respectively, which are not typically associated with any of the terms in the compound. In the conceptual combination literature, scholars generally distinguish between 'emergent attributes' such as those above and 'extension based emergent attributes' (Hampton, 1997; Rips, 1995). Extension based emergent attributes are those attributes (properties or features) which arise via the identification of an already existing category. Take for instance the property 'talk' as emerging in interpreting the compound 'pet bird'. Although the property is emergent in that 'talking' is not a feature typical of birds or of pets, this property is recoverable by identifying a well-known category constructing so a subset of birds (i.e. parrots).

I want to claim that many of what the literature has referred to as 'emergent properties' are just sets of assumptions which are inferentially derived as implications in interpreting an utterance. It is because they are implications and because implications are not only derived in the comprehension of metaphorical speech, that emergent properties can be found not only in the interpretation of figurative language but of literally intended

utterances (e.g. as in 'oxford graduate factory worker' and 'rugby player who knits'). I also want to argue that since emergent properties are implications derived by processing an utterance as a whole, they do not need to be associated with any individual term in particular. They simply need to be derived inferentially at some point in interpretation from the combination of at least two premises.

Relevance Theory, I will argue, can account for the derivation of emergent properties by showing how, guided by more or less precise expectations of optimal relevance, the hearer uses the set of assumptions made accessible by the encoded concepts as mere input to pragmatic inference. The number of inferential steps and the kind of inferential routes followed during the comprehension process may result in the derivation of a range of implications based on premises which are not stored ready-made in the encyclopaedic entries of the encoded concepts. These implications may depart only slightly or quite considerably from assumptions stored in the encyclopaedic entry of the encoded concept, the result being one of the situations shown in (11) (see Vega Moreno, 2004). The following table illustrates how this inferential process may go for the comprehension of the 'butcher' example in (19):[2]

(19)

(a) S has said to H 'that surgeon is a butcher'	*Decoding of S's utterance.*
(b) S's utterance is optimally relevant to H	*Expectation raised by the recognition of S's utterance as a communicative act, and acceptance of the presumption of relevance it automatically conveys.*
(c) S's utterance will achieve relevance by justifying on his immediately preceding comment that he wants the surgeon dismissed	*Expectation raised by (b), together with the fact that such a justification would be most relevant to H at this point.*
(d) The fact that a surgeon has operated in	*First assumption to occur to H which,*

[2] I am not claiming here that this is necessarily the sequence in which comprehension occurs. According to Relevance Theory, mutual adjustment takes place in parallel, rather than in sequence. This table is modelled on those in Wilson and Sperber (2000, 2002).

a grossly incompetent way is a good reason for wanting him dismissed	*together with other appropriate premises such as those below, might satisfy expectation (c).*
(e) A competent surgeon makes incisions in order to preserve life, using high levels of precision, delicacy, foresight and planning to avoid risks	*First accessible assumptions from the encoded concept SURGEON which might combine with (d) and other assumptions to satisfy expectation (c).*
(f) A butcher cuts dead meat in a way that falls far short of the high levels of precision, delicacy, foresight and planning to avoid risk required in a competent surgeon	*First accessible assumptions from the encoded concept BUTCHER which might combine with (d), (e) and a suitably enriched interpretation of (a) to satisfy the expectation in (c).*
(g) The surgeon is a BUTCHER* (where BUTCHER* denotes people who make incisions in a way that falls far short of the levels of precision, delicacy, foresight and planning to avoid risk required in a competent surgeon)	*First enriched interpretation of (a) which might combine with (d), (e) and (f) to satisfy the expectation in (c). Created by pragmatic adjustment of encoded concept by backward inference.*
(h) The surgeon operated in a way that falls far short of the high levels of precision, delicacy, foresight and planning to avoid risk required by his job	*Conclusion derived by H from combining (f) and (g). Accepted as an implicature.*
(i) Surgeons who make incisions in a way that falls short of the levels of precision, delicacy, foresight and planning required may cause serious damage to someone in their care	*Next most accessible assumption from encoded concept SURGEON which might combine with (h) to help satisfy the expectation in (c).*
(j) The surgeon is a BUTCHER* (where BUTCHER* denotes people who make incisions of a certain type which falls far	*Further enrichment of (a) which might combine with available assumptions to satisfy the expectation in (c). Pragmatic*

short of the levels of precision, delicacy, foresight and planning to avoid risk required in a competent surgeon doing his job, and cause damage to someone in their care)[3]	*adjustment of the encoded concept* BUTCHER.
(k) The surgeon who operated on the speaker's wife caused serious damage to someone in his care through his lack of precision, delicacy, foresight and planning	*Conclusion derived by H from combining (i) and (h). Accepted as an implicature.*
(l) A surgeon who falls far short of required standards and causes damage to his patient as a result is grossly incompetent	*Next most accessible contextual assumption from encoded concept* SURGEON *which might combine with (k) to help satisfy the expectation in (c).*
(m) The surgeon is a BUTCHER* (where BUTCHER* denotes people who make incisions in a way that falls far short of the levels of precision, delicacy, foresight and planning to avoid risk required in a competent surgeon doing his job, cause damage to someone in their care, and are grossly incompetent)	*Further enrichment of (a) which might combine with available assumptions to satisfy the expectation in (c). Pragmatic adjustment of the encoded concept* BUTCHER.
(n) The surgeon who operated on the S's wife was grossly incompetent	*Conclusion derived by H from combining (l) and (k). Accepted as an implicature*
(o) Grossly incompetent surgeons deserve to be dismissed	*Contextual assumption treated as implicit premise*
(p) The surgeon who operated on the S's wife deserves to be dismissed	*Conclusion inferred from (o) and (n). One of several possible weak implicatures of S's utterance*

[3] The concept BUTCHER* as presented here would denote anyone (not necessary surgeons) who make cuts of this type. In any case, it is important to bear in mind, that the hearer of the utterance does not find out what the actual denotation of the concept BUTCHER* constructed during the interpretation process would be until he arrives at an interpretation of which satisfies his expectations of relevance.

The husband's utterance, in the circumstances described, raises in the doctor (the hearer) certain expectations of relevance which he expects to satisfy in processing that utterance. At the moment of the utterance, the hearer has certain highly accessible assumptions, such as the assumption that the speaker's wife has suffered as a result of her operation, that the speaker must be terribly upset about this, that he must be extremely angry with the surgeon and with the hospital, that he and his wife would probably like to make some kind of formal complaint, and so on. In processing the utterance in this context, some of the assumptions made accessible via the hearer's concept of a butcher, such as (f) for instance, become more accessible than others. Following a path of least effort, he starts considering these assumptions in their order of accessibility and adding them to the context in the hope of deriving a set of implications that will satisfy the expectations of relevance raised by the utterance. Because of the presence in memory of the concepts SURGEON and BUTCHER as well as the set of assumptions above, the assumptions that being a surgeon requires high levels of precision, delicacy, foresight, etc., and that butchers do not have these qualities, may be highly accessible to the hearer at the time. He adds these assumptions to the context and derives the implication that the surgeon fell far short of the required standards in performing his job. This piece of information may trigger further inferences. For instance, combining the information that the surgeon fell far short of the standards required by his job with the assumption that someone in his care was damaged may lead to the conclusion that he was careless, negligent and liable for sanction, e.g. dismissal or prosecution. The "emergent properties" 'being careless', 'negligent', 'liable to sanction', etc., are thus no more than implications derived inferentially, which would be potentially treated by the hearer as implicatures of the utterance.

It is worth noticing that since utterance interpretation involves a process of mutual adjustment of explicit content, context and implicatures, it follows that as the hearer derives the implications above, the concept conveyed by the word 'butcher' is continuously adjusted in order to warrant the derivation of these implicatures. Processing continues, with the hearer accessing and combining assumptions, deriving implications and adjusting explicit content until he arrives at a combination of explicit

context, context and implicatures that satisfies his expectation of relevance, at which point he stops. As a result of arriving at this combination, the hearer would have constructed a new ad hoc concept BUTCHER* which denotes the set of people who fall short of the standards of precision, delicacy and foresight required in making an incision, causing damage to humans beings in their care, and being liable for sanction as a result. It is this concept (or one roughly similar in import) that is taken to be a constituent of the explicature of the utterance and that warrants the derivation of the implicatures above. Because the encoded concept is merely a starting point for inference, there is no reason why it should not be adjusted to a point where the entities it is normally used to denote fall outside the denotation of the new ad hoc concept that results.

My suggestion is then that the reason why modern theories of metaphor cannot provide an explanation for the emergence problem, and so cannot provide a successful account of metaphor interpretation, is partly that they lack an inferential comprehension procedure. Any account of metaphor which sees interpretation as involving simply the selection and attribution of properties without an inferential process operating in between cannot possibly account for metaphor interpretation successfully.

4. Relevance Theory and the Transformation Problem

It might be argued that, even if the relevance-theoretic approach to metaphor interpretation can account for a subset of emergent properties, it cannot account for how all emergent properties are derived. It might be claimed, for instance, that it cannot account for the transformation problem presented in the previous chapter, as in (20)-(21):

(20) *Julie:* I am afraid about the divorce. My husband's lawyer is a shark.

(21) *Mary:* Are you sure your husband does not mind looking after the children the whole weekend?

Jane: Don't worry about it. He is a teddy bear!

The degree and type of 'aggressiveness' predicated of the speaker's husband's lawyer should be applicable to people in general and typical of lawyers in particular, and so different from the degree and type of aggressiveness that characterises real sharks. Similarly, the property 'being soft' or 'cuddly' in the encyclopaedic entry of the concept TEDDY BEAR can only be taken to apply loosely to the speaker's husband. These examples show how interpretation cannot reduce to the selection and attribution of encyclopaedic properties of the metaphor vehicle to the metaphor topic: some transformation often needs to be involved. The question is: how can an inferential approach to metaphor interpretation account for these examples? Although I agree that some inferential approaches to communication (e.g. standard Gricean ones) have problems in dealing with them, I want to show how Relevance Theory is not one of them (see also Vega Moreno, 2004).

I have described how Relevance Theory is built within a picture of cognition which assumes that people are capable of forming and communicating not only an indefinite range of thoughts but also an indefinite number of concepts which are constituents of those thoughts. These unlexicalised (and often one-off) concepts are formed by selecting different bits of information in memory and adjusting existing stable concepts in order to yield appropriate implications. What I want to point out here is that there is no apparent reason why, in arriving at a hypothesis about the combination of explicit content, context and implicatures that the speaker might have intended to convey, pragmatic adjustment should operate only on the set of encoded concepts. Instead, a considerable amount of pragmatic adjustment may involve the narrowing and broadening of concepts which are not encoded by the utterance but are rather constituents of contextual assumptions and implications being considered during the interpretation process.

The idea that concepts which are not linguistically encoded but are considered during inferential comprehension may be adjusted during this process sheds interesting light on the transformation problem. Consider Black's example 'man is a wolf'. In interpreting this utterance, the hearer, following a path of least effort, may start considering assumptions associated with the encoded concept WOLF in the order in which they occur to him. He takes each of these assumptions as a premise and adds it to

the context hoping to derive a set of implications that may help to satisfy his expectations of relevance. The expectation that the implications the speaker intended to convey will be consistent with his assumptions about men is likely to guide the interpretation and motivate the adjustment of concepts which figure in the encyclopaedic assumptions made accessible by the encoded concept WOLF. A highly accessible assumption that the hearer may consider from his knowledge of wolves is that 'wolves are aggressive'. The concept AGGRESSIVE as it applies to wolves may need to be adjusted on line to warrant the derivation of a range of implications that apply appropriately to men. For instance, wolves exhibit a rather physical type of aggression which may involve attacking and killing their prey. Presumably this is not the type of aggression the speaker wants to attribute to man: there is an element of hyperbole involved.

Let's suppose that the metaphor above is uttered in a situation where people have been discussing the difficulty of keeping up with a competitive lifestyle. In this situation, the hearer has access to a range of assumptions which he can take as potential implications or implicatures of the utterance (e.g. the implication that men are competitive creatures, that they may undermine others to gain success, etc.). These implications may be used via backward inference to adjust the concept WOLF. In this way, the concept AGGRESSIVE as it applies to wolves would be adjusted on-line to a point where it warrants the derivation of the expected type of implicatures which apply to competitive men. The same fine-tuning process may operate in exploiting some other assumptions about wolves that the hearer may consider during the comprehension process. He may, for instance, access the assumption that wolves are predators, which may enable him to derive a range of implications (e.g. they attack other creatures; they only consider their own survival, etc.). He takes these implications, together with assumptions about men and business life, as input in order to construct, by mutual adjustment, a hypothesis about the speaker's meaning. In the process, the concept PREDATOR as it applies to wolves would be adjusted to yield a new concept PREDATOR*, which applies to men with competitive, aggressive, selfish behaviours, men who are quite willing to ruthlessly exploit and manipulate people in a weaker position in order to serve their own ends, thus warranting the derivation of a set of implicatures

which help to satisfy his expectation of relevance. (e.g. men are competitive, undermine others to achieve their own success, look for vulnerable people in the system and exploit them mercilessly, etc.).

Different expectations of relevance generated by different utterances lead to a concept being adjusted in different ways. Consider the same word *aggressive*, but now applied to sharks. In processing the metaphor 'my lawyer is a shark', and on the assumption that the speaker is happy with his lawyer and confident he is good at his job, the concept AGGRESSIVE as applied to sharks may be adjusted to denote a kind of (positive) aggressiveness that involves energy and courage. However, processing the metaphor in a different situation, say, where the speaker is afraid of his lawyer's tactics, the concept AGGRESSIVE would be adjusted to denote a kind and level of (negative) aggressiveness which involves intentional emotional damage to others. These concepts which figure as constituents of the hearer's thoughts about these lawyers differ so from each other, and from the concepts which figure in his thoughts about wolves and men above.

According to Relevance Theory, the ad hoc concepts that result from adjusting the encoded concepts during the interpretation process are taken to be constituents of the explicature of the speaker's utterance. In this way, the explicatures of the above metaphorical utterances would include the concepts WOLF* and SHARK* as constituents. These may be one-off concepts which warrant the derivation of the particular implicatures required to satisfy the hearer's expectations of relevance. What I have tried to show here is that deriving these implicatures involves a certain amount of pragmatic fine-tuning of other concepts. That is, in constructing the concept intended as a constituent of the explicature, other concepts intended as constituents of the implicatures of the utterance are also adjusted during comprehension. As a result, the encyclopaedic entry of the concept SHARK*, created on line, would include the assumptions that sharks* are AGGRESSIVE* and PERSISTENT*. The encyclopaedic entry of the ad hoc concept WOLF* would include the assumptions that wolves* are AGGRESSIVE** and PREDATORS*. Since the concepts SHARK* and WOLF* are constituents of the explicature of the utterance, adding these assumptions to the context inferentially warrants

implicatures that help to satisfy the hearer's expectation of relevance and so to achieve relevance in the expected way.

It is important to notice that the adjustment of concepts which are not linguistically encoded is not unique to metaphor. Instead, it is a natural by-product of the mutual adjustment process that takes place in understanding virtually every utterance, whether literally, loosely or hyperbolically used. Consider, for instance, the examples in (22)-(24):

(22)	a. The sofa is soft	Explicature: THE SOFA IS SOFT*
	b. Baby skin is soft	Explicature: BABY SKIN IS SOFT**
	c. The cat is soft	Explicature: THE CAT IS SOFT***
(23)	a. I love the touch of this sofa	Implicature: THE SOFA IS SOFT*
	b. I love the touch of baby skin	Implicature: BABY SKIN IS SOFT**
	c. I love the touch of cat's fur	Implicature: CAT'S FUR IS SOFT***[4]
(24)	a. My hair is too long	Implicature: S WANTS TO CUT* HER HAIR
	b. The cake is ready!	Implicature: H CAN NOW CUT** THE CAKE
	c. The grass has grown fast	Implicature: S WANTS H TO CUT*** THE GRASS

In (22)-(24), the encoded concept SOFT is adjusted on line to denote a different type of softness on each occasion. In (22), the adjustment of the concept SOFT results in the construction of an ad hoc concept which is taken to be a constituent of the explicature. (23) shows how the ad hoc concept formed may be a constituent of one of the premises considered during interpretation and yielding potential implicatures. On some occasions, the (unlexicalised) concept which is a constituent of this premise may have been stored in the encyclopaedic entry of the encoded concept (e.g. the encyclopaedic assumption that baby's skin is soft refers to a particular type of softness, in (23b)). In other cases, it has to be constructed on-line (e.g. assumptions about the softness of the sofa in (23a)).

[4] The idea here is that the hearer of these utterances would probably infer that the reason why the speaker likes the touch of that particular item is because it is soft. The property 'being soft' would need to be adjusted on-line on each occasion.

(24) presents a similar case in which different ad hoc concepts CUT*, CUT**, CUT*** are constituents of assumptions which the speaker might have intended to convey as implicatures of her utterance. In order to arrive at the thought the speaker intended to convey, the hearer needs to adjust a concept which has not been linguistically encoded.

A possible objection to this account is that it may create problems for the stability of concepts. If new ad hoc concepts can be constructed at will, how is their denotation to be fixed and remain constant across times? Although this is a problem for all accounts that appeal to ad hoc concepts, I do not think it seriously undermines the account proposed here. The claim is not that the hearer is entitled to adjust any concept that is a constituent of any thought that occurs to him, or that there are no stable concepts at all. What is being claimed is that the formation of ad hoc concepts takes place naturally during comprehension because of the need to arrive at the right combination of explicit content, context and cognitive effects. The construction of an ad hoc concept therefore takes place under severe pragmatic constraints. The denotation of the resulting concept must be such that it warrants the derivation of the expected implications, and this should be enough to explain how its denotation differs from the denotation of the stable concept from which it is derived.

4.1 The Bulldozer Case

Having looked at the pervasiveness of pragmatic adjustment in utterance interpretation, and at the different inferential steps that may be involved in processing a metaphor, I will try to show how Relevance Theory might account for the problematic metaphorical example 'my boss is a bulldozer' in (16). In understanding this utterance, as in understanding any utterance, the hearer takes any assumptions made accessible by the encoded concept as potential inputs to an inferential process designed to make the utterance relevant in the expected way. Following a path of least effort, he considers potential contextual assumptions about bulldozers in their order of accessibility. These may include the assumption that bulldozers are machines and that they remove obstacles that stand in their way. Although these assumptions may not contribute to the derivation of contextual implications, they can nevertheless be used as a starting point to derive hypotheses about what the speaker might have intended to convey. That is, they can be

used as premises in an inferential process which may involve several inferential steps, and several instances of pragmatic fine-tuning, before the resulting implications may be plausibly taken to apply to the speaker's boss. There is no single processing route a hearer must take in interpreting this utterance. Different hearers, and the same hearer at different times, may derive different implications by following relatively different inferential paths. I consider below some ways in which the comprehension of (16) might proceed.

One inferential route the speaker may take in interpreting this utterance is the following. He may start by accessing from the encoded concept BULLDOZER the encyclopaedic assumption that machines are inanimate objects, from which it follows that they are insensitive to human feelings. He may also access some contextual assumptions from other sources which are likely to be highly accessible at the time. These may include the assumption that people, including bosses, are generally expected to have some sensitivity to the feelings of the people they work with, etc. These assumptions might combine with the utterance to yield the implication that his boss is inadequate and unpleasant to work with because he is insensitive to the feelings of others. In order to warrant the derivation of implicatures along these lines, the hearer must adjust the encoded concept BULLDOZER by backward inference into a new concept, BULLDOZER*, that denotes a set of people who are inadequate as bosses and unpleasant to work with because of an insensitivity to human feelings. If the resulting combination of explicit content, context and cognitive effects satisfies the hearer's expectations of relevance, comprehension would stop. If not, he would continue using highly accessible assumptions to derive further possible implications, and adjust the encoded concept accordingly. He may for instance take the assumption that bulldozers make it impossible to hear the voices of anyone standing near them, and present a danger to those who come too close, and combine them with some other encyclopaedic assumptions about the relationship between bosses and employees which he has stored in memory. He may derive from this combination a further array of implications the speaker might have intended to convey as weak implicatures of her utterance (e.g. that the boss does not listen or allow discussion, that he presents a threat, etc.).

Thus, although the property of being a machine cannot be directly attributed to the topic of the metaphor or be associated with the new ad hoc concept BULLDOZER*, it can be used as a starting point for an inferential process whose output is a set of assumptions which may indeed be accepted as part of the speaker's meaning, and which may end up being stored in the encyclopaedic entry of the new concept. Or, to put it differently, even though the properties 'being ruthless' or 'being insensitive' cannot be found in our encyclopaedic entry for the concept BULLDOZER, some encyclopaedic assumptions in that entry may be used, together with other background assumptions, as starting point for inferring the implications the speaker might have intended to convey on that occasion. The inferential process may involve several steps, which take the constructed ad hoc concept further and further away from the encoded concept, and allow a single inferential comprehension procedure to account for all the cases of metaphor interpretation falling somewhere along the continuum illustrated in (11).

Another inferential route a hearer might take in interpreting the utterance in (16) might be the following. The hearer may consider the highly accessible encyclopaedic assumptions that bulldozer remove obstacles that stand in their way, and that they are big pieces of machinery difficult to move and control. From the information that bulldozers are big, heavy and smash anything that stands in their way, the hearer may be able to derive a number of implications some of which may be taken to apply to the speaker's boss in some sense. For instance, the boss may not represent a physical danger to people or obstacles who stand in his way, but he may cause mental damage and ignore mental objections and arguments against his proposals. Arriving at these implications involves some adjustment to the concepts stored in the encyclopaedic entry for bulldozer. Thus by using a wider range of contextual assumptions about bosses, and by backward inference from the range of expected effects, the hearer may pragmatically reinterpret a range of assumptions applying to real bulldozers (e.g. that they destroy things around them, that smaller entities around them are vulnerable to their power, etc.) to the point where they warrant the expected implications. The output of this process would be a range of implications which may themselves combine with further background assumptions about bosses to yield some weaker implicatures and further implications (e.g. that the employees are afraid of the boss, of talking to him, of sharing

their own thoughts with him; that they feel oppressed and frightened of being reprimanded, humiliated, dismissed etc.).

What I am proposing in effect is that just that as the concept AGGRESSIVE, accessed from the encyclopaedic assumptions 'sharks are aggressive' and 'wolves are aggressive' had to be adjusted in interpreting metaphors such as 'my lawyer is a shark' or 'man is a wolf', so the assumption that 'bulldozers remove obstacles that stand in their way', accessed from the encyclopaedic entry for bulldozer may need to be broadened during the interpretation of (16). The resulting concept, say [TO REMOVE OBSTACLES THAT STAND IN ONE'S WAY]*, denotes a type of removal, a type of obstacles and a range of situations in which obstacles are removed which can warrant the expected set of implications. Thus, just as the concepts encoded by an utterance may need to be adjusted during the comprehension process, so may the assumptions made accessible by these encoded concepts. The more creative the metaphor, the more adjustment is likely to be required. As Sperber and Wilson (1986/1995: 236) put it:

> "In the richest and most successful cases, the hearer or reader can go beyond just exploring the immediate context and the entries for concepts involved in it, accessing a wide area of knowledge, adding metaphors of his own as interpretations and possible developments he is not ready to go into, and getting more and more very weak implicatures, with suggestions for still further processing"

The process of mutual adjustment of explicit content, context and cognitive effects, guided at every point by the hearer's more or less precise expectations of relevance, and often involving the pragmatic adjustment of concepts that figure as constituents of the assumptions in the encyclopaedic entries of the encoded concepts, continues until the hearer arrives at a combination that satisfies his expectations of relevance. Providing that the hearer's expectations of relevance are satisfied by the type of implications above, the hearer would have constructed an ad hoc concept BULLDOZER* which denotes a set of people who are insensitive to the feelings of others, ignore their suggestions and objections, are fixated on their own goals at the expense of others, are a danger to those who oppose them, etc. This concept, which figures as a constituent of the explicature of (16), may be required on later occasions, in interaction with different people, a point where it becomes a full-fledged concept stored in memory

by many individuals, and may even become lexicalised as an additional meaning of the word *bulldozer.* In this way, a one-off concept may become easier to access, cheaper to process and hence more likely to be reused (for discussion, see section 6 in this chapter, and particularly, chapters 6 and 7 and also Sperber and Wilson, 1998)).

5. Relevance Theory and Cognitive Approaches to Metaphor

Approaches to metaphor can be characterised by their commitment to or rejection of a number of claims, such as those listed below:

(25) i. *The Literal Priority Claim:* metaphorical meanings are derived after literal meanings.
ii. *The Comparison Claim*: metaphors convey (implicit) similes.
iii. *The Cognitive Dispensability Claim:* metaphors are ornamental and can be paraphrased without loss.
iv. *The Two-Domain Claim:* metaphor comprehension involves the alignment, comparison or interaction of two domains (represented by the metaphor topic and the metaphor vehicle).

Although deeply rooted in standard approaches, I have shown how the literal priority claim is currently rejected by virtually every recent cognitive approach to metaphor. These approaches have also moved away from the idea that metaphors are merely ornamental and cognitively dispensable, towards the new idea that they are cognitively significant devices whose meaning cannot be paraphrased without loss. However, the fact that metaphor is cognitively significant is often presented more as an intuition about metaphorical force than as an explicit account of metaphorical meaning analysable in terms of, say, propositional content and non-propositional effects. With regard to the comparison claim, although most recent accounts of metaphor reject the view that all there is to metaphor interpretation is the derivation of an implicit simile, many current approaches still claim that transforming the metaphor into a comparison is an important

step in comprehension (e.g. Gentner, 1983; Gentner and Wolff, 1997; Miller; 1993; Ortony , 1979).

Metaphor researchers since Aristotle have explicitly or implicitly committed to what I have referred to here as the 'two domain claim'. Early and contemporary approaches alike have assumed that metaphor interpretation is primarily a matter of establishing a relation between two distinct concepts or domains, typically represented by the (metaphor) topic and the (metaphor) vehicle. The relation taken to hold between them is often referred to as the (metaphor) ground (Richards, 1936). According to these approaches, a metaphorical interpretation often arises as a result of the tension caused by bringing together two distinct concepts or domains (e.g. the domains of people and pigs in interpreting 'my flatmate is a pig'). Understanding a metaphor involves alleviating this tension by identifying an appropriate relation between the metaphor topic and the metaphor vehicle (e.g. dirtiness), which would constitute the metaphor ground. Different scholars have proposed different ways to achieve this (e.g. by establishing a comparison, a mapping or an interaction between domains) with different views taking metaphor interpretation to involve different processes (e.g. the activation and suppression of features from topic and vehicle; the alignment of functional relations; the instantiation of a conceptual metaphor; the construction of a further conceptual structure, etc.) (see Gentner, 1983; Glucksberg and Keysar, 1990; Grady, Oakley and Coulson, 2000; Lakoff and Johnson, 1980; Tourangeau and Sternberg, 1981; Ortony, 1979).

The relevance-theoretic approach to metaphor defended here departs from the four standard claims in (25). In my view, these claims are implausible, and wrongly treat the phenomenon of metaphor comprehension as requiring special interpretation procedures not required for utterance interpretation more generally. If we do not derive a comparison in processing an utterance that uses a word hyperbolically (e.g. 'I am starving') and we do not establish an interplay between two contrasting domains in understanding a literal utterance (e.g. 'she works with computers), or an approximation (e.g. 'I earn 800 pounds a week') why should we do so in understanding a metaphorical utterance? If an alternative account can be developed using no special procedures, and if this account can deal adequately with the full range of cases of metaphor, it should be preferred on grounds of simplicity.

Relevance Theory, I have shown, challenges the Literal Priority claim by arguing that the hearer of a metaphorical utterance (like the hearer of any other utterance) aims to find an interpretation that is relevant in the expected way rather than one that is literally true. This is not because of any maxim or any norm that a speaker may follow or not: the pursuit of relevance is a result of the way our cognitive systems have evolved, shaped by millions of years of evolutionary pressure to become increasingly efficient at processing information. Arriving at an interpretation of metaphor which satisfies the hearer's expectations of relevance need not require a comparison, alignment or mapping between domains. I have tried to show that treating metaphor as involving a comparison or interaction between just two concepts or domains (represented by the metaphor topic and metaphor vehicle) is problematic for several reasons. First, it is difficult to see how information from other sources (e.g. other available contextual assumptions, expectations about the speaker's meaning) may be brought into the interpretation process. Second, if comprehension involves an interaction or mapping between two domains, there is a risk of circularity: the properties which the topic helps select in the vehicle are the properties attributed to the topic by the vehicle. Third, as noted above, looking at metaphor as involving interplay between two concepts isolates the interpretation of metaphors from the interpretation of other utterances, including those literally or loosely intended, which is undesirable unless it can be demonstrated that there is a sharp distinction to be made.

In fact, many philosophers and linguists have questioned the distinction between literal and non-literal language. According to Rumelhart, attempting to draw a clear-cut distinction between literal and non-literal language is like attempting to distinguish formal and informal language (Rumelhart, 1979/1993). Atlas (2005: 3-4) rejects the view that there is "some fundamental cleavage between language that is figurative or literary, containing terms used metaphorically, and language that is standard and ordinary, containing terms used literally". However, studies of metaphor interpretation have not really succeeded in showing how metaphor can be dealt with as part of a continuum between literal and figurative language. Even Glucksberg (2001) treats metaphor as something that has to be distinguished from other types of category extension.

The relevance-theoretic view that utterances are understood by considering possible hypotheses about explicit content, context and cognitive effects in their order of accessibility fits much better with the idea of a continuum. Every utterance, whether literally, loosely, hyperbolically or metaphorically intended, is processed using the same comprehension procedure, involving the same mutual adjustment process and guided by the more or less precise expectations of relevance. Whether the interpretation which satisfies those expectations would be classified in folk-linguistic terms as literal, approximate, hyperbolic, metaphoric, etc. does not affect the interpretation process. Every step in inferential comprehension involves adding highly accessible assumptions to the context and deriving contextual implications (or other cognitive effects). This process, unlike the selection and mapping of features from one domain to another, is not liable to the criticism of circularity. Finally, the relevance-theoretic account captures the role that context and pragmatic expectations play in the interpretation process. Thus, the information that affects the interpretation of a metaphor does not come only from the concepts encoded by the topic and the vehicle, but from many different sources. Any accessible information from any source may affect the interpretation of the utterance and fine-tuning of the encoded concepts.

5.1 Relevance Theory and the Class-Inclusion Theory

Much of the previous chapter focused on Glucksberg's Class-Inclusion Theory of metaphor. The reason for this choice was that, despite possible drawbacks I highlighted (in section 5.3), the Class-Inclusion theory marked a considerable step forward in the experimental literature, independently introducing and testing hypotheses that have much in common with the relevance-theoretic approach. In this section, I look at some of the similarities and differences between these two approaches.

Despite their important differences, there is a crucial common feature Relevance Theory and Class-Inclusion Theory share which sets them apart from much of other research on metaphor. This is the idea, explicit in later relevance-theoretic accounts, and inferable from the Class-Inclusion model, that the content of a metaphorical word contributes to the explicit content of the utterance (but see also Nogales, 1999 and

Recanati, 1995, 2003). [5] Recall that according to both approaches, what is explicitly asserted by the speaker in producing the nominal metaphor *'my lawyer is a shark'* is not that his lawyer is a literal shark, but that he is a SHARK*: that is, that he is a member of the ad hoc category of 'people and animals who are aggressive, to be feared, etc.' This view of metaphor as contributing to the explicit content and thus the truth-conditional content of the utterance is quite innovative and different from standard semantic and pragmatic models.

Although early and modern psycholinguistic approaches to metaphor, including the Class-Inclusion approach, have not been much concerned with the semantic-pragmatic distinction, this distinction has been essential to research on linguistics. The question of whether metaphor falls within the scope of semantics or pragmatics has given rise to considerable discussion and research (see Nogales, 1999; Kittay, 1987; Vicente, 1992, 1993, 1995, 1997). Most accounts of metaphor within the linguistic (and philosophy) tradition fall into two main groups: those which see metaphor as a matter of semantics and those which argue that it is a largely pragmatic phenomenon. On the one hand, a reason for viewing metaphor as a matter of semantics has been the intuition that metaphorical meaning contributes to the truth-conditional content of an utterance. So 'John is a lion' would be true if and only if John is a certain kind of person (brave, courageous, willing to fight, etc.). On the other hand, metaphor is typically context-dependent: so metaphor is not about what words mean (e.g. *lion*) but what speakers mean when using them in a particular context (Davidson, 1977). If metaphor is a matter of speaker meaning rather than sentence meaning, then it should belong to the domain of pragmatics (Grice, 1975/1989; Levinson, 1983). The problem is that metaphor seems to be both truth-conditional and context-dependent, and so seems to belong simultaneously to both semantics and pragmatics. Continuing with the traditional division of labour seeing semantics as the study of truth-conditional content and pragmatics as the study of context-dependent meanings, rules out any possibility of providing an adequate theory of metaphor.

[5] As pointed out in Chapter Two, Recanati (1995, 2003) has also defended a similar. He takes metaphor comprehension to involve the operation of pragmatic processes at local level modulating word senses the output of which modulation would contribute to the truth-conditional content of the utterance (e.g. the machine *swallowed* the credit card).

The problems that arise in these discussions are not very different from those that arise in trying to define literal meaning and distinguish literal from non-literal language (see Ariel, 2002).[6] Literal meaning has generally been seen as encoded, context-independent and truth-conditional. The problem is that no proposition conveyed by an utterance meets all these conditions. In this dissertation, I have shown how Relevance Theory has defended the need to maintain a distinction between encoded meaning and communicated meaning. What is encoded is not the same as what is communicated, but is only an instrument for communication, providing input or evidence to be used in inferring communicated content, both explicit and implicit. The pragmatic enrichment of a logical form into a fully propositional form is generally context-dependent and the output of the enrichment process is generally a proposition which is explicitly conveyed and whose truth conditional content determines whether the utterance is true or false. Deriving a literal or a metaphorical interpretation often requires a great deal of inferential elaboration, including the pragmatic fine-tuning of encoded concepts which contributes to the level of explicit content. This view, which is shared by Relevance Theory and the Class-Inclusion account (and Recanati, 1995, 2003), captures the intuition that metaphor is both truth-conditional and context-dependent.

Although both the Class Inclusion Theory and Relevance Theory claim that understanding a metaphor consists in constructing an unlexicalised ad hoc concept which denotes an ad hoc category, it is important to note the processes which these theories see as involved in interpreting a metaphorical utterance, and so in constructing the unlexicalised concept, are quite different. Recall that on the Class Inclusion view, ad hoc category formation results from the alignment of topic dimensions and vehicle properties (Glucksberg, 2001; Glucksberg, McGlone and Manfredi, 1997). This alignment results in a selection of dimensions and properties which allows the hearer to construct an ad hoc concept denoting the first superordinate category which can include both the topic and the vehicle as members and of which the vehicle is generally a (proto)typical member. So in understanding a metaphor such as 'my lawyer is a shark', for instance, the hearer would select some dimensions made accessible by the topic

[6] For discussions on literal meaning see also Bezuidenhout and Cutting, 2002b; Carston, 2002b; Dascal, 1987, 1989; Gibbs, 1989, 2002a, 2002b; Gibbs et al. 1993; Lakoff, 1986; Recanati, 1995, 2003; Rumelhart, 1979/1993; Searle, 1979a, 1979b, 1980.

'lawyer' (e.g. skill, character) and some properties made accessible by the vehicle 'shark' which can assign value to those dimensions (e.g. 'sharks are aggressive', 'sharks attack others') which are used to construct the first superordinate category of which both the literal referent of the vehicle and the topic are members (e.g. the category of 'entities which are aggressive and attack others').

The relevance-theoretic account does not share any of these claims. First, it does not claim the ad hoc category formed in interpreting a metaphorical utterance is constructed simply by accessing a subset of assumptions or properties associated with the vehicle (or the topic) taken literally (without using them as premises in an inferential process). Instead, it proposes that these assumptions are used as premises in deriving contextual implications. Second, it does not claim that the literal referent of the metaphor vehicle is always a member (or a prototypical member) of the resulting ad hoc category, or that this ad hoc category is always the first superordinate category to which both topic and vehicle could belong. In fact, I have shown that since different chains of inference might be involved in understanding a metaphor, the resulting ad hoc category may exclude certain members of the denotation of the encoded concept. In other cases, it may exclude all the members of the denotation of the encoded concept, so that the literal referent of the metaphor vehicle is not only not a prototypical member of the resulting ad hoc category, but not a member at all (e.g. 'that surgeon is a butcher', 'Mary is a block of ice'). So the relevance-theoretic approach, unlike the Class-Inclusion approach, accounts for the full range of metaphors on the continuum in (11).

Lexical pragmatic adjustment, and particularly lexical broadening as understood in Relevance Theory, is therefore rather different from the process of ad hoc category construction described in the Class-Inclusion model. In the last chapter, I argued that an important drawback of the Class-Inclusion Theory, and of modern approaches to metaphor more generally, is that it lacks adequate inferential machinery to account for how the appropriate ad hoc concept is constructed, and therefore how the metaphor interpretation process actually works. The processes envisaged by the Class-Inclusion model are not inferential, and mostly involve the activation in memory of a range of dimensions and properties associated with metaphor topic and metaphor vehicle. The lack of appropriate pragmatic constraints and the lack of a distinction between

contextual assumptions and contextual implications, I have shown has been very problematic for modern metaphor research. It has prevented the Class-Inclusion approach, for instance, from giving a full account of a good number of metaphors such as those presented in the previous chapter, some repeated here in (26) and (27):

(26) a. (Of a surgeon who has been negligent) That surgeon is a butcher.
b. (Of a pianist who has played terribly) The pianist butchered the sonatas.
c. (Of a teacher who fails most of the class) That teacher is a butcher.
d. (On a gruesome crime scene) This man is a butcher!

(27) a. John is an iron bar.
b. My lawyer is a shark.

The examples in (26), I have argued, are problematic for the Class Inclusion Theory because, even if we select the "right" dimensions from the topic and the "right" properties from the vehicle, the combination of these two does not guarantee a successful interpretation. The speaker of (26a) and (26b), for instance, must have certainly have intended the hearer to entertain a much more fine-grained interpretation than merely the assumption that the surgeon in (26a) and the pianist in (26b) are not skilful. Different metaphor topics, such as those in (26a)-(26c), may give access to a similar set of dimensions which are paired with a single vehicle (butcher), yet they allow for a wide array of different and fine-grained interpretations. The Class-Inclusion Theory, as it stands, cannot account for these different interpretations. The examples in (27), I argued, present even more serious cases. With these, a single sentence can be used to convey an indefinite range of metaphorical meanings, each one requiring the derivation of a certain type of implications and the construction of a different ad hoc concept on the fly. It is not clear from the Class-Inclusion theory, and interaction theories more generally, how a single combination of topic and vehicle (e.g. lawyer-shark), can allow for the different combinations of dimensions and properties which are necessary for the various possible interpretations of the sentences in (27).

Unlike the Class-Inclusion theory, the relevance-theoretic account of metaphor does not assume that arriving at an interpretation of a metaphorical utterance depends on the identification of a certain combination of dimensions and properties. Neither does Relevance Theory claim that the first superordinate category which includes both topic and vehicle need to be the one the speaker intended. According to Relevance Theory, the interpretation process depends on the expectations of relevance raised in the hearer by a certain utterance (with different utterances generating different expectations). Assumptions activated by the encoded concept are considered in their order of accessibility until those particular expectations are satisfied. A single word or phrase, say 'butcher', 'shark' or 'iron bar' may thus be used to convey a wide array of different unlexicalised concepts (e.g. BUTCHER*, BUTCHER**, SHARK*, SHARK**, etc.). The different fine-tunings of the concepts encoded by these words in memory are a function of different accessibility orderings and different expectations of relevance, both highly sensitive to contextual specifics. In the process of mutual parallel adjustment, the hearer's expectations of relevance add an extra degree of activation to some encyclopaedic assumptions, making certain hypotheses about implicatures highly accessible, which in turn leads, by backwards inference, to enrichments of the explicit content in a particular (optimally relevant) direction.

5.2 Experimental Evidence

An assumption common to both feature-matching models and attribution views is that presenting subjects with a metaphor topic or metaphor vehicle in advance should help with comprehension. Since these models treat interpretation as involving the accessing, and matching or attribution of some of the properties associated with these terms, allowing the hearer to think about and access those properties before interpreting the metaphor should speed the comprehension process. Although these predictions are confirmed by experimental research (see Wolff and Gentner, 2000), they do not offer a way of distinguishing feature-matching models from property-attribution models. To resolve this problem, McGlone and Manfredi (2001) (see also Glucksberg et al., 2001) presented subjects not with a topic and a vehicle alone, but with selected properties of each of these, some of these properties are relevant to the interpretation of the metaphor,

while others are irrelevant. The metaphor 'my lawyer is a shark', for instance, was preceded by one of the following:

(28) a. Properties of the topic which are relevant (e.g. 'lawyers can be ruthless').
b. Properties of the topic which are irrelevant (e.g. 'lawyers can be married').
c. Properties of the vehicle which are relevant (e.g. 'sharks can be ruthless').
d. Properties of the vehicle which are irrelevant (e.g. 'sharks can be blue')

The finding that people took relatively longer to interpret a metaphor when it was preceded by information irrelevant to the interpretation, as in (28b) and (28d), than when it was preceded by relevant information, as in (28a) and (28c), was only to be expected. A more interesting finding was that people took longer to process the target metaphor when it was preceded by an assertion about an irrelevant property of the literal referent of the metaphor vehicle, as in (28d), than when followed by an irrelevant property of the literal referent of the metaphor topic, as in (28b). Although irrelevant properties of the literal referent of the topic (e.g. 'lawyers can be married') did not facilitate comprehension, it did not interfere with it either. By contrast, irrelevant properties of the literal referent of the vehicle (e.g. 'sharks can be blue') not only did not facilitate comprehension but actually interfered with it.

McGlone and Manfredi argue that this asymmetry cannot be accounted for by matching models, but is to be predicted by models which take the concept pragmatically conveyed by the vehicle term to denote a different category from the one denoted by the literal use of the word. Although these results are seen as supporting the Class-Inclusion theory, I want to argue they are also consistent with the relevance-theoretic approach. Both the Class-Inclusion theory and Relevance Theory assume, in their own ways, that the concept expressed by the word 'shark' in the metaphor 'my lawyer is a shark' is not the concept encoded by the word (SHARK) but a much broader concept (SHARK*) which denotes the ad hoc category of 'people and animals which are aggressive, obstinate, destructive, etc'. Both theories predict that presenting subjects only with properties of sharks which are also properties of the members of this wider category, should facilitate comprehension. Presenting subjects with properties of real sharks (e.g. sharks can be

blue', 'sharks are good swimmers') which are not also properties of the members of the intended ad hoc category should not only not facilitate comprehension but interfere with it.

It is interesting that this interference effect was not found in metaphor topics. One reason for this may be that although in processing the metaphor only a subset of the encyclopaedic assumptions associated with the encoded concept LAWYER are actually considered in processing the metaphor, the resulting concept is still intended in its literal sense. Presenting subjects with other properties of lawyers (e.g. that they are married, that they use briefcases) which are not relevant to the interpretation of the metaphor may not facilitate its interpretation but need not necessarily cause problems with it either.

Similar results were found by Gernsbacher, Keysar and Robertson (2001), who studied the enhancement and suppression of information during metaphor processing. In one experiment, they asked subjects first to read a literal or a metaphorical sentence, e.g. 'that large hammerhead is a shark' or 'that defence lawyer is a shark', and then to judge whether a certain statement, e.g. 'sharks are tenacious' or 'sharks are good swimmers', was true. They found that people took considerably less time to verify the target statement 'sharks are tenacious' after reading the metaphorical sentence 'lawyers are sharks' than after reading the literal statement 'hammerheads are sharks'. They also found that subjects took longer to judge whether the statement 'sharks are good swimmers' was true after reading the metaphor than after reading the literal sentence.

The authors took these findings to suggest that interpreting a metaphor involves, on the one hand, enhancement (i.e. increased activation) of the set of properties which are relevant to the interpretation (e.g. tenacity) and, on the other hand, suppression of a set of attributes that are not relevant to that interpretation (e.g. being a good swimmer). More generally, these findings are consistent with the selective processing hypothesis underlying the relevance-driven account of comprehension and with the Encoding Specificity Hypothesis presented in Chapter One, as they show how only selected bits of information from the encoded concepts are considered in processing an utterance. Only those bits of information present at the moment of processing ('encoding', in one sense of that term) will be stored in working memory and are likely to have an effect on incoming information (e.g. by having a high degree of activation) and to act as good

retrieval cues for the item being processed (e.g. the utterance); information which did not play a role in deriving the intended interpretation should be an ineffective retrieval cue.

6. Creative and Standardised Loose uses

We may assume that loosely used expressions (e.g. metaphorical expressions) range from the very creative to the rather conventional or standardised. At one end of the spectrum of creativity we can place the poetic metaphors of great poets. At the other end, we may place a number of dead (or nearly dead) metaphors which have arguably developed an extra sense in the language ('the *leg* of the table', 'the *foot* of the mountain', 'a syntactic *tree*', etc.). Lying somewhere in-between, we may find a number of ordinary but novel metaphorical and hyperbolic uses, as well as an array of relatively familiar and widely used metaphorical and hyperbolic expressions. These include not only nominal metaphors (e.g. 'John is a *shark*') and verbal metaphors (e.g. 'I had to *swallow* her criticism') but also a range of phrasal metaphors which are more or less standardised in the language (e.g. *to swim in money, to swim with sharks, to hold all the aces, to break someone's heart, to break the ice*, etc.).

Relevance Theory has argued since its very beginnings that there is no theoretical reason to assume that different metaphorical uses, which vary in their degree of creativity, will be understood in different ways, following different comprehension procedures (Sperber and Wilson, 1986/1995, 1987, 1990, 1991). On this approach, the differences between relatively creative and relatively conventional metaphorical uses is captured in terms of the strength of the explicatures and implicatures which the hearer takes the speaker to have intended to convey and the amount of processing required to derive them. Generally speaking, relatively standardised metaphors are seen as giving access in memory to a few highly salient contextual assumptions, which lead to the derivation of a few strong implicatures. Understanding the metaphor 'John is a pig' in context of stereotypical assumptions may, for instance, involve the derivation of a few strongly communicated implicatures (e.g. John is dirty, John is disgusting). The more creative the metaphor, however, the less salient the set of contextual assumptions the

speaker may have intended to convey, and the weaker the implicatures derived from the combination of these assumptions and explicit content. Shakespeare's metaphor 'Juliet is the sun', for instance, may allow the hearer/reader to derive some of a wide array of weak implicatures with roughly similar import (e.g. implicatures about Juliet being a bringer of light, radiance, innocence, beauty, etc.):

> "The surprise or beauty of a successful creative metaphor lies in this condensation, in the fact that a single expression [...] will determine a very wide range of acceptable weak implicatures" (Sperber and Wilson 1986/1995: 237)

> "It is the range and indeterminacy of the implicatures which gives the metaphor its poetic force" (Pilkington, 2000: 102)

One important feature of poetic metaphors is so that their comprehension does not involve considering simply the immediate context: the hearer is expected to explore a complex network of contextual assumptions and chains of inference, yielding a wide array of weakly communicated assumptions which give the metaphor its poetic flavour. This is captured in the quote from Sperber and Wilson (1986/1995: 236) presented earlier (p. 133) and is also inherent in Pilkington's work:

> "[A] good poetic metaphor, by activating a wide network of contextual assumptions prior to the interpretation of the metaphorical utterance itself, may give greater direction to the interpretation of metaphors, enabling them to be read in a richer, more creative way than would otherwise be possible. The success of a poetic metaphor depends not only (if at all) on its originality, but on the creation of a context which encourages and guides the exploration of the encyclopaedic entries of the concepts involved." (Pilkington, 2000: 103).

In line with this assumption, Pilkington (2000) has argued that a metaphor presented in isolation may not be perceived as poetic yet it may be if embedded in a wider text. My interest in Pilkington's proposal is particularly in the role that the direction of the search for relevance seems to play in constructing a possible context and interpreting the metaphor. The reader of a poem, for instance, seems to travel around a network of contextual assumptions, picking a subset of accessible ones which

direct him into a certain inferential path and so into a certain interpretation of the poem and of the metaphors in the poem. The greatness of a good creative metaphor may lie in that different hearers, or the same hearer at different times, may pick different subset of (weakly manifest) contextual assumptions along the comprehension process directing attention to different inferential paths and so to the derivation of slightly or considerably different interpretations. Since each of these inferential routes explore a wide array of weakly communicated assumptions, the reader will often derive a wide array of weak implicatures and implications which compensate the processing effort invested.

In interpreting an utterance, a poem or any ostensive act, the hearer's expectations of relevance direct the hearer/reader towards a certain combination of assumptions and a certain inferential route yielding a range of implications which the speaker might have intended to convey. In the case of a very creative metaphor, the responsibility for considering those assumptions, following that inferential route and deriving that range of implications is largely left to the hearer/reader. Different readers, or the same reader at different times, may therefore consider different sets of weakly manifest assumptions and derive different interpretations. As Pilkington points out, preceding context may help guide the hearer into a particular inferential route. Another way of guiding the hearer is to set up strong expectations of relevance by using relatively standardised forms. The more familiar a hearer is with a certain metaphorical expression, the more precise are his expectations of relevance and the more salient are the set of intended assumptions and inferential route to follow in processing it.

The responsibility for constructing the intended interpretation of relatively familiar metaphors in everyday conversation is thus not just down to the hearer; the speaker also shares a lot of the responsibility. A speaker aiming at optimal relevance should have some idea about which sort of assumptions will be most accessible to the hearer at the time, and which sort of implications he will be likely to derive, and should formulate the utterance accordingly. In producing the utterance 'my boyfriend is a baby', for instance, the speaker should predict what sort of assumptions and computations the hearer is likely to consider in processing that utterance. Had she intended him to consider a different range of assumptions, computations and derive a different range of

implications than those normally derived in processing the familiar metaphor in stereotypical circumstances, she should have produced another linguistic stimulus.

We may assume that the less creative and more standardised the metaphor, the more precise the hearer's expectations of relevance and the narrower the search space for the construction of an optimally relevant interpretation will be. Rather than needing to explore a wide range number of encyclopaedic entries in considerable depth, the hearer of a standardised metaphor may simply consider one or two strongly manifest assumptions and derive one or two strongly communicated implicatures, as in the 'John is a pig' example above. The speaker's use of a relatively standardised form is thus likely to put the hearer on the right track, allowing different hearers to consider roughly the same set of assumptions as premises and derive roughly the same implications.[7]

6.1 Inferential Routes and Pragmatic Routines

According to Relevance Theory, our cognitive systems have evolved towards greater efficiency, developing heuristics for directing the individual's attention, effort and cognitive resources to the optimal processing of those bits of information which are likely to yield the greatest cognitive effects for the investment of the least processing effort. The formulation of an utterance or text is expected to guide the hearer towards the intended interpretation with hypotheses about the intended meaning considered in their order of accessibility. As we have seen, the selective processing of available information often results in the construction of new mental representations (e.g. one-off concepts and one-off thoughts). I want to argue that recurrent selective processing of a familiar

[7] Pilkington (2000) suggests that standardised metaphors may be understood by accessing a set of metarepresented assumptions *en bloc* from the encyclopaedic entry of the encoded concept (e.g. the concept PIG). Although I share the view that some set of assumptions may be immediately activated in the encyclopaedic entry of the encoded concept in processing a familiar metaphor, I don't see a particular reason why they must be metarepresented. Familiar metaphors, and even familiar idiomatic expressions (to be discussed in the next chapter) may be standardised in a language and give access repeatedly to the same set of assumptions without these needing to be metarepresentationally used. These figurative uses contrast in this way with other set of standardised figurative and formulaic uses such as proverbs or sayings (e.g. *a bird in the hand is worth two in the bush*) which are generally used to communicate metarepresented thoughts (i.e. thoughts which are attributed to a cultural or linguistic community). The study of how these metarepresented thoughts are spread in a culture is certainly an interesting line of research (see Sperber, 1990, 1992, 1994a, 1996, 1997 on reflective beliefs and the epidemiology of representations).

stimulus may lead to the development of a more or less automatic cognitive procedure or inferential route to process this stimulus.

In interpreting an utterance, the hearer is seen as following a path of least effort in order to arrive at the right combination of explicit content, context and implicatures. The most accessible assumptions are considered first, and the least costly inferential routines are used. Each act of utterance interpretation is creative, to some extent, in that each time the hearer constructs a different combination of explicit content, context and implicatures. Let's assume, nevertheless, that in processing a certain word or phrase, the hearer may repeatedly, consider roughly the same contextual assumptions, derive roughly the same implications and enrich the explicit content in roughly the same ways. Frequency of use reduces processing effort. We may, therefore, assume according to the view of cognition and communication defended in Relevance Theory, that after enough exposure to this word or phrase, the hearer's effort, attention and cognitive resources will be automatically allocated to these assumptions and implications narrowing the search space considerably. I will illustrate this point with examples (29)-(31):

(29) I cannot go to the cinema with you. I am seeing my doctor.
(30) I am going to church.
(31) Peter and Mary slept together.

In the absence of information that would indicate otherwise, the hearer of these utterances would generally assume that the speaker is rather strongly communicating that she is having a medical appointment in (29), that she is attending mass in (30), and that Peter and Mary had sex in (31). Notice that these are not the only possible interpretations. For example, the speaker may be seeing her doctor to play tennis, going to church to pick up some leaflets and intended to convey that Peter and Mary shared a bed. A speaker aiming at optimal relevance should choose a linguistic stimulus which will put the hearer on the right inferential track, allowing him to derive the intended cognitive effects for no gratuitous processing effort. In the case of (29)-(31), a speaker aiming at optimal relevance should have foreseen that the hearer would be likely to

arrive at the interpretations described above. Had she intended the hearer to take a different inferential route, she should have chosen a different linguistic stimulus.

It is not necessary to say that the hearer of (29) has stored a separate meaning in memory for the expression 'to see one's doctor', but we may want to claim that his familiarity with this expression might direct him to processing it in a particular way. This might be done along several different lines. We might assume that frequent processing of this combination of words in stereotypical contexts would result in certain encyclopaedic assumptions from the encoded concepts getting a higher degree of activation. Another possibility is that familiarity with the combination may direct the hearer to fine-tune the encoded concept SEE in a particular way (to denote a particular type of seeing - e.g. visiting through appointment). Finally, since the expression is often used to convey roughly the same sort of implications (e.g. about medical appointments), the hearer may have quick access to these implications and take them as part of the speaker's meaning. Since the comprehension process is one of mutual parallel adjustment, we may not need to choose between these possibilities. Instead, the hearer's attention may automatically consider a particular combination of highly accessible hypotheses about explicit content, context and implicatures, which are further adjusted with available assumptions until the hearer arrives at a combination which satisfies his expectations of relevance.

The more familiar the hearer is with a particular combination of words (e.g. to sleep together) the more likely is that the most highly activated assumptions will be those which had been considered or derived in processing the string on previous occasions. If the expectations of relevance raised by the speaker's utterance are satisfied after processing these highly accessible assumptions, he will stop. If not, he will continue and consider the next most accessible hypothesis, processing the concepts in the utterance at a deeper level until he arrives at a satisfactory interpretation. It is this tendency of the mind, in its search for relevance, to use assumptions and inferential processes which have been used in processing the stimuli on previous occasions that I will refer to as a pragmatic routine.[8]

[8] The literature on pragmatic routines is very limited. The closest proposal to that defended here and developed in the following chapter are that of those who argue that some indirect requests (e.g. 'can you pass me the salt?') or familiar proverbs (e.g. *too many cooks spoil the broth*) may be understood by

Pragmatic routines are a kind of cognitive procedure that might be expected to develop given the Cognitive Principle of Relevance.[9] Use of available pragmatic routines is encouraged by the Communicative Principle of Relevance since they make an utterance particularly easy to process in a way that is likely to satisfy expectations of relevance. A speaker aiming at optimal relevance should expect a hearer familiar with a particular use or particular expression to follow the path of least effort it makes accessible. If she intends him to consider some other range of assumptions and/or derive some other range of implications, she should choose a different linguistic stimulus which would direct the hearer's inference in a different direction.

One important point to consider in the processing of novel metaphorical uses is that once an inferential route is taken, the hearer/reader can establish connections between concepts that may have not been established before, and construct an ad hoc concept which might have never constructed before. In talking of poetic metaphors, Pilkington comments:

> "A creative metaphor combines insight with depth. The reader must be able to construct the links between concepts where links are not well-established. If the context is too readily accessible, on the one hand, or if the search through context leads nowhere, on the other, then the metaphor fails as a creative metaphor" (Pilkington, 2000: 104).

In understanding a novel metaphor (e.g. 'Juliet is the sun'), heard for the first time, the hearer/reader needs to find his way to a plausible interpretation. This may involve the construction of a particular type of context, the enrichment of the explicit content in a particular way and the derivation of a particular range of implications. Repeated processing of the same metaphor may result in the hearer considering roughly the same contextual assumptions (e.g. the assumption that the sun is source of life) and following roughly the same inferential steps to enrich the explicit content (e.g. by constructing the ad hoc concept SUN*) and derive roughly the same sort of implications. This may lead to development of a pragmatic routine for the processing of this expression, which would

---

compiling or short-circuiting the inferential steps involved in comprehension (for research on this see Bach, 1995a, 1995b/1998; Bach and Harnish, 1979; Groefsema, 1992; Harnish, 1993, 1995; Morgan, 1978; Sadock, 1972, 1993; Unger, 1996).

[9] See Sperber (forthcoming) for interesting ideas on the relation between the development of cognitive procedures and modularity.

allocate his attention and resources in a similar way to the ways in which it has been successfully allocated on in previous occasions.

Sometimes, in processing familiar metaphorical uses such as 'John is a pig' or 'you are a baby', the hearer may consider highly accessible hypotheses about the sort of contextual assumptions (e.g. assumptions about dirt or need of care), the sort of enrichment at explicit level (e.g. the construction of the concepts PIG* or BABY*) and the sort of implicatures which the loose use of these words is generally used to convey in similar situations (e.g. that John is dirty, that the hearer is immature, etc.). It is this highly accessible inferential route towards a particular combination of explicit content, context and cognitive effects which is most accessible by the hearer that he will consider first in processing the utterance. This combination is, however, merely a template for comprehension and will need to be further fine-tuned in context in every occasion so to satisfy the more or less precise expectations of relevance generated by the particular utterance. A result of this fine-tuning, the hearer may construct slightly different interpretations and construct slightly different concepts at different times (e.g. BABY*, BABY**, BABY***). Repeated exposure to one particular broadening of the encoded concepts may result in this concept being itself stored in memory as a stable concept which the hearer may use thereafter as a shortcut to construct other one-off ad hoc concepts in interpreting further utterances.

7. Conclusion

Modern psycholinguistic views of metaphor, in taking a constructivist approach to metaphor interpretation, see the comprehension of metaphor as involving no other cognitive processes than those involved in the comprehension of literal utterances. Evidence for this approach, as we have seen, is that metaphorical utterances do not take longer to understand than literal utterances, provided they are properly contextualised, and that metaphorical interpretations, like literal interpretations, are derived non-optionally.[10] If metaphor interpretation involves the same cognitive processes as the

[10] Although generally taken as supporting the claim that understanding literal and metaphorical utterances involves the same cognitive processes, this in itself is not enough to guarantee such a conclusion. The problem with psycholinguistic models is that they do not normally account for how literal

understanding of literal utterance an adequate general framework for utterance interpretation should be able to accommodate metaphor comprehension comfortably, and an adequate account of metaphor comprehension should require only processes involved in the comprehension of non-metaphorical utterances. Whereas standard pragmatic models fail to satisfy the first of these requirements, early and modern psycholinguistic models of metaphor interpretation typically fail to satisfy the second of these conditions. In this chapter, I have shown that, unlike traditional Gricean pragmatic views and modern cognitive approaches to metaphor, Relevance Theory is able to integrate the comprehension of metaphor within a general view of communication and cognition. Looking at metaphor as a lexical-pragmatic phenomenon, the relevance-theoretic approach is not only consistent with the experimental findings on metaphor processing but also meshes well with the wider field of creative cognition.

This chapter has illustrated once more how the selective processing of information associated with the encoded concepts often results in the construction of new concepts during the interpretation process. The inferential route taken in constructing these concepts and the steps of inference involved in their derivation are constrained at every point by pragmatic principles and expectations of relevance. Deriving an appropriate interpretation for a metaphorically intended utterance (as for any utterance) involves some degree of creativity as new representations are inferentially formed and new implications derived. The repeated processing of familiar metaphors, may at some point, result in selective processing of the set of assumptions and computation generally involved in processing the expression. These assumptions and computations may become more and more salient to the hearer directing his attention to a familiar inferential route and minimising the processing effort invested in interpreting the utterance.

This process whereby people develop inferential routes to understand familiar figurative uses will be explored in considerable detail in Chapter Six. While these last two chapters have studied the comprehension of relative ordinary metaphorical uses, the next two chapter looks into even more standardised figurative uses as they study how

---

interpretations are derived, so it is difficult to see whether different processes or mechanisms are involved. The account presented by Relevance Theory, and defended here, shows explicitly how the same comprehension procedure is involved in understanding both literal and metaphorical utterances.

people communicate with idiomatic expressions (e.g. *to break the ice, to spill the beans, to hold all the aces*). Traditionally characterised as dead metaphors, idioms have often been seen as no longer involving the types of interpretation processes required for metaphor interpretation. So if metaphor interpretation involves the type of inferential mechanisms outlined here, we might conclude that the comprehension of idioms does not involve similar processes. I believe that this intuitive view of idiom comprehension cannot, in fact, be sustained and I will challenge the claim that all there is to idiom comprehension is the retrieval of a ready-made meaning from memory, without any processing of information associated with the concepts encoded by the individual words that make it up.

Chapter 5

Analysability in Idiom Comprehension

"Idiomaticity is important for this reason, if for no other, that there is so much of it in every language."

Uriel Weinreich (1969: 23)

1. Introduction

The idea that the meanings of the constituents of most idioms are available during processing and contribute to understanding these expressions has figured prominently in modern (psycholinguistic) research (Cacciari and Tabossi, 1993; Everaert et al., 1995; Gibbs, 1994a; Glucksberg, 2001; Glucksberg and Cacciari, 1989; Nunberg, 1978; Nunberg et al., 1994.; Wasow et al., 1983). Contrary to the standard assumption that idioms are no more than long words with no internal composition, the idea that idioms at are least partly compositional or analysable has been taken to provide a more adequate explanation of the linguistic behaviour of idioms and of how people interact with these expressions. A wave of interesting and fruitful experiments has been carried out to test these hypotheses and a number of approaches to idiom representation and processing have been offered to support their findings (e.g. Cacciari and Glucksberg, 1991; Cacciari and Tabossi, 1988; Gibbs, 1990; Titone and Connine, 1999). Although this 'compositional' view of idioms has revolutionised the study of idiomaticity, I believe that the notion of 'compositionality' or 'analysability' in which modern approaches are grounded is not always one and the same. Since the degree of analysability of idioms is taken to affect how these expressions are acquired, used, processed and interpreted, there is an urgent need to clarifying this inconsistency. This chapter provides an overview of existing research on idioms as well as trying to shed light on the inconsistent terminological and theoretical issues inherent in early and modern approaches.

2. Idioms: Arbitrariness or Compositionality?

Two principles lie at the core of linguistic theory: the Principle of Arbitrariness (of the sign) (Saussure, 1916/1975) and the Principle of Compositionality. The principle of arbitrariness posits that the relation between a word (signifier) and its meaning (signified) is arbitrary in that there is nothing in the form of a word, say *table*, that gives us a clue about what that word means or the entities it is used to denote. A single concept, say TABLE, can indeed be expressed by different words in different languages (e.g. *mesa* in Spanish, *Tish* in German). The principle of arbitrariness applies widely, with only a few exceptions, such as the existence of onomatopoeic words for which a phonological resemblance to the entities denoted can be established. Whereas the meaning of words needs to be arbitrarily stipulated in memory, the meanings of phrases and sentences do not. This is because the meaning of a phrase or sentence can be derived compositionally by combining the meaning of its individual words according to the morpho-syntactic rules of the language. However, the existence of idioms such as those in (1) challenges this elegant picture.[1]

(1) *To hold all the aces, to speak one's mind, to break the ice, to lay the cards on the table, to pull s.o.'s leg, to give a hand, to stab s.o.'s back, to miss the boat, to pull strings, to be on cloud nine, to change one's mind, to lose one's train of thought, to hit the sack, to kick the bucket, to come out of the blue, to break s.o.'s heart, to spill the beans, to have one's feet on the ground, to turn over a new leaf, to be the icing on the cake, to keep s.o. at arm's length, to be the last straw (that broke the camel's*

[1] A considerable amount of work has been done within lexicology to define idioms so as to distinguish them from other formulaic uses such as irreversible binomials (e.g. *salt and pepper*), conventional similes (e.g. *as white as a sheet*), proverbs (e.g. *a stitch in time saves nine*), routine formulae (e.g. 'How do you do?'), indirect requests ('can you...?'), compounds (e.g. *spick and span*) among others (see Fernando, 1996; Fernando and Flavell, 1981; Makkai, 1972; Moon, 1998a; Weinreich, 1969). It is outside the scope of my work here to discuss these definitions. I have presented instead some typical examples in (1) for the reader to get an intuitive grasp of the type of phenomena to be discussed. It is important to notice that the terminology used in the literature and in everyday speech can be confusing. The term 'idiomaticity' is often used to refer to the whole class of formulaic uses including those in (1). The term 'idiomatic' is often used in everyday speech to refer to anything that sounds 'native-like', such as the correct use of prepositions. Whenever I use the word 'idioms', 'idiomatic expressions' or 'idiom strings' I am referring only to a subset of formulaic uses, such as those in (1). Whenever I refer to a meaning as idiomatic or to the study of idiomaticity, I will be referring to the meaning and study of these strings and not to those of any other phraseological unit.

back), to cost an arm and a leg, to go over the line, to fill the bill, to chew the fat, to add fuel to the fire, to get out of the frying pan into the fire, to be in the same boat.

The reason idioms challenge compositionality is that combining the meanings of the parts of an idiom according to the morpho-syntactic rules of the language yields in a literal meaning rather than an idiomatic meaning of the string. Lack of compositionality has indeed generally been considered an essential property of idioms and a good indicator of idiomaticity.

> "I shall regard an idiom as a constituent or a series of constituents for which the semantic interpretation is not a compositional function of the formatives of which it is composed" (Fraser, 1970: 22).
>
> "[They] are idiomatic in the sense that their meaning is non-compositional" (Chomsky, 1980: 149).
>
> "Idioms [...] do not get their meanings from the meanings of their syntactic parts" (Katz, 1973: 358).

It might be argued that one way to bridge the gap between the literal and the idiomatic meaning would be to treat these expressions as metaphorical, and so to assume that the gap is bridged by pragmatic inference. This possibility has been discarded traditionally on the basis that the constituent words in the idioms are semantically empty. Although a motivation for why the idiom means what it does may have been available at some point in time, there is now no apparent synchronic reason why people would describe dying as kicking a bucket or revealing hidden information as spilling beans. The standard way of dealing with idioms has been to treat them as lexical items whose meaning cannot be compositionally derived or pragmatically inferred, but is simply arbitrarily stipulated in memory and retrieved as a unit in interpretation.

There are two types of evidence in favour of the view of idioms as lexical items. First, psycholinguistic research has shown that idiom strings are processed considerably faster than literal strings of the same length (e.g. *spill the beans* is understood faster than 'reveal a secret'). Also, an idiom is understood faster when processed in a context in which its idiomatic meaning is intended than when processed in a context in which its literal meaning is intended (e.g. *kick the bucket* is understood faster when used to mean

'die' than when used to mean 'hit a pail with one's foot') (Gibbs, 1980; Ortony et al., 1978). Second, in assuming that a single semantic representation is assigned to idioms as a whole and not to their individual components, traditional scholars expected idioms to behave linguistically as lexical items and so not to allow internal transformation (Chomsky, 1980; Cruse, 1986; Fraser, 1970; Katz, 1973). Examples such as those in (2)-(6) have often been used to support this claim:

(2) **The bucket was kicked* by John.
(3) **The bucket*, John *kicked* yesterday.
(4) *I think he will *kick the rusty bucket* soon.
(5) *The *breeze was being shot.*
(6) *Mary was *chewing the fat* and Peter was *chewing it* too.

Although the assumption that idioms are lexical items with no internal composition has been deeply rooted in linguistic research for years, modern research on idioms has begun to question this characterisation. One motivation for this challenge has been the observation that many idioms do, in fact, allow a considerable degree of internal transformation, illustrated in the English and Spanish examples in (7)-(13):

(7) Many *strings were pulled* but he was not elected.
(8) He is very stubborn, but in the end he will have to *change his* square *mind* and accept the deal.
(9) I have been a real idiot. They'll never offer me that job now. That was *the last boat and I missed it*
(10) The *strings* he said he would *pull* for you
(11) Ya se que estás ocupado pero de vez en cuando *una mano* en la casa me podrías *echar.*
I know you are busy but every now and then *a hand* in the house you could *throw* at me. (literal translation (LT))
Spanish idiom: *to throw a hand* – equivalent English idiom: *to give a hand*

(12) No me gustan las bromas, así que la próxima vez, *el pelo* se lo vas a *tomar* a tu padre!

I don't like jokes, so next time, *the hair* you are going to *take* to your father! (LT)

Spanish idiom: *to take someone's hair* – meaning: to tease someone.

(13) Durante la reunión *se pusieron* todas *las cartas sobre la mesa.*

During the meeting all *the cards were laid on the table.*

Spanish idiom: *poner las cartas sobre la mesa* – equivalent English idiom: *to lay one's cards on the table.*

The standard linguistic argument that idioms have no internal structure and so behave linguistically as lexical items may be able to explain why some idiomatic uses, such as those in (2)-(6), are not acceptable yet it fails to explain why others, such as those in (7)-(13), are. The examples in (7)-(13) show that at least some idioms appear to behave linguistically as phrases and not as long words. As a result, they are capable of undergoing a number of syntactic transformations such as passivisation, as in (7), topicalisation, as in (10), internal modification, as in (7), (8) and (13) or ellipsis, as in (9), without losing their idiomatic meaning. The importance of transformations like these is that they are not external and so affect the string as a unit, but internal and so affect only a constituent of the idiomatic phrase, which can be focused, modified or even omitted.

A number of scholars have struggled for years to establish the syntactic rules underlying idiomatic usage - rules that will explain, for instance, why idioms such as *kick the bucket* or *break a leg* do not passivise whereas idioms such as *pull strings, spill the bean* and *lay one's cards on the table* do (e.g. Chafe, 1968; Fraser 1970; Katz, 1973; O'Grady, 1998; Reagan, 1987). Despite these efforts, few generalisations have been forthcoming. Idioms have turned out to be a rather heterogeneous class with different idioms allowing different numbers and kinds of transformation. The transformational capacity of idioms seems to range from minimal (morphological) variation (e.g. tense marking '*he kicked the bucked*') to almost full syntactic flexibility as in (14):

(14) a. This is the most interesting episode, the one when *the beans are finally spilled* (passive)

b. The soldiers will *spill the beans* under pressure but I am not sure the high commanders will *spill them* that easily (omission and anaphoric reference)

c. Despite the torture, he didn't *spill a single bean* (internal modification)

d. She won't be pleased by just anything you say, you'll have to *spill all the beans*, no more secrets! (internal modification)

e. He is a very reserved person. I am sure, sooner or later, he'll open up and will tell me more about his current life, but not about what happened between him and his brother. *Those beans*, I am certain he'll never *spill*, not even to me (focus)

It is important to notice that even if a theory of idiom transformation assumes that different syntactic rules apply to different idioms, it will still have to explain how native speakers acquire those rules. It will have to explain for instance how native English speakers know that a sentence such as 'the *law needs to be laid down* soon' has an acceptable idiomatic reading while a sentence such as '*the fat will be chewed* tonight' does not even if they have not encountered any of these idioms in these forms ever before.

The difficulty of accounting for the syntactic versatility of idioms in purely syntactic terms has led some scholars to consider the phenomena as semantically motivated. Newmeyer (1972), for instance, proposed that idioms have the same semantic structure as their equivalent literal paraphrases and that this has an effect on their syntactic behaviour. In his view, the reason why an idiom such as *kick the bucket* does not passivise is that the verb 'die' is intransitive and so does not passivise. The reason why the idiom *spill the beans* passivises is that the verb 'reveal' is transitive and allows passivisation. However, Nunberg (1978: 212) saw several important problems with this proposal. Not only can an idiom be paraphrased in many ways, but also two idioms with similar meanings may actually have very different syntactic behaviour. So, even though *kick the bucket* and *give up the ghost* can be roughly paraphrased as 'die', the latter but

not the former idiom allows passivisation without disruption of the idiomatic reading. Compare example (2) with (15):

(15) Once *the ghost has been given up*, there is nothing medical science can do (Nunberg, 1978: 212)

In order to solve the mystery of the heterogeneous transformational potential of idioms, Nunberg (1978) proposes to look not at syntax or semantics alone, but at the relation which language users establish between the two. To do this, he argues, one first has to abandon the standard assumption that a single arbitrary semantic representation is assigned in memory to the string as a whole. Rather than looking at idioms as a homogeneous non-compositional group, Nunberg proposes to view them as lying along a continuum of compositionality or analysability varying in the extent to which their individual parts have identifiable idiomatic referents (Nunberg, 1978; see also Nunberg et al., 1994; Wasow et al., 1983). The degree of compositionality of idiom strings is believed to play a major role in the linguistic behaviour of these expressions.

3. Idioms as (Partly) Analysable Phrases

A basic assumption underlying Nunberg's approach (Nunberg, 1978) is that people may be able to provide a post-hoc motivation for why the idiom means what it does e.g. why *kick the bucket* is used to refer to the act of dying, *spill the beans* to the disclosing of secrets, *hit the sack* to the act of going to bed, etc., as well as assume everyone else makes sense of the meaning of the idiom in just the same way. The set of beliefs that people take to license idiom meaning need not be that which motivated the meaning of the idiom back in history but just a set of assumptions which helps them to make (synchronic) sense of its current meaning.[2]

One way people have of making sense of the meaning of idioms is to see whether individual words in the idiom have the same meanings as they do in some non-idiomatic contexts. Nunberg points out that English native speakers may assume that the word *kick*

[2] The next chapter explores this idea in considerable depth.

in the idiom *kick the bucket* is the same verb *kick* which appears in phrases such as 'the man kicked the ball' and the word *hit* in *hit the panic button* to be the same word *hit* which appears in expressions such as 'the man hit the wall', which can be roughly paraphrased as 'strike'. Stating the 'literal' meaning of an idiom depend on its use on a particular occasion. On some occasions, people may take a single word such as *hit* to mean 'strike', as in *hit the panic button;* on other occasions, they make take the same word to mean roughly 'collide with something violently', as in *hit the sack;* or to mean 'make contact with', as in 'to hit bottom'. He proposes that the assumptions that people make about the contribution of the use of *hit* in each of these phrases would have an effect on the way they use these expressions, and so on their syntactic behaviour. The same is believed to be the case for *kick* in *kick the bucket.* Taking the word *kick* in *kick the bucket* to refer to a one-off act, for instance, may constrain the syntactic behaviour of the idiom, preventing it from being used idiomatically in the progressive form, e.g. 'he was *kicking the bucket* for weeks'.[3]

A crucial assumption of Nunberg's model is that in providing a rationale for the meaning of an idiom, people generally attempt to decompose the meaning of the idiom into the meanings of its parts in such a way that the constituents of the idiom are mapped to the elements in the relation the idiom refers to.[4] In learning that the expression *pop the question* refers to the act of proposing marriage, for example, people generally take the verb *pop* to refer to the act and manner of uttering something and the noun phrase *the question* to refer to the particular question being asked. Similarly, in learning that the expression *spill the beans* refers to the act of revealing hidden information, people take the verb *spill* to denote the act of revelation and *the beans* to

[3] It is difficult to know whether the idiom *kick the bucket* is taken to refer only to abrupt deaths because of the constraints imposed by the meaning of *kick*, or whether people have taken the verb *kick* to be making some kind of contribution to the meaning of the idiom (e.g. assumptions about abruptness) because the idiom refers to abrupt deaths. Whatever the direction of the inference, the main point is that there seems to be a correlation, however minimal, between a) the assumptions people make about idiom meaning and b) the syntactic or linguistic behaviour of these strings.

[4] It is worth pointing out that the notion of decomposition here is very different from the notion of decomposition defended in some approaches to lexical semantics. Decompositional lexical semantics assumes that the meaning of a word e.g. bachelor can decompose into certain meaning primitives e.g. SINGLE, ADULT, MALE, which define the word. The claim of the decomposition view of idioms is different. The idea here is not that the meaning of an idiom is decomposed into meaning primitives, but that its meaning is analysable into its constituent parts so that each part of the idiom can refer to an element in the idiom's denotation (e.g. in *pop the question*, *pop* is taken to refer to the act of uttering and *question* to the type of question being uttered, namely a marriage proposal).

refer to the information being revealed. Intuitions about the 'literal' meaning of idioms such as these, like intuitions about the 'literal' meanings of idioms such as *kick the bucket* or *hit the sack* above, are thought to depend on the hypotheses people make about how the idiomatic use is compositionally derived (Nunberg, 1978: 216). And the hypotheses people make about how an idiom is compositionally derived depend on the beliefs which people take to license idiom meaning.

According to Nunberg, the reason why idioms such as *kick the bucket* and *give up the ghost* have different linguistic behaviour, as in (2) versus (15), even though they have roughly the same meaning, is that people take the meanings of these idioms to be licensed by substantially different beliefs. The idiom *kick the bucket* is taken to be licensed by the normal assumption that the act of dying is a one-place relation that refers to a certain change in the state of a person. Since it is a one-place relation, the individual parts of the idiom cannot be assigned independent referents. There are, however, less conventional but also widely accepted ways in which people often characterise death, which involve a two-place relation, e.g. between a person and his spirit, between a soul and a body, a person and God. The meanings of the idioms *to give up the ghost* or *to shuffle off this mortal coil, to meet one's maker,* etc., may be licensed by these beliefs, which allow people to perceive these expressions as relatively decomposable. For the idiom *give up the ghost*, the verb *give up* is taken to refer to the act of abandoning or rendering, and the noun phrase *the ghost* to a person's spirit or soul. According to Nunberg, it is difficult to see how the expression *kick the bucket,* denoting an abrupt death, could be seen as invoking a two-place relation. Had the expression been *kick this bucket* or *kick this vale*, people could make sense of the idiom in such a way as to take the verb *kick* to denote the act of abandoning something and the noun phrase *this bucket* or *this vale* to refer to the thing being abandoned e.g. this world. Licensing idiom meaning in this way would result in people perceiving the expression as relatively decomposable, and so relatively syntactically flexible (e.g. 'the day will come when *this bucket we will finally kick*').

In contrast to traditional models which treat idioms as belong to a homogeneous non-compositional group, this approach sees idioms as lying along a continuum of compositionality or analysability varying in the extent to which people can see how the

constituents in the idiom might refer to a constituent of the state of affairs the idiom refers to. At one end of the spectrum of compositionality, there are what Nunberg refers to as 'non-decomposable idioms'. These are expressions for which people cannot assign idiomatic referents to the individual constituents in the idiom, but only to the string as a whole. They include idioms such as *kick the bucket* or *chew the fat,* for which a single semantic representation (e.g. referring to the act of dying suddenly or talking informally, respectively) is mapped to the idiom as a unit. At the other end of the spectrum, there are idioms which Nunberg refers to as 'normally decomposable'. These are idioms which people can identify as referring to an open relation Rxb for which they can easily see how the constituents of the idiom map to the elements in this relation. The idiom *pop the question,* for instance, is said to refer to a relation between a person and a proposal. This idiom is normally decomposable to the extent that people can see that the verb *pop* refers to this relation and the noun phrase *the question* to the proposal. Unlike idioms such as *kick the bucket* or *chew the fat* idioms such as *pop the question, break the ice, lay down the law, pass the buck, spill the beans, draw the line, foot the bill, pull strings,* etc. are relatively normally decomposable, as people can see how each of the parts can refer to a constituent of the referent of the idiom.

The main aim of the decomposition view has been to argue that the syntactic behaviour of idioms is not a random phenomenon, but depends to a considerable extent on the idiom's degree of decomposability (Nunberg, 1978; Nunberg, Sag and Wasow, 1994; Wasow, Sag and Nunberg, 1983). The easier it is for people to identify the constituents in an idiom as having independent referents, the easier it is to see how the individual parts of the idiom contribute to the overall idiomatic interpretation, and the more likely that those parts can be focused (e.g. in passives), substituted or modified. In this way, the reason why idioms such as *pull strings* and *spill the beans* can be internally modified, as in (7), (10) and (14), whereas idioms such as *kick the bucket* or *chew the fat* in (2)-(4) and (6) cannot, is said to be because people can generally assign independent idiomatic meaning to the constituent parts of the former (*spill* –> 'reveal', *the beans* –> 'hidden information'; *pull* -> 'exertion of force', *strings* –> 'influence') but not of the latter. According to this view, the traditional assumption that idioms have no internal composition and so behave linguistically as lexical items cannot be maintained for

decomposable idioms, even if it can still be claimed to hold for non-decomposable idioms such as *kick the bucket* or *chew the fat*. The failure of traditional models to spot the transformation potential of idioms may well have been due to their tendency to focus on a few 'non-decomposable' idioms. The possibility that many idioms are decomposable opens a new window through which to look at idioms and their transformation potential.

> "[W]e will say that an idiomatic transitive VP is decomposable just in case it is used to refer to a state or activity such that it would be normally believed that that activity could be identified as an open relation Rxb, such that the object NP of the idiom refers to b, and the verb to R [...] We can assume then, that speakers will not passivise idiomatic VP's that are not decomposable. If there is no assurance that hearers will be able to identify the referent of the focussed NP, then it makes no sense to focus on it" (Nunberg, 1978: 221, 225)

Somewhere between normally decomposable idioms and nondecomposable idioms are what Nunberg refers to as 'abnormally decomposable' idioms. Unlike normally decomposable idioms, an abnormally decomposable idiom is not licensed by conventions whereby each of its constituents can be used to refer to the constituents of its referent; instead this relation is mediated by a conventional metaphor. People can see, for instance, how the idiom *hit the panic button* is used to refer to someone's becoming panicked only by thinking about the act of literally hitting a panic button that people can perform in dangerous circumstances. According to Nunberg, abnormally decomposable idioms should not be expected to be syntactically productive and should be particularly odd in passives. The reason for this is that "the object NP of abnormally decomposable idioms does not itself refer to some component of the idiomatic referent, but only to a component of the relation by means of which that referent is conventionally identified" (Nunberg, 1978: 228). [5]

[5] In his analysis of the effect of decomposition on the syntactic behaviour of idioms, Nunberg (1978) was mostly interested in the use of focus in passives. As the quote indicates, only parts of idioms to which people can assign independent idiomatic meanings (i.e. parts with identifiable idiomatic referents) should be capable of being focused. Nondecomposable idioms are odd in the passive, precisely because people cannot assign meaning to idiom parts in this way (e.g. **the bucket was kicked* by him). Abnormally decomposable idioms are often odd in the passive too (e.g. '?It is better that he does not find out his daughter is pregnant or *the panic button will be hit*). The reason this time is that the parts of idioms (e.g. the NP) do not refer directly to an element in the idiomatic referent but only to a (metaphorical) relation

To sum up, Nunberg and colleagues have suggested that the phenomenon of idiomaticity has generally been overgrammaticalised. Syntactic and semantic rules have been proposed to account for regularities which result from independent pragmatic processes and discursive functions (Nunberg et al., 1994: 494). A better solution, they argue, is to consider the possibility that grammar may freely perform a range of syntactic operations, such as those in (2)-(15).[6] The acceptability or unacceptability of these sentences when understood as idiomatic would need to be explained not on syntactic but on functional grounds. It would depend on the beliefs which people take to license the idiom's meaning and which make language users perceive a certain mapping between (the elements) in the syntactic structure of the idiom and (the elements in) the denotation of the idiom. People generally agree in their intuitions about which transformations apply to which idioms, even if they have never encountered those idioms in those forms before, because they use roughly the same assumptions to license the meanings of familiar idioms.

4. Psycholinguistic Research on the Analysability of Idioms

The idea that the constituent words of most idioms have identifiable meanings which contribute to the overall idiomatic meaning of the expression (the 'compositional hypothesis') has been very influential, not only within linguistics (e.g. Everaert et al., 1995; Geeraert, 1995; Nunberg, 1978; Nunberg, Sag and Wasow, 1994; Wasow, Sag and Nunberg, 1983) but also, and particularly, within psychology (e.g. Cacciari and Glucksberg, 1991; Cacciari and Tabossi, 1993; Gibbs, 1994a; Gibbs and Nayak, 1989; Glucksberg, 1993, 2001; Glucksberg and Cacciari, 1989; Titone and Connine, 1999). In fact, the motivation for this research, and for the challenge to the standard characterisation of idioms as long words, comes not only from the observation that idioms allow a considerable amount of syntactic transformation, as in (7)-(13), but also

---

by means of which that referent is typically identified. According to this view "only (but not all) ... normally decomposable idioms can be passivised" (Nunberg, 1978: 228).

[6] The idea that the study of idioms and the restrictions on their transformation is not a simply matter of syntax has been considered by a number of researchers (e.g. Ackema and Neeleman, 2001; Flores d'Arcais, 1993; Geeraerts, 1995; Peterson and Burgess, 1993; Van de Voort and Vonk, 1995) and it is the view I assume in this work.

from observations about their possibilities of semantic modification or lexical substitution, as in (16)-(21):

(16) He is very stubborn, but in the end he will have to *change his square mind* and accept the deal. (from (8))

(17) During the meeting *all the cards were laid on the table.* (from (13))

(18) Despite the torture, he didn't *spill a single bean.*

(19) He absolutely hates me, so if it is true he has found out about my affair, he must now be in my house *pouring the beans* to my wife.

(20) OK there! Now you are *barking up the right tree!*

(21) Sin darme cuenta, *me metí de cabeza en la boca del lobo.*
Without realising, I *got headfirst into the wolf's mouth.*
Meterse en la boca del lobo (to get into the wolf's mouth) – to get into a problematic or dangerous situation.

These examples show how individual parts of an idiom, as opposed to the string as a whole, can be modified in conversation so as to convey a meaning different from the meaning the speaker would have conveyed had he used the original form. This flexibility shows once more that idiom strings typically behave very much like phrases rather than lexical items. Taken together, the examples in (7)-(13) and (16)-(21) highlight the need to move away from the standard assumption that all idioms are non-compositional strings and towards the view that the internal semantic composition of most idioms affects the way people interact with these expressions.

A wide range of experimental work has been carried out in psycholinguistics to show that the degree of analysability of idioms plays an important role in their acquisition, use and interpretation. It has been shown, for instance, that analysable idioms are easier to acquire by children and faster to process by adults than unanalysable idioms (Gibbs, 1987, 1991; Levorato and Cacciari, 1992, 1995, 1999). Analysable idioms show a higher potential for internal lexical and syntactic transformation than unanalysable idioms (Gibbs and Gonzales, 1985; Gibbs and Nayak, 1989; Gibbs, Nayak and Cutting, 1989). Modern research agrees that the reason for these differences is that

for analysable idioms, but not for unanalysable idioms, people can see how the individual words in the idiom have a meaning which contributes to the overall idiomatic interpretation. Although the compositional hypothesis that most idioms are partly analysable has revolutionised the study of idiomaticity and is now deeply rooted in modern (psycholinguistic) research on idioms, I will argue here that it is not always clear that scholars have the same notion of compositionality or analysability in mind.

4.1 The Role of Analysability in Idiom Use and Interpretation

We have seen that the compositional view of idioms on which virtually every modern approach to idioms is grounded is based on two assumptions. First, it assumes that people have consistent intuitions about the degree of idiom analysability or compositionality: that is, the degree to which people can see how the internal semantics of the idiom contributes to the overall idiomatic interpretation. Second, it assumes that the degree to which people perceive an idiom as compositional (analysable) has an effect on how they use it and so on the acceptability of idioms which have undergone syntactic or semantic transformation. In order to test the first of these assumptions, Gibbs and Nayak (1989) provided subjects with a list of idioms, each one with a paraphrase of their meaning (e.g. *kick the bucket* -> to die; *spill the beans* -> to reveal a secret; *break the ice* -> to start a conversation, etc.) and asked them to divide the members of the list into two groups according to whether or not the individual words made "some unique contribution" to the phrase's idiomatic meaning. Borrowing Nunberg's terminology, the authors referred to the idioms whose parts made such a contribution as 'decomposable' (or 'analysable'), and to the idioms whose parts made no such contribution as 'non-decomposable' (or 'non-analysable'). The idiom *pop the question* was described as decomposable because the individual constituents *pop* and *question* were said to be semantically related to the meanings of 'utter' and 'marriage proposal'; and so taken to contribute individually to the phrase's idiomatic meaning. The idiom *kick the bucket* was described as non-decomposable because its individual constituents *kick* and *bucket* were not seen as making an independent contribution to the idiomatic meaning.

Given this classification, subjects were asked to take the set of idioms they had judged as decomposable and split them into two further groups. In one group they were asked to include idioms whose individual words were "directly" related in meaning to their idiomatic interpretations ('normally decomposable' idioms'). In the other group, they were asked to include idioms whose individual constituents were related to their figurative meaning in a less direct and more metaphorical way ('abnormally decomposable idioms'). The idiom *pop the question* was provided as an example of 'normally decomposable' idioms because the word *pop* is closely related to the idea of sudden asking and the word *question* to the type of question being asked (i.e. marriage proposal). The idiom *spill the beans* was offered as an example of the abnormally decomposable class because, according to the authors, although it is easy to perceive a close relationship between the meaning of *spill* and the act of revealing, there is a much less direct, metaphorical, relation, between the meaning of *beans* and its idiomatic referent (secrets). The two experiments outlined here resulted in the following classification (see also Gibbs, Nayak, Bolton and Keppel, 1989):

(22)

Nondecomposable	Normally decomposable	Abnormally decomposable
Chew the fat	*Break the ice*	*Bury the hatchet*
Cool one's heels	*Button one's lip*	*Carry a torch*
Give the sack	*Clear the air*	*Cook one's goose*
Give the bounce	*Hit the sauce*	*Crack the whip*
Go for broke	*Lay down the law*	*Get down to brass tacks*
Kick the bucket	*Let off steam*	*Lay an egg*
Knock on wood	*Lose one's grip*	*Line one's pocket*
Make the scene	*Miss the boat*	*Pass the buck*
Pack a punch	*Play with fire*	*Promise the moon*
Play the field	*Pat on the back*	*Push the panic button*
Raise the roof	*Pop the question*	*Spill the beans*
Shoot the breeze	*Rack one's brains*	*Steal one's thunder*

The results of the experiments showed that subjects were highly consistent in their classifications. On the basis of these findings, Gibbs and Nayak (1989) proposed what they referred to as the 'Idiom Decomposition Hypothesis'; which, like Nunberg's

decomposition view, posits that people have strong intuitions about the way in which idiom parts contribute to the overall idiomatic interpretation of the string.

Having found that people have common intuitions about how the parts of idioms contribute to the overall idiomatic meaning, Gibbs and colleagues aimed to test whether these intuitions would have an effect on people's judgements about the syntactic flexibility of idioms (as claimed by Nunberg and colleagues) as well as their potential for lexical transformation. Gibbs and Nayak (1989) therefore presented subjects with sets of idiom variants based on the idioms in the list in (22), which had been transformed in various ways (e.g. they had undergone tense marking, adverb insertion, adjective insertion, passivisation, or action nominalisation). Each of these variant forms was presented along with a paraphrase of the idiomatic meaning of the original idiom. Subjects were then asked to rate the similarity of meaning between each pair (e.g. the similarity between 'the teacher will *lay down* the school *law* if the children make too much noise' and 'the teacher will give strict school orders if the children make too much noise'). Ratings were made on a 1-7 point scale with 7 indicating high similarity and 1 indicating no similarity.

As predicted by the authors, the results showed that subjects gave higher ratings to idioms which had been judged in the previous experiment as normally decomposable (average of 5.08) than to those previously judged nondecomposable (average 4.60). The authors took this difference as relatively significant and as consistent with the Idiom Decomposition Hypothesis. They argued that the reason why normally decomposable idioms can appear in a variety of syntactic forms is precisely because people can see how the parts of the idiom contribute independently to the overall idiomatic interpretation. They can see, for instance, how the verb *lay down* is used to refer to the act of establishing something and the noun phrase *the law* is used to refer to the rules being established. Nondecomposable idioms, however, are said to be syntactically frozen because the individual parts of these idioms do not contribute independently to the overall idiomatic interpretation.

A very similar experiment was carried out to test whether people's intuitions about the degree of compositionality of idioms also affects their judgements about lexical transformation. Gibbs, Nayak, Bolton and Keppel (1989) presented subjects with

variants of the idioms in (22), this time with some type of lexical substitution (e.g. *fasten one's lip, sink the hatchet*), each idiom variant was presented along with a paraphrase of the idiom's original meaning, and the task was to rate how similar the meaning of the new form was to the meaning of the original form. Different conditions were tested: a) one in which the verb was substituted by a semantically related word (e.g. *kick* by *punt* in *kick the bucket* or *hold* for *carry* in *carry a torch)*; b) another in which the noun in the NP was substituted by a related noun (e.g. *kick the pail, carry a light*); and c) a condition in which both noun and verb were substituted (e.g. *punt a pail, hold a light*). Ratings were made on a 1-7 point scale, with 7 indicating high similarity and 1 indicating no similarity:

(23)

CHANGE	Normally decom	Abnorm decomp	Non-decomp	Mean rating
No change	5.44	5.42	5.35	5.38
Verb	3.42	3.61	3.21	3.46
Noun	4.10	3.67	3.20	3.69
Both	3.23	3.10	2.50	2.94
Mean rating	4.08	3.95	3.58	

As expected by the authors, the results showed that ratings for nondecomposable idioms were lower than for normally and abnormally decomposable idioms in each condition and overall. Mean ratings for each type of lexical substitution showed that better ratings were provided when at least one of the parts of the idiom remained unchanged. The most significant difference seems to be between normally decomposable idioms and nondecomposable idioms when the noun phrase of the original expression is replaced and when both the noun phrase and the verb are different. Overall, the table shows that the ratings between the different idiom types within the same condition are quite close, nevertheless, the authors took these results as significant and once more interpreted them as evidence in favour of the Idiom Decomposition Hypothesis. That is, it is claimed that the reason why people rate variants such as *bury the axe* as more acceptable than variants such as *chew the lard* or *kick the pail* is that for the former, but not the latter, they can see how the meanings of the individual words in the original idiom make a meaningful contribution to the overall idiomatic meaning.

A further line of research has investigated how the degree of analysability of idioms affects the ease/difficulty of comprehension and acquisition. According to traditional models, children should learn idioms by arbitrary association of meaning to a certain linguistic form; however, experimental research has shown that this is not generally the case. Gibbs (1991) found that, in making sense of the meaning of an idiom, children generally attempt to do some compositional analysis. In testing children, of ages five to ten, he found that younger children understood decomposable idioms better than nondecomposable idioms. When an idiom is analysable (decomposable), children are seen as assigning independent meanings to its individual parts (e.g. assigning meanings to *lay down* and *the law* respectively) and as combining these parts of the idiom and their meanings to form the overall idiomatic interpretation. According to Gibbs, children find it difficult to understand non-decomposable idioms because the overall idiomatic meaning cannot be determined by a compositional analysis of its parts.

The same argument is also said to apply to adult comprehension. Gibbs, Nayak and Cutting (1989) found that, contrary to the prediction of traditional models, people take longer to process nondecomposable idioms than either decomposable idioms or literal control strings. However, decomposable idioms are processed faster than both non-decomposable idioms and literal control strings. The authors argued here too that decomposable (analysable) idioms are more easily understood because people can identify the idiomatic meanings of the parts that compose the idiom. Since each of the components of a decomposable idiom has an independent meaning which contributes to the overall idiomatic meaning, these expressions are processed "in a compositional manner where the semantic representation of each component is accessed and combined according to the syntactical rules of the language" (Gibbs, 1994a: 284). A compositional analysis of non-decomposable idioms, however, does not provide information about the idiomatic meaning of these expressions and is indeed generally inconsistent with this meaning; that is why these idioms take longer to process.

4.2 Comments on Experimental Research

The results of the above experiments, as of many others in the literature, have been generally accepted (by psycholinguistic) research on idioms, without further question, as

providing evidence that people share intuitions about the degree of analysability of idioms, and that these intuitions affect the way they use and understand idioms. In much of this research, the terms 'analysability', 'compositionality' and 'decomposition' seem as closely linked; the degree of analysability/compositionality of an idiom is seen as essentially dependent on the degree to which the overall idiomatic meaning can decompose into the meanings of its parts (i.e. on the ability of people to assign independent idiomatic meanings to the parts of the idiom).

Gibbs and colleagues have defined the notion of decomposition in terms of semantic fields:

> "One way of characterising idiom decomposition is in terms of semantic *fields* [...] For example, the individual parts of *pop the question* must be in the same semantic field, or conceptual domain, as their idiomatic referents "propose" and "marriage" for this idiomatic phrase to be, at least, partially decomposable" [...] However, the individual components of phrases such as *kick the bucket* or *chew the fat* are not in the same semantic fields as their respective figurative referents (i.e. "to die" or "to talk without purpose") and should not be viewed as semantically decomposable" (Gibbs and Nayak, 1989: 107).

This quote is puzzling. An idiomatic referent cannot be in a conceptual domain, because it is not a conceptual entity but an object in the world. To make sense of the quote we must take Gibbs and Nayak to be claiming that the meanings of the words *pop* and *propose* are semantically related. Analysability in this model can then be seen as the degree to which people can assign independent idiomatic meanings to the idiom's parts and/or the extent to which the idiom's overall meaning is semantically related to the literal meaning of its component words. This lack of a distinction between decomposition (the extent to which the parts of the idiom are mapped to elements in a representation of the state of affairs the idiom refers to), on the one hand, and meaning relatedness (the extent to which the literal meanings of the words in the idiom are related to the overall idiomatic meaning and its parts), on the other, is very common in the psycholinguistic literature on idioms, particularly in the work of Gibbs and colleagues.

It is precisely because analysability is seen as linked to the combination of these processes that it is not always clear from much experimental work what it is that really

affects people's use and understanding of idioms. For instance, do subjects understand analysable idioms faster than unanalysable idioms because, in processing decomposable idioms, they select the figurative meanings assigned to the parts of the idiom and combine them compositionally according to the rules of the language? Or is it because, in understanding analysable idioms, the initial literal processing of the words in the idiom can be integrated into the derivation of the idiomatic interpretation? Do people accept syntactically transformed idioms because they can assign independent idiomatic meanings to the parts regardless of their literal meaning? Or is syntactic transformation essentially dependent on the semantic contribution of literal word meanings to idiom meaning? And what about lexical transformation?

One interesting result of the experiments on syntactic and lexical transformation described above is the different ratings given to abnormally decomposable idioms when they had undergone syntactic and lexical transformation. Whereas the average rating for abnormally decomposable idioms was the same as that for nondecomposable idioms in the syntactic task (4.62), it was close to the average rating for normally decomposable idioms in the lexical task (3.95 vs 4.08). How can the finding that abnormally decomposable idioms are simultaneously syntactically frozen and lexically flexible fit the compositional hypothesis? In explaining the results of the syntactic experiment, Gibbs and Nayak (1989) argued (as in Nunberg, 1978) that people find it difficult to modify the abnormally decomposable idioms because their individual parts do not refer directly to their referents, but do so only by means of a certain metaphorical relation. In explaining the results of the lexical experiment, Gibbs, Nayak, Bolton and Keppel (1989) claimed that abnormally decomposable idioms are lexically flexible because people perceive a certain metaphorical relation between the idiom components and the overall idiomatic meaning. I believe that this internal contradiction is not trivial and indicates that syntactic and lexical flexibility may depend on quite different processes which a single notion of 'analysability' or 'compositionality' is not able to capture.

I suggest that the potential for syntactic transformation of idioms may depend on whether people can identify idiomatic referents for the individual parts of the idiom and not so much on whether the independent idiomatic meanings assigned to those parts are related to the literal meanings of the words or the idiom as a whole. After all, idioms

such as *spill the beans* are very syntactically flexible, as in (14), with the word *beans* being modified, focused and omitted even though its literal meaning is related neither to the idiomatic meaning 'hidden information' nor to the overall idiomatic meaning of the expression. It is the link that people can establish between elements of the syntactic structure of the idiom and elements of the relation the idiom is used to denote (e.g. between the act of dying suddenly, revealing a secret or making peace and the syntactic elements V and NP of the expressions *kick the bucket*, *spill the beans* and *bury the hatchet*) that is a crucial determinant of syntactic flexibility.

By contrast, the acceptability or unacceptability of a lexical substitution such as those presented in the experiment above need not depend so much on whether the elements of the idiom map to elements in the denotation as on whether people can establish a relatively motivated and transparent relation between the literal meaning of the words in the idiom and the meaning of the idiom as a whole. We can substitute the word *pour* for *spill* in *spill the beans*, as in (19), but not *peas* for *beans*, because only the meaning of the word *spill* is relatively transparently related to the meaning of the idiom, whereas the meaning of the word *beans* is not. The fact that abnormally decomposable idioms can undergo a certain degree of lexical transformation suggests people can establish a relatively motivated relation between the meanings of the words in these idioms (or the way they combine) and the overall idiomatic interpretation. Provided we have an appropriate context, we may be able to *hit, push* or *press the panic button* or even *hide, burn,* and *sink the hatchet.*

It is also worth noting that the inconsistency of the results of these experiments may also be due to the vagueness of the instructions given in the classification task. In the classification experiment, subjects were asked to divide a list of idioms into normally decomposable, abnormally decomposable or nondecomposable depending on how "saliently" the parts of the idiom contributed to the idiomatic meaning. So idioms whose individual parts made a "unique contribution" to the overall idiomatic meaning were classified as decomposable. If this contribution was "direct" they were further classified as normally decomposable; if it was less direct e.g. "metaphorical", they were classified as abnormally decomposable. But what is meant by a "unique" contribution? or by a contribution being "direct"? Does it mean that normally decomposable idioms are those

for which at least one of the parts can be interpreted literally? Would an idiom for which one can assign independent referents to the idiom parts, but whose parts can only be figuratively interpreted count as normally decomposable?

In providing an example of an abnormally decomposable idiom, Gibbs and Nayak (1989) chose the expression *spill the beans* because, according to them, although the word *spill* makes a direct contribution to the overall idiomatic meaning, the noun phrase *the beans* does not. Instead, it makes a less direct contribution, which they characterise as "metaphorical". One problem with this is that, for most people at least, if there is any metaphorical contribution of the meanings of the words in the idiom *spill the beans*, it would more likely to be provided by the word *spill*, not *beans*! A second, more crucial, problem is that it is not clear whether this idiom would actually count as abnormally decomposable, at least according to Nunberg's definition of abnormally decomposable idioms.

According to Nunberg, an idiom may be normally decomposable even if the rationale for using one constituent is relatively obscure. This applies to idioms such as *spill the beans* or *paint a pretty picture.* What matters for an idiom to be normally decomposable is that people can easily see how the verb in these idioms refers to the (mode of) transmission of information and the noun phrase to the information being transmitted, even if there is not a clear relation between the literal meanings of the noun phrases and their idiomatic meanings. Although these idioms are normally decomposable, and therefore likely to be syntactically flexible, one element of the idiom (e.g. *the beans*) is mapped to an element of the relation the idiom is used to denote (e.g. hidden information) merely by elimination (e.g. if the idiom denotes the revelation of secrets and the verb *spill* refers to the act of revealing, then the noun phrase must refer to the secrets being revealed). In seeing analysability (compositionality) of idioms as depending on both the identification of idiomatic referents for idiom parts and the semantic relation between the literal meanings of these parts and the idiomatic meanings assigned to them, Gibbs and colleagues end up with a very heterogeneous group of abnormally decomposable idioms; these include idioms for which the elements in the idiom map onto the elements in the relation the idiom is used to denote either directly

(e.g. *spill the beans, pass the buck*), or indirectly, and by mediation of a further (metaphorical) relation (e.g. *hit the panic button, carry a torch*).

It is also worth noting that there is an important mismatch between the instructions subjects receive in the classification task and the assumptions underlying the experiments which use the resulting classification as data. Although the subjects are apparently instructed to classify idioms according to the semantic contribution made by constituent words, the experimenters take the syntactic and lexical flexibility of idioms to depend rather on whether the parts of the idiom can be assigned independent referents. In other words, while subjects may assume that an idiom is decomposable to the extent that the 'literal' meanings encoded by the constituent words can be seen as to making some meaningful contribution (literal or figurative) to the idiomatic meaning, experimenters take the idioms that subjects judged as decomposable to be those whose parts have independent idiomatic referents, and therefore carry some independent idiomatic meaning. This mismatch is particularly evident in the syntactic task: as for Gibbs and colleagues (as for Nunberg and colleagues) what matters for an idiom to be syntactically flexible is whether or not the constituent parts of the idiom has an identifiable idiomatic referent. Only parts with easily identifiable referents can be focused, modified or omitted.

But given these confusing instructions, why did the results show that subjects were consistent in their classifications? One possible reason is that the idioms were not presented in isolation or in a discourse context but were presented together with a paraphrase of their meaning (e.g. *bury the hatchet* -> to resolve a dispute; *play with fire* -> to experiment with danger; *rack one's brains* -> to search one's memory). The results of the classification task, and therefore judgements on the analysability of idioms may have been partly constrained by the presence of this paraphrase. The vagueness of the instructions, the presence of paraphrases and the unclear notion of analysability which the experimenters aimed to test might have affected the results in significant ways. For example, subjects classified the idiom *promise the moon* as abnormally decomposable, even when it is possible to see how *promise* is literally intended, how 'promising the moon' is an example of promising something impossible and how each word in the idiom can map onto an element in the denotation (e.g. *promise* –> promise, *the moon* –>

something unachievable). They classified idioms such as *spill the beans* or *pass the buck* as abnormally decomposable, even though it is easy to assign independent idiomatic referents to each part of the idiom. Also, they classified idioms such as *let off steam* (denoting the act of getting angry) and *button one's lips* (referring to the act of keeping quiet) as normally decomposable, while the equally metaphorical idiom *push the panic button* was judged as abnormally decomposable. Relatively transparent idioms such as *speak one's mind*, which roughly describes the act of saying what one really thinks, were classified as non-decomposable while relatively opaque idioms such as *play the market* were classified as normally decomposable (see Gibbs, Nayak and Cutting, 1989). None of these results is straightforwardly predictable on the account given.

Scholars working on idioms have often accepted experimental results such as those of Gibbs and colleagues above without questioning the inconsistencies noted here or re-examining the vague notion of analysability or compositionality on which these experiments are based. I believe there is an urgent need both to provide a theoretically well-developed notion of analysability and to revise and refine existing experimental work on the role of analysability in idiom use, processing and interpretation.

5. The Nature of Compositionality

The core assumption underlying modern research on the compositionality of idioms is given in (24):

(24) Idioms lie along a continuum of compositionality or analysability depending on the extent to which people (while processing an idiom) can see how the meanings of the individual parts contribute to the overall idiomatic meaning.

This assumption is often formulated vaguely enough to allow for at least two different interpretations, as in (25) and (26):

(25) An idiom is analysable to the extent that people can see how the individual parts of the idiom can be assigned independent referents and so can be seen as carrying

independent idiomatic meanings. It is this idiomatic meaning assigned to idiom parts that contributes to the overall idiom meaning. (Analysability as Decomposition).

(26) An idiom is analysable to the extent that people can see how the encoded 'literal' meanings of the constituent words (and the way these meanings combine) are relatively transparently related to the overall idiomatic meaning of the expression. It is the encoded 'literal' meanings of the idiom parts that contribute (literally or figuratively) to the overall idiomatic meaning. (Analysability as Word Meaning Contribution)

According to the first reading in (25), the degree of compositionality or analysability of an idiom such as *pass the buck* would depend on the extent to which people can map the elements in the syntactic structure of the idiom (V(NP)) to elements in the denotation of the idiom (the verb denotes the act of handing something on and the noun phrase denotes responsibility for a problem). This idiom can be judged as highly compositional in that the meaning of the idiom can be decomposed in such a way that each of the elements in the syntactic structure can be assigned a referent in the denotation, and hence an independent idiomatic meaning. According to the second reading in (26), the degree of compositionality or analysability of the same idiom *pass the buck* would depend not on whether each element in the idiom maps to an element in the denotation, but on the extent to which people can perceive a clear (literal or figurative) relation between the 'literal' meanings of the constituent words and the idiomatic meaning of the idiom as a whole. In this case, the idiom *pass the buck* could only be judged as partially analysable or compositional, because only one of the constituent words in the idiom (*pass*) can be seen as making a transparent contribution to the overall idiomatic meaning.

So, which of these two notions of analysability do scholars have in mind? I believe that an important problem in modern research on idioms is that they have generally failed to recognise (25) and (26) as distinct criteria. Instead, the term 'compositionality' or 'analysability' is generally used in a vague way, sometimes

suggesting the criteria in (25), sometimes suggesting the criterion in (26), and, often, a combination of the two (e.g. the experiments by Gibbs and colleagues above). I will argue that this is not just a matter of terminology but a serious problem with important theoretical implications.

5.1 Composition and Decomposition

The roots of the problem may lie in the original formulation of the compositional hypothesis. In attempting to explain the syntactic behaviour of idioms, Nunberg and colleagues proposed that idioms vary in the extent to which people can perceive them as decomposable. That is, in the extent to which people can see how the elements in the idiom (e.g. V and NP) refer to distinct components of the idiomatic referent (e.g. to the relation of exploiting, on the one hand, and to the connections exploited, on the other, for *pull strings*). According to this view, an idiom is compositional to the extent that it is decomposable. If people can assign independent idiomatic meanings to the parts of the idiom, they should be able to see how these meanings combine to yield the overall interpretation:

> "We claim that the pieces of an idiom typically have identifiable meanings which combine to produce the meaning of the whole. Of course, these meanings are not the literal meanings of the parts. Rather, idiomatic meanings are generally derived from literal meanings in conventionalised, but not entirely arbitrary ways." (Wasow, Sag and Nunberg, 1983: 109).

The notion of 'compositionality' proposed by these scholars is rather different from the one proposed in traditional accounts of idioms. In claiming that idioms are not compositional, traditional approaches simply pointed out that a standard compositional analysis of the idiom will result in a ('literal') meaning other than the one conveyed when the string is idiomatically used. The compositional view of idioms defended by Nunberg and colleagues claims that for a phrase to be noncompositional two conditions have to be met: first, its meaning cannot be derived by composing the 'literal' meanings of its parts, and second, this meaning, once known, cannot be divided into its constituent parts (Nunberg et al., 1994: 496). It is this latter notion of compositionality via decomposition that these scholars argue applies to idioms.

We can take this notion of compositionality as roughly synonymous with decomposition to be the one described in (25). The problem is that, although its output is ultimately a formal structure, assigning idiomatic meanings to idiom parts depends on a very general (pragmatic) process of making assumptions about why an idiom means what it does. Providing a rationale in terms of the assumptions and conventions that license the meaning of an idiom generally involves exploring the meanings of the constituent words, a process roughly like the one described in (26). Often, the easier it is for people to establish a transparent relation between the literal meaning of the words in the idiom and the idiomatic meanings they are used to convey, the easier it is to perceive the idiom as decomposable. In this way, people may perceive the idiom *pop the question* as decomposable partly because they can establish a quite transparent relation between the meanings *pop* and *utter* and *question* and *marriage proposal.* Similarly, they may perceive the idiom *spill the beans* as decomposable because of the transparent relation between the meanings 'spill' and 'reveal'. It may well have been this interaction between decomposition and word meaning contribution that has prevented modern scholars from seeing (25) and (26) as involving two separate processes and which has created so much confusion in the use of terms 'compositionality' and 'analysability'.

The following hypotheses seem thus to underlie the original compositional view as proposed in Nunberg (1978):

(27) People search for assumptions or beliefs which provide a rationale for why an idiom means what it does.

(28) To provide a rationale for idiom meaning, people see whether the constituent words can give access to some relevant background information, or can be seen as used with the same meaning as in other (non-idiomatic) contexts.

(29) The assumptions we form in (27) and (28), together with syntactic information about the constituent word, determine how people map the elements of the idiom onto elements in the denotation.

(30) Which mapping people construct between syntax and semantics, as in (29), has an important effect on which syntactic transformations they find acceptable.

Although the original compositional hypothesis is built around all these assumptions, Nunberg and colleagues seem to assume that the compositionality of an idiom essentially depend on the process in (29): that is, on the mapping between the syntactic structure of an idiom and the semantic representation assigned to that idiom. Evidence for this is that they assume the meanings assigned to idiom parts may be literal (e.g. the meaning assigned to *miss* in *miss the boat*), or metaphorical (e.g. the meaning assigned to *spill* in *spill the beans*) but may also be unrelated and assigned simply by elimination (e.g. the meaning assigned to *the beans* in *spill the beans* or to *buck* in *pass the buck*). Although it may well be true that the closer the literal meaning of the words is to the idiomatic meaning they are thought to convey (e.g. between 'pop' and 'utter'), the easier it is to perceive the idiom as decomposable, what makes an idiom compositional or decomposable is essentially the fact that each of the elements in the syntactic structure can be assigned an independent referent.

5.2 Analysability as Transparency

Some scholars define the degree of analysability of idioms not so much in terms of the mapping between elements in the idiom and elements in the denotation, but rather in the terms of transparency of the relation between word meaning and idiom meaning. This approach is typical of experimental research on the acquisition of idiom meanings:

> "Semantic analysability, therefore, defines the extent to which a speaker can make sense of the meaning of an idiom on the basis of the information conveyed by the meaning of the constituent words and by the idiom's semantic structure (be it metaphorical, metonymical, analogical etc.)" (Levorato and Cacciari, 1999: 53)

Research in child language acquisition has shown that the degree of analysability of idioms plays an important role in how children acquire these expressions (Cacciari and Levorato, 1998; Gibbs, 1987; 1991; Levorato and Cacciari, 1992, 1995, 1999; Nippold and Rudzinski, 1993). In one experiment, based on the notion of analysability above, Levorato and Cacciari (1999) presented children, aged seven to nine, with a context (e.g. about a child on his first day at school) which ended in a familiar idiom, followed by a

question about the meaning of the idiom (e.g. 'what does it mean that the character... *broke the ice*?). The children were given a multiple choice of answers to this question. One answer was based on the literal meaning of the idiom (e.g. he broke a piece of ice), another was based on its idiomatic meaning (e.g. he made friends with his class mates) and a third was an associate answer.[7] (e.g. he told his mother everything). In a second experiment, in order to distinguish the role of context from the role of analysability, they repeated the experiment presenting the idioms in isolation.

The results showed that analysable idioms were understood better than unanalysable idioms in all conditions. Older children chose 91.5% of correct idiomatic answers for analysable idioms presented in context and 91.8% of analysable idioms presented out of context. The performance of older children was poor for unanalysable idioms, for which they gave 74.7% of correct responses in the context condition and 62.2% in the out-of-context condition. Seven year olds also performed better with analysable than with unanalysable idioms. Providing 74.7% of correct idiomatic responses for analysable idioms in context and 62.2% for analysable idioms out of context. Unanalysable idioms were again harder to understand for these younger children, with 49.6% of correct responses in the context condition and a mere 42% in the out-of-context condition.

Taken together these findings suggest that, contrary to standard assumptions, children do not only rely on contextual clues to understand the meaning of an idiom, but also use (internal) semantic cues embedded in the idiom. Differences in the results for analysable and unanalysable idioms in the out-of-context condition suggest in fact that context and semantic analysability play important but independent roles in acquisition. The clearer the context and the more transparent the relation between the literal meaning of the idiom and its idiomatic meaning, the easier it is for children to understand these expressions. One interesting result to emerge from the findings above is that, although both context and word meaning can potentially be explored in interpretation, they only seem to be explored when needed. The younger the child, the more support from context

[7] An associate answer was taken to express a meaning which was plausible in the context and semantically appropriate but different from idiomatic meaning.

and internal semantics is needed.[8] Older children, however seem to rely on contextual cues only when the idiom being understood is not analysable.

A final interesting finding was that when younger children were unable to provide an idiomatic answer for analysable idioms presented out of context, they preferred semantically associative answers (21%) over literal ones (6.8%). This pattern was also found in older children when presented with unanalysable idioms out of context. In this condition, when unable to provide an idiomatic answer, they generally preferred semantically associated responses (25.2%) over literal interpretations (12.6%). This suggests that children can generally grasp roughly the figurative meaning of the expression even if they do not grasp their exact meaning.

5.3 Decomposition and Transparency

The existence of so-called abnormally decomposable idioms seems by itself to cry out for a distinction between decomposition and transparency. For most, if not all, expressions falling into this category, it is generally possible to establish a non-arbitrary relation between the meanings of the constituent words (or meaning resulting from their combination) and the meaning of the idiom as a whole, even if one cannot see how each of the elements in the idiom's syntactic structure maps onto an element in the idiom's denotation. Examples of this include the expression *bury the hatchet,* which denotes an action that ends a dispute, *hit the nail on the head,* which denotes a successful performance, *turn over a new leaf,* which roughly means to forget about the past and start again, etc. The idiom *bury the hatchet* for instance, cannot be perceived as decomposable because the constituent elements do not map onto elements in the idiom's denotation (e.g. we cannot say that the hatchet refers to the dispute and the burying of it to the end of the dispute). The meaning of the idiom is, however, not arbitrary: our

[8] In studying the acquisition of figurative competence, Levorato has proposed the Global Elaboration Model (e.g. Levorato, 1993). This model proposes that in acquiring an idiom children explore component word meanings in order to make sense of it. As they grow older, they gradually learn to integrate these meanings with the meaning of the expression as a whole and with the wider context. For younger children, processing occurs in a rather local manner in which every word is considered in sequence; as they grow older, they learn to combine chunks of discourse until they acquire the ability to take the more global meaning for the wider text. According to the Global Elaboration Model, the development of a child's figurative competence is parallel to the development of a child' s linguistic (and metalinguistic) abilities, such as the ability to build up the meaning of a text and to go beyond linguistic meaning to the meaning the speaker intended to convey.

knowledge that hatchets are weapons of war, or that the burying of a hatchet can be seen as the end of a fight, may help us to perceive the idiom as partly transparent. A relatively transparent motivation for the overall idiom meaning can be inferred even if we are not familiar with the ancient tradition of burying a hatchet after a battle as a symbol of peace. It is this motivated relation between the literal meaning of the string and its idiomatic meaning that allows for some lexical transformation such as that discussed above (e.g. in the appropriate context, one may *bury/sink/hide/unbury/grasp the hatchet,* etc.).

Another important reason to keep decomposition and word meaning contribution apart is that people always seem to attempt to carry over some assumptions associated with the ('literal' meanings of the) words in the idiomatic meaning even if this idiom cannot decompose. We have seen, for instance, some scholars have proposed that the assumption that the verb *kick* in *kick the bucket* denotes a sudden, abrupt (or one-off) action may constrain the way people use this expression even if the meaning of the whole is rather opaque (Cacciari and Glucksberg, 1991; Gibbs, 1994a; Hamblin and Gibbs, 1999; Nunberg, 1978). We have also seen that a single word (e.g. *hit)* may make different contributions to different idioms, affecting how the idioms are used and understood thereafter. We can argue then that certain contribution of word meaning to idiom meaning (however minimal or opaque) happens regardless of whether or not the idiom is perceived as decomposable.

We may conclude from the evidence presented here that the process described in (25), which I shall refer to as 'mapping' or 'decomposition', is ultimately a formal process which establishes a relation between two objects: (the elements of) a syntactic structure (e.g. V and NP) and (a representation of) a denotation. Sometimes, a relation can be established between each of the elements in the idiom and each of the elements in the denotation. This is the case of normally decomposable idiom such as *spill the beans* or *pull strings*. On other occasions, a relation can only be established between the idiom as a whole and the denotation. This is the case not only of nondecomposable idioms such a *to kick the bucket* or *to chew the fat* but also, I have argued, of many of the so-called abnormally decomposable idioms such as *to push the panic button* or *to bury the hatchet*. We may conclude also that the process in (26), which I shall refer to as

'word meaning contribution' or 'transparency', is essentially a pragmatic process whereby people infer a relation between two types of meaning: the literal meanings of the words in the idiom (or the literal meaning of the expression as a whole) and the figurative meaning of the expression as a whole. For some idioms, one of the words in the idiom makes a more transparent contribution to overall idiom meaning than the rest, as shown in (31). The contribution of word meaning to idiom meaning may be relatively literal, as in (32), or metaphorical, as in (33). It may be restricted to one word in the idiom, as in (32)-(33), or to the compositional analysis of the whole string, which can be taken figuratively, e.g. metaphorically, as in (34), or hyperbolically, as in (35). Finally, in some cases, no transparent contribution is provided, as in (36):

(31) *Spill the beans, break the ice, bark up the wrong tree, fed up to the back teeth.*
(32) *Promise the moon, cost an arm and a leg, start from scratch, miss the boat.*
(33) *Pull strings, take under one's wing, pluck up courage, blow one's stack.*
(34) *Back to square one, sit on the fence, stab someone in the back, hit the target.*
(35) *Cry one's eyes out, eat one's heart out, not sleep a wink, cost the earth.*
(36) *Kick the bucket, chew the fat, break a leg, shoot the breeze.*

We may assume (partly in line with Nunberg and colleagues (Nunberg, 1978; Nunberg et al., 1994) that the decomposition process generally occurs as a by-product of a much more general pragmatic process whereby people make assumptions about idiom meaning, word meaning and the way they interact and integrate with each other. Whereas work on the syntax of idioms should be mostly concerned with the formal mapping (i.e. decomposition) resulting from this wider pragmatic process, work on the pragmatics of idioms should take the opposite direction. The crucial aim of a pragmatic approach to idioms, such as the one I will present in the next chapter, is to account for how hearers recognise the assumptions the speaker may have intended to convey in using an idiom. The crucial aspect of idioms that set them apart from lexical items is that some of these assumptions may be recoverable from the concepts encoded by the individual words in the string. It is analysability as understood in these terms that ought

be the focus of theoretical and experimental research on the (on-line) comprehension of idioms. Much of the next chapter is dedicated to bringing light into this issue.

6. Analysability and the Processing and Representation of Idioms

Notice, now, that which notion of analysability we take to apply to idioms may suggest, directly or indirectly, a particular manner of processing or representation. If we assume, as in (25) above, that the parts of an idiom contribute to the overall idiomatic meaning by virtue of the idiomatic meanings assigned to them via decomposition, this may suggest that processing this idiom involves selecting and combining these idiomatic meanings to derive the figurative interpretation. So deriving the meaning of the idiom *spill the beans,* for instance, would involve selecting and combining the meanings REVEAL and HIDDEN INFORMATION associated with the parts of the idiom. This view of processing in turn suggests a certain manner of representation. It suggests that the idiom must be stored in such a way that a different semantic representation is assigned to each constituent word. Sometimes, it seems that this is what Nunberg, Gibbs and colleagues have in mind:

> "We claim that the pieces of an idiom typically have identifiable meanings which combine to produce the meaning of the whole" (Wasow, Sag and Nunberg, 1983: 109)

> "Because the individual components in decomposable idioms contribute systematically to the figurative meaning of these phrases, people may process idioms in a compositional manner where the semantic representation of each component are accessed and combined according to the syntactical rules of the language" (Gibbs, 1993: 64).

Alternatively, if we assume, as in (26) above that the parts of the idiom contribute to the overall idiomatic meaning by virtue of their literal meanings and not the idiomatic meanings assigned to them, the compositional process would take place as normal, selecting and combining the 'literal' meanings encoded by the constituent words until the meaning associated with the string as a whole is retrieved from memory. Here, there is no need to assume that the parts of the idiom are assigned independent idiomatic

representations in memory. Some experimental research on idiom processing seems to point in this direction.

6.1 The Activation of Idiomatic Meaning

Although the idea that idioms lie on a continuum of compositionality has been supported by experimental results showing that people have strong intuitions about the internal composition of idioms, there is still one important finding which supports the traditional view of idioms as lexical items. This is the finding that idiom strings are generally understood as fast as, or faster than, literal phrases of the same length and complexity (Gibbs, 1980; McGlone, Glucksberg and Cacciari, 1994; Swinney and Cutler, 1979). Two noncompositional models of idiom processing have been proposed to account for the ease of idiom processing. I refer to them as the Simultaneous Processing model (Swinney and Cutler, 1979) and the Figurative First model (Gibbs, 1980). Swinney and Cutler (1979) suggest that two processes take place while an idiom is being processed: a computational process which derives the literal (compositional) meaning of the string and a retrieval process which retrieves the idiomatic meaning of the idiom from memory. These processes take place in parallel as the early parts of the string are heard and continue until one of the meanings is accepted as the intended one. The authors argue that, since retrieval and acceptance of the idiomatic meaning generally takes place before a complete literal interpretation has been derived, idioms are often understood faster than literal phrases of the same length and complexity. Gibbs (1980) proposes an alternative explanation for the speed of idiom processing. In his view, the idiomatic meaning of an idiom can be retrieved directly from memory, completely bypassing the literal meaning.

In order to test the plausibility of these two models and of the view that idioms may be stored as lexical units, Cacciari and Tabossi (1988) carried out a series of cross-modal priming experiments. In one of these, they asked subjects to listen to a list of neutral sentences ending in an expression that was either literally or idiomatically intended (e.g. 'after the excellent performance, the tennis player was *in seventh position/ heaven*'). Subjects had to perform a lexical decision task on a word presented at the end of the string. The target word might be related to the literal meaning of the last word in

the idiom (e.g. *saint*), to the figurative meaning of the idiom as a whole (e.g. *happy*) or unrelated (e.g. *umbrella*). The results suggested that the literal meaning of the last word in an idiom is activated immediately, and remains activated after 300ms. The idiomatic meaning, however, only shows signs of activation 300ms after the sentence had been heard.

The authors took these results to challenge both the Simultaneous Processing model and the Figurative First processing model. The findings clearly show that the literal meanings of the words in the idiom cannot be bypassed, contrary to Gibbs's claim. The automaticity of linguistic processing makes it impossible for people not to activate the literal meanings of the words they hear (this has been demonstrated repeatedly for ambiguous words, for which, even in linguistic contexts that are strongly biased in favour of a single sense, all senses are initially activated (see Swinney, 1979, etc.)). Furthermore, contrary to the prediction of both the Simultaneous Processing and the Figurative First models, the idiomatic meaning of the idiom was not activated immediately. In order to account for these results, Cacciari and Tabossi propose to modify the Simultaneous Processing model. While Swinney and Cutler (1979) assume the competition between the literal and figurative meanings of the idiom starts as soon as the first word is heard, Cacciari and Tabossi argue that this competition may start much later, namely at the point where the idiom meaning becomes available to the hearer. The point at which an idiom is recognised as a unit in the absence of contextual bias is referred to by these scholars as the 'idiom key'.

In order to determine this point, Tabossi and Zardon (1993) carried out a number of further experiments. In one of these, subjects were asked to listen to a sentence and perform a lexical decision task on a word with a meaning related to the overall idiomatic meaning of the string. Target words were presented after the verb (control position), after the first content word of the idiom or after the second content word. So in listening to a sentence such as 'after the tournament, the tennis player *was in seventh heaven*, target words might appear after the verb 'be', after the word 'seventh' or after the word 'heaven'. Two types of idioms were studied: idioms whose key occurred early in the string, such as *setting his mind at rest'* ; where the key comes after the first content word (target word e.g. 'resigned') and idioms whose key occurred later in the string, such as

hit the nail on the head (target word e.g. 'accurate'). The results showed that for idioms with an early recognition key, people were significantly faster at recognising the target word both after the first content word (638ms) and after the second content word (626ms) than after the verb (672ms). For idioms with a later recognition key, however, reaction times were roughly the same whether the target word appeared after the verb (637ms) or after the first content word (631ms), both reaction times were significantly slower than the reaction times obtained when the target word was presented after the second content word (605ms). These results confirm the finding that overall idiom meaning is not immediately activated as the first word in the idiom is heard, but only at a later point. They also show that in the absence of a biasing context, the point of idiom recognition coincides with the point of idiom uniqueness; that is, the point at which the idiom can be (uniquely) recognised.

6.1.1 The Configuration Hypothesis

Cacciari and Tabossi (see also Glucksberg, 2001) took their findings that the overall idiom meaning is not immediately available to be inconsistent with Marslen-Wilson's cohort model of lexical processing (Marslen-Wilson, 1987). This model was based on the assumption that as soon as the first syllable of a word is heard, the entire cohort of words in the hearer's lexicon which begin with that syllable are immediately activated. So on hearing /pen/, words such as 'pen', 'pencil', 'penny', etc. are activated in the hearer's mind. As further information is processed (e.g. /penta .../), the number of candidates is reduced until reaching a point is reached where only one possible candidate is left (e.g. 'pentagon'). Since idiom strings, unlike lexical items, do not show signs of activation as soon as the first syllable or even the first word is processed, Cacciari and Tabossi concluded that idioms may not, after all, be stored or represented in memory as lexical items. Instead, they proposed that idioms may be stored as configurations.

According to the Configuration Hypothesis (Cacciari and Glucksberg, 1991; Cacciari and Tabossi, 1988; Glucksberg, 2001), idioms are represented and processed no differently from other memorised strings, such as songs lyrics, titles, or riddles. A characteristic of memorised strings is that they are not generally recognised as soon as the first word of the sequence is heard, but only after the hearer has processed enough

input to recognise the string as a familiar configuration. Thus, on hearing, for instance, 'oh beautiful for spacious skies...', North Americans are said to be normally capable of recognising the sequence as a token of the American song 'America the beautiful'. This recognition activates information associated in memory with this particular song, which is used to derive an interpretation. So if the sequence is uttered in a context in which an American has done something heroic, the hearer would understand the speaker to be praising the bravery and glory of the American spirit. If the sequence is uttered in a situation in which the hearer and speaker are walking through narrow dirty streets in New York, the speaker would be understood as using the song sarcastically. The Configuration Hypothesis claims that idioms are processed in just this way. They are processed literally, with each of their individual words being activated until the hearer arrives at the 'idiom key', which allows him to recognise the string as a stored configuration. At this point, the hearer has access both to the meanings of the words that have been activated in the process and the meaning of the idiom as a whole. Since idioms are composed of words that have an entry in the lexicon, these meanings have to be activated in processing the string: whether they are actually used to derive a literal interpretation of the string is another matter (Blasko and Connine, 1993; Cacciari and Glucksberg, 1991; Cacciari and Tabossi, 1988; Flores D'Arcais, 1993; Tabossi and Zardon, 1993).

6.2 Activation and Integration in Processing

The findings that the meaning of an idiom is not activated as soon as the first word is heard may be taken to suggest that idioms are not stored as lexical items: however, it does not tell us much about how hearers actually derive the idiomatic interpretation. These experimental results are consistent with the standard assumption that idioms have a single stipulated meaning in memory, which is different from the literal meanings of its parts. Even if idioms are stored as configurations, the overall meaning associated with a configuration need not be semantically motivated (as in opaque idioms, for example). All the configuration model can tell us for instance about the comprehension of an opaque idiom such as *kick the bucket* or a relatively transparent idioms such as *to be on the seventh heaven* is that, in the absence of biasing context, the idiomatic meaning of

these expressions is activated after the last word of the idiom is heard. It does not say how the utterance is interpreted once this meaning is activated or whether the degree of transparency (analysability) of the idiom has an effect on the derivation of the intended interpretation.

The issue of how the meaning of the words in idiomatic expression affects interpretations, we have seen, has figured prominently in modern research on idioms. It would be desirable so that on-line experiments on idiom processing would reflect this effect. One would expect, for instance, to find differences in processing between transparent and opaque idioms. For the former but not for the latter, the assumptions which are considered in processing the literal meanings of the words may be consistent with the idiomatic meaning and so would not need to be rejected. There is a need in the literature to provide experimental research that shows how the transparency of idioms affects the on-line interpretation of idioms.

Pursuing similar lines of thought, Titone and Connine (1994) carried out eye-tracking study which measured people's eye movements when reading sentences containing idioms. They presented people with sentences which were biased towards the literal or the idiomatic meaning of the expressions. The biasing context could appear either proceeding or following the idiom (e.g. 'after being ill for months, she finally *kicked the bucket*' or 'she finally *kicked the bucket*, after being ill for months') and the idioms varied in their degree of transparency. The results showed that both idiomatic and literal meanings of an idiom were activated immediately yet the degree of transparency of the idiom had an effect on the duration of eye-fixation in the ambiguous regions. Results showed that the duration of eye-fixation in ambiguous regions was longer for opaque idioms than for transparent idioms. The authors took these findings to suggest that subjects find it easier to disambiguate the expressions and assign the intended meaning in context when the idiom is transparent because both the idiomatic and literal meanings are related in meaning. Subjects were taken to spend longer disambiguating the idiom and providing an adequate interpretation to the string when the idiom was opaque because the literal and idiomatic meanings of opaque idioms are considerable distinct. We may take these findings to suggest that in reading an utterance containing an idiom, subjects aim to integrate assumptions from the 'literal' encoded

concepts of the constituent words of the idiom with the overall interpretation, may this be literal or idiomatic. I will look at this idea in the next chapter.

7. Conclusion

In this chapter, I have presented and discussed early and modern approaches to idioms, and analysed the notion(s) of 'compositionality' and 'analysability' used in these approaches. I have argued for a move away from approaches based on a single notion of analysability, towards a model designed to capture the independence and interaction of decomposition and transparency. From a pragmatic perspective at least, I shall argue that the crucial issue is not so much whether a perfect one-to-one relation can be established between elements in the syntactic structure of an idiom and elements in the denotation of the idiom, but whether people perceive the meaning of the idiom as not entirely arbitrary, and as at least partly motivated by the 'literal' meanings of its parts. In the next chapter, I present a relevance-theoretic analysis of idioms designed to shed light on the mechanisms which guide and constrain the interaction between word meaning and idiom meaning in interpretation.

Chapter 6

Idioms, Transparency and Pragmatic Inference

> "Conventional phrases are a sort of fireworks, easily let off, and liable to take a great variety of shapes and colours not all suggested by their original form"
> (Charles Dickens. *David Copperfield.* Chapter 41)

1. Introduction

The traditional approach to idioms as lexical items has two main implications. On the one hand, idioms are not expected to behave linguistically as phrases but as long words, and so they are not expected to allow internal transformation. On the other hand, as a lexical item, the meaning of an idiom is not seen as derived from the meanings of its parts, but as an arbitrary form-meaning pair which is stored as such in memory. In Chapter Five I showed how modern approaches to idioms have attempted to combat the view of idioms as lexical items, proposing instead that idioms are at least partly compositional or analysable. In challenging the view of idioms as long words, these approaches have brought into question, directly or indirectly, the implications about arbitrarinesss and lack of transformational capacity mentioned above. I have argued in Chapter Five that this double reaction against traditional models has introduced some confusion into the literature, with the compositionality or analysability of idioms sometimes linked to the ability of speakers or hearers to assign meanings to the parts of the idiom (whether this meaning is motivated or not), and sometimes linked to their ability to infer a relation between the literal meaning of the words in the idiom (and the way they compose) and the idiomatic meaning of the overall expression. This chapter is mostly concerned with this last ability. That is, it is concerned with what I refer to as the degree of 'transparency' of idioms.

Traditional scholars are right in saying that the meaning of an idiom is not the result of a compositional analysis of the literal meanings of its parts. However, the fact that the meanings of idioms are not compositionally derived does not necessarily entail that they must be arbitrarily stipulated in memory. It may still be possible to infer some

meaningful relation between the literal and idiomatic meanings, and this relation may be exploited in using and interpreting idioms in everyday conversation. Although, explicitly or implicitly, current psycholinguistic models agree that the comprehension of idioms depends to a certain extent on the existence of a non-arbitrary relation between the meanings of the constituent words and the overall idiomatic meaning, they do not examine in any detail the pragmatic processes that enable these meanings to interact in on-line comprehension.[1] Complementing experimental research such as that presented in the previous chapter, I will provide a relevance-theoretic account of idioms which aims to fill this gap.

The pragmatic approach to idioms I propose here is grounded on two main assumptions. On the one hand, I accept the relevance-theoretic idea that there is not a clear-cut distinction between literal and loose interpretations, but a continuum of cases. On the other hand, I suggest that most idioms lie along that continuum of looseness and as a result they vary in the extent to which the overall idiomatic meaning can be inferred from the meanings of the parts and their manner of combination (i.e. in their degree of transparency). The relevance-driven comprehension procedures that apply to utterances containing idioms will constrain the direction of the inference process and the depth to which the encoded 'literal' meaning is processed. Repeated processing of familiar idioms may result, I propose, in the development of a pragmatic routine which directs the hearer along a certain inferential route, and towards shallow processing of the encoded concepts.

2. Idioms, Metaphors and Unfamiliar Words

Much of the literature on metaphor has focused on the interpretation of nominal metaphors of the sort *X is Y*. Much ordinary speech, however, involves the comprehension of verbs, compounds and even whole phrases which are metaphorically intended, such as those in (1)-(6):

[1] Few pragmatic approaches to idioms have indeed been proposed in the literature. None of which, to my view, are fully theoretically and cognitively appealing (see Coulmas, 1981b; Eizaga Rebollar, 2002; Strässler, 1982).

(1) My father is *glued* to the computer for hours.

(2) His conscience has *rotted* after so many years in power.

(3) I know I cannot be *first violin* but I deserve a better job in the company than this.

(4) I'll do my best to convince them of the advantages of the new product but I will not *clean anybody's shoes.*

(5) Since I work at university, I feel I am *swimming with sharks.*

(6) Supervisions are mentally stimulating. I like the way my supervisor *squeezes my brain* to make me solve the problems I encounter.

Comprehension of the utterances in (1)-(6) would proceed in just the same way as the comprehension of any utterance. Following a path of least effort, the hearer takes the encoded concepts as a starting point for deriving the speaker's meaning. A consequence of taking the path of least effort is that he often finds his expectations of relevance satisfied after considering only a subset of the encyclopaedic assumptions associated with the encoded concepts. This relatively shallow processing generally results in the construction of an ad hoc concept on-line. In (1), for instance, the hearer might construct an ad hoc concept GLUED*, which is broader than the encoded concept in that it denotes situations in which someone is very close to and inseparable from something, even if not actually glued to it. In (2), he might construct an ad hoc concept ROTTED*, which is broader than the encoded concept in that it denotes certain states of moral corruption or degeneration as well as those involving flesh, vegetation, etc. In (3), the encoded 'literal' meaning of the compound *first violin* may be broadened to create an ad hoc concept [FIRST VIOLIN]*[2], which denotes people who have an important leadership role and who enjoy the praise and success proper to that position. Finally, the hearer of (4), (5) and (6) might understand the phrases *to clean someone's shoes*, *to swim with sharks* and *to squeeze someone's brain* as conveying the phrasal ad hoc concepts [TO CLEAN SOMEONE'S SHOES]*, [TO SWIM WITH SHARKS]* and [TO SQUEEZE SOMEONE'S BRAIN]*. These concepts may denote not (just) the situations described by the expressions taken literally but something more general: in (4), a situation in which people are degraded, humiliated, doing something unpleasant and below their social

status; in (5), a state of affairs in which people are in an environment where they feel uncomfortable, unsafe, surrounded by entities which they do not trust and which they fear will harm them; and in (6), a state of affairs in which people are forced to think very hard.

It could be argued that there are at least two ways in which the hearer of such examples may enrich the encoded sentence meaning to warrant the expected cognitive effects. On the one hand, he may adjust the individual concepts encoded by the words in the utterance, creating unlexicalised concepts which are broad enough to warrant the expected cognitive effects (e.g. SWIM*, SHARK*, GLUED*). On the other hand, he may understand a combination of words or a whole phrase as expressing a concept whose denotation is broad enough to warrant these effects. I believe that, providing that the resulting enriched proposition makes the utterance relevant as expected, it does not matter which way the inferential fine-tuning goes. In (1)-(3), for instance, it may not matter whether the hearer constructs the ad hoc concepts GLUED*, ROTTED*, and [FIRST VIOLIN]* or the phrasal ad hoc concepts [TO BE GLUED TO X]*, [TO HAVE ROTTED]* and [TO BE FIRST VIOLIN]*. Different people may enrich the explicit content in different ways, constructing different concepts, all of which may yield roughly the same implications for roughly the same processing effort, and hence make roughly the same contribution to relevance.

On other occasions, however, as in (4)-(6), in order to arrive at the intended implications, the hearer would need to add to the context both assumptions retrieved or derived from the encyclopaedic entry of the encoded concepts (e.g. the assumption that sharks are dangerous creatures) and assumptions derived from the compositional meaning of the phrase (e.g. the assumption that swimming with sharks is a dangerous activity). Since the pragmatic adjustment of the encoded concepts alone may not warrant the intended implications (e.g. the implications that the speaker feels unsafe, uncomfortable in his job, etc.), pragmatic adjustment at word level may need to be complemented with pragmatic adjustment at phrase level. As a result, the hearer may construct a phrasal concept (e.g. [TO SWIM WITH SHARKS]*) whose denotation is broader than that of the compositional meaning of the phrase. This fine-tuning process

[2] Brackets indicate that the whole string is metaphorically interpreted, and not just the last word.

operating at phrase level can be seen as taking place in the interpretation of (4)-(6). Thus, comprehension of these metaphorical uses may involve the construction of a phrasal ad hoc concept which would be understood as a constituent of the proposition expressed, as in (7)-(9). Moreover, such constituents would contribute to the truth-conditional content of the utterance, as shown in (10)-(11) where the truth or falsity of the proposition expressed in each case depends on whether it is the literal or ad hoc phrasal concept that is intended:

(7) THE SPEAKER WILL NOT [CLEAN ANYBODY'SHOES]*

(8) THE SPEAKER IS [SWIMMING WITH SHARKS]*

(9) THE SPEAKER'S SUPERVISOR [SQUEEZES HER BRAIN]*

(10) If your supervisor squeezes your brain, she'll go to jail

a. SQUEEZE X'S BRAIN – true

b. [SQUEEZE X'S BRAIN]* – false

(11) If your supervisor squeezes your brain, you'll produce a good thesis

a. SQUEEZE X'S BRAIN – false

b. [SQUEEZE X'S BRAIN]* – true

In the examples discussed so far, both here and in previous chapters, I have argued that the formation of the new concepts is a by-product of the pragmatic fine-tuning of the encoded concepts during the mutual adjustment process. This is not to say that all the information used in the construction of an ad hoc concept is retrieved ready-made from the encyclopaedic entry of the encoded concept. Instead, as I have tried to show, these assumptions are simply the starting point for an inferential process in which they are combined with other accessible hypotheses about explicit content, context and cognitive effects to yield implications which may themselves be used as premises for further inference. The resulting ad hoc concepts may warrant a diverse range of implications derived from different combinations of these assumptions (as we have seen with examples such as 'that surgeon is a butcher'). Although the same comprehension procedure and mutual adjustment process are at work in interpreting every utterance

(whether literally, approximately or metaphorically intended), the hearer is not always familiar with the concepts encoded by the speaker's words and may therefore have no direct access to the encyclopaedic information normally associated with these concepts, and no direct knowledge of their denotations. On these occasions, arriving at the intended combination of explicit content, context and cognitive effects may involve the construction of an ad hoc concept. However, this concept cannot be formed by pragmatic fine-tuning of the concept encoded by the unfamiliar word, but must be accessed indirectly, using hypotheses about the speaker's meaning derived from other sources during the interpretation process. Consider (12) and (13), for instance:

(12) The plane could not take off and we had to return to the airport. They said it was due to technical problems: one of the *turbines* was damaged and needed to be replaced.

(13) My knee is really bad. The results from the *arthroscopy* show there is hardly any cartilage left.

In engaging in ordinary conversation, listening to the news, speeches, etc. we often encounter words which we do not understand but whose meaning we can work out from other clues. Considering hypotheses about explicit content, context and cognitive effects, the hearer of (12) and (13) may be able to assign some tentative content to the concept conveyed by the unknown word. The new concept may contribute to relevance in the expected way, by warranting the derivation of appropriate cognitive effects. In (12), for instance, the hearer may construct the hypothesis that a turbine is a part of a plane. If the resulting interpretation satisfies his expectations of relevance, he will be entitled to assume that his hypothesis was correct. The word *turbine* may be (incompletely) understood as expressing a certain concept, say TURBINE*, which denotes a certain kind of plane part. In the same way, the hearer of (13) may be able to construct the hypothesis that an arthroscopy is a certain kind of medical procedure

which may be used on knees, and construct a partly understood concept, say ARTHROSCOPY*.[3]

I want to propose that unfamiliar idioms lie somewhere in between novel metaphorical expressions (such as those in (1)-(6)) and unknown lexical items (such as those in (12)-(13)), varying in the extent to which their meanings can be fully inferable, partly inferable or not inferable from the encoded 'literal' meaning of the string on a first encounter.[4] A feature of idioms, even the most opaque ones, is that, unlike lexical items, they are generally composed of words which are familiar to the hearer. The degree of transparency of an idiom would be determined by the extent to which some of the encyclopaedic information made accessible by these words can actually help the hearer to derive an appropriate overall interpretation. At one end of the transparency spectrum, we find very opaque idioms, for which none of the encyclopaedic assumptions made accessible by the words in the string (separately or in combination) helps with the identification of the speaker's meaning, as in (14)-(15):

(14) *Jason:* The old man did not want to sell his house so the council waited until he *kicked the bucket* to get hold of everything he owned.

(15) *Tom:* Where is Vanessa?

Carol: I don't know, she is probably *chewing the fat* with her friends somewhere.

Using encyclopaedic assumptions about buckets or the kicking of buckets would not help the hearer of (14) to derive implications about dying, and using encyclopaedic assumptions about fat or the chewing of fat would not allow the hearer of (15) to derive

[3] Sperber and Wilson (see Sperber, 1997; Sperber and Wilson, 1998) have developed a more detailed account based on the distinction between intuitive concepts and attributive or reflective concepts. Many concepts encoded by unfamiliar words are initially attributively (or reflectively) understood (e.g. "arthroscopy, whatever people mean by that"). After enough exposure, they may get an intuitive grasp of the concepts and no longer need to attribute their content to anyone else (e.g. the concepts encoded by the words 'bread', 'chair', etc.). Some concepts always remain attributive, at least for the majority of language users (e.g. the concepts encoded by the words 'witch', 'neuron', 'planet', etc.). Although I agree with this account, my aim here is simply to show how hypotheses about explicit content, context and cognitive effects constructed during the mutual parallel adjustment process play an important role in assigning meaning to unfamiliar words on hearing them for the first time. The resulting concept is indeed unlikely to be a full-fledged concept but merely attributively or reflectively understood.

implications about talking. Although the first time these opaque idioms are encountered, the hearer may explore these assumptions in an attempt to assign a plausible interpretation to the string, he would need to reject them for not helping to make the utterance relevant in the expected way.[5] On some occasions, the hearer may be able to assign some tentative content to the idiom along the lines suggested above for the unknown words in (12)-(13). On some other occasions, explicit learning would be needed.

We have seen how, contrary to standard assumptions, modern research on idioms has shown that opaque idioms are rare and that most idiomatic expressions enjoy at least some degree of transparency, as with the examples in (16):

(16) *to hold all the aces, to speak one's mind, to lay one's cards on the table, to stab s.o. in the back, to miss the boat, to pull strings, to be on cloud nine, to change one's mind, to have one's feet on the ground, to turn over a new leaf, to be the icing on the cake, to keep s.o. at arm's length, to be the last straw (that broke the camel's back), to cost an arm and a leg, to go over the line, to fill the bill, to add fuel to the fire, to get out of the frying pan into the fire, to be in the same boat, to lose a train of thought, to slip one's mind,* etc.

Accessing some encyclopaedic assumptions associated with the concepts encoded by the words in these strings, or derivable from their combination, generally helps a hearer unfamiliar with the expressions to infer an appropriate idiomatic interpretation. Consider (17) and (18):

(17) There is no way I will get the job. Peter, however, *holds all the aces.*

(18) John is a very disloyal person I would not be surprised if he *stabs you in the back.*

---

[4] I believe this claim holds whether the hearer is unfamiliar or familiar with the expression. Although this section is mostly concerned on how idioms may be acquired, as the chapter progresses, I will show that familiarity with an idiom does not necessarily make comprehension less inferential.

[5] As suggested in the previous chapter, even in these very opaque cases, the hearer may still assume that some of the encyclopaedic assumptions associated with the encoded concepts contribute to the overall idiomatic interpretation (e.g. the assumption that kicking is an abrupt act, for the idiom *kick the bucket*).

The idioms in (17)-(18) are relatively transparent, in that even when a hearer is unfamiliar with their meaning, he can easily infer it given the encoded 'literal' meaning. Comprehension would proceed along the same lines as for the phrasal metaphors in (4)-(6) above. The hearer, following a path of least effort, would consider assumptions made highly accessible by the encoded concepts (e.g. the assumption that the ace is the highest and therefore best card; that those who hold all the aces are likely to win; that stabbing in the back is a cowardly act of betrayal, etc.), and fine-tune the encoded 'literal' meaning of the phrase until he arrives at an interpretation that satisfies his expectations of relevance. The output of this pragmatic adjustment process would generally be a phrasal ad hoc concept broader in its denotation than the compositional meaning of the phrase (e.g. [TO HOLD ALL THE ACES]* would denote situations in which someone has every chance of winning something; the concept [TO STAB IN THE BACK]* would denote acts of betrayal). It is this ad hoc concept that the hearer would take to be part of the speaker's meaning.

At some point on this spectrum of transparency are what we can call 'partially transparent' idioms. These are idioms for which encyclopaedic assumptions associated with some, but not all, of the encoded 'literal' concepts may help to derive an appropriate idiomatic interpretation. This may happen because the hearer is not familiar with one of the words in the string, as with the Spanish idioms *pensar en* (think of) *las musarañas*, meaning roughly 'to be absent minded' or *meterse* (to get into) *en un embolao,* meaning roughly 'to get into trouble'. Many people who use the former idiom do not know that musarañas are a type of mice (shrews). To make sense of the expression, they simply take it to refer to something not particularly significant. The rationale for this is that thinking about something insignificant would prevent the hearer from paying attention to what really matters causing him to be absent-minded.[6] Notice, too, that even when the hearer is familiar with the words in the idiom, the information associated in memory to the concept it encodes may not help him to make sense of the idiomatic meaning. This is the case, for instance, with the phrase *the beans* in the

[6] Because of the phonological relation between the words *musarañas* and *arañas* (spider), I originally took the word to refer to some spider-like animals. I made sense of the idiom by assuming that thinking about these small, arguably meaningless, creatures would stop someone from concentrating on more important things, causing him to be absent-minded.

English idiom *to spill the beans* or *the buck* in *to pass the buck.*[7] In making sense of the expression, the hearer may sometimes add to his encyclopaedic entry for the encoded concept the sort of assumptions that would make the expression transparent in the expected way. On learning that the Spanish idiom *acostarse con las gallinas* (*to go to bed with the hens*) means 'to go to bed very early', for instance, the hearer may assume that hens go to bed very early even if this assumption was not there to begin with. Similarly, on learning that the English idiom *to pass the buck* means 'to pass over a problem', the hearer may infer that a buck is sometimes a problem or a burden which one may want to get rid of.

My main concern in this work is not so much with opaque idioms but with idioms which enjoy at least some degree of transparency. I will argue that these expressions are initially understood very much like metaphors: by exploring the encyclopaedic entries of the encoded 'literal' concepts, the hearer looks for implications that would make the utterance relevant in the expected way. Repeated processing of the same expression may result in the hearer using roughly the same encyclopaedic assumptions and deriving roughly the same implications on numerous occasions and so that the expression would become a kind of standardised loose use. The most essential feature of idioms, as I will argue throughout the chapter, is in fact this ability to move back and forth between literalness and looseness, creativity and standardisation.

3. Making Sense of Idioms

Although pragmatics is concerned with the on-line comprehension process and not with the historical events that gave rise a certain expression or meaning, looking at how the meanings of words, and of idioms, have evolved over time may shed some light on issues which are central to pragmatics. The idea that many idioms started out as literal utterances that underwent, over time, a metaphorisation process, provides an interesting

[7] Some scholars (e.g. McGlone, Glucksberg and Cacciari, 1994) have proposed that the idiom *spill the beans* is more transparent than an alternative expression *spill the mud,* might be suggesting that the phrase *the beans* does make some contribution to idiom meaning (e.g. in that beans, like secrets, are many and countable).

test case for the relevance-theoretic claim that understanding metaphorical meaning involves pragmatic broadening of the encoded meaning.

Tracking back the events that gave birth to various idiomatic expressions researchers have found that many idioms were originally intended literally (Dunkling, 1998; Parkinson, 2000). Thus, *spill the beans* was originally used to refer to the spilling of beans, *barking up the wrong tree* was used to refer to hunting dogs barking at trees where there was no prey, and so on. Some studies suggest that the expression *spill the beans* might have originated as part of a game played in rural fairs in America. This game involved contestants guessing the number of beans in a jar. The correct number was revealed by spilling the beans after bets had been made. Asking someone to (literally) spill the beans was asking them to reveal the concealed information (Dunkling, 1998). If the idiom did originate in this way, we can assume the expression must have been repeatedly used to convey roughly the same implications (e.g. the implication that the speaker wants someone to reveal some hidden information, that performing the act of spilling beans would result in the revelation of this information, etc.). We can see also how the popularity of this game at the time might have led to the loose use of the expression *spill the beans* to convey some of these implications. That is, the expression could have been used at the same time both literally, to refer to the spilling of real beans during the game, and by extension, to refer to other types of events which involved the revelation of some hidden information.

From a relevance-theoretic perspective, the comprehension of *spill the beans* when used in this loose way would be no different from the comprehension of approximations and metaphorical uses. The literal compositional meaning of the expression would give access to a range of encyclopaedic assumptions; during on-line interpretation, these assumptions would be considered, in their order of accessibility, in the search for implications that would satisfy the expectations of relevance raised by the utterance. These expectations may be satisfied by a loose interpretation on which the expression *spill the beans* is taken to refer to events in which information is revealed, with the literal spilling of beans in the game-setting as a special case.

The Spanish idiom *tirar la casa por la ventana* (*to throw the house out of the window*) seems to have followed the same broadening process over time. The origin of

the idiom goes back to the end of the XVIIIth and beginning of the XIXth centuries when it was traditional in Spain for people who won the lottery to throw their furniture and old possessions out of their windows so as to show off their wealth and indicate that they were about to commence a new life of luxury (Buitrago, 2002).[8] Knowing this, we may assume that people living in Spain at that time must have used the expression "literally" to convey roughly the same implications.[9] These might have included the implication that the people who throw their possessions out of their windows are wasting or losing money, that they would need to spend a great deal of money buying new things, that this expenditure of money is unnecessary, that they are not behaving in a very sensible or discriminating way, etc. Familiarity with the expression (in its literal sense) may have led native speakers to start using the expression loosely, to refer to situations in which someone is spending or losing large amounts of money somehow unreasonably.

For both the English and the Spanish expressions, what might have started as a novel extension of the literal meaning of a familiar phrase may have eventually turned into a standardised metaphor which was repeatedly used to convey roughly the same implications. These standardised uses with broader meanings may have continued even after the allusion to the original tradition was no longer comprehensible to most people. Even without knowing the origin of the expression and the sort of implications it was originally used to convey, modern speakers still use it to convey roughly the same implications, as in (19) and (20):

(19) Nobody but you knows what is happening in that office. Only you can *spill the beans.*

(20) La boda va a ser perfecta, mi padre esta decidido a *tirar la casa por la ventana.*

---

[8] According to Buitrago (2002), the tradition is still observed in some places in North Italy, where people throw old things into the street at the end of the year as a sign of a better start for the new year.

[9] This idiom is interesting because, although one can say that, at the time, the expression was literally used, a strictly literal interpretation could not have been intended, since it is physically impossible to throw a house through one window of the house. In deriving the "literal" interpretation of the utterance, hearers would need to enrich their interpretation in such a way as to take the word *house* to refer not to the house as such but only to the things inside the house, or to some things inside the house.

The wedding is going to be perfect, my father is determined to *throw the house out of the window.*

In this way, many current idiomatic expressions might have moved over time from a literal or approximate use, to a novel loose use (e.g. as metaphor or hyperbole), to a standardised loose or idiomatic use. Other idioms, however, might never have been literally intended but have started out as loose uses, which would later become standardised in the language (e.g. *to slip one's mind, to lose a train of thought, to change one's mind, to burst into tears, to cry one's eyes out,* etc.). This diachronic movement from a literal meaning to a loose interpretation which is later standardised is often also seen in the acquisition of these expressions by modern language users. In the next section, I will suggest that in making sense of unknown idiomatic expressions, people generally process them very much as they process the metaphorical uses in (4)-(6). That is, they take the encoded 'literal' meaning of the phrase as input to inferring a broader (looser) meaning. After enough exposure, this broader meaning may become standardised as an idiomatic meaning.

3.1 Synchronic Rationale

That current native speakers are not aware of the link between the present meaning of many idioms and their original use does not necessarily mean that they perceive the idiom as opaque, or that its meaning is now arbitrarily learned. An idiom is (relatively) transparent to an individual if he can infer at least some of the idiomatic meaning from the encoded 'literal' meaning. In acquiring the meaning of an idiom, as in understanding a literal or metaphorical use of a word, the hearer may use encyclopaedic information made accessible by the encoded concepts merely as a starting point for inferring the speaker's meaning. Although the assumptions used in the interpretation process may differ from those which gave rise to the idiom, they may allow the hearer to provide some kind of "synchronic rationale" for why the idiom means what it does and so may allow him to perceive the idiom as relatively transparent.

In a series of interesting experiments, Keysar and Bly (1995, 1999) tested whether a single string may be perceived as transparent by subjects who assumed it had a certain

meaning as well as by subjects who assumed it had the opposite meaning. They presented subjects with some unfamiliar idioms in a scenario biasing the interpretation towards either the expression's real idiomatic meaning or a meaning which was the opposite of this meaning. For instance, some subjects were encouraged to think that the expression *to applaud to the echo* meant 'to demonstrate high acclaim' (original meaning), whereas others were encouraged to believe that it meant 'to criticise or ridicule'. Similarly, some subjects were presented with a context in which the expression *to play the bird with the long neck* meant 'to be looking out for someone or something' (original meaning) whereas others were presented with a context in which it meant 'to avoid encounters'. After reading the text, subjects were asked to guess the meaning of the expression by choosing between the two meanings above (and an unrelated meaning). The results showed that subjects systematically chose the idiomatic meaning which was consistent with the overall context. Crucially, when asked to predict what meaning they would predict an overhearer would assign to the expression when it was presented to them in isolation, they systematically reported that he would take the idiom to have just the same meaning as they chose.

The authors took these findings to suggest that subjects in both scenarios had constructed a story that had allowed them to make sense of the expression and perceive it as relatively transparent. Thus, subjects who were presented with the expression *to applaud to the echo* in a context in which it could plausibly mean 'to demonstrate high acclaim' may have focused on the word *applaud,* because of its positive connotations and its association with the demonstration of high acclaim. By contrast, subjects who were encouraged to believe that the expression meant 'to criticise or ridicule' may have focused on the word *echo* and on its negative connotations. Similarly, people who thought that the expression *to play the bird with the long neck* meant 'to be looking out for someone or something' may have assumed that the long neck allowed the bird to look around, just as it does for giraffes. By contrast, those who thought the expression meant 'to avoid encounters' may have assumed the long neck allowed the bird to hide its head in the sand in the way ostriches do.

My interpretation of Keysar and Bly's findings is that, when presented with the idiomatic expression in a biasing context, the subjects should have had quite precise

expectations of relevance. Comprehension would then have involved a considerable amount of backward inference in which attention would be selectively allocated to encyclopaedic assumptions from the encoded concepts that might warrant the expected implications. The same process can be seen operating in (21) and (22):

(21) The young lawyer gave an excellent performance. I wouldn't have been surprised if the audience had stood up and *applauded to the echo.*

(22) The young lawyer gave an appalling performance. I wouldn't have been surprised if the audience had stood up and *applauded to the echo.*

In (21), the information that the lawyer gave an excellent performance may direct the hearer's attention to some aspects of his encyclopaedic knowledge about applauding, and applauding loudly enough to produce an echo, which are consistent with this assumption and so help to achieve relevance in the expected way. He may assume, for instance, that the echo is the outcome of intensive energetic applause. In (22), the information that the lawyer gave an appalling performance, however, may guide the hearer in a different direction. He may consider the hypothesis, for instance, that the type of applause described is a sign of sarcastic mockery. In each case, the more or less precise expectations of relevance generated by the speaker's utterance, and the accessibility of interpretive hypotheses, constrain the direction of the inference, allocating the hearer's attention and processing resources in different ways. In each case, the hearer, in the search for an optimally relevant interpretation, would be encouraged to supply a different subset of encyclopaedic assumptions, and derive different ad hoc concepts and implications. The hearer of (21) may construct a phrasal ad hoc concept [APPLAUD TO THE ECHO]* which denotes acts of giving intense praise, while the hearer of (22) may construct a phrasal ad hoc concept [APPLAUD TO THE ECHO]** which denotes acts of ridicule. Provided that the content assigned to these concepts is partly recovered from the meanings of the words in the idiom (e.g. by considering assumptions associated with applause and with the production of an echo while applauding), the expression may be perceived as at least partly transparent (as Keysar and Bly's findings suggest).

Contrary to the standard view that idioms are understood as lexical items, a wide range of on-line and off-line experimental research on acquisition of idioms has shown that the interpretation of unknown idioms seems to be affected not only by contextual cues (as in the above examples) but, crucially, by the internal semantics of the string (e.g. Cacciari, 1993; Cacciari and Levorato, 1989, 1991, 1998; Levorato, 1993; Levorato and Cacciari, 1992, 1995, 1999; Flores d'Arcais, 1993; Forrester, 1995; Gibbs, 1991). For instance, results from the on-line experiments carried out by Cacciari and Levorato (1999), which are outlined in the previous chapter, show that context and word meaning play two independent but interactive roles in idiom comprehension. These experiments showed that idioms which are at least partly transparent are easier to understand than opaque idioms, because people can use the meanings of the individual constituents of in the expression and the structure of the phrase as clues to the overall idiomatic meaning. Context does play an important role, and is exploited by children in acquisition as they develop the ability to integrate linguistic and extralinguistic knowledge in the processing of an utterance or a text (Levorato, 1993; Levorato and Caccirari, 1992, 1995),[10] but it cannot by itself explain how children make sense of idiomatic expressions. Idiom acquisition seems to depend on the accessibility of information both from the context in which the expression is processed and from the concepts encoded by the words in the idiom.

A number of off-line studies have been carried out to analyse the strategies which people use in interpreting unknown idiomatic strings (see Cacciari, 1993). Typical answers included one from a child who reported having made sense of the Italian idiom *to be on the seventh heaven/sky,* meaning 'to be extremely happy', in the following way: "we all know that heaven/sky is wonderful, so if there was a seventh one, can you imagine?" This and other answers provided by both children and adults suggests that

[10] The approach to the development of idiomatic competence and figurative competence more generally proposed by these authors is much more complete, as they study the different stages in comprehension and how the development of more general linguistic and cognitive abilities affect the comprehension of figurative speech: from the abandonment of a literal strategy to the development of metalinguistic skills (see Cacciari and Levorato, 1998). A very interesting line of research which falls outside the scope of this thesis would be to attempt to integrate this approach with Sperber's ideas on the different stages a child goes through in development of his metarepresentational abilities (Sperber, 1994b but see also Wilson, 2000) so as to analyse not just the comprehension of idioms in normal development

people systematically use the encoded 'literal' meanings in working out the meaning of the overall string. A similar conclusion can be drawn from other experiments, such as that by Forrester (1995), which shows that people understand unknown idiomatic expressions by treating them as if they were metaphorically intended.

Although the approach to idioms defended in this work is largely theoretical, it is consistent with experimental research on idiom acquisition reported in the literature. Here, I find Keysar and Bly's findings particularly interesting, because they seem to suggest not only that people explore context and the literal meaning of the words in interpreting an idiom, but that they may do this in different ways, so that a single expression can be taken to convey different meanings. Shedding light on the processes which direct the hearer towards a certain interpretation is crucial to understanding how people acquire idiomatic expressions. I would like to suggest that selection of the contextual assumptions (from the concepts encoded by the words in the idiom, the rest of the utterance and from background knowledge) which people use to make sense of idiomatic expressions, and the direction of the inferential process in which these assumptions are used as premises, are constrained at every point by the hearer's search for an optimally relevant interpretation. It is selective, relevance-driven processing which allows the hearer to construct a number of different interpretations for a single phrase on different occasions. In line with this, we can imagine, for instance, how a phrase such as *to burn the house* may be potentially used to convey a wide range of loosely intended (and equally transparent) meanings, any of which may be standardised in the English language, as in (23):

(23) a. To do something big, wild.

e.g. Tom's party is going to be great. He has promised *to burn the house.*

b. To abandon everything, give up.

e.g. I know you failed the exam, but there is no reason *to burn the house,* you need to keep trying.

c. To do something drastic with negative consequences.

---

but also how these expressions would be understood by people whose metarepresentational abilities are impaired (e.g. Happé, 1993, 1994).

e.g. I know that the company is going through difficult times and some changes are needed, but what you are proposing (firing half of the staff) is to *burn the house*! People won't accept that.

d. To incur big expenses.

e.g. Since she is the only daughter, her father will *burn the house* to give her the best wedding ever.

e. To give up something valuable for a good reason/cause.

e.g. When you have to decide between a life of luxury and the love of your life, you don't mind *burning the house.*

The different expectations of relevance generated by these utterances may not only add an extra layer of activation to some of the encyclopaedic assumptions associated with the encoded concepts, but may also, crucially, encourage the hearer to use these assumptions as premises in following a certain inferential route. By inferential "route", or "path", I mean a combination of selected assumptions and computations used to derive a certain set of implications. In (23b), certain contextual assumptions (e.g. the assumption that burning a house would result in the destruction of something valuable) may be added as a premise to a context of selected contextual assumptions (e.g. assumptions about surrender). The resulting contextual implications may also provide input to further inference processes, yielding further implications (e.g. implications about abandoning something valuable one has worked hard for, etc.). Since the hearer may consider only a subset of encyclopaedic assumptions from the encoded concepts and process them following a different inferential route, the same expression can be loosely used to convey a wide range of different meanings. Furthermore, since each of these meanings would be (slightly or considerably) different extensions of the compositional meaning of the phrase, only a subset of the implications derivable from the compositional meaning may be understood as part of the speaker's meaning in each case. The point here is that although the hearer may follow a number of different inferential routes to make the idiom transparent, only one of these may be standardised in the language he uses. The (initially one-off) inferential route built to understand this

particular use may therefore be systematically (re)constructed as the hearer encounters the expression in further utterances.

3.2 The Contribution of Word Meaning

We have seen throughout this thesis that the concept encoded by a word can give access in memory to a wide array of different encyclopaedic assumptions, some of which may be added to the context in order to derive the intended interpretation. The argument I am proposing here is that the same process may be used in interpreting idiomatic utterances where different encyclopaedic assumptions associated with an encoded concept may be used to infer the overall idiomatic meaning. The concept encoded by the word *wing*, for instance, gives access in memory to a range of encyclopaedic assumptions, different subsets of which are selected in processing different idioms. So, in processing the idiom *to clip someone's wings,* the hearer may consider the assumption that wings are used to fly, which itself may provide input for further inference. In processing the idiom *to take someone under one's wing*, however, he may consider the encyclopaedic assumption that birds protect their young with their wings, and this, again, may provide input for further inference. Although the same word is used in both expressions, selective processing of encyclopaedic assumptions yields a range of different implications in each case (e.g. implications about freedom of movement, or lack of it, and implications about love, help and protection). It is the derivation of these implications (as well as other implications which these assumptions make accessible) which allow the utterance to achieve relevance in the expected way and lead the hearer to perceive the idiom as relatively transparent.

Sometimes a word appears in several different idioms in which it is understood in roughly the same ways. The word *ace*, for instance, gives access to the encyclopaedic assumption that aces are the highest (and therefore best) cards. This assumption would be highly accessible to the hearer in processing any of the following idiomatic expressions *to be an ace at something, to hold all the aces, to have an ace up one's sleeve, to have an ace in one's hand* and *play your ace*. As suggested for the examples *to applaud to the echo* and *to burn the house* above, a single assumption such as this may be combined with a range of different contextual assumptions in every case. It may be

combined, for instance, with assumptions derivable from the compositional meaning of the phrase (e.g. assumptions about holding aces, having aces up one's sleeve, etc.). This combination may therefore yield different implications on each occasion (e.g. implications about luck, cheating, etc.). These implications may be seen as part of the speaker's meaning or used to derive a further set of implications that the speaker might have intended to convey. Again, it is the choice of an inferential route which takes as input contextual assumptions made accessible by the words in the string and yields those implications as output which allow the hearer to perceive the string as relatively transparent.

3.3 Conclusions on Acquisition

If the arguments defended here are right, what is important in acquiring an idiom such as those presented here is not whether the hearer becomes aware of the historical story underlying the original use of the phrase, but whether he can construct an inferential path by which the overall idiomatic meaning can be at least partly inferred from encyclopaedic information associated with the encoded concepts, either alone or in combination. My suggestion is that, although both the Spanish and the English idioms in (19) and (20) may be unfamiliar to a hearer who does not know the story that motivated their original use, he may still be able to perceive the idioms as transparent or relatively transparent. Taking the encoded 'literal' meaning of the Spanish idiom *to throw the house out of the window* as input to pragmatic inference, he may be able to derive a number of implications (e.g. about the intentional destruction of valuable property and the waste of money that this entails) which may result in a loose interpretation of the string (e.g. one in which the phrase is taken to denote actions in which money is wasted in a rather crazy manner). The idiomatic meaning of the English expression *spill the beans,* however, is not as easily inferable from the meaning of the words in the string. I have suggested that to most English native speakers, the expression would be only partly transparent. This is because, although the concept encoded by the word *spill* may be loosely understood as indicating the act of letting something out or even revealing something, there is nothing in the hearer's encyclopaedic knowledge of beans that would allow him to derive implications about secrets or the revelation of hidden information.

These examples may be taken to suggest that there are, after all, two ways in which an idiom may be transparent. On the one hand, an idiom may be 'directly transparent', in that the hearer can see how some of the assumptions associated with the encoded concepts or derived from their compositional meaning are conveyed by an idiomatic use of the string (e.g. with the Spanish idiom *to throw the house through the window* or the English idioms *to hit the nail on the head, to miss the boat, to give up the ship, to hold all the aces,* etc.). On the other hand, the idiom may be 'retrospectively transparent' if the hearer can only identify encyclopaedic assumptions that would make the idiom relatively transparent AFTER a potential meaning for the expression has been constructed (e.g. *to spill the beans, to pass the buck, to hit the sack, pensar en las musarañas, meterse en un embolao,* etc.) (as defended in Nunberg, 1978; and see also Cacciari, 1993). I have claimed here that since comprehension typically involves mutual parallel adjustment, both processes may take place in parallel. That is, the hearer may use highly accessible encyclopaedic assumptions to derive implications, as well as consider highly accessible hypotheses about implications and use them to enrich the explicit content and the context by backward inference. These hypotheses about the speaker's meaning are tested in their order of accessibility until the hearer arrives at a combination of explicit content, context and cognitive effects which satisfies his expectations of relevance. I have suggested that this mutual adjustment process takes place not only in understanding the alternative meanings of an idiom (e.g. *applaud to the echo, burn the house*), and in understanding an idiom regardless of its degree of transparency (e.g. *kick the bucket, spill the beans, hold all the aces*), but in interpreting virtually every utterance, whether it contains unknown words or words which are literally or loosely used.

The examples presented here suggest that transparency and opacity are not fixed properties of idioms, but dimensions along which they can be characterised by a particular person on a particular occasion (e.g. at a particular point in time). Whether an idiom is perceived by an individual as more or less transparent at a certain moment would largely depend on the assumptions available to him at the time, and their degree of accessibility. Generally, the greater the number of implications which are also

derivable from the literal meaning of the phrase, the easier the derivation of the (loose) interpretation will be and the more transparent the idiom seems to the hearer.

I have suggested here, for instance, that the idiom *spill the beans* is generally perceived as only partly transparent by current native English speakers, even though it may have seemed rather more transparent to people in the past. The reason for this difference is that language users may once have had access to encyclopaedic information about the popular game the expression refers to, and to the sort of implications which the idiom was used to convey in those situations. It is through our knowledge of everyday affairs, including current sports and games, that many idioms are perceived as transparent by the modern language user, as transparent as the idiom *spill the beans* might once have been to English native speakers. Consider, for instance, the English idioms in (24):

(24) *to hold all the aces, to be an ace, to hide an ace up one's sleeve, to lay the cards on the table, the ball is in your court, to throw in the towel/the sponge, to be a team player, to bark up the wrong tree, to call a spade a spade, to flog a dead horse, to start/get the ball rolling, to keep the ball rolling, to hit below the belt, to be off the hook, to jump the gun, to be back to square one,* etc.

The idioms in (24) vary in their degree of transparency. What makes some of them quite transparent to current native speakers is the easy access they provide to encyclopaedic assumptions about the type of activity, sport or game which the expression alludes to, and the ability to derive implications using these assumptions. Since different people have different knowledge and experience, an idiom which is transparent to one person may remain opaque to others. For those familiar with ball games (e.g. tennis) and with boxing, for instance, it might be easy to supply the assumption that when the ball is in our court, it is our turn to act, and that throwing in the towel is a sign of surrender. These people would see the idiomatic expressions *the ball is in your court* and *to throw in the towel* as considerably more transparent than those people who do not have access to these assumptions or have not been able to infer their relation to the idiomatic meaning.

Acquiring native speaker command of an idiom, then, should involve on the one hand, seeing how the sort of implications the idiom is used to convey can be inferred from the encoded 'literal' meaning of the string, and on the other hand, fine-tuning these implications so that they accord with those derived by other members of our linguistic community. In line with this idea is the finding by Keysar and Bly (1999) that the more the subjects gain familiarity with an expression (e.g. by using it in novel utterances), the more confident they become about the meaning they have assigned to it, and the more reluctant they are to accept that the idiom is transparent with a different meaning. Thus, just as current English speakers find it hard to see how the expression *to spill the beans* can be used to mean 'to keep a secret', so do people who acquire the meaning of the idiom *applaud to the echo* in a context biased towards the interpretation 'to ridicule' find it hard to see how this expression could be used to convey the (actual) meaning 'to demonstrate high acclaim'. These results suggest that not only do people form hypotheses about the relation of word and idiom meaning in making sense of the idioms in context, they also seem to retain these assumptions and assume that other people interpret idiom transparency in similar ways.[11]

My proposal is thus that, in interpreting an utterance containing an unfamiliar idiom, as in interpreting any other utterance, the hearer takes the encoded conceptual representation as the starting point for inference. Following a path of least effort, he adds associated encyclopaedic assumptions to the context in order of accessibility, taking a particular inferential route whose output should be the range of implications the speaker might have intended to convey. This selective relevance-oriented process directs the hearer to a combination of assumptions and computations which should help him infer the speaker's meaning. The reason for spelling out the processes I see as involved in selecting these inferential routes is that I believe they play a fundamental role in the comprehension of idiomatic expressions as the hearers become more familiar with them. In the next section, I will argue that repeated processing of an idiom may recurrently

[11] The account on implications defended here has important consequences also for second language learning and issues on translation. As pointed out by Gutt (1991) a good translation of standardised figurative uses like idioms or proverbs is that which uses the expression in the other language which conveys roughly the same implications. The English idiom *to cost an arm and a leg* would so be adequately translated into the Spanish *to cost an eye of your face* and into the Italian *to cost an eye of your head* which are used in roughly the same contexts to convey roughly the same sort of implications.

direct a hearer along the same inferential route, which may at some point develop into a full-fledged pragmatic routine (also in Vega Moreno, forthcoming).

4. Familiar Idioms: Representation and Processing

The crucial thing about idioms is that they are generally used to convey roughly the same meaning in different situations. We may conclude from this that, in interpreting an utterance containing an idiom, a hearer aiming to satisfy his expectations of relevance would repeatedly follow roughly the same inferential route. That is, he would consider roughly the same (highly accessible) encyclopaedic assumptions, use them as premises in the same inferential computations, derive roughly the same implications and enrich the proposition expressed by adjusting the encoded concepts in roughly the same way. Thus, although the expectations of relevance raised by different utterances would often be satisfied by different combinations of explicit content, context and cognitive effects, processing an idiom may repeatedly direct the hearer to the same sort of hypotheses about the speaker's meaning. In line with the relevance-oriented view of cognition and communication defended in Relevance Theory, I want to suggest that the more a certain inferential route is followed, the more accessible and cheaper in processing terms it will become. Having helped to achieve relevance on previous occasions, it is likely to become highly accessible for use on subsequent similar occasions.

The upshot of repeated use may be that the hearer develops a pragmatic routine for the processing of some familiar idioms: that is, on hearing the idiom his attention and processing resources would be automatically directed along the same inferential route which has been followed in processing the string on previous occasions. Let's illustrate how this account would work for the comprehension of the idiom *to hold all the aces*, as in (25)-(26):

(25) There is no way I will get the job. Peter, however, *holds all the aces.*

(26) We will not know for sure who will win this year's general elections until the votes are counted, but everybody knows that Clinton *holds all the aces.*

Rather than needing to dig into the encyclopaedic entries of the encoded concepts or search for the best inferential route as might happen on first encounter (as claimed in presenting example (17)), a hearer familiar with the idiom *to hold all the aces* may find his attention and processing resources automatically directed towards the sort of contextual assumptions and implications that have generally led to a successful outcome in processing this idiom, such as those in (27):

(27) Assumptions: An ace is the best (highest) card
Someone who holds all the aces is likely to win
Someone who holds all the aces is a very lucky person
Etc.

Using these selected assumptions in processing different utterances containing the idiom (such as those in (25) and (26)) would generally yield roughly the same implications, with the compositional meaning of the phrase being adjusted in roughly the same ways (i.e. so as to warrant the derivation of these implications). In (25) and (26), this may involve broadening the compositional meaning of the phrase so that the resulting phrasal ad hoc concept denotes situations in which some individual is in a winning position, as in (28):

(28) [TO HOLD ALL THE ACES]* denotes situations in which some individual has every chance of winning or succeeding at something.

Since the hearer would generally broaden the meaning of the phrase in roughly the same ways across different occasions of use, it may be more economical to store this broader concept in memory rather than constructing it ad hoc. The sort of assumptions which are used in broadening the original concept may thus end up being stored as part of the encyclopaedic entry of the new concept, as in (29):

(29) Conceptual Address [TO HOLD ALL THE ACES]*[12]

Lexical entry: syntactic $_{vp}$[VNP] and phonological information

Encyclopaedic entry: assumptions about the state of affairs the concept denotes.

If someone holds all the aces then he is in a winning position

If someone holds all the aces then he has great chances of success

If someone holds all the aces then he is very lucky

Etc.

Provided that the hearer of (25)-(26) has the concept in (29) stored in memory, he should be able to access it at some point in processing the utterances in (25)-(26). As in processing any other encoded concept, he would use some encyclopaedic assumptions it makes accessible and look for enough implications to satisfy his expectations of relevance.

Whether the idiom has an independently stored conceptual address, as in (29), or whether it has only an associated pragmatic routine which is not yet lexicalised, the comprehension process would be roughly the same: selected encyclopaedic assumptions associated with the expression would be used to infer a range of implications, which may themselves be used as input to derive further utterance-specific implications. In (25), for instance, the assumption that Peter has great chances of success might be combined (among others) with the contextual assumption that Peter and the speaker are applying for a job to yield a range of implications which may make the utterance relevant in the expected ways (e.g. that Peter has every chance of getting the job, that Peter has a better chance of getting the job than the speaker, etc.). In (26), the assumption that Clinton has every chance of winning something may be combined with the assumption that he is a candidate in the US general elections to yield the implication that he has every chance of winning the US general elections. [13]

[12] The asterisk (*) here does not indicate that the meaning is created ad hoc but that it is a pragmatic adjustment of the compositional meaning which happens to have been stored in memory.

[13] Eizaga (2002) has proposed an approach to idioms from Relevance Theory which defends the following two ideas. On the one hand, following ideas (defended in passing) by Pilkington (2000) and Papafragou (1996) about standardised metaphors and metonymies (respectively), she argues that many idioms which are not yet lexicalised and some idiom variants are understood via the activation of some set of mutually manifest hypotheses or metarepresented assumptions which are repeatedly accessed in processing the string. She refers to this process as a 'generalised pragmatic routine'. This idea, though

I have suggested here that in interpreting an utterance containing an idiom, the hearer would generally take the compositional meaning of the string as input to pragmatic inference. For relatively transparent idioms, though not for opaque idioms, the hearer might develop a pragmatic routine that allows him to speed through the familiar inferential steps involved in inferring the idiomatic meaning. It is worth noticing that this process is not very different from the one I suggested in Chapter Four for the comprehension of familiar non-idiomatic expressions (e.g. 'to see one's doctor', 'to sleep with someone') and of standardised metaphorical uses (e.g. 'to be a pig'). In fact, some people may never store a separate conceptual address for certain idiomatic expressions, but may process them very much like familiar metaphorical expressions. Idioms may therefore be seen as lying along a continuum depending on whether they are processed by an individual as standardised loose uses, which are not yet lexicalised, or as standardised loose uses which have been assigned their own conceptual address and encyclopaedic entry. My claim here is that the comprehension of a relatively transparent idiom is not essentially different in these cases (although the processing effort factor may vary). In fact, different individuals may represent the same idiom differently and still communicate efficiently.

Although different idioms may be represented in different ways by different individuals we may assume that many familiar idiomatic expressions end up having a stable conceptual address which is accessed at some point in comprehension. For opaque idioms, the meaning assigned to the idiomatic string may be quite arbitrarily stipulated, as claimed by traditional accounts. However, most idioms are likely to be stored as standardised loose uses of one type or another. For these idioms, the activation and accessibility of assumptions associated with the concepts encoded by other constituents in the string need not disrupt the interpretation of the idiom but will often be consistent with it. One of the crucial features of (relatively transparent) idioms is in fact that they

---

interesting and similar to that defended here, is not properly addressed. In fact, it is simply assumed rather than developed. On the other hand, she proposes that a lexicalised idiom often gives access to both conceptual information and to procedures which are used in interpreting the string and variants of the string. Although again an interesting idea as it stands, it is not adequately developed. In fact, the notion of 'procedure' in which it relies seems to have resulted from a combination of relevance-theoretic ideas (e.g. on procedural meaning, pragmatic routines and development of cognitive procedures) and, generally from a misinterpretation of these ideas.

allow hearers to move along the continuum from literalness to looseness and metaphoricity as they adjust the compositional meaning of the phrase and consider the idiomatic meaning of the string during the interpretation process.

4.1 Activation and Interpretation

In the previous chapter, I presented some experimental evidence that the meaning of familiar idiomatic expressions is not immediately activated as the first word in the string is heard but becomes active at a later point (see Cacciari and Tabossi, 1988; Tabossi and Zardon, 1995). In an utterance ending in an idiomatic phrase which is plausible on both literal and idiomatic interpretations, the idiomatic meaning was activated only after the point of idiom uniqueness (i.e. the point at which the string can be uniquely identified as an idiom) had been reached. Although for highly predictable idioms (e.g. *set his mind at rest*), the meaning of the idiom was activated slightly earlier than for less predictable idioms (e.g. *to hit the nail on the head*), in neither case was the idiomatic meaning accessed after the first content word. Familiarity with the string, and particularly the presence of a biasing context, has been shown to affect the point of idiom activation (or idiom recognition) so that it may no longer coincide with, but actually precede, the point of idiom uniqueness (e.g. Flores d'Arcais, 1993).[14]

Cacciari and colleagues have generally analysed these findings as showing that the processing of an idiom remains literal until the idiomatic expression (or configuration) is activated by arriving at the idiom key, at which point both the idiomatic and the literal meaning compete until one of them is chosen. This approach, I argued in the previous chapter, is an updated version of the Simultaneous Processing model defended in the late seventies by Swinney and Cutler (1979), who suggested that the literal and idiomatic

[14] Some scholars have observed that idioms are not pronounced just as their literal counterparts. Van Lancker and colleagues (e.g. Van Lancker and Canter, 1981; Van Lancker and Canter and Terbeek, 1981; Van Lancker and Kemper, 1987) for instance recorded people producing text in a phrase was either literally or idiomatically used. He then isolated the part where the idiom was produced and presented it to other set of subject who were asked to identify whether the string was being literally or idiomatically used. He found that subjects were extremely good at guessing just from the pronunciation of the phrase whether it had been literally or idiomatically intended. Ashby (forthcoming) has proposed the way in which idioms are pronounced is directly related to the compositionality of the string. He argues for instance that bizarre effects result from focusing on non-compositional (as opposed to compositional) parts of the idiom. We may conclude from this research that phonological information may help the activation of idiomatic meanings, disambiguate the phrases and even give an extra push to the hearer to follow a certain inferential route.

meanings of an idiom are processed in parallel until the hearer makes a choice. Although the above experiments shed interesting light on the point at which an idiom is activated and so becomes accessible to the hearer, they do not say much about how that idiom is actually interpreted. They do not explain how a hearer might decide which interpretation the speaker might have intended on that particular occasion, or whether he establishes a relation between the compositional and idiomatic meaning of the phrase in constructing this interpretation. For instance, they don't discuss whether the assumptions the hearer considers in processing the string 'literally' before the idiomatic meaning is accessed are integrated into the comprehension process or simply rejected at that point.

I believe that talking of 'literal' processing or 'idiomatic' processing in the way these scholars do is not entirely accurate. In line with Relevance Theory, I want to suggest that processing should not be seen as literal, metaphorical or idiomatic but simply as relevance-driven. The rather selective, and therefore initially quite shallow, relevance-oriented processing of an utterance would lead the hearer to consider only highly accessible encyclopaedic assumptions from the encoded concepts in looking for implications. At some point, the concept encoded by the idiom may itself be accessed, with some highly accessible encyclopaedic assumptions associated with this concept being added to the context to derive further implications. Whether the contextual assumptions already present in the context are strengthened by the new information, combine with it to yield implications or are rejected and eliminated for not contributing to relevance in the expected ways would vary from idiom to idiom, depending on their degree of transparency and so therefore on whether the idiomatic meaning can be inferred, or partly inferred, from the compositional meaning of the phrase. Let's look at some examples:

(30) *Janet:* Is your boyfriend coming to the party?

Jenny: I am afraid not. He is spending Sunday with his mother, as always. I cannot stand the way he *is tied to his mother's apron strings.*

(31) *Sue:* I really love that dress but *it costs an arm and a leg.*

(32) *Tim:* We have been very affected by the accident but I think it is time we *turn over a new leaf* and get on with our lives.

(33) *Jason:* The old man did not want to sell his house so the council waited until he *kicked the bucket* to get hold of everything he owned.

(34) *Tom:* Where is Vanessa?

Carol: I don't know, she is probably *chewing the fat* with her friends somewhere.

(35) *Joe:* Don't take it seriously. I am sure he was only *pulling your leg.*

Following a route of least effort, the hearer of (30)-(32) may start considering a certain subset of encyclopaedic assumptions made accessible by the encoded concepts or the compositional meaning of the phrase (e.g. in (30) the assumptions that tying involves attachment or in (32) that turning over involves a change of position). This relatively shallow processing may yield a range of implications which can be seen as part of the speaker's meaning or can be used to derive the sort of intended implications. In other words, since the idiomatic meaning of relatively transparent idioms like these can be inferred or partly inferred by adjusting the encoded concepts, using selected encyclopaedic assumptions associated with these concepts in order to derive implications may be a step towards a certain (loose) interpretation which will be later fine-tuned towards an idiomatic interpretation.

By contrast, in processing the utterances in (33)-(35), the shallow processing of the encoded concepts may lead to the hearer initially considering some contextual assumptions (e.g. about chewing, kicking, pulling, etc.) and deriving tentative implications, which will later be rejected and eliminated as not contributing to relevance in the expected ways. Since the idiomatic meanings of these opaque idioms cannot be inferred or partly inferred from the encoded meanings of their parts, accessing selected assumptions associated with these concepts would not help to derive the implications or other cognitive effects. We can therefore conclude that although hearers follow the same comprehension procedure in interpreting any idiomatic expression (and indeed any utterance), the selection of tentative contextual assumptions and implications needs to be adjusted in deriving the idiomatic interpretation, or to be rejected as making no contribution to this interpretation.

It would be interesting to conduct on-line experiments which might test this claimed difference. Although the finding that analysable idioms are understood faster

than unanalysable idioms (e.g. Gibbs, 1991) is consistent with the view of idioms presented here, it is important to bear in mind that the notion of analysability or compositionality used in these experiments does not always coincide with the notion of transparency I have defended in this chapter and the previous one (i.e. transparency as the extent to which the idiomatic meaning can be inferred from the encoded 'literal' meaning of the string). The closest the literature has come to testing the distinction I have shown between transparent and opaque idioms has been the eye-tracking experiments carried out by Titone and Connine (1994, 1999) outlined in the previous chapter. These experiments, let's recall, presented ambiguous idioms in a context biased towards their literal or their idiomatic meaning. The aim was to test duration of eye-fixation in the ambiguous regions of the idiom while the text was read. The results showed that the duration of eye fixation was longer for opaque idioms than for transparent idioms. Given the relevance-theoretic approach to idiom comprehension outlined above, we can assume that the shallow processing of the encoded concepts leads the hearer to start enriching the proposition expressed in a certain way and, to start deriving a tentative set of implications. In the case of relatively transparent idioms, but not in the case of opaque idioms, this enrichment and these implications may be compatible with both a literal and an idiomatic interpretation. Further processing of the utterance may direct the hearer to fine-tune the meaning of the expression in either direction, so that a literal or an idiomatic interpretation is finally derived.

5. Pragmatic Adjustment

So far, I have claimed that the stored meaning of an idiomatic expression is generally a broadening of the compositional meaning. I want to show now that this broader (metaphoric or hyperbolic) concept may itself sometimes need to be pragmatically adjusted into a new ad hoc concept which contributes to the truth-conditional content and warrants the expected implications, consider (36)-(43):

(36) Since the Paddington derailment, trains run *at a snail's pace.*

(37) Since she broke her hip, my grandma walks *at a snail's pace.*

(38) My husband is very handy. He painted the house *in the blinking of an eye.*

(39) My husband got dressed *in the blinking of an eye.*

(40) Mi padre *ha perdido la cabeza* por esa mujer. Está locamente enamorado.
My father *has lost his head* for this woman. He is madly in love.

(41) Mi padre *perdió la cabeza* cuando le dijimos que mi hermano había muerto.
My father *lost his head* when I told him my brother had died.

(42) Mi padre *ha perdido la cabeza,* no reconoce a nadie
My father has *lost his head.* He cannot recognise anyone anymore.

(43) Mi padre *ha perdido la cabeza.* Quiere dejar su trabajo en el banco para hacerse marinero.
My father has *lost his head.* He wants to give up his job in the bank to become a sailor.

Let us assume along the lines argued above that the expression *at a snail's pace* encodes a concept [AT A SNAIL'S PACE]*, which denotes states of affairs in which something happens very slowly. Similarly, the expression *in the blinking of an eye* might encode a concept [IN THE BLINKING OF AN EYE]*, which denotes states of affairs in which something happens very fast, and the expression *to lose one's head* might encode a concept [TO LOSE ONE'S HEAD]*, which denotes situations in which someone has lost the capacity to reason. What the examples in (36)-(43) seem to indicate is that these concepts often have to be pragmatically enriched in order to warrant the derivation of the expected cognitive effects.

In (36), for instance, the concept [AT A SNAIL'S PACE]* may need to be adjusted to a point where it warrants the conclusion that trains are running at many fewer kilometres per hour than before, whereas in (37) it would need to be adjusted to a point where it warrants the conclusion that the speaker's grandma walks much more slowly than an average adult. It is the concept resulting from this adjustment that seems to contribute to the truth-conditional content of the utterance. In (38), for instance, the proposition expressed would be judged true if the speaker's husband took only three hours to paint a three-bedroom house. The same would not hold for (39), however, as taking three hours to get dressed is quite a long time. In this case, the concept encoded

([IN THE BLINKING OF AN EYE]*) would need to be adjusted to a point where it warrants the conclusion that the speaker's husband got dressed in just a few minutes, enabling them to leave promptly.

The examples in (40)-(43) suggest that the Spanish phrase *perder la cabeza*, or its rough English equivalent *lose one's head,* can also be used in different utterances to convey slightly different meanings, and so to yield a different range of conclusions. It may be used, for instance, to convey a concept which denotes the state of being deeply in love, as in (40); the state of being in despair, as in (41); the state of being mentally disabled, as in (42); the state of being a bit mentally unstable, as in (43). Comprehension of these utterances may involve both pragmatic broadening of the compositional meaning of the phrase and pragmatic narrowing of the metaphorical meaning stored in memory. Here, as always, pragmatic adjustment leads to the construction of an ad hoc concept which yields enough implications (e.g. about the type and degree of mental instability and inability to reason) to satisfy the hearer's expectations of relevance.

Recalling the arguments of the last chapter, it is important to notice that the Spanish expression *perder la cabeza,* or its rough English equivalent *to lose one's head,* may be seen as being either as decomposable or nondecomposable. If we assume that the word *lose* describes the act of losing something and the phrase *one's head* describes the capacity for reasoning, the idiom would be seen as decomposable, compositional or analysable. Alternatively, if we take the whole expression to denote the state of being somewhat insane or the process of going crazy, the expression would be perceived as nondecomposable, or at least abnormally decomposable. Whatever position we adopt, the compositionality approach cannot explain satisfactorily how an individual would interpret the utterances in (40)-(43) and, crucially, how he would assign a (slightly) different interpretation in every case. The approach defended here can provide the beginning of an adequate answer to these questions by claiming that relatively transparent idioms generally encode a standardised loose (e.g. metaphoric or hyperbolic) meaning which can be pragmatically adjusted to yield a range of slightly different meanings. Often, the pragmatic narrowing and pragmatic broadening operate simultaneously in the construction of a hypothesis about the speaker's meaning.

It might be argued, in line with my proposal in Chapter Four, that the interpretation of idioms, such as those in (36)-(43), may involve the pragmatic adjustment of some of the encyclopaedic assumptions associated with the concept encoded by the idiomatic string. In discussing metaphors such as 'my boss is a wolf', let's recall, I suggested that the encyclopaedic property of BEING AGGRESSIVE as applied to wolves may be pragmatically adjusted during the interpretation process so as to apply to warrant the derivation of implications that apply to men. With regard to idioms such as those in (36)-(43), we might consider the possibility, for instance, that the encyclopaedic property of MOVING SLOWLY made accessible by the idiom *at a snail's pace* or the encyclopaedic property HAPPENING FAST made accessible by the idiom *in the blinking of an eye* may need to be adjusted in context so as to warrant the derivation of the expected implications. The different positive adjustments of these concepts will be linked to different inferential routes and different cognitive effects. These cognitive effects may lead by backward inference to the construction of different ad hoc concepts. In (37)-(40), the resulting concepts might denote the particular type and degree of speed required to make the utterance satisfy the hearer's expectations of relevance.

If this approach is along the right lines, then, at some point in interpreting an utterance containing an idiom, the hearer would have access to the concepts encoded by the words in the idiom, the concept encoded by the idiom as a whole, and often also to a pragmatic routine for bridging the gap between the compositional meaning and the idiomatic meaning. Following a path of least effort in confirming hypotheses about the speaker's meaning, he may move back and forth between the compositional and idiomatic interpretations until he arrives at a particular interpretation which satisfies his expectations of relevance. The selective relevance-oriented comprehension procedure would favour the most accessible hypotheses at every point where he has to make a choice. The processing of the concept encoded by the idiom and, particularly, the concepts encoded by the words in the string, would therefore be relatively shallow, with only highly accessible encyclopaedic assumptions being considered.

5.1 Word Meaning and Idiom Meaning

It has been pointed out that idioms generally make use of what Coulmas (1981a, 1981b) refers to as 'idiom-prone lexemes'. These may be light verbs, or verbs which, in relevance-theoretic terms, are seen as encoding pro-concepts, which need to be enriched on each occasion of use (e.g. Sperber and Wilson, 1998). Examples include verbs such as *put, take, make, do, have, be,* etc. in English, and *poner* (put), *coger* (take), *hacer* (make/do), *tener* (have), *ser* (be – permanent state), *estar* (be – temporary state), etc. in Spanish. Coulmas (1981b) suggests that the presence of 'idiom-prone lexemes' may be taken by hearers unfamiliar with the string as indicating that an idiomatic expression is being used. However, my interest in these verbs is closer to that of Nunberg (1978), who argued that language users often take a single word (e.g. *hit*) to have different meanings in different idioms (e.g. *to hit the sack, to hit the panic button, to hit the road,* etc.). I want to argue here that in interpreting an idiom, people generally fine-tune the encoded concepts so that the whole expression can be loosely interpreted in a particular way. This is true, not only of the verbs mentioned above but of virtually every verb and every word in an idiom. To illustrate my argument here are some examples:

(44) a. PUT: *put words in someone's mouth, put the lid on something, put the genie back in the bottle, put a finger on the wound, put your life in someone's hands, put the cat among the pigeons, put on a brave face.*

b. BREAK: *break the ice, break someone's heart, break the news, break ground, break one's back, break a leg, break the bank.*

c. BITE: *bite the hand that feeds you, bite the bullet, bite someone's head off, bite the dust, bite your tongue.*

Interpreting the above expressions would involve some degree of pragmatic fine-tuning at both word and phrase level. At the level of the word, the hearer of the idioms in (44) may need to fine-tune the meanings of the words *put, break* and *bite* narrowing or broadening them in appropriate ways. Understanding the expressions *bite one's tongue, bite the bullet* or *bite the hand that feeds you,* for instance, may involve the hearer narrowing the concept encoded by the word *bite* so as to warrant some

implications normally derived from literal uses of the word. Some of these implications might then be seen as part of the idiomatic meaning that the speaker intended to convey by that particular utterance.

Thus, the meaning of relatively transparent idioms is often inferable by taking the individual words in the string as literally, approximately or loosely intended, and by taking the whole phrase as conveying a loose (metaphoric or hyperbolic) meaning. In the course of comprehension, some of the assumptions more accessible by the encoded concepts, and the sort of implications they can be used to convey may be accepted as part of the idiomatic interpretation, placing idioms somewhere along a continuum from literalness to looseness. So, in saying that someone has *missed the boat,* the speaker may convey that he has missed something (literal or approximate use of *miss*); in saying that someone is *giving up the ship,* she conveys that something is being abandoned (literal or approximate use of *give up*); in saying that someone is *crying their eyes out* she conveys that someone is extremely upset (hyperbolic or metaphoric use of *crying*); in saying that someone has *broken the ice,* she may want to convey that something (e.g. silence or tension) is being overcome (approximate or metaphoric use of *break*), and so on. Crucially, pragmatic fine-tuning would also operate at phrase level, so that the whole string can be loosely (e.g. hyperbolically or metaphorically) understood, as in the idioms *to lose one's nerve, to bury the hatchet, to jump down someone's throat, to slip one's mind* and *to change one's mind.*

One of the reasons why idioms are of interest to pragmatics is in fact that the encyclopaedic assumptions associated with the encoded concepts point the hearer towards the right fine-tuning required to make the utterance relevant in the expected ways. Acquiring an idiom involves sorting out the best encyclopaedic assumptions and best inferential route involved in fine-tuning the idiom in the right direction. In interpreting a familiar idiom one simply follows this familiar inferential route in order to construct a hypothesis about the speaker's meaning. In the next section I will look at how people interpret variants of idiomatic strings and at the role that accessible encyclopaedic assumptions and inferential routes may play in the comprehension process.

6. Interpreting Idiom Variants

Evidence of the intimate relation between the literal and the idiomatic meaning of idioms is provided by examining real occurrences of these expressions in everyday use. Corpus research has shown that many idioms allow a considerable degree of lexical flexibility, as illustrated by the examples below:

British English - most examples from corpus research in Moon (1998a; 1998b):

(45) *Have/keep/be with your feet on the ground*

(46) *Get/have/keep your eye on*

(47) *Burn your boats/bridges*

(48) *Hit the roof/ceiling*

(49) *Take the biscuit/cake*

(50) *Throw/toss/chuck in the towel/ the sponge*

(51) *Sweeten/sugar the pill*

(52) *Lower the guard/let your guard down*

(53) *Drag your feet/heels*

(54) *Take something with a grain/pinch of salt.*

(55) *Come/fall down to earth*

(56) *Get out of bed on the wrong/right side*

(57) *Fight/defend tooth and nail*

(58) *Walk/tread on air*

(59) *Start/keep the ball rolling*

Spanish idioms (examples from dictionary of Spanish idioms, Buitrago, 2002)

(60) *Poner/colocar/tener (a alguien) contra las cuerdas* (to put/place/have (someone) against the ropes).

(61) *(No) echar/lanzar campanas al vuelo* (to (not) throw/throw bells flying -> (not) to announce good news, generally a bit too soon).

(62) *Abrir/cerrar el pico/la boca* (to open/close one's beak/mouth -> to speak/to shut up)

(63) *Dejar/quedarse/estar en la estacada* (to leave/to be left/be in the stockade -> to be abandoned when needing help the most)

Although the syntactic flexibility of idioms has been of some interest to linguists, examples of lexical flexibility such as those presented here have not received much attention. The existence of lexical variability in idiomatic strings is only to be expected given the approach defended here. I have argued throughout this work that most idiomatic strings are to some extent transparent, in that their meaning can be at least partly inferred from the encoded concepts and the compositional meaning of the phrase. Selective processing of the assumptions more accessible by the encoded concepts yields some implications which may be attributed as part of the speaker's meaning. Altering the words used may therefore lead the hearer to consider slightly or substantially different implications.

Given this account, there may be at least two reasons for a speaker to use a lexical variant. First, one word may be substituted for another because it makes roughly the same contribution to the overall meaning yielding roughly the same implications and does not cause the hearer any more processing effort. Second, one word may be substituted for another because it gives access to different assumptions, which yield implications not derivable (or not derivable with the same degree of strength) from the original form and which offset any extra effort required to derive them. Let's consider these possibilities in turn.

The nearly synonymous Spanish verbs *echar, tirar, lanzar* (all of which can be translated into English by the verb *throw)* may be used indifferently in some idioms, as in (64):

(64) a. *Echar/tirar la casa por la ventana* (to throw the house out of the window -> to make big expenses in a not very sensitive way)

b. *Echarlo/tirarlo/lanzarlo todo por la borda* (to throw everything overboard -> to ruin everything)

c. *Echar/lanzar campanas al vuelo* (to throw bells flying -> to announce good news, generally a bit too soon)

d. *Echar/lanzar el anzuelo* (to throw the hook -> to do something to trick someone)

My suggestion is that the reason why some of these verbs may be used interchangeably in (64) is that, in these cases, the concepts they encode make roughly the same contribution to the meaning of the idiomatic expression and the relevance of the utterance. For instance, in (64a), the verbs *echar* and *tirar* may be used to convey narrowed concepts which share roughly the same denotation: that is, they are used to indicate roughly the same type of throwing and thus warrant roughly the same implications. The words themselves are not exact synonyms (which are difficult to find in a language). The concepts they encode would therefore give access to different encyclopaedic assumptions, and would denote slightly different types of throwing. The verb *lanzar* (like the English verb *fling*), for instance, indicates a certain kind of gesture and a certain kind of movement through the air that is not required by the verb *tirar*. In idiomatic uses intended to indicate this particular type of throwing (e.g. the flinging of something in the air, as in (64c) and (64d)), the verb *lanzar* is often preferred.

I would claim that the use of an idiom variant is motivated by the sort of assumptions made most accessible by the encoded concepts, and the way in which these assumptions contribute to (or modify) the overall idiomatic interpretation. Since different words may encode concepts which make roughly the same contribution to overall meaning, all these uses may become standardised, or even lexicalised, as in (47)-(49), (51), (58), (62), (64), etc. Sometimes, different speakers, dialects or cultural groups prefer one use to another (as in American and British variants). Provided that these uses are equally easy to process, they may all achieve relevance in roughly the same way, yielding roughly the same implications.

The choice of one word (e.g. one verb) rather than another may be intended to point the hearer towards slightly different encyclopaedic assumptions and so to direct him along a (slightly) different inferential route. In (60), for instance, while the three verbs give access to a range of similar implications (e.g. about impotence, danger, etc.), they also yield some rather different implications. A change of perspective, for instance (e.g. whether someone is *put/placed/thrown against the ropes* or *has someone against*

the ropes) may lead to certain specific implications. Also, the use of the verb *place* versus *throw* may yield different implications (e.g. about aggressiveness) or the same implications of differences of strength. Different implications may be conveyed by saying that someone '*has her feet on the ground'* rather than in saying that she 'needs to *keep her feet on the ground'*. Different surface forms may encourage the hearer to narrow the interpretation by focusing on different aspects of the situation the idiom is generally used to describe, as in utterances such as: 'I will *set the ball rolling*', '*the ball is rolling'*, 'we need *to keep the ball rolling'*, all three of which may be used in the same conversation, or even the same utterance. On some occasions, the implications derived from a variant may even be the opposite of those which would have been derived from the more standard use, as in: 'he *got out of bed on the right side'*. In processing this utterance, the hearer may use the highly accessible assumption that right is the opposite of wrong, and adjust the implications so that they are the opposite to those that would have been conveyed by the original form (i.e. *to get out of bed on the wrong side*). The substitution of a word by an antonym is indeed a common type of idiom variant, as in (57) and (62).

I want to propose that idiom variants lie along two continua which often run in parallel. On the one hand, they differ in the extent to which the assumptions they make accessible alter the interpretation slightly or substantially. On the other hand, they differ in the extent to which the particular meaning they convey is one-off, or standardised, (and even lexicalised). At one end of these spectra of novelty, we find cases in which the original idiom meaning is modified with both the surface form and their interpretation being rather standardised, as in (64). At the other end, there are cases in which the modification is greater and may involve the hearer treating the variant as an allusion to the stored representation, as in (65)-(73). A wide range of cases, such as many of those in (45)-(63) and those in (74)-(79) may fall somewhere in-between:

(65) He is very stubborn, but in the end he will have to *change his square mind* and accept the deal.

(66) I am fed up with this situation, let's just *throw the cards on the table* once and for all!

(67) During the meeting *all the cards were laid on the table.*

(68) Despite the torture, he didn't *spill a single bean.*

(69) He absolutely hates me, so if it is true he has found out about my affair, he must now be in my house *pouring the beans* to my wife.

(70) OK there! Now you are *barking up the right tree!*

(71) (Teacher to student) I think *we are barking at different trees.* (attested)

(72) (Teacher to student) We *are on different trains of thought.* (attested)

(73) Sin darme cuenta, *me metí de cabeza en la boca del lobo.*
Without realising, I *got head first into the wolf's mouth.*
Meterse en la boca del lobo (to get into the wolf's mouth) –> to get into a problematic or dangerous situation.

Other attested variants in the corpus (from Moon, 1998a) include:

(74) *Add fuel to the fire/throw fuel on the fire.*

(75) *Put someone off the scent/throw someone off the scent.*

(76) *To pass the buck/the buck stops here/the buck passes somewhere.*

(77) *Another nail in the coffin/a final nail in the coffin/to nail down the coffin/to drive the first nail into the coffin.*

(78) *The writing is on the wall/to see the writing on the wall.*

(79) *To let the cat out of the bag/the cat is out of the bag.*

Psycholinguistic research on lexical flexibility has generally been most interested in cases falling somewhere towards the creative end where the substitution makes a rather clear difference to meaning, as in (65)-(73). These are sometimes referred to as cases of semantic flexibility (see Cacciari and Glucksberg, 1991; McGlone, Glucksberg and Cacciari, 1994). My aim here is to propose a unified approach to idiom comprehension which can account for idiom variants falling anywhere along the continua just proposed. Having described how rather standardised and 'semi-standardised' idiomatic variants are understood, I will go on to consider more creative cases, such as those in (65)-(73).

The comprehension of more creative variants proceeds, I argue, in the same way as the comprehension of less creative cases with the concepts encoded by the words in the string and the string as a whole simply taken as cues to infer the speaker's intended meaning. Let's look at the example in (65). The speaker in (65) seems to have blended two different standardised metaphorical uses: *to have a square mind* and *to change one's mind.* Although different inferential routes may be taken to process this expression, let's assume the hearer accesses the concepts encoded by these familiar phrases following a path of least effort, and starts accessing some of their associated encyclopaedic assumptions. He may, for instance, consider assumptions about rigidity of thinking associated either to the loose use of the word *square* or to the phrase *square mind* and combine them with encyclopaedic assumptions from the concept [TO CHANGE ONE'S MIND]*. The result may be the derivation of a range of implications which cannot be derived, or cannot be derived with the same strength, from using the original form *to change one's mind.* These include, for instance, implications about the degree of stubbornness of the person being talked about, the amount of effort that needs to be invested in making him consider alternative lines of thinking, etc.

Let's now look at (69). The similarity in form of the phrase *pour the beans* with the familiar phrase it alludes to (*spill the beans*) together with the expectations of relevance generated by the speaker's utterance may make the concept [SPILL THE BEANS]* highly salient at some point in interpretation. Both this concept, as it becomes available, and the concepts encoded by the words in the utterance would be taken by the hearer as cues to infer the ad hoc concept the speaker intended to express. Following a path of least effort, the hearer may consider, for instance, the assumption that the act of pouring generally entails some degree of intentionality (e.g. pour water in a glass) and combine it with assumptions accessed from the concept [SPILL THE BEANS]* (e.g. 'if someone spills the beans then someone reveals hidden information'). The result from this combination may be the derivation of a range of implications, such as the implication that the revelation of the hidden information was intentionally performed, and the construction of an ad hoc concept [TO POUR THE BEANS]* which warrants the derivation of these implications. Although these implications might have also been derivable from the speaker's use of the original idiom in this utterance, they would not

have been derived with the same degree of strength. The use of the idiom variant thus leads the hearer in the right direction towards the derivation of the intended implications.

As proposed in Chapter Five, the existence of idiom variants is good evidence that the meaning of idiomatic expressions is not entirely arbitrary but at least partly inferable from the meaning of their parts. In (71), for instance, substituting the word *different* for the word *wrong* may result in the hearer not accessing the assumption that someone has made a mistake even though this is an assumption that would have been highly accessible had the speaker used the original form. Similarly, the hearer in (72) may not derive implications about something being forgotten even though these implications would have been rather strongly implicated had the speaker used the idiom in its original form (*to lose one's train of thought*). These examples suggest, once more, that in interpreting an idiom, the hearer processes the encoded concepts in the string, at least to a certain degree. A change in the constituents that compose the expression may lead the hearer not to consider some of the assumptions which would have been highly accessible had the original form being produced.

My proposal, then, is that all cases of idiom variants, from the rather straightforward to the more creative, are understood by the same comprehension procedure just described, where the concepts encoded by the words in the string and the whole phrase the variant alludes to are used as input to pragmatic inference. As in the comprehension of any utterance and any word, the comprehension of idiom variants follows a path of least effort with the hearer investing only the necessary effort in processing the encoded concepts. The assumptions resulting from the rather shallow processing of the string can often be integrated with other contextual assumptions being considered in processing the utterance to derive the set of intended implications. As a result of this process, a novel ad hoc concept is generally formed by adjusting the concept associated with the original idiom in the ways, and to the point where, it can warrant the derivation of these implications.

Generally, the more a variant departs from the original form, the more processing effort the hearer may need to invest in searching for the right set of assumptions and inferential route. The acceptability of the novel use would depend on whether this investment of extra processing effort is offset by extra or different cognitive effects,

which could not have been more economically conveyed by the use of another string. The reason why idiom variants such as those in (80)-(82) may not be generally acceptable is that the encoded concepts (GROUND, LENTIL and BULL) would not normally give access to encyclopaedic assumptions which can help in deriving extra or different cognitive effects.

(80) During the meeting, *the cards were laid on the ground.*

(81) Despite the torture, he didn't *spill the lentils.*

(82) Sin darme cuenta, *me metí en la boca del toro.*

Without realising, I *got into the bull's mouth.*

Meterse en la boca del lobo (to get into the wolf's mouth) – to get into a problematic or dangerous situation.

Combining these assumptions with the other assumptions made accessible by the concept encoded by the whole string may indeed yield some implications, but not of a type that the speaker could possibly have intended to offset the extra effort involved. Examples like these may be classified as errors typically made by children and second language learners.

7. Conclusion

Throughout these last two chapters, we have seen how idioms often behave like other linguistic expressions in being able to take a variety of forms by undergoing syntactic and semantic transformations. In fact, some experimental findings suggest that idioms are linguistically processed just like any non-idiomatic strings. On the one hand, it has been shown that the literal meanings of the words in an idiom are immediately activated as the idiom is heard, and remain activated during interpretation (Cacciari and Tabossi, 1988). On the other hand, it has been found that the syntactic parsing of idioms takes place as normal, even after the string is recognised as idiomatic (Flores d'Arcais, 1993). One position that has been taken in the literature, and which I have implicitly supported here, is that whether a syntactic or semantic variant of an idiom is understood as

idiomatic or is taken literally is not a matter of grammar but involves a decision taken on pragmatic grounds (e.g. Ackema and Neeleman, 2001; Flores d'Arcais, 1993; Geeraerts, 1995; Peterson and Burgess, 1993; Van de Voort and Vonk, 1995). Another assumption which has been explicitly or implicitly made in the literature, including the work presented here, is that the relation between the literal compositional meaning and the idiomatic meaning is not entirely arbitrary.

Supporters of the Configuration Hypothesis have generally been characterised as pursuing this last idea. Cacciari and Glucksberg (1991), for instance, attempt to capture the degree of compositionality or analysability of idioms in the following typology:

(83) a. **Analysable-Opaque idioms:** the relation between the idiom's elements and the idiom meaning may be opaque, but the meanings of the words can still constrain the interpretation and use of the string (e.g. *kick the bucket*, where *kick* denotes an abrupt action).

b. **Analysable-Transparent idioms:** there is a clear semantic relation between the elements of the idiom and the components of the idiom's meaning (e.g. *break the ice* where *break* corresponds to the idiomatic sense of changing mood and *the ice* to the idiomatic sense of social tension).

c. **Quasi-Metaphorical idioms:** these are idioms for which the literal referent is itself an instance of the idiomatic meaning (e.g. *give up the ship* describes both a prototypical example of the act of surrendering and the act of surrendering in general).

Although this typology does not make a clear distinction between decomposition and transparency, it seems to be grounded on the assumption that hearers are virtually always capable of inferring a relation (however minimal) between the meanings of the words in the idiom and the meaning of the overall idiomatic expression. However, this idea (originally in Nunberg, 1978) is no more than an interesting intuition unless a proper inferential account can be provided. This account would need to explain how some of the assumptions made accessible by the encoded 'literal' meanings of the words can be selected and used as premises in an inferential process whose output

would be the idiomatic interpretation. Much of this chapter has been dedicated to examining this process in detail.

Psycholinguistic researchers on idioms have often concluded that the fact that the meanings of the words in an idiom are activated and remain activated during comprehension (e.g. Cacciari and Tabossi, 1988) suggests that these activated meanings can be used in processing the string in both its original and variant forms (e.g. McGlone, Glucksberg and Cacciari, 1994). The main problem with existing psycholinguistic approaches, including the Configuration Hypothesis (e.g. Cacciari and Tabossi, 1988; Cacciari and Glucksberg, 1991), is that they don't actually show how word meanings contribute to the derivation of the intended interpretation. What the psycholinguistic literature on idioms desperately needs is strong pragmatic constraints and powerful inferential machinery which could bridge the gap between the compositional meaning of the idiom and its idiomatic meaning and help to determine the extent and depth to which the encoded concepts are processed in comprehension. The main aim of the pragmatic account of idioms I have presented here is to complement experimental research by accounting for just this process.

Notice that the typology in (83) suggests that the approach defended by Cacciari and Glucksberg (1991) shares with the relevance-theoretic approach to idioms defended here the idea that some idioms (which they refer to as 'quasi-metaphorical') are understood by taking the idiom meaning to be an extension of the compositional meaning of the phrase. Having proposed the Class-Inclusion theory of metaphor, Glucksberg's aim here is to extend this analysis to the comprehension of certain idioms (Glucksberg, 1991, 1993, 2001). However, there are two important problems with his approach. On the one hand, lacking adequate inferential machinery to explain how hearers construct the idiomatic meaning from the compositional meaning, this approach to idioms suffers from the same shortcomings as the Class-Inclusion theory of metaphor (discussed in Chapter Three). On the other hand, by assuming that only quasi-metaphorical idioms, and not any of the other idiomatic strings in (83), are understood by a process of meaning extension, these scholars fail to account for the continuum of cases which their own approach to the degree of analysability, as in (83), is designed to capture. A better approach, such as the one presented here, should aim to provide an

inferential account where the whole continuum of idioms (from the rather transparent to the very opaque) are understood using the same comprehension procedure. In this chapter, I have tried to present such an account.

We may conclude from the arguments presented in these two chapters that only an adequate inferential approach to idioms can actually account for what is, arguably, their most crucial feature: that is, that their idiomatic meaning is not entirely arbitrarily stipulated but partly derivable from the compositional meaning of the phrase. The relevance-theoretic approach to idioms I have presented here is designed to complement the experimental literature by showing how the selective relevance-driven processing of the encoded concepts guides the hearer at every point to follow the most accessible inferential route in deriving the overall (idiomatic) meaning. The different layers of inference which make up this inferential route may be gone through very fast in the interpretation of familiar idioms, or be modified in the interpretation of idiom variants. In either case it is the set of assumptions and computations which the hearer uses in interpreting the string which help to bridge the gap between the encoded 'literal' meaning of the phrase and the resulting idiomatic interpretation and which may help the hearer perceive the expression as relatively transparent.

Chapter 7

Conclusions:

Creativity and Convention beyond Figurative Speech

1. Introduction

Much of the work in this thesis has been dedicated to studying the creative processes involved in the comprehension of ordinary utterances, particularly where this involves the construction of new mental representations. I have shown how selective processing by a relevance-driven cognitive system often results not only in the creation of new mental representations (e.g. one-off concepts and one-off thoughts) but also in the development of pragmatic procedures for the efficient construction of those representations. In this chapter, I want to look at the wider cognitive picture so as to show that these processes are not only at work in the comprehension of idioms or familiar metaphors, or even in the comprehension of everyday utterances. Instead, the selective processing and development of inferential routines involved in utterance interpretation reflect a tendency of the mind to process information in the most productive way and to develop methods for minimising the effort required to process familiar stimuli.

2. The Psychology of Routines

2.1 Controlled and Automatic Processing

First proposed by Schneider and Shiffrin (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977, but see also Shneider and Chein, 2003), a distinction is generally made in cognitive psychology between controlled and automatic processing. Controlled processing is characterised as a slow, conscious and voluntary process, which requires high levels of attention and a considerable amount of time and effort. Automatic processing, by contrast, is fast, involuntary, unconscious, effortless and requires little or no attention. The most widely discussed example of a process that starts out being fully

controlled but becomes automatised with practice is the ability to read. Humans are not born with the ability to read; they need to develop it. As children, we are taught techniques to use in reading. The process at this early stage is slow, conscious and requires a considerable amount of effort and attention. After enough practice, it becomes automatised and so can be performed fast, effortlessly, unconsciously and involuntarily.

Evidence for the automaticity of reading skills is provided by the Stroop effect discussed briefly in chapter two. The best-known experiment involved presenting subjects with a list of names of colours (e.g. GREEN, BLUE, RED) in a colour that is either consistent with this name (e.g. the word *green* written in green) or inconsistent with it (e.g. the word *green* painted in blue). Subjects were asked to name the colour in which the word was written. No difficulty in performing the task was found in the case where the colour of the word and the colour named by the word coincided. However, when they differed, interference from reading was found. Subjects found it hard to name the colour of the words because they involuntarily and unconsciously read the words.

A common explanation for the Stroop effect has been that word reading is automatic while colour naming is controlled (Posner and Snyder, 1975). A problem with this conclusion is highlighted by MacLeod and Dunbar (MacLeod and Dunbar, 1988). These scholars carried out an experiment in which subjects had to use colour names to refer to a small number of shapes (e.g. 'green' for a square shape). After 288 trials, they could perform this naming task fast and without much effort. Having established these arbitrary name-shape pairs, they were presented with the same shapes either in a colour which was consistent with their name or in another colour. When the arbitrarily stipulated colour and the real colour were the same subjects had no problems in saying the name arbitrarily assigned to the shape. However, when the arbitrarily stipulated colour and the real colour differed, they could not help naming the real colour of the shape they were looking at. These results show that the conclusion drawn from Stroop's experiment, that colour naming is a controlled process, cannot be right. Colour naming in MacLeod and Dunbar's experiment proved to be a rather automatic process which subjects could not help but perform fast, effortlessly, unconsciously and involuntarily. In order to resolve this inconsistency, one might need to treat controlled and automatic

processing as extreme points on a continuum, with different skills enjoying different degrees of automaticity. Cohen, Dunbar and McClelland (1990) propose a gradual approach to the controlled-automatic dichotomy, in which reading is seen as more automatised than colour naming, and colour naming as more automatised than shape naming. This hierarchy of automaticity explains why colour naming appears to be a controlled process in Stroop's experiment but an automatic process in McLeod and Dunbar's.

2.2 Automaticity and Expertise

The development of high levels of automaticity often seems to run parallel to the development of expertise. Examples of fast, almost unconscious and effortless processing include the way that doctors interpret x-rays accurately and bird watchers identify species after only a quick glance. A common psychological explanation for these phenomena is that experts often develop (over time, and after enough practice) the ability to detect a set of richer perceptual features than the novice can detect. This ability to spot fine-grained perceptual features seems to result from repeated selective processing. Medical trainees and bird watchers learn to focus on a set of selected (diagnostic) features when looking at x-rays and birds, respectively. Paying systematic attention to a selected set of features drawn from those available at the time often results in the development of a 'perceptual' procedure which pins down, fast and effortlessly, the set of potentially relevant features (e.g. dark spots on a person's lungs) which will allow a fast and accurate diagnosis. Recurrent selective processing may lead to the development of an ability to detect fine-grained distinctions between apparently similar stimuli, and to do so after very limited (perceptual) exposure to these stimuli.[1]

Experimental research has shown that the development of expertise need not depend so much on a natural gift or high levels of intelligence as on a great deal of dedication and practice (Ericksson, Krampe and Tesch-Romer, 1993). In aiming to discover the ingredients needed to achieve mastery in chess, DeGroot (1965, 1966) presented expert and average players with a chess board in one of two conditions. In one

[1] See Horsey (2002) for an interesting study on the relation between the development of heuristics to detect diagnostic features (e.g. in bird watching, chicken sexing, etc.), the ability of humans to detect members of a category and the relation between intuitive and reflective concepts.

condition, the pieces were presented in a real configuration (e.g. as in the middle of a game). In the other condition, the pieces were placed at random. In each case, subjects were allowed to look at the board for just 5 seconds and asked to reconstruct what they had seen. In the configuration condition, chess masters were capable of reconstructing more than twenty pieces, while weaker players could only remember four or five pieces, these last numbers being the ones expected given the capacity of our working memory. Interestingly, when presented with the same pieces at random rather than in a configuration, chess experts and weaker players were equally bad, being only able to recognise four to five positions. What these findings suggest is that the short-term memory span of expert players is capable of storing only as much information as that of an average player. So why did they outperform average players in the configuration condition? One possible answer is that chess masters have developed the ability to interpret stimuli at a higher level. Just as children at some point develop the ability to see words as units rather than as individual letters, chess masters develop the ability to see several chess pieces together as forming patterns which they have encountered before. That is, the reason why they can remember an average of twenty positions in the configuration condition is that they are remembering an average of four or five patterns of four or five pieces each.

Systematic exposure to the game often results in chess masters storing thousands of configurations which they have encountered in real situations, along with knowledge of how to act when encountering these configurations (Newell and Simon, 1972; Simon and Gilmartin, 1973). We can assume that repeated exposure to certain patterns leads chess players to perceive a higher-level configuration which is recognised as a problem-solving task (i.e. a task which consists in solving a problem), and that experience in solving the problem leads to the development of a certain procedure which guides the way they act when presented with these familiar configurations. An important point to notice from the experiment described above is that for chess masters to reconstruct the board, they must not only be able to perceive a higher-level configuration in the

available stimuli, but must also be simultaneously aware of the lower-level elements (chess pieces) that make up each of those configurations.[2]

The evidence presented here suggests that selective attention plays an important role in the development of fine-grained abilities such as those involved in the detection of diagnostic features and the construction of higher-level representations, and of mental routines for processing this selected information. In the next section, I will look at the stages in this development: that is, at the different levels of controlled processing appropriate to novices, and the different levels of automaticity involved in the acquisition of expertise.

2.3 Stages in Expertise Development and Degrees of Automaticity

We have seen that, although a certain task (e.g. reading, analysing x-rays, playing chess) may involve a considerable amount of effort, time and attention, systematic experience with this task often leads to the development of a range of procedures which direct our attention to potentially relevant information. That is, based partly on previous experience, our minds automatically select information which is likely to be relevant to us in performing a familiar task, and direct our attention towards it. Selective attention often results in increasingly automatic processing. As noted above, controlled, slow and conscious processing and fast, effortless and largely unconscious processing are at two ends of a continuum. The development of everyday tasks in infancy or in adulthood (e.g. reading, writing, riding a bike, tying one's shoelaces, etc.) involves movement along this continuum in the direction of increasing automaticity. Looking at research on acquisition of cognitive and motor skills may, therefore, help us understand the stages a novice goes through in becoming an expert, and hence the stages involved in the movement from controlled to automatic processing.

Following Fitts and Posner (1967), many scholars have regarded the development of a new skill as involving three main stages. The first is the 'cognitive stage', in which people are seen as memorising a set of facts about the task to be performed. So in

[2] The similarities with language and the comprehension of familiar strings of words such as idiomatic phrases are important. As I have shown in Chapter Six, the hearer of a familiar idiomatic phase generally constructs a concept at phrasal level while still having access to the concepts that compose this phrase and to an inferential routine which links them both.

learning to ride a bike, the learner memorises where to place the feet, how to hold the brakes, etc. In learning to drive a car, he memorises how to change to each different gear, etc. This information is generally explicitly taught to the learner, who stores each piece of information individually and accesses it slowly, investing a considerable amount of effort and conscious attention. It is also generally agreed that the information, at this stage, is still in declarative form and is therefore accessed separately and not yet compiled into a procedure.[3] After enough time and practice, the learner enters what is described as the 'associative stage'. At this point, he has noticed and corrected errors made during earlier stages of the learning process, and would avoid following paths that have previously led to an unsuccessful result. Crucially, he has learned to strengthen the connections between the steps involved in the process, which is now in a procedural form. Once he has developed a particular procedure for performing the task, and has gained enough practice with this procedure, he may reach the 'autonomous stage'. There the procedure becomes increasingly automatised, so that the time, effort, and degree of conscious attention required to perform the task are gradually reduced. Reaching this stage makes it possible for the learner to perform another activity at the same time (e.g. talking to a friend while riding a bike, thinking about train schedules while tying his shoelaces, etc.).

According to this model, skill acquisition depends very much on the move from controlled to automatic processing. In learning how to drive, how to find the way to a friend's house, how to cook an omelette, etc. we go through a series of steps which get compiled into a procedure. After considerable experience, this procedure is performed fast and almost unconsciously, with no need to pay attention to the individual representations that make it up. Karmiloff-Smith (1996) suggests that although the movement from conscious to unconscious processes plays an important role in a person's development, child development sometimes proceeds in the opposite direction, involving a move from unconscious automatic processing to conscious reflection.

Indeed, the two types generally complement each other. To show this, Karmiloff-Smith considers the development of piano playing. In the first few stages of developing

[3] By declarative knowledge, cognitive psychologists generally refer to the description of facts. By procedural knowledge, they refer to the description of a series of steps to be performed in doing something.

this skill, people may be seen as going through the stages proposed by Fitts and Postner. At first, in learning to play a song, each individual note is represented and attended to separately, as the information is still in declarative form. At some point, patterns made up of several notes begin to emerge, until eventually the whole song is played as a piece. When the performance is generated by procedural knowledge, it is difficult, if not impossible, for the learner to start playing the piece in the middle, or to continue playing after being interrupted. Instead, the whole piece must be run off as a unit. Many average players do not get past this stage. According to Karmiloff-Smith, true mastery involves a process of what she calls 'representational re-description', in which the pianist again becomes aware of the internal composition of the piece: that is, of the notes and chords and the order in which they are to be played. This awareness enables him to manipulate this information and so produce variations on a piece, introduce sections from other pieces, etc.[4] According to this view, then, two types of processes are involved in the development of this skill: a process of proceduralisation and a process of explicitation:

> "[D]evelopment and learning, then, seem to take two complementary directions. On the one hand, they involve the gradual process of proceduralization (that is, rendering behaviour more automatic and less accessible). On the other hand, they involve a process of "explicitation" and increasing accessibility (that is, representing explicitly information that is implicit in the procedural representations sustaining the structure of this behaviour)" (Karmiloff-Smith, 1996: 17)

According to Karmiloff-Smith, children's development crucially involves the latter process: that is, it involves the ability to "turn implicit information into explicit knowledge" (ibid).[5] Having performed a certain task successfully largely unconsciously (e.g. producing a certain linguistic form, or managing to get a toy working, etc.), a child

[4] As shown for chess players, reaching this state of mastery allows the individual to have simultaneous access to the higher representation and to the components that form that representation.

[5] Although the distinction between explicit and implicit learning is not synonymous with the distinction between controlled and automatic processing, it is highly related to it and often runs in parallel. Explicit learning takes place consciously as we are taught something. Explicit knowledge is knowledge which results from explicit learning: e.g. the knowledge that there are five continents. Many of the things we know, we have acquired implicitly, without conscious awareness. A child's acquisition of the grammar of his mother tongue is generally seen as a typical case of implicit learning. In contrast, learning the grammar of a foreign language via explicit teaching at school is generally characterised as an instance of explicit learning.

is seen as often attempting to finding out the stages he followed to succeed in the performance and learns as a result. It is this balance, between a) trying to find a solution to a task by searching for the right procedure to perform it (e.g. processing an utterance, solving a problem) and b) trying to see how a possible solution might have been derived, that seems to play a role in a number of cognitive tasks during child development.

2.4 Conclusions on the Psychology of Routines

The work on psychology presented in this section, like that presented in Chapter One, provides good evidence for the relevance-theoretic claim that our minds have evolved in the direction of increasing efficiency and tend to allocate attention and processing resources to the information which seems most relevant at the time. In this chapter, I have shown that systematic focus on selected information and procedures which are likely to contribute to relevance in processing a familiar stimulus (e.g. performing a familiar task) may often result in the development of a certain routine. This goal-directed attention to selected features and procedures often results in the emergence of fine-grained abilities and the construction of higher-level representations which allow the individual to process and categorise stimuli fast and accurately.

It therefore seems plausible to conclude that the development of cognitive routines for processing familiar stimuli is quite widespread, and does not happen only in utterance interpretation. Like the construction of ad hoc concepts and ad hoc categories, the emergence of cognitive procedures is a by-product of the tendency of the human mind to maximise relevance. It is important to notice, though, that while performance on other cognitive tasks (e.g. those presented above) involves movement along the continuum from controlled to automatic, utterance comprehension is an automatic process from the start. The recurrent processing of a familiar stimulus leads to increasingly automatic selection of the sort of assumptions and procedures likely to yield a satisfactory interpretation. The development of a pragmatic routine for the processing of familiar constructions (e.g. the use of formulaic language), may allow the hearer to speed up (or shortcut) the steps in the inferential process, minimising the processing effort invested in deriving implications.

3. Creativity and Convention in Language

According to Chomsky, the use of language is a creative act. This creativity is two-fold. On the one hand, natural languages are generative in that we can, on the basis of a finite number of words and rules, construct an indefinite number of sentences, most of which had never been produced before. On the other hand, in creating these sentences, we can convey an indefinite number of thoughts. "An essential property of language is that it provides the means for expressing indefinitely many thoughts" (Chomsky, 1965: 6). However, these assumptions about generativity contrast with the systematic use of pre-fabricated chunks of language during spoken and written discourse. The pervasiveness of formulaic language has sometimes led scholars, particularly lexicologists, to challenge the Chomskyan view of language (Moon, 1998; Wray, 2002):

> "Communicative competence is not a matter of knowing rules for the composition of sentences and being able to employ such rules to assemble expressions from scratch as and when occasion requires. It is much more a matter of knowing a stock of partially pre-assembled patterns, formulaic frameworks, and a kit of rules, so to speak, and being able to apply the rules to make whatever adjustments are necessary according to contextual demands" (Bolinger, 1975: 297).

The idea seems to be that even when it is possible for us to construct a novel sentence or a novel phrase to convey a particular meaning, we often choose to rely on formulaic language which we have used and processed repeatedly. So, although the English language allows speakers to form perfectly grammatical sentences such as 'have a wonderful anniversary of your birth' or 'have a nice remembrance of the day your mother delivered you', it is rather unlikely that native speakers would choose to congratulate a friend on his birthday by uttering these expressions. Instead, they would prefer to use the very familiar string 'happy birthday' or even (the less frequent) 'many happy returns (of the day)'. Repeatedly heard and repeatedly used constructions such as these are often preferred not only because they are easier for the speaker to produce and for the hearer to understand but also because they are typically judged as more 'idiomatic' (i.e. native-like). Choosing other expressions (e.g. those above) instead may be seen as an indication that the speaker doesn't have a good command of the English

language. She may be a foreigner translating into English the expression standardly used in her mother tongue to congratulate people on their birthdays.

A number of scholars have argued that the pervasive use of formulaic language is largely for economy reasons (Kuiper, 1996; Nattinger and DeCarrico, 1992; Perkins, 1999; Wray, 2002). "The main reason for the prevalence of formulaicity in the adult language system appears to be the simple principle of economy of effort." (Perkins, 1999: 56). Pursuing the idea that our cognitive systems have developed in the direction of efficiency and minimisation of processing effort, it makes sense to think that language users may end up repeatedly using a small set of familiar expressions to convey a certain set of implications. If both a familiar expression and a novel one would achieve the same cognitive effects, but using the familiar expression minimises the hearer's processing effort, a speaker aiming at optimal relevance should choose the former. Formulaic expressions arise naturally in every language and in every culture, showing the tendency of language users to store ready-made expressions and use them to communicate. The existence of "conventionalised" expressions such as routine formulae, including similes (e.g. *as white as a sheet*), compounds (*spick and span*), proverbs (*a stitch in time saves nine*) or idiom strings (*break the ice*) has been taken as evidence that, although linguistic communication may involve a great deal of creativity, it also involves a great deal of convention or standardisation. An adequate account of utterance use and interpretation should allow for a balance between creativity and convention, and show how it affects the way people communicate their thoughts and interpret the utterances of others.

3.1 Internal Composition

One interesting observation about language use is that many of the words and word strings used in everyday language are not internally analysed by the hearer during the interpretation process. These uses are often referred to as formulaic (Wray, 2002). On one occasion, Kellogg, the cereal manufacturer, set up a test in which people were asked what Rice Krispies were made of, and found that people were generally surprised to be told they were actually made of rice! According to Wray (2002: 3), the reason for this lack of awareness is that people do not generally derive the meaning of the phrase on the basis of the meanings of its component words but seem to bypass those meanings

completely. The example is far from unique: many of the words and word strings we use every day are composed of words which have a meaning of their own but which we rarely pay conscious attention to. I have observed in Spain how restaurant menus often translate Spanish dishes strictly literally: so 'ropavieja' (meat with potatoes and chickpeas) is translated as 'dirty clothes' and 'papas arrugadas' (salted boiled potatoes) as 'wrinkled potatoes'. The word-for-word translation sounds shocking to Spaniards, simply because we do not generally compute the literal meaning of these phrases. Many of the words we use on a daily basis in our native languages have meanings other than the ones their linguistic form seems to indicate:

> "It is now time to face the fact that English is a crazy language [...]. In the crazy English language, the blackbird hen is brown, blackboards can be green or blue, and blackberries are green and then red before they ripen. [...] To add to this insanity, there is no butter in buttermilk, no egg in eggplant, no grape in grapefruit, no bread in shortbread, neither worms nor wood in wormwood, neither mush nor room in mushroom, neither pine nor apple in pineapple, neither peas not nuts in peanuts, and no ham in hamburger. [...] And we discover even more culinary madness in the revelations that sweetmeat is made of fruit, while sweetbread, which isn't sweet, is made from meat" (Lederer, 1989)[6]

Generally speaking, I would think these formulaic uses may be placed alongside opaque idioms or relatively opaque idioms, in that the meaning of the parts does not contribute to the derivation of the intended meaning. The depth of processing of the components in a transparent idiom (e.g. *miss the boat)* may be deeper than the processing of the components of words such as 'eggplant' or 'Rice Krispies' because they help with deriving the intended meaning in a way that the meaning of 'egg' or 'rice' do not.

Although, puns, cartoons and jokes may exploit the double senses of the words in the quote above, they are often not exploited in everyday conversation. The reason for this may not be that speakers and hearers are unaware that the words are composed of meaningful units, but that the processing of these units may not be necessary to derive the interpretation the speaker intended to convey by using those words on a particular occasion. In the previous chapters, I showed how the relevance-driven

comprehension of utterances generally involves the selective and relatively shallow processing of available information, so that only those activated assumptions which are likely to contribute to the relevance of the utterance would be processed. Since hearers often arrive at a satisfactory interpretation after processing the words in an utterance in a relatively shallow manner (e.g. after recognising that 'blackbird' is a type of bird, and 'blueberry' a type of berry), they may not need to consider further information made accessible by these words, their homonyms or the words they are composed of. On some occasions, however, as in understanding a pun, hearers may need to look more deeply into the encoded concepts, or other concepts activated by those concepts, so as to derive the expected cognitive effects. They may then become aware of the different levels of meaning that words and word strings make available. The extra effort invested in exploring the internal composition of these words and expressions is, on these occasions, rewarded with extra effects, not generally derived in normal conversation.

We may conclude from the arguments presented here that having evolved towards increasing efficiency, our minds tend to minimise processing effort by allocating attention and cognitive resources to the selected bits of information which are potentially relevant at the time and to process them in the most relevance-enhancing way. Familiarity with a stimulus may direct the mind to follow a certain familiar route in processing it, minimising the processing effort invested and the depth to which the stimulus is processed. The use of familiar constructions in ostensive verbal communication and the comprehension of these constructions follow these general cognitive patterns.

4. Figurative Language and Human Cognition

Our conversations, like our literature, are flooded with figurative expressions which may be used to illustrate the balance between creativity and convention characteristic of language use. Creative metaphors may be placed at one end of the spectrum of creativity as a good poetic metaphor can convey a wide array of very fine-grained thoughts and emotions. At the other end of the spectrum, we find a variety of dead metaphors and

[6] Extract from Richard Lederer's *Crazy English* available from the web at

opaque idiomatic expressions which are no longer metaphorically interpreted. Lying somewhere in-between, there is a wide range of ordinary metaphorical and idiomatic uses which vary in their degree of creativity or standardisation. Much of this dissertation has been concerned with the on-line comprehension of this last range of uses. In this section, I want to look at the wider cognitive picture to see whether the use and interpretation of these expressions is motivated and constrained by cognitive principles or patterns which have not yet been discussed.

4.1 Conceptual Metaphor Theory

Throughout this thesis, I have presented and discussed not only my own relevance-theoretic approach to the comprehension of metaphorical and idiomatic uses, but also a number of modern approaches to figurative language. However, I have not, yet, discussed what to many is THE theory of metaphor which is capable of accounting for both poetic and ordinary metaphorical uses. This is the Conceptual Metaphor theory, proposed by Lakoff and Johnson (1980) and developed by Lakoff and colleagues (e.g. Lakoff, 1987; Lakoff and Turner, 1989; Gibbs, 1990, 1992, 1993, 1994a, 1994b, 1995; Gibbs et al. 1997). The main reason for this is that, unlike the pragmatic approach defended here, the Conceptual Metaphor theory does not primarily aim to account for the on-line interpretation of (metaphorically intended) utterances, but rather to provide an explanation for the underlying cognitive patterns which make speakers produce these utterances: that is, for speaking in the way they do. In this section, I will briefly present some of the core ideas of this theory and consider how they might be integrated into the relevance-theoretic approach defended in this dissertation.

The most important claim of the Conceptual Metaphor theory is that metaphor is an aspect not of language but of cognition: "our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature" (Lakoff and Johnson, 1980: 3). According to this approach, people construct many concepts, particularly abstract concepts, by mapping their knowledge of more concrete domains onto their knowledge of more abstract domains. To give a few examples, people are seen as conceptualising life in terms of journeys, love in terms of journeys, minds in

---

http://www.pw1.netcom.com/~rlederer/arc_ceng.htm

terms of containers, anger in terms of heated fluid in a container, and so on. The mapping involves setting up systematic correspondences between the elements of these domains. So, for the conceptual metaphor LOVE IS A JOURNEY, the lovers correspond to travellers, the road they travel to their relationship, the physical obstacles on the path to the emotional obstacles they encounter in their relationship, etc. The claim is that hundreds of metaphorical conceptualisations or conceptual metaphors such as LOVE IS A JOURNEY, ANGER IS A HEATED FLUID IN A CONTAINER or ARGUMENT IS WAR are stored in long-term memory and motivate our use of language, for example our use of expressions such as those in (1)-(3):

(1) 'Our marriage is off to a great start', 'their relationship is at a crossroads', 'I am lost in this relationship', 'we are going nowhere', 'we are on the rocks', 'there are no obstacles our love cannot overcome', ' we are back on track again'.
(2) *Hit the ceiling, blow one's stack, get hot under the collar, lose your cool, flip your lid, get steamed up.*
(3) 'Your claims are indefensible', 'he attacked every weak point in my argument', 'his criticisms were right on target', 'I demolished his argument', 'I've never won an argument with him', 'you disagree?, Okay, shoot!', 'If you use that strategy, he'll wipe you out', 'he shot down all of my arguments' (From Lakoff and Johnson, 1980: 4)

According to cognitive linguists, people in a certain culture talk of love in terms of journeys because they actually THINK of love in terms of journeys. Thus, the fact that Americans use the linguistic expressions in (1)-(3) is seen not as a matter of chance but as a reflection of underlying metaphorical patterns independently existing in their minds, which motivate and constrain the way they think and, therefore, the way they speak.

The analysis of linguistic expressions has in fact been taken as the main source of evidence for the Conceptual Metaphor approach: "primarily on the basis of linguistic evidence, we have found that most of our ordinary conceptual system is metaphorical in nature" (Lakoff and Johnson, 1980: 4). Although widely influential, the Conceptual

Metaphor theory has also been widely criticised (e.g. Cacciari and Glucksberg 1995a, 1995b; Glucksberg and McGlone, 1999; Glucksberg et al. 1992, 1993; Keysar and Bly, 1999; Keysar et al. 2000; McGlone, 1996, 2001; Murphy, 1996, 1997). The methodology just described, whereby cognitive linguists make strong claims about the structure of our minds from our use of certain expressions has been particularly controversial. McGlone (2001: 95) has, for instance, pointed out that the Conceptual Metaphor account is trapped in a circular argument: we talk of life in terms of journeys because we think of life in terms of journeys and we know we think of life in terms of journeys because we speak of life in terms of journeys.

The mental images triggered by figurative expressions have been taken by supporters of the Conceptual Metaphor theory as a second main source of evidence in favour of this theory (e.g. Gibbs and O'Brien, 1990; Lakoff, 1987). According to cognitive linguists, it is because people have tacit knowledge of the conceptual metaphors that underlie certain figurative expressions that they can communicate with these expressions and form the mental images they do. This tacit knowledge of conceptual metaphors is indeed seen as "most easily uncovered through a detailed examination of speakers' mental images" (Gibbs and O'Brien, 1990: 37).

In Chapters Five and Six, I have shown how modern research on idioms has supported the view that the meaning of idioms is not entirely arbitrary. What Conceptual Metaphor theorists, particularly Gibbs and colleagues (Gibbs, 1990, 1992, 1993, 1994a, 1995; Gibbs and Nayak, 1991; Gibbs and O'Brien, 1990; Gibbs et al., 1997; Nayak and Gibbs, 1990), propose is that the meaning of idioms is partially motivated by conceptual metaphors pre-existing in our minds. The meaning of *spill the beans,* for instance, is seen as motivated by the conceptual metaphors THE MIND IS A CONTAINER and IDEAS ARE PHYSICAL ENTITIES, which map information from the source domain (i.e. our knowledge of spilling beans) to a target domain (i.e. our knowledge of revealing secrets). In order to test whether the mental images triggered by this idiom are consistent with this claim, Gibbs and O'Brien (1990) asked subjects to form a mental image of the idiom *spill the beans,* and of the literal paraphrase 'reveal a secret', and to answer a number of questions about these images (e.g. about the causes, manner, intentionality, and consequences of the action being performed). The results

showed that subjects formed very similar mental images and gave consistent answers to the questions only in the idiom condition, but not when they formed an image of the literal paraphrase. So for the idiom *spill the beans*, subjects systematically reported that the beans were initially in a container, that they were supposed to be kept in that container, that the spilling appears to be accidental, that spilled beans are rarely in a neat pile, and are not easy to retrieve. Gibbs and O'Brien (see also Lakoff, 1987) took these findings to suggest that people's tacit knowledge of the conceptual metaphors motivating the meaning of the idiom constrain the mental images they form. Since literal paraphrases are not constrained by conceptual metaphors, no consistency in the mental images they trigger was to be expected. In other words, the authors assumed that subjects think of minds in terms of containers and of ideas in terms of physical entities, and so are able to map the container in which beans are kept onto the human mind, and the beans onto ideas being revealed thus forming the mental images they did.

Even if we accept the controversial view that the mental images triggered by a figurative expression actually reveal something about our cognitive structure or the cognitive processes at work when we communicate with these expressions, there is not enough evidence for the assumption that our consistency in forming mental images is necessarily due to pre-existing conceptual metaphors which structure our thought. Cacciari and Glucksberg (1995a, 1995b) have noticed, for instance, that some of the mental images triggered by idioms are not based on their idiomatic but on their literal meanings, and so they may not actually shed much light on how these expressions are understood or what motivates their idiomatic meanings. Perhaps the most interesting finding from the work of Gibbs and O'Brien is the one mentioned above that the mental images people form are rather consistent only when triggered by idioms but not by their paraphrases. We may explain this by pointing out that there is a wide variety of stereotypical ways in which information may be revealed, and so a number of images which people may form on hearing the phrases 'to reveal information' or 'to reveal a secret' (e.g. they may imagine someone talking, someone sending a letter, etc.). By contrast, there are not that many stereotypical ways in which someone may spill beans (or spill coffee, milk or soup), and this may itself narrow down the range of possible images. The literal meaning of the phrase (and the words in that phrase) may indeed

constrain the image we form. The assumptions that the beans were supposed to be kept in a container, that the spilling appears to be accidental, that spilled beans are rarely in a neat pile, which people claimed to have had in mind in forming the image may well have been derived from their encyclopaedic knowledge of spilling, and in particular the spilling of food. We might thus predict that subjects will form relatively consistent images for literal phrases such as 'spill the coffee' or 'spill the peas', and that these images will not be particularly different from those triggered by the idiomatic string.

We might go further and suggest that people are consistent in the mental images they form for idiomatic expressions because these images are constrained not only by the literal meaning but also by the idiomatic meaning of the string. For instance, those assumptions derived from the literal meaning of the phrase which are also consistent with the overall meaning might generally be preferred, encouraging the hearer to derive certain implications which might be reinforced by a certain mental image (e.g. implications about negative consequences of the spilling and the revelation). Similarly, the answers which subjects give to specific questions about the mental image (e.g. about the intentionality, manner or consequences of the specified action) might be affected not only by their knowledge about the (literal) spilling of food or beans but by their knowledge about the idiom and its meaning: that is, by their knowledge about the secret nature of the information being kept and the consequences of revealing this information. My conclusion is that, although mental images triggered by idioms may sometimes help hearers to establish relations or bridge the gap between the encoded 'literal' meaning of the string and its idiomatic meaning, there is not enough evidence to suggest that they "uncover" metaphorical schemas which structure our thoughts.

As noted above, one important concern about the Conceptual Metaphor theory as a theory of how people communicate with metaphors is that it is not clear what role conceptual metaphors play, if any, in the on-line comprehension of metaphorical uses (Glucksberg, Brown and McGlone, 1993; Glucksberg and Keysar, 1990; Glucksberg, Keysar and McGlone, 1992; McGlone, 1996, 2001). This problem has been acknowledged by supporters of the theory:

> "Even though my work is quite suggestive of the possibility that people make use of various kinds of conceptual knowledge when understanding

> idioms, it is inappropriate to conclude from my data that people normally and automatically instantiate conceptual metaphors when understanding language. It might very well be the case that people tacitly recognize that idioms have meanings that are motivated by different kinds of conceptual knowledge. But this does not mean that people always tap into this conceptual knowledge each and every time they hear certain idioms. It might even be true that people rarely make use of this conceptual knowledge during ordinary language understanding" (Gibbs, 1994a: 306)

The suggestion that conceptual metaphors in fact play no role in the interpretation of idioms makes the Conceptual Metaphor theory unsafe as a theory of how people communicate with metaphors. In an attempt to solve this problem, Gibbs and colleagues (Gibbs et al., 1997) carried out a series of on-line experiments designed specifically to test whether conceptual metaphors are instantiated in immediate idiom comprehension. In one of these experiments, subjects were asked to read simple stories (one line at a time) on a computer screen. These stories ended in either an idiomatic expression, a literal paraphrase of the idiom, or a control sentence. A lexical decision task was presented immediately after the idiom. The target word was either related to conceptual metaphors motivating the preceding idiom or unrelated to these conceptual metaphors. Some of the material used included examples (4) and (5):

(4) *He blew his stack* (idiom phrase)
He got very angry (literal phrase)
He saw many dents (control phrase)
Heat (related target)
Lead (unrelated target)
ANGER IS HEATED FLUID IN A CONTAINER (conceptual metaphor)

(5) *It was a shot in the arm* (idiom phrase)
It was very encouraging (literal phrase)
She thought he was lying (control phrase)
Drug (related target)
Drag (unrelated target)
ENCOURAGEMENT IS GIVING SOMEONE A DRUG (conceptual metaphor)

Two main hypotheses underlie this experiment. The first is that conceptual metaphors motivate the meaning of the idioms tested, but not of the literal phrases or control sentences; the second is that people access the conceptual metaphor motivating the idiom when interpreting the string. In line with the authors' predictions, the results showed that after an idiom had been processed reaction times in the lexical decision task were considerably faster when the target word was related to the conceptual metaphor that motivated its meaning (752ms) than when the target word was unrelated to the conceptual metaphor (929ms). This difference didn't show up for literal or control strings. Reaction times for the target word related to the conceptual metaphor were also faster in the idiom condition (752ms) than in the literal (921ms) or control condition (986ms).

One main problem with this experiment is that the priming (or lack of priming) of the target items may be a by-product of semantic relatedness rather than of the instantiation of conceptual metaphors. After all, the meaning of the word *heat* is related to the literal meaning of the phrase *to blow one's stack,* but not to literal meaning of the phrase 'to be very angry'. Similarly, the meaning of the word *drug* is related to the literal meaning of the word *shot* but not of the meaning of the phrase 'it was very encouraging'. Also, the meaning of the unrelated target items used was unrelated not only to the conceptual metaphors, arguably, motivating the meaning of the string but to the literal and idiomatic meanings of the phrase. Thus, it is not surprising that they were not primed by hearing the idiomatic phrase.

To test whether the priming of a word was the result of semantic relatedness or of the instantiation of conceptual metaphors, the authors presented subjects with the idioms at the end of stories which biased the hearer towards their literal meanings. Since, according to these authors, the literal meaning of the phrase is not motivated by a conceptual metaphor, no priming of the target word should take place in this condition. In line with their predictions, the results showed that target words related to the conceptual metaphors were processed faster than unrelated targets when the phrase was idiomatically intended (899ms vs 914ms) but not when it was literally intended (933ms vs 879ms). A strange unpredicted finding reflected in the immediate preceding numbers

which was repeated across both experiments but not discussed by the authors was that in the literal and control conditions, unrelated target words were recognised considerably faster than related target words. As the above results show, the difference was generally significantly greater than the difference in reaction times for target words in the idiomatic and literal phrases. In other words, reaction times for the word *drag* were faster than for the word *drug* after processing the phrases 'it was very encouraging' and 'she thought he was lying'; and the difference between both was greater than that between the reaction times for the word *drug* after subjects had processed the idiomatic phrase *it was a shot in the arm* the one hand, and the literal and control phrases above, on the other. This clearly makes no sense: it suggests a certain element of randomness in subjects' answers which casts doubt on the psychological validity of the experiment.

Although the authors took the results from these experiments to support their view that the conceptual metaphors that motivate the meaning of an idiom are instantiated during its on-line comprehension, there is no clear evidence for this. In fact, despite their to attempt to show that conceptual metaphors play a role in on-line comprehension, Gibbs and his colleagues themselves do not seem totally convinced:

> "Our findings show that conceptual metaphors can under some circumstances be quickly accessed during immediate idiom comprehension. This conclusion does *not* mean that pre-existing metaphorical concepts are *automatically* accessed each time an idiom is encountered in discourse. Although people may have quick access to conceptual metaphors during some aspects of idiom processing, this does not imply that idiom comprehension depends on the activation of these conceptual metaphors" (Gibbs et al. 1997: 149)

To conclude my discussion on the Conceptual Metaphor theory, I would like to point out an interesting, yet worrying, paradox inherent in this theory. One important claim in the Conceptual Metaphor theory, as I have shown, is that conceptual metaphors motivate our use of many figurative expressions but not our use of literal expressions. This claim is interesting for at least two reasons. First, because many of the expressions they see as motivated by conceptual metaphors are not generally perceived as figurative but as relatively literal (e.g. 'I am low' - SAD IS DOWN; 'we have passed the due date' - TIME IS A PASSING MOTION; 'I see what you mean' - KNOWING IS SEEING). Second, and

crucially, because arguing that conceptual metaphors motivate only the use and understanding of figurative and not of literal uses clearly leads to the conclusion that literal and non-literal language are essentially different. Although cognitive linguists have repeatedly claimed that there is not a clear distinction between literal and non-literal uses of language (which seems to be supported by the idea that relatively literal expressions such as 'I am low' or 'I see what you mean' are motivated by conceptual metaphors), the version of the Conceptual Metaphor theory they defend, and the experiments it gives rise to, are grounded on this distinction.

4.2 Conceptual Metaphor Theory and Relevance Theory

Despite the criticism just presented, there is something in the core of the Conceptual Metaphor theory that seems to be worth retaining. People do seem capable of making analogical mappings between distinct domains. Among our many cognitive abilities, there is indeed the ability to exploit resemblances and draw analogies. Many of the analogies we draw to make sense of experience are one-off, yet some are repeatedly formed and may therefore become more or less standardised. Possibly one of the analogical mappings we recurrently exploit in thinking and in expressing our thoughts (e.g. by talking, painting, etc.) takes our knowledge about our own body as the source domain. In this way, we imagine extraterrestrials to have human traits (e.g. heads, limbs, etc.), we establish resemblances between the lower and upper parts of things (e.g. tables, chairs, mountains, etc.) and the lower and upper parts of our bodies (e.g. head, legs, feet), and so on. Analogies between lives and journeys are also repeatedly exploited in literature, and even in everyday speech (e.g. it is common in psychotherapy), with different authors instantiating it in different ways. The Spanish poets Antonio Machado and Jorge Manrique, for instance, exploit this analogy differently in the poems below:

"Caminante son tus huellas
el camino y nada más;
caminante, no hay camino,
se hace camino al andar.
Al andar se hace camino,
y al vover la vista atrás

se ve la senda que nunca
se ha de volver a pisar" [7]

Antonio Machado

"Nuestras vidas son los ríos,
que van a dar a la mar
que es el morir." [8]

Jorge Manrique

Thus, humans have not only the ability to construct analogical mappings between domains but also the ability to recurrently form a mapping which has proved particularly useful (e.g. between the human body and non-human entities or between lives and journeys). A set of more or less standardised analogies which are repeatedly used in thinking and in communication may be highly accessible to an individual at a given time, just as pieces of encyclopaedic information which are repeatedly considered in processing a word are highly accessible when that word is processed in stereotypical situations. Because they are highly accessible, they may affect our choices as speakers, and so in some sense constrain the way we speak; they may also direct the hearer's attention towards a particular inferential route for use in deriving the expected cognitive effects. However, the fact that people use analogies in everyday thinking and communicating does not mean that our thought is analogically structured. Also, the fact that creating or understanding a metaphor may sometimes involve the use of one-off or standardised analogical mappings does not mean that these mappings constrain the way we think, talk or communicate, or that all metaphors involve such mappings.

I would like to suggest, in fact, that what cognitive linguists refer to as conceptual metaphors are not really metaphors at all. Many are simply more or less standardised analogies (or similes) which people may exploit in conversation and which readers may construct or retrieve from memory in understanding a novel metaphor, a text or a poem. Life is not necessarily conceptualised in terms of journeys even if it may be compared to journeys on certain occasions as the two have some aspects in common (e.g. they

[7] Free literal translation: "walker, your footprints are the path and nothing else; walker, there is no path, the path is made as you walk. As one walks, the path is made, and as one looks back one sees the path that one is never to walk again."

have a beginning and an end; there have obstacles along the way, etc.). Life is like a journey in some respects, yet it is not a journey: it does not take place on a bus, on a train, or a boat, we don't have to buy tickets to embark on, etc. Conceptualising life as a journey would imply we cannot think of life without thinking of journeys. This is clearly not the case.[9]

The linguistic expressions which cognitive linguists take to be motivated by a certain group of conceptual metaphors may in fact instantiate not these conceptual metaphors but some acquired or innate knowledge about the world. According to cognitive linguists, the idioms *blow one's stack, get hot under the collar, lose your cool, flip your lid,* and *get steamed up* are motivated by the conceptual metaphor ANGER IS HEATED FLUID IN A CONTAINER, whereas the idioms *jump down someone's throat* and *bite someone's head off* are motivated by the conceptual metaphor ANGER IS ANIMAL BEHAVIOUR. I don't see these expressions as instantiating a conceptual metaphor or analogical mapping between the domain of anger and the domains of heated fluid and animal behaviour, as cognitive linguists suggest. Instead, they may be seen as giving access to two highly accessible assumptions about human nature. The first group of idioms gives access to knowledge about how our body behaves when we get angry (e.g. the body heats up, the heartbeat speeds up, the veins get tight, etc.), and the second group gives access to knowledge about how people behave when they get angry (e.g. they become aggressive, they lose control of their actions, they shout, hit things, beat others, etc.). It is the accessibility of this information which makes it possible to use and understand these expressions, to perceive them as relatively transparent, and to bring consistency among them. All this can be acknowledged without claiming that we (unavoidably) think in metaphors.

I began this section by pointing out that the Conceptual Metaphor theory was initially proposed not as a way of explaining how people understand metaphors but as explaining why they use the metaphors they do. By contrast, Relevance Theory, as a theory of human communication, was proposed to account for how people understand linguistic utterances, including those which are metaphorically intended. An adequate account of metaphor should shed light not only on the on-line comprehension of

[8] Free literal translation: "our lives are the rivers, that end in the sea, which is death."

particular uses but also on the wider cognitive picture. I hope to have made it clear enough in this thesis that Relevance Theory can fulfil both goals. Unlike Conceptual Metaphor theory, Relevance Theory does not assume that metaphors are special and therefore understood by special cognitive machinery not required for producing and understanding literal expressions. Instead, Relevance Theory claims that the comprehension of literal and loose uses (of which metaphor is just a subtype) involves the same communicative principle and comprehension procedure, both grounded on a fundamental feature of human cognition: its search for relevance.

Some scholars have considered ways in which Conceptual Metaphor theory and Relevance Theory may complement each other. Gibbs and Tendahl (forthcoming) argue that a word may give access in memory not just to a range of encyclopaedic assumptions, as claimed in Relevance Theory, but to conceptual metaphors which motivate the figurative use of the word. These conceptual metaphors may be considered in their order of accessibility by the relevance-theoretic comprehension procedure during utterance interpretation. For instance, in processing the utterance 'Mark is a bulldozer' the hearer may access the conceptual metaphors PEOPLE ARE MACHINES, THE MIND IS A MACHINE, THINKING IS A MECHANICAL ACTIVITY, etc. Carston and Wilson (2005) propose that analogies such as 'the mind is like a machine', 'people are like machines', 'ect' may be stored in long term memory and may be used as contextual assumptions used during the interpretation. They disagree with Gibbs and Tendahl and other cognitive linguists about the view that "conceptual metaphors" are genuine metaphors which structure our thoughts. Instead, they see them as relatively standardised similes or comparisons, which people may or may not use in interpretation, and whose accessibility does not fundamentally alter the nature of the interpretation process. The proposal I have defended here follows directly from their view.

The goal of pragmatics is to bridge the gap between encoded 'literal' meaning and communicated meaning. In bridging this gap, the hearer considers encyclopaedic assumptions made accessible by the encoded concepts, including more or less standardised similes. Though on many occasions the hearer may not need to access these similes in order to derive the expected effects, they may help to speed up the

[9] Many thanks to Deirdre Wilson for enlightening discussions on these issues.

interpretation process by narrowing the search space and suggesting positive implications. As I claimed for pragmatic routines, accessing a standardised analogy may minimise the processing effort required to derive the expected effects by directing the hearer in an appropriate way. This weaker version of the claims underlying Conceptual Metaphor theory can integrate well with the relevance-theoretic approach to metaphor comprehension without the need to alter the fundamental claim that there is no clear-cut distinction between literal and loose uses, which are understood by the same comprehension procedure. I would like to conclude by suggesting that strong pragmatic constraints and the search for optimal relevance guide the choice of assumptions, analogical mappings and cognitive procedures, and the depth to which this information is processed in interpretation.

5. Conclusion

In this thesis, I have argued that the tendency of the mind to minimise the mental effort invested in deriving cognitive effects and, particularly, in constructing hypotheses about a communicator's intentions, often results in the selective and relatively shallow processing of available information. Selective processing may lead to the construction of new representations (e.g. ad hoc concepts) and, eventually, result in the development of cognitive procedures (e.g. those required to form those representations). The relevance-theoretic framework and the principles of human cognition and communication on which it is grounded provide the necessary tools to explain the mechanisms underlying selective processing, and so the creativity and routinisation which result from it. The relevance-theoretic account of the interpretation of loosely intended expressions I have defended in this dissertation integrates well with this view of cognition and communication. I have shown how the selective processing and depth of processing of encoded 'literal' meaning involved in interpreting novel as well as standardised metaphorical and idiomatic uses are strongly constrained at every point by considerations of relevance. Strong pragmatic principles, powerful cognitive abilities and the tendency of our minds to search for relevance balance (in on-line processing and

throughout time and evolution) the degree of creativity and convention involved in language use and verbal communication.

References

Ackema, P. and Neeleman, A. 2001. Competition between Syntax and Morphology. In J. Grimshaw, G. Legendre and S. Vikner (eds.) *Optimality-Theoretic Syntax*. 29-60. MIT Press, Cambridge, MA.

Allott, N. and Rubio Fernández, P. 2002. This paper fills a much-needed gap. In P. Afuta, A. El Ghali, and F. Toussenel (eds.) *The Proceedings of the Atelier des Doctorants en Linguistique*. Universite Paris, Paris.

Anderson, J. 2000. *Cognitive Psychology and its Implications*. Worth Publishers, New York.

Ariel, M. 2002.The demise of a unique concept of literal meaning. *Journal of Pragmatics* 34: 361-402.

Aristotle. IVth BC. *Poetics*. Translated in 1996 by M. Heath. Penguin Books, New York.

Aristotle. IVth BC. *Rhetoric*. Translated in 1946 by R. Roberts. Oxford University Press, Oxford.

Ashby, M. Forthcoming. Prosody and idioms in English. Paper revised and resubmitted to *Journal of Pragmatics*.

Atlas, J. 2005. *Logic, Meaning, and Conversation: Semantical Underdeterminacy, Implicature, and their Interface*. Oxford University Press, Oxford.

Bach, K. 1994. Conversational impliciture. *Mind and Language* 9: 124-62.

Bach, K. 1995a. Standardization vs. Conventionalization. *Linguistics and Philosophy* 18: 677-686.

Bach, K. 1995b/1998. Standardization revisited. In A. Kasher (ed.) *Pragmatics: Critical Concepts*. Vol IV: 712-722. Routledge, London and New York.

Bach, K. 1997. The semantics-pragmatics distinction: what it is and why it matters. *Linguistische Berichte* 8: 33-50. Special Issue on Pragmatics.

Bach, K. and Harnish, R. 1979. *Linguistic Communication and Speech Acts*. MIT Press, Cambridge, MA.

Baddeley, A. 1978. The trouble with levels: a reexamination of Craik and Lockhart's framework for memory research. *Psychological Review* 85 (3): 139-153.

Baddeley, A. 1997 (ed.) *Human Memory: Theory and Practice*. Psychology Press, Hove, UK.

Barclay, J.R., Bransford, J.D., Franks, J.J, McCarrell, N.C., and Nitsch, K. 1974. Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behaviour* 15: 667-679.

Barsalou, L.W. 1981. Determinants of graded structure in categories. Dissertation, Stanford University.

Barsalou, L.W. 1982. Context-independent and context-dependent information in concepts. *Memory and Cognition* 10: 82-93.

Barsalou, L.W. 1983. *Ad hoc* categories. *Memory and Cognition* 11: 211-227.

Barsalou, L.W. 1985. Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11: 629-654.

Barsalou, L.W. 1987. The instability of graded structure: implications for the nature of concepts. In U. Neisser (ed.) *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization.* 101-140. Cambridge University Press, New York.

Barsalou, L.W. 1989. Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou and A. Ortony (eds.) *Similarity and Analogical Reasoning.* 76-121. Cambridge University Press, Cambridge.

Barsalou, L.W. 1991. Deriving categories to achieve goals. In Bower, G. (ed.) *The Psychology of Learning and Motivation.* Harcourt Brace Javanovick, New York.

Barsalou, L.W. 1993. Flexibility, structure, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols. In A.C. Collins, S.E. Gathercole, and M.A. Conway (eds.) *Theories of Memory.* 29-101. Lawrence Erlbaum, London.

Barsalou, L.W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577-660.

Barsalou, L.W. and Bower, G. 1980. A priori determinants of a concept's highly accessible information. Paper presented at the *American Psychological Association Convention,* Montreal.

Barsalou, L.W. Salomon, K. and Wu, L. 1999. Perceptual simulation in conceptual tasks. In M. Hiraga, C. Sinha, and S. Vilcox (eds.) *Cultural, Typological and Psychological Perspectives in Cognitive Linguistics. The Proceedings of the*

4th Conference of the International Cognitive Linguistics Association. Vol. 3. John Benjamins, Amsterdam.

Barsalou, L.W., and Sewell, D.R. 1984. Constructing representations of categories from different points of view. *Emory Cognition Project Technical Report #2,* Emory University.

Barsalou, L.W., Medin, D. 1986. Concepts: fixed definitions or context-dependent representations? *Cahier de Psychologie Cognitive* 6: 187-202.

Barsalou, L.W., Sewel, D.R. and Spindler, J.L. (ms) Flexibility and stability in concepts.

Barsalou, L.W., Yeh, W., Luka, B., Olseth, K., Mix, K. and Wu, L. 1993. Concepts and Meaning. In K. Beals, G. Cooke, D. Kathman, K.E. McCullough, S. Kita, and D. Testen (eds.) *Chicago Linguistics Society 29: Papers from the Parasession on Conceptual Representations.* 23-61. University of Chicago, Chicago Linguistics Society, Chicago.

Becker, A. 1997. Emergent and common features influence metaphor interpretation. *Metaphor and Symbolic Activity,* 12 (4): 243-259.

Bezuidenhout, A. and Cutting, J.C. 2002. Literal meaning, minimal propositions and pragmatic processing. *Journal of Pragmatics* 34: 433-456.

Black, M. 1962. Metaphor. In M. Black. *Models and Metaphors.* 25-47. Cornell University Press, Ithaca, New York.

Black, M. 1979. More about metaphor. In A. Ortony (ed.) *Metaphor and Thought.* 19-43. Cambridge University Press, New York.

Blasko, D. and Connine, C.M. 1993. Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Metaphor and Cognition* 9 (2): 295-308.

Blutner, R. 1998. Lexical pragmatics. *Journal of Semantics* 15: 115-162.

Blutner, R. 2002. Lexical semantics and pragmatics. *Linguistische Berichte* 10: 27-58.

Bobrow, S. and Bell, S. 1973. On catching on to idiomatic expressions. *Memory and Cognition* 1: 343-346.

Bolinger, D. 1975. *Aspects of Language.* (2nd edition). Harcourt Brace Jovanovich, New York.

Bowdle, B. and Gentner, D. 1999. Metaphor comprehension: from comparison to categorisation. *Proceedings of the Twenty-Fist Conference of the Cognitive Science Society* 90-95.

Brown, S. and Craik, F. 2000. Encoding and Retrieval of Information. In E. Tulving and F. Craik (eds.) *The Oxford Handbook on Memory.* 93-107. Oxford University Press, New York.

Buitrago, 2002. *Diccionario de Dichos y Frases Hechas.* Espasa, Madrid.

Cacciari, C. 1988. The effect of context on idiom comprehension. *Bulletin of the Psychonomic Society* 26 (6): 525-525.

Cacciari, C. 1993. The place of idioms in a literal and metaphorical world. In C. Cacciari, and P. Tabossi (eds.) *Idioms: Processing, Structure, and Interpretation.* 27-52. Lawrence Erlbaum, Hillsdale, New Jersey.

Cacciari, C. and Glucksberg, S. 1991. Understanding idiomatic expressions: the contribution of word meanings. In G.B. Simpson (ed.) *Understanding Word and Sentence.* 217-240. Elsevier, Amsterdam.

Cacciari, C. and Glucksberg, S. 1994. Understanding figurative language. In M.A. Gernsbacher (ed.) *Handbook of Psycholinguistics.* 447-477. Academic Press, New York.

Cacciari, C. and Glucksberg, S. 1995a. Imaging idiomatic expressions: literal or figurative meanings? In M. Everaert et al. (eds.) *Idioms: Structural and Psychological Perspectives.* 43-56. Lawrence Erlbaum, New Jersey.

Cacciari, C. and Glucksberg, S. 1995b. Understanding idioms: do visual images reflect figurative meanings. *European Journal of Cognitive Psychology* 7 (3): 283-305.

Cacciari, C. and Levorato, M.C. 1989. How children understand idioms in discourse. *Journal of Child Language* 16: 387-405.

Cacciari, C. and Levorato, M.C. 1991. Spilling the beans on children's comprehension and production of idioms. *Bulletin of the Psychonomic Society* 29(6): 498-498.

Cacciari, C. and Levorato, M.C. 1998. The effect of semantic analysability of idioms in metalinguistic tasks. *Metaphor and Symbol.* 13(3): 159-177.

Cacciari, C. and Tabossi, P. 1988. The comprehension of idioms. *Journal of Memory and Language* 2: 668-683.

Cacciari, C. and Tabossi, P. 1993 (eds.) *Idioms: Processing, Structure and Interpretation.* Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Cacciari, C., Levorato M.C. and Cicogna, P. 1997. Imagination at work: conceptual and linguistic creativity in children. In M. Ward. et al. (eds.) *Creative Thought: an Investigation of Conceptual Structures and Processes.* 179-208. American Psychological Association, Washington, DC.

Camac, M.K. and Glucksberg, S. 1984. Metaphors do not use associations between concepts, they are used to create them. *Journal of Psycholinguistic Research* 13: 443-455.

Carston, R. 1988. Implicature, explicature and truth-theoretic semantics. In R. Kempson (ed.) *Mental Representations.* 155-181. Cambridge University Press, Cambridge. Reprinted in S. Davis (ed.) 1991. *Pragmatics: a Reader.* 33-51. Oxford University Press, New York.

Carston, R. 1996. Enrichment and loosening: complementary processes in deriving the proposition expressed? *UCL Working Papers in Linguistics* 8: 205-232. Reprinted 1997 in *Linguistische Berichte* 8: 103-127. Reprinted in 1997 *Linguistische Berichte* 8: 103-127. Special Issue on Pragmatics.

Carston, R. 1998a. Pragmatics and the Explicit/Implicit Distinction. PhD Dissertation. University College, London. University of London.

Carston, R. 1998b. The semantics/pragmatics distinction: a view from relevance theory. *UCL Working Papers in Linguistics* 10: 53-80. Reprinted in K. Turner (ed.) 1999. *The Semantics/Pragmatics Interface from Different Points of View* (CRiSPI 1). Elsevier Science.

Carston, R. 2000. Explicature and semantics. *UCL Working Papers in Linguistics* 12: 1-44. Reprinted in S. Davis & B. Gillon (eds.) 2004. Semantics: A Reader. Oxford: Oxford University Press.

Carston, R. 2002a. *Thoughts and Utterances.* Oxford, Blackwell.

Carston, R. 2002b. Linguistic meaning, communicated meaning and cognitive pragmatics. *Mind and Language* 17: 127-148.

Carston, R. 2002c. Metaphor, ad hoc concepts and word meaning – more questions than answers. *UCL Working Papers in Linguistics* 14: 83-105.

Carston, R. and Wilson, D. 2005. The pragmatics of metaphor: the 'emergent property' issue. Paper presented at *Metaphor: an Interdisciplinary Conference* University of London, January 2005.

Chafe, W. 1968. Idiomaticity as an anomaly in the chomskyan paradigm. *Foundation of Language* 4: 109-127.

Chomsky, N. 1965. *Aspects of the Theory of Syntax.* MIT Press, Massachusetts.

Chomsky, N. 1980. *Rules and Representations.* Columbia University Press, New York.

Clark, E. 1987. The principle of contrast. In B. MacWhinney (ed.) *Mechanisms of Language Acquisition.* Lawrence Erlbaum Associates, Inc. Hillsdale, New Jersey.

Clark, E. and Clark, H. 1979. When nouns surface as verbs. *Language* 55: 767-811.

Clark, H. 1983. Making sense of nonce sense. In G. Flores d'Arcais and R. Jarvella (eds.) *The Process of Understanding Language.* 297-332. Wiley, New York.

Clark, H. and Gerrig, R. 1983. Understanding old words with new meanings. *Journal of Verbal Learning and Verbal Behaviour* 22: 591-608.

Cohen, J. Dunbar, K. and MacClelland, J. 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review* 97: 332-361.

Collins, A.M. and Loftus, E.F. 1975. A spreading activation theory of semantic processing. *Psychological Review* 82: 407-428.

Coulmas, F. (ed.) 1981a. *Conversational Routine: Explorations in Standarized Communication Situations and Prepatterned Speech.* Mouton Publishers, The Hague, the Netherlands.

Coulmas, F. 1981b. Idiomaticity as a problem of pragmatics. In H. Parret (ed.) *Possibilities and Limitations of Pragmatics.* 139-151. John Benjamins, Amsterdam.

Craik, F. 2002. Levels of processing: past, present ... and future? *Memory* 10 (5/6): 305-318.

Craik, F. and Lockhart, R. 1972. Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behaviour* 11: 671-684.

Craik, F. and Tulving, E. 1975. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104: 208-294.

Cremark, L. and Craik, F. (eds.) 1979. *Levels of Processing in Human Memory.* Hillsdale, New Jersey.

Cruse, A. 1986. *Lexical Semantics.* Cambridge University Press, Cambridge.

Dascal, M. 1987. Defending literal meaning. *Cognitive Science* 11: 259-281.

Dascal, M. 1989. On the roles of context and literal meaning in understanding. *Cognitive Science* 13: 253-257.

Davidson, D. 1977. What metaphors mean. In S. Sacks (ed.). *On Metaphor*. 29-45. University of Chicago Press, Chicago.

DeGroot, A. 1965. *Thought and Choice in Chess*. The Hague, Mouton.

DeGroot, A. D. 1966. Perception and memory versus thought: some old ideas and recent findings. In B. Kleinmuntz (ed.) *Problem Solving: Research, Method and Theory*. 19–50. Wiley, New York.

Dickens, C. 1850. *David Copperfield.* Original manuscript.

Dunkling, L. 1998. *Dictionary of Curious Phrases.* Harper Collins, Glasgow.

Eizaga Rebollar, B. 2002. Procesamiento de Modismos: un enfoque Psicopragmático PhD Thesis, Universidad de Cádiz, Spain.

Erickson, T. and Mattson, M. 1981. From words to meaning: a semantic illusion. *Journal of Verbal Learning and Verbal Behaviour* 20: 540-552.

Ericsson, K. A., Krampe, R. T., and Tesch-Romer, C. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100(3): 363-406.

Estes, Z. and Glucksberg, S. 1999. Relevance and feature accessibility in combined concepts. In M. Hahn and S.C. Stones (eds.) *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society.* 149-154. Erlbaum, Mahwah, NJ.

Estes, Z. and Glucksberg, S. 2000a. Interactive property attribution in concept combination. *Memory and Language* 28(1): 28-34.

Estes, Z. and Glucksberg, S. 2000b. Similarity and attribution in concept combination: reply to Wisniewski. *Memory and Cognition* 28: 39-40.

Everaert, M., Van den Linden, E., Schenk, A. and Scheuder, R. (eds.) 1995. *Idioms: Structural and Psychological Perspectives*. New Jersey, Lawrence Erlbaum.

Fauconnier, G. and Turner, M. 1998. Conceptual integration networks. *Cognitive Science* 22(2): 133-187.

Fauconnier, G. and Turner, M. 2002. *The Way We Think.* Basic Books, New York.

Fernando, C. 1996. *Idioms and Idiomaticity.* Oxford University Press, Oxford.

Fernando, C. and Flavell, R. 1981. *On Idiom: Critical Views and Perspectives.* Vol. 5. Exeter Linguistic Studies, University of Exeter, UK.

Finke, R.A., Ward, T.B. and Smith, S.M. 1992. *Creative Cognition: Theory, Research, and Application.* MIT Press, Cambridge, MA.

Fitts, P. and Posner, M. 1967. *Human Performance.* Brooks Cole, Belmont, CA.

Flores d'Arcais, G. 1993. The comprehension and semantic interpretation of idioms. In C. Cacciari and P. Tabossi (eds.) *Idioms: Processing, Structure, and Interpretation.* 79-98. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Fodor, J. 1998. *Concepts: where Cognitive Science Went Wrong.* Oxford, Clarendon Press.

Forrester, M. 1995. Tropic implicature and context in the comprehension of idiomatic phrases. *Journal of Psycholinguistic Research* 24: 1-22.

Franks, B. 1995. Sense generation: a 'quasi-classical' approach to concepts and concept combination. *Cognitive Science* 19: 441-505.

Fraser, B. 1970. Idioms within transformational grammar. *Foundations of Language* 6 (1): 22-42.

Garner, W. 1974. *The Processing of Information and Structure.* Wiley, New York.

Geeraerts, D. 1995. Specialization and Reinterpretation in Idioms. In M. Everaert, E. van der Linden, A. Schenk and R. Schreuder (eds.) *Idioms: Structural and Psychological Perspectives.* 57-73. Erlbaum Associates, Inc. Publishers, Hillsdale, New Jersey.

Gentner, D. 1983. Structure-mapping: a theoretical framework for analogy. *Cognitive Science* 7: 155-170.

Gentner, D. 1989. The mechanisms of analogical reasoning. In S. Vosniadou and A. Ortony (eds.) *Similarity and Analogical reasoning.* 199-241. Cambridge University Press, Cambridge.

Gentner, D. and Bowdle, B. 2001 Convention, form and figurative processing. *Metaphor and Symbol* 16 (3): 223-247.

Gentner, D. and Wolff, P. 1997. Alignment in the processing of metaphor. *Journal of Memory and Language* 37: 331-355.

Gentner, D., Bowdle, B., Wolff, P. and Boronat, C. 2001. Metaphor is like analogy. In Gentner, D., Holyoak, K. and Kokinov, B. (eds.) 2001. *The Analogical Mind: Perspectives from Cognitive Science.* 199-253. Cambridge MA, MIT Press.

Gernsbacher, M.A., Keysar, B., Robertson, R. and Werner, N. 2001. The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language* 45: 433-450.

Gibbs, R. 1980. Spilling the beans on understanding and memory for idioms in conversation. *Memory and Cognition* 8: 449-456.

Gibbs, R. 1986. Skating on thin ice: literal meaning and understanding idioms in conversation. *Discourse Processes* 9: 17-30.

Gibbs, R. 1987. Linguistic factors in children's understanding of idioms. *Journal of Child Language* 14: 587-60.

Gibbs, R. 1989. Understanding and literal meaning. *Cognitive Science* 13: 243-251.

Gibbs, R. 1990. Psycholinguistic studies on the conceptual basis of idiomaticity. *Cognitive Linguistics* 1: 417-451.

Gibbs, R. 1991. Semantic analyzability in children's understanding of idioms. *Journal of Speech and Hearing Research* 34: 613-620.

Gibbs, R. 1992. What do idioms really mean. *Journal of Memory and Language* 31: 485-506.

Gibbs, R. 1993. Why idioms are not dead metaphors. In C. Cacciari and P. Tabossi (eds.) *Idioms: Processing, Structure, and Interpretation.* 57-75. Lawrence Erlbaum, Hillsdale, New Jersey.

Gibbs, R. 1994a. *The Poetics of Mind.* Cambridge University Press, Cambridge.

Gibbs, R. 1994b. Figurative Thought and Figurative Language. M. A. Gernsbacher (ed.) *Handbook of Psycholinguistics.* 411- 446. Academic Press, San Diego.

Gibbs, R. 1995. Idiomaticity and human cognition. In M. Everaert et al. (eds.) *Idioms: Structural and Psychological Perspectives.* 97-116. Lawrence Erlbaum, New Jersey.

Gibbs, R. 1999a. Interpreting what the speaker say and implicate. *Brain and Language* 68: 466-485.

Gibbs, R. 1999b. Speaker's intuitions and pragmatic theory. *Cognition* 69: 355-359.

Gibbs, R. 2002a. A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics* 34(3): 457-486.

Gibbs, R. 2002b. Experimental psychology and pragmatics. *ICU Open Lectures on Cognitive Pragmatics.* ICU, Japan.

Gibbs, R. and Gonzales, G. 1985. Syntactic frozenness in processing and remembering idioms. *Cognition* 21: 241-259.

Gibbs, R. and Moise, J. F. 1997. Pragmatics in understanding what is said. *Cognition* 62: 51-74.

Gibbs, R. and Nayak, N. 1989. Psycholinguistic studies on the syntactic behaviour of idioms. *Cognitive Psychology* 21: 100-138.

Gibbs, R. and Nayak, N. 1991. Why idioms mean what they do. *Journal of Experimental Psychology: General* 120: 93-95.

Gibbs, R. and O'Brien, J. 1990. Idioms and mental imagery: the metaphorical motivation for idiomatic meaning. *Cognition* 36: 35-68.

Gibbs, R. and Tendahl, M. Forthcoming. Cognitive effort and effects in metaphor comprehension: relevance theory and psycholinguistics. To appear in *Mind and Language.*

Gibbs, R., Bogdonovich, J., Sykes, J. and Barr, D. 1997 Metaphor in idiom comprehension. *Journal of Memory and Language* 37: 141-154.

Gibbs, R., Buchalter, D, Moise, J. and Farrar, W. 1993. Literal meaning and figurative meaning. *Discourse Processes* 16: 387-404.

Gibbs, R., Nayak, N. and Cutting, C. 1989. How to kick the bucket and not decompose: analysability and idiom processing. *Journal of Memory and Language* 28: 576-593.

Gibbs, R., Nayak, N., Bolton, J. and Keppel, M. 1989. Speaker's assumptions about the lexical flexibility of idioms. *Memory and Cognition* 17(1): 58-68.

Gildea, P. and Glucksberg, S. 1983. On understanding metaphor: the role of context. *Journal of Verbal Learning and Verbal Behaviour* 22: 577-590.

Gineste, M.D., Indurkhya, B. and Scart, V. 2000. Emergence of features in metaphor comprehension. *Metaphor and Symbol* 15 (3): 117-135.

Glass, A., Holyoak, K. and Santa, J. 1979. *Cognition.* Addison-Wesley, Massachusetts.

Glucksberg, S. 1991. Beyond literal meanings: the psychology of allusion. *Psychological science* 2: 146-52.

Glucksberg, S. 1993. Idiom meaning and allusional content. In C. Cacciari, and P. Tabossi (eds.) *Idioms: Processing, Structure, and Interpretation.* 3-26. Lawrence Erlbaum, Hillsdale, New Jersey.

Glucksberg, S. 2001. *Understanding Figurative Language: from Metaphors to Idioms.* Oxford University Press, Oxford.

Glucksberg, S. and Cacciari, C. 1989. Contributions of word meanings to idiomatic meanings. *Bulletin of the Psychonomic Society* 27(6): 531-531.

Glucksberg, S. and Keysar, B. 1990. Understanding metaphorical comparisons: beyond similarity. *Psychological Review* 1: 3-18.

Glucksberg, S. and Keysar, B. 1993. How metaphors work. In A. Ortony. *Metaphor and Thought.* 401-424. Second Edition. Cambridge University Press, New York.

Glucksberg, S. and McGlone, M.S. 1999. When Love is not a journey: what metaphors mean. *Journal of Pragmatics* 31: 1541-1558.

Glucksberg, S. Brown, M.E and McGlone, M.S. 1993. Conceptual metaphors are not automatically accessed during idioms comprehension. *Memory and Cognition* 21(5): 711-719.

Glucksberg, S., Gildea, P and Bookin, H.A. 1982. On understanding nonliteral speech: can people ignore metaphors? *Journal of Verbal Learning and Verbal Behaviour* 21: 85-98.

Glucksberg, S., Keysar, B. and McGlone, M. 1992. Metaphor understanding and accessing conceptual schema: reply to Gibbs. *Psychological Review* 99: 578-581.

Glucksberg, S., Manfredi, D. and McGlone, M. 1997. Metaphor comprehension: how metaphors create new categories. In T. Ward, S. Smith and J. Vaid (eds.) *Creative Thought: an Investigation of Conceptual Structures and Processes.* 327-350. American Psychological Association, Washington, D.C.

Glucksberg, S., McGlone, M. and Manfredy, D. 1997. Property Attribution in Metaphor comprehension. *Memory and Language* 36: 50-67.

Glucksberg, S., Newsome, MR. and Goldvarg, Y. 2001 Inhibition of the literal: filtering metaphor–irrelevant information during metaphor comprehension. *Metaphor and Symbol* 16(3/4): 277-293.

Goodman, N. 1972. Seven strictures on similarity. In N. Goodman. *Problems and Projects.* 22-32. Bobs-Merrill, Indianapolis.

Grady, J., Oakley, T. and Coulson, S. 2000. Blending and Metaphor. In G. Steen and R. Gibbs (eds) *Metaphor in Cognitive Linguistics.* 101-124. John Benjamins, Philadelphia.

Grice, H.P. 1975/1989. *Studies in the Way of Words.* Harvard UP, Cambridge, MA.

Groefsema, M. 1992. Can you pass the salt?: a short-circuited implicature? *Lingua* 87: 103-135.

Gutt, E.A. 1991. *Translation and Relevance.* Blackwell, Oxford.

Hahn, U. and Chater, N. 1997. Concepts and similarity. In K. Lamberts and D. Shanks. *Knowledge, Concepts and Categories.* 43-93. Psychology Press. Have, UK

Hamblin, J. and Gibbs, R. 1999. Why you can't kick the bucket as you slowly die: verbs in idiom comprehension. *Journal of Psycholinguistic Research* 28(1): 25-39.

Hampton, J. 1997. Emergent attributes in combined concepts. In T. Ward, S. Smith and J. Vaid (eds.) *Creative Thought: an Investigation of Conceptual Structures and Processes.* 83-110. American Psychological Association, Washington, D.C.

Hannon, B. and Craik, F. 2001. Encoding specificity revisited: the role of semantics. *Canadian Journal of Experimental Psychology* 55(3): 231-243.

Happé, F. 1993. Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition,* 48: 101-119.

Happé, F. 1994. Understanding minds and metaphors: insights from the study of figurative language in autism. *Metaphor and Symbolic Activity* 10: 275-295.

Harnish, R. 1993. Communicating with proverbs. *Communication and Cognition* 26 (3/4): 265-290.

Harnish, R. 1995. Proverbs and pragmatics I: Prolegomena to a comparison of English and Hungarian. In K. Sornig, D.W. Halwachs, C. Penzinger and G. Ambrosch. *Linguistics with a Human face.* Grazer Linguistische Monographien 10: 135-146.

Horsey, R. 2002. The art of chicken sexing. *UCL Working Papers in Linguistics* 14: 107-117.

Inhoff, A.W. Lima, S.D. and Carroll, P.J. 1984. Contextual effects on metaphor comprehension. *Memory and Cognition* 12: 558-563.

Jackendoff, R. 1995. The boundaries of the lexicon. In M. Everaert, E.J. Van der Linden, A. Schenk and R. Schreuder (eds.) *Idioms: Structural and Psychological Perspectives.* 133-165. Lawrence Erlbaum, New Jersey.

Karmiloff-Smith, A. 1996. *Beyond Modularity: a Developmental Perspective on Cognitive Science.* MIT Press, Cambridge MA.

Katz, J. 1973. Compositionality, idiomaticity, and lexical substitution. In S. Anderson, P. Kiparsky. (eds.) *A festschrift for Morris Halle.* 357-376. Holt, New York.

Keysar, B. 1989. On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language* 28: 375-385.

Keysar, B. and Bly, B. 1995. Intuitions of the transparency of idioms: can one keep a secret by spilling the beans? *Journal of Memory and Language* 34: 89-109.

Keysar, B. and Bly, B. 1999. Swimming against the current: do idioms reflect conceptual structure? *Journal of Pragmatics* 31: 1559-1578.

Keysar, B., Shen, Y., Glucksberg, S. and Horton, W. 2000. Conventional language: how metaphorical is it? *Journal of Memory and Language* 43: 576-593.

Kittay, E. 1987. *Metaphor: its Cognitive Force and Linguistic Structure.* Oxford University Press, Oxford.

Kuiper, K. 1996. *Smooth Talkers: the Linguistic Performance of Auctioneers and Sportscasters.* Lawrence Erlbaum, New York.

Lakoff, G. and Johnson, M. 1980. *Metaphors We Live by.* University of Chicago Press, Chicago.

Lakoff, G. and Turner, M. 1989. *More than Cool Reason: a Field Guide for Poetic Metaphor.* University of Chicago Press, Chicago.

Lakoff, G., 1987. *Women, Fire and Dangerous Things: what Categories Reveal about the Mind.* University of Chicago Press, Chicago.

Lakoff, G.1986. The meanings of literal. *Metaphor and Symbolic Activity* 1: 291-296.

Lasersohn, P. 1999 Pragmatic halos. *Language* 75: 522-551.

Lederer, R. 1989. *Crazy English.* Pocket Books, New York.

Levinson, S. 1983. *Pragmatics.* Cambridge University Press, Cambridge.

Levinson, S. 2000. *Presumptive Meanings.* MIT Press, Cambridge, MA.

Levorato, M.C. 1993. The acquisition of idioms and the development of figurative competence. In C. Cacciari and P. Tabossi (eds.) *Idioms: Processing, Structure, and Interpretation.* 101-123. Lawrence Erlbaum, Hillsdale, New Jersey.

Levorato, M.C. and Cacciari, C. 1992. Children's comprehension and production of idioms: the role of context and familiarity. *Journal of Child Language* 19: 415-433.

Levorato, M.C. and Cacciari, C. 1995. The effects of Different tasks on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology* 60: 261-283.

Levorato, M.C. and Cacciari, C. 1999. Idiom comprehension in children: are the effects of semantic analysability and context separable? *European Journal of Cognitive Psychology* 11 (1): 51-66.

Lewis, D. 1979/83 Scorekeeping in a language game. In D. Lewis 1983 *Philosophical Papers.* Volume I. 233–249. Oxford University Press, Oxford. Originally published in *Journal of Philosophical Logic* 8, *1979.*

Locke, J. 1690. An essay concerning human understanding. Original manuscript

MacLeod, C. and Dunbar, K. 1988. Training and Stoop-like interference: evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14: 126-135.

Makkai, A. 1972. *Idiom Structure in English.* Mouton, The Hague.

Malgadi, R.G. and Johnson, M. 1976. Modifiers in metaphors: effects of constituent phrase similarity on the interpretation of figurative sentences. *Journal of Psycholinguistic Research* 55: 95-107.

Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word recognition. *Cognitive Psychology* 25: 71-102.

McGlone, M.S. 1996. Conceptual metaphors and figurative language understanding: food for thought. *Journal of Memory and Language* 35: 544-465.

McGlone, M.S. 2001. Concepts as metaphors. In S. Glucksberg. *Understanding Figurative Language: from Metaphors to Idioms.* 90-107. Oxford University Press, Oxford

McGlone, M.S. and Manfredi, D. 2001. Topic-vehicle interaction in metaphor comprehension. *Memory and Cognition* 29(8): 1209-1219.

McGlone, M.S., Glucksberg, S. and Cacciari, C. 1994. Semantic productivity in idiom comprehension. *Discourse Processes* 17: 167-190.

Memory 2002. Vol 10 (5/6). Special issue on Levels of Processing.

Meyer, D.E. and Schvaneveldt, R.W. 1971. Facilitation in recognising pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* 90: 227-234.

Miller, G. 1993. Images and Models, Similes and Metaphors. In A. Ortony (ed.) *Metaphor and Thought.* 357–400. Cambridge University Press, Cambridge.

Moon R. 1998a. *Fixed Expressions and Idioms in English: a Corpus-based Approach.* Clarendon Press, Oxford.

Moon, R. 1998b. Frequencies and forms of phrasal lexemes in English in A. Cowie (ed.) *Phraseology: Theory, Analysis and Applications.* 79-100. Oxford University Press, Oxford.

Morgan, J. 1978. Two types of convention in speech acts. In P. Cole (ed.) *Syntax and Semantics 9: Pragmatics.* Academic Press, NewYork.

Murphy, G. 1996. On metaphorical representation. *Cognition* 60: 173-204.

Murphy, G. 1997. Reasons to doubt the present evidence for metaphoric representation. *Cognition* 62: 99-108.

Nattinger, J.R. and DeCarrico, J.S. 1992. *Lexical Phrases and Language Teaching.* Oxford University Press, Oxford.

Nayak, N. P. and Gibbs, R. W. 1990. Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology: General* 119: 315-330.

Newell, A. and Simon, H. 1972. *Human Problem Solving.* Prentice-Hall, New York.

Newmeyer, F. 1972. The insertion of idioms. In *Papers from the Eighth Regional Meeting.* 294-302. Chicago Linguistics Society, Chicago.

Nippold, M.A. and Rudzinski, M. 1993. Familiarity and transparency in idiom explanation: a developmental study of children and adolescents. *Journal of Speech and Hearing Research* 36: 728-737.

Nogales, P. 1999. *Metaphorically Speaking.* CSLI/Cambridge University Press, Stanford.

Nunberg, G. 1978. *The Pragmatics of Reference.* Indiana University Linguistics Club, Bloomington.

Nunberg, G., Sag, I. and Wasow, T. 1994. Idioms. *Language* 70(3): 491-538.

O'Grady, W. 1998. The syntax of idioms. *Natural Language and Linguistic Theory* 16: 279-312.

Ortony, A. (ed.) 1993. *Metaphor and Thought.* Second Edition. Cambridge University Press, New York.

Ortony, A. 1979. Beyond literal similarity. *Psychological Review* 86: 1671-180.

Ortony, A., Schallert, D., Reynolds, R. and Antos, S. 1978. Interpreting metaphors and idioms: some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behaviour* 17: 465-477.

Papafragou, A. 1996. On metonymy. *Lingua* 99: 169-195.

Parkinson, J. 2000. *Catchphrase, Slogan and Cliché.* Michael O'Mara Books Limited, London.

Perkins, M.R. 1999. Productivity and formulaicity in language development. In M. Garman, C. Letts, B. Richards, C. Shelleter and S Edwards (eds.) *Issues in Normal and Disordered Child Language: from Phonology to Narrative..* 51-67. University of Reading, Reading.

Peterson, P. and Burgess, C. 1993. Syntactic and semantic processing during idiom comprehension: neurolinguistic and psycholinguistic dissociations. In C. Cacciari and P. Tabossi. (eds.) *Idioms: Processing, Structure, and Interpretation.* 201-225. Lawrence Erlbaum, Hillsdale, New Jersey.

Pilkington, A. 2000. *Poetic Effects: a Relevance Theory Perspective.* John Benjamins, Amsterdam.

Posner, M.I., and Snyder, C.R. 1975. Attention and cognitive control. In R.L. Solso (ed.) *Information Processing and Cognition: the Loyola Symposium.* 55-85. Lawrence Erlbaum, Hillsdale, New Jersey.

Pustejovsky, J. 1995. *The Generative Lexicon.* Cambridge, MA: MIT Press.

Ratcliff, R. and McKnoon, G. 1981. Does activation really spread? *Psychological Review* 88: 454-462.

Reagan, R. 1987. The syntax of English idioms: can the dog be put on? *Journal of Psycholinguistic Research* 16(5): 417- 441

Recanati, F. 1989. The pragmatics of what is said. *Mind and Language* 4: 295-329.

Recanati, F. 1995. The alleged priority of literal interpretation. *Cognitive Science* 19: 207-232.

Recanati, F. 2003. *Literal Meaning.* Cambridge University Press, Cambridge.

Richards, I.A. 1936. *The Philosophy of Rhetoric.* Oxford University Press, London.

Rips, L.J. 1995. The current status of research on concept combination. *Mind and Language* 10: 72-104.

Ritchie, D. 2003. Categories and similarities: a note on circularity. *Metaphor and Symbol* 18(1): 49-53.

Rosch, E. 1978. Principles of categorization. In E. Rosch and B. Lloyd (eds.) *Cognition and Categorization.* 28-46. Lawrence Erlbaum, New Jersey.

Rubio Fernández, P. 2005. *Cognitive Mechanism and Pragmatic Processes in Lexical Interpretation: the On-line Construction of Concepts.* PhD Thesis. Cambridge University.

Rumelhart, D. 1979/1993. Some problems with the notion of literal meanings. In A. Ortony (ed.) *Metaphor and Thought.* 71-82. Cambridge University Press, Cambridge.

Sadock, J. M. 1972. Speech act idioms. In P. M. Peranteau, J. N. Levi, and G. C. Phares (eds.). *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society.* 329-339. University of Chicago, Chicago.

Sadock, J.M. 1993. Figurative speech and linguistics. In A. Ortony (ed.) Metaphor and Thought. (2 Ed.) 42-57. New York: Cambridge University Press.

Sandford, A and Sturt, P. 2002. Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive sciences* 6(9): 382-386.

Sandford, A. 2002. Context, attention, and depth of processing during interpretation. *Mind and Language* 17: 188-206.

Saussure, F. de. 1975. *Course in General Linguistics,* translated from the French (1916) by Wade Baskin. Peter Owen, London.

Schenider, W. and Chein, J. 2003. Controlled and automatic processing: behaviour, theory and biological mechanisms. *Cognitive Science* 27: 525-559.

Schneider, W. and Shiffrin, R. 1977. Controlled and automatic human information processing: I detection, search and attention. *Psychological Review* 84: 1-66.

Searle, J. 1979a. Literal meaning. In J. Searle. *Expression and Meaning.* 117-136. Cambridge University Press, Cambridge.

Searle, J. 1979b. Metaphor. In Searle, J. 1979. *Expression and Meaning.* 137-161. Cambridge University Press, Cambridge.

Searle, J. 1980. The background of meaning. In J. Searle et al. (eds.) *Speech Act Theory and Pragmatics.* 221-232. Reidel Publishing Co., Dordrecht, Netherlands.

Shiffrin, R.M., and Schneider, W. 1977. Controlled and automatic information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84: 127-190.

Simon, H. and Gilmartin, K. 1973. A simulation of memory for chess positions. *Cognitive Psychology* 5: 29-46.

Smith, E. and Medin, D. 1981. *Categories and Concepts*. Harvard University Press, Cambridge, MA.

Smith, L and Samuelson, L. 1997. Perceiving and remembering: category stability, variability and development. In K. Lamberts and D. Shanks (eds.) 1997. *Knowledge, Concepts and Categories.* 161-196. Psychology Press, Hove, UK.

Smith, S.M., Ward, T.B. and Finke, R.A. (eds.) 1995. *The Creative Cognition Approach.* MIT Press, Cambridge, MA.

Sperber, D. Forthcoming. Modularity and relevance: how can a massively modular mind be flexible and context-sensitive? To appear in P. Carruthers, S. Laurence and S. Stich (eds.) *The Innate Mind: Structure and Content.*

Sperber, D. 1982. Apparently irrational beliefs. In S. Lukes and M. Hollis (eds.) *Rationality and Relativism.* 149-180. Blackwell, Oxford.

Sperber, D. 1990. The epidemiology of beliefs. In C. Fraser and G. Gaskell (eds.). *The Social Psychological Study of Widespread Beliefs.* 25-44. Clarendon Press, Oxford.

Sperber, D. 1994a. The modularity of thought and the epidemiology of representations. In L. Hirschfield and S. Gelman (eds.) *Mapping the Mind: Domain Specificity in Cognition and Culture.* 39-67. Cambridge University Press, Cambridge.

Sperber, D. 1994b. Understanding verbal understanding. In J. Khalfa. *What is Intelligence?* 179-198. Cambridge University Press, Cambridge.

Sperber, D. 1996. *Explaining Culture: a Naturalistic Approach.* Blackwell, Oxford.

Sperber, D. 1997. Intuitive and reflective beliefs. *Mind and Language* 12: 67-83.

Sperber, D. 2000. Metarepresentations in an evolutionary perspective. In D. Sperber *Metarepresentations: an Interdisciplinary Perspective.* 117-137. New York, Oxford University Press.

Sperber, D. and Wilson, D. 1986/1995. *Relevance: Communication and Cognition.* Blackwell, Oxford.

Sperber, D. and Wilson, D. 1987. Précis of *Relevance. Behavioral and Brain Sciences* 10(4): 697-710.

Sperber, D. and Wilson, D. 1990. Rhetoric and relevance. In J. Bender and D. Wellbery. *The Ends of Rhetoric: History, Theory, Practice.* 140-155. Stanford University Press, Stanford.

Sperber, D. and Wilson, D. 1991. Loose talk. In S. Davis (ed.) *Pragmatics: a Reader.* 540-549. Oxford University Press, Oxford.

Sperber, D. and Wilson, D. 1998. The mapping between the mental and the public lexicon. In P. Carruthers and J. Boucher (eds.) *Thought and Language.* 184-200. Cambridge University Press, Cambridge.

Sperber, D. and Wilson, D. 2002. Pragmatics, modularity and mind-reading. *Mind and Language* 17(1/2): 3-23.

Sternberg, R. (ed.) 1999. *Handbook of Creativity.* Cambridge University Press, New York.

Strässler, J. 1982. *Idioms in English: a Pragmatic Analysis.* Tübingen: Gunter Narr.

Stroop, J.P. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18: 643-662.

Swinney, D. and Cutler, A. 1979. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour* 18: 523-534.

Tabossi, P. and Zardon, F. 1993. The activation of idiomatic meaning in spoken language comprehension. In C. Cacciari and P. Tabossi. (eds.) *Idioms: Structure, Processing and Interpretation.* 145-162. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Titone, D. and Connine, C. 1994. The comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 20(5): 1126-1138.

Titone, D. and Connine, C. 1999. On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics* 31(3): 1655-1674.

Tourangeau, R. and Rips, L. 1991. Interpreting and evaluating metaphors. *Journal of Memory and Lan*guage 30: 452-472.

Tourangeau, R. and Sternberg, R. 1981. Aptness in metaphor. *Cognitive Psychology,* 13: 27-55.

Tulving, E. 1979. Relations between encoding specificity and levels of processing. In L. Cermark and F. Craik (eds.) *Levels of Processing in Human Memory.* 405-428. Hillsdale, New Jersey.

Tulving, E. 1983. *Elements of Episodic Memory.* Oxford University Press, New York.

Tulving, E. and Thomson, D. 1973. Encoding Specificity and retrieval processes in episodic memory. *Psychological Review* 80: 352-373.

Tvesky, A. 1977. Features of similarity. *Psychological Preview* 84: 327-352.

Unger, C. 1996. Proverbs and Relevance. Unpublished manuscript

Van der Voort, M. and Vonk, W. 1995. You don't die immediately when you kick an empty bucket: a processing view on semantic and syntactic characteristic of idioms. In M. Everaert et al. (eds.) *Idioms: Structural and Psychological Perspectives.* 283-300. Lawrence Erlbaum, New Jersey.

Van Lancker, D. and Canter, G.J. 1981. Idiomatic versus literal interpretations of ditropically ambiguous sentences. *Journal of Speech and Hearing Research* 24: 64-69.

Van Lancker, D. and Kemper, D. 1987. Comprehension of familiar phrases by left- but not right-hemisphere damaged patients. *Brain and Language* 32: 265-277.

Van Lancker, D., Canter, G.J. and Terbeek, D. 1981. Disambiguation of ditropic sentences: acoustic and phonetic cues. *Journal of Speech and Hearing Research* 24: 330-335.

Vega Moreno, R.E. 2004. Metaphor interpretation and emergence. *UCL Working Papers in Linguistics 16: 297-322.*

Vega Moreno, R.E. Forthcoming. Idioms, transparency and pragmatic inference. *UCL Working Papers in Linguistics 17*

Vicente, B. 1992. Metaphor, meaning and comprehension. *Pragmatics* 2: 49-62.

Vicente, B. 1993. En torno a la semántica y la pragmática de la metáfora: una revisión crítica. *Pragmalingüística* 1: 307-334.

Vicente, B. 1995. *Mecanismos Semántico-Pragmáticos en el Análisis de la Metáfora.* Universidad del País Vasco, Servicio de Publicaciones, País Vasco.

Vicente, B. 1996. On the semantics and pragmatics of metaphor: coming full circle. *Language and Literature* 5: 195-208.

Ward, T., Smith, S. and Vaid, J. (eds.) 1997. *Creative Thought: an Investigation of Conceptual Structure and Processes.* American Psychological Association, Washington, DC.

Wasow, T., Sag, I. and Nunberg, G. 1983. Idioms: an interim report. In S. Hattori, K. Inoue (eds.) *Proceedings of the XIIIth International Congress of Linguistics,* Tokyo.

Weinreich, U. 1969. Problems in the analysis of idioms. In J. Puhvel (ed.) *Substance and Structure of language.* University of California Press, Berkley.

Wilson, D. 1995. Is there a maxim of truthfulness? *UCL Working Papers in Linguistics* 7: 197-212.

Wilson, D. 1998. Linguistic structure and inferential communication. In B. Caron (ed.) *Proceedings of the 16th International Congress of Linguists.* Elsevier Sciences, Oxford.

Wilson, D. 2000. Metarepresentation in linguistic communication. In D. Sperber. (ed.) *Metarepresentations.* 411-448. Oxford University Press, Oxford.

Wilson, D. 2002. Relevance Theory: from the basics to the cutting edge. *ICU Open lectures on cognitive pragmatics.* ICU, Japan.

Wilson, D. 2003. New directions for research on pragmatics and modularity. *UCL Working Papers in Linguistics* 15: 105-127. Reprinted in 1995. *Lingua* 115: 1129- 1146.

Wilson, D. 2004. Relevance and lexical pragmatics. *UCL Working Papers in Linguistics* 16: 343-360. Printed also in 2003. *Italian Journal of Linguistics/Rivista di Linguistica*, 15(2): 273-291. Special Issue on Pragmatics and the Lexicon.

Wilson, D. and Sperber, D. 1981. On Grice's theory of conversation. In P. Werth (ed.) *Conversation and Discourse.* 155-178. Croom Helm, London.

Wilson, D. and Sperber, D. 1985. On choosing the context for utterance interpretation. In J. Allwood and E. Hjelmquist (ed.) *Foregrounding Background.* 51-64. Bokforlaget Doxa AB, Lund, Sweden.

Wilson, D. and Sperber, D. 1993a. Linguistic form and relevance. *Lingua* 90: 1-25.

Wilson, D. and Sperber, D. 1993b. Pragmatics and time. *UCL Working Papers in Linguistics* 5: 277-298. Reprinted in R. Carston and S. Uchida (eds) 1998. *Relevance Theory: Applications and Implications.* John Benjamins: 1-22

Wilson, D. and Sperber, D. 2000. Truthfulness and relevance. *UCL Working Papers in Linguistics* 12: 215-254. Reprinted in 2002. *Mind* 111(443): 583-632.

Wilson, D. and Sperber, D. 2002. Relevance Theory. *UCL Working Papers in Linguistics* 14: 249-290. A version of this paper was reprinted in G. Ward and L. Horn (eds.) 2004. *Handbook of Pragmatics*. 607-632. Blackwell, Oxford.

Wolff. P. and Gentner, D. 2000. Evidence for the role-neutral initial processing of metaphor. *Journal of Experimental Psychology: Learning, Memory and Cognition* 29: 529-541.

Wray, Alison. 2002. *Formulaic Language and the Lexicon.* Cambridge University Press, Cambridge.

Zeldin, T. 1998. *Conversation: How Talk Can Change Your Life.* Harvill Press, London.